

A person wearing a high-visibility orange and yellow safety vest is kneeling on a grey carpeted floor, operating a mobile robot. The mobile robot has a black base with two large, treaded wheels wrapped in clear plastic. A laptop is mounted on the robot, displaying a blue interface. A VR headset is mounted on a vertical pole above the robot. In the foreground, a small, black, four-legged quadruped robot is visible. The background shows a white wall and a white cabinet.

Proceedings of the 41th International Symposium on Automation and Robotics in Construction

Lille, France, June 3-5, 2024

Proceedings of the 41th International Symposium on Automation and Robotics in Construction

ISSN (for the proceedings series): 2413-5844

ISBN (for this issue of the proceeding series): 978-0-6458322-1-1.

The proceeding series is Scopus indexed.

Scopus[®]

All papers are available in the IAARC Website and presentations in the IAARC YouTube Channel



Copyrights reserved.

© 2024 International Association on Automation and Robotics in Construction

This work including all its parts is protected by copyright. Any use outside the narrow limits of copyright law without the consent of the individual authors is inadmissible and punishable. This applies in particular to reproductions, translations, microfilming and saving and processing in electronic systems.

The reproduction of common names, trade names, trade names etc. in this work does not justify the assumption that such names are to be regarded as free within the meaning of the trademark and trademark protection legislation and can therefore be used by everyone, even without special identification.

Cover design: Zhong Wang, Muhammed Adil, Vicente A. Gonzalez; Image: Robotic and XR systems for remote and autonomous inspection in construction developed by the Infrastructure Human Tech (IHT) Lab at the University of Alberta (www.iht-lab.com)

Editorial Board

Editors in Chief

Gonzalez, Vicente A. (University of Alberta)
Zhang, Jiansong (Purdue University)
García de Soto, Borja (New York University Abu Dhabi)
Brilakis, Ioannis (University of Cambridge)

Editors (Area Chairs)

Al Balkhy, Wassim (Ecole Centrale de Lille)
Brosque, Cynthia (Stanford University)
Bugalia, Nikhil (Indian Institute of Technology Madras)
Cai, Jiannan (The University of Texas at San Antonio)
Chen, Qian (The University of British Columbia)
Ducoulombier, Laure (Bouygues Construction)
Gan, Vincent (National University of Singapore)
Gheisari, Masoud (University of Florida)
Herrera, Rodrigo F. (Pontificia Universidad Católica de Valparaíso)
Hu, Rongbo (Kajima Corporation)
Hu, Yuqing (The Pennsylvania State University)
Lee, Gaang (University of Alberta)
Isaac, Shabtai (Ben-Gurion University of the Negev)
Iturralde, Kepa (Technical University of Munich)
Liang, Ci-Jyun (Stony Brook University)
Mei, Qipei (University of Alberta)
Mantha, Ram Kumar (University of Sharjah)
Martinez, Eder (University of Applied Sciences and Arts Northwestern Switzerland)
Murguia, Danny (University of Cambridge)
Ng, Ming Shan (Kyoto Institute of Technology)
Wu, Lingzi (University of Washington)
Zhu, Zhenhua (University of Wisconsin-Madison)
Zou, Yang (University of Auckland)

Local Organizing Committee

Lafhaj, Zoubeir (Chair)
Al Balkhy, Wassim
Rebai, Slim
Boutabba, Assia
Hmana, Yasmine
Zerrari, Reda
Linner, Thomas
M.S.Ng, Charmaine
Hu, Rongbo
Karmaoui, Dorra
Bogucka, Agnieszka
Ghattassi, Imen
El Babidi, Saad
Ayadi, Syrine
Coveli, Anita

Sponsors:



TEAMOTY



Collaboration:



Reviewers

Abdelmegid, Mohammed
Adán, Antonio
Afsari, Kereshmeh
Agrawal, Ajay Kumar
AlBalkhy, Wassim
Albeaino, Gilles
Alberdi Pagola, Pablo
Alduais, Mohammed
Alotaibi, Emran
Alwisy, Aladdin
Ambrosino, Michele
Amoudi, Omar
Anandh, K.S.
Ang, Karyne
Ariyachandra, Mahendrini
Armeni, Iro
Asmone, Ashan
Atencio Castillo, Edison
Bae, Juhyeon
Baek, Francis
Bartels, Niels
Behne, Ann-Kathrin
Berti, Nicola
Bhadaniya, Parth
Bielski, Jessica
Blankenbach, Jörg
Borrmann, Andre
Bosché, Frédéric
BouHatoum, Makram
Brännström, Mattias
Brosque, Cynthia
Bruttini, Alessandro
Bugalia, Nikhil
Bulgakov, Alexey
Buuveibaatar, Munkhbaatar

Cao, Jianpeng
Carbonari, Alessandro
Caro, Stéphane
Chang, Yuntsui
Chen, James
Chen, Jingdao
Chen, Qian
Chen, Weiwei
Cheng, Jiun-Yao
Cisterna, Diego
Collins, Fiona
Cooperman, Alissa
Correa, Fabiano
Danel, Thomas
Devkota, Ajit
Ding, Shiqi
Dong, Beixuan
Ducoulombier, Laure
Eiris, Ricardo
ElMenshawy, Mohamed
Eskudero, Ibon
Esser, Sebastian
Fan, Chao
Fawad, Muhammad
Follini, Camilla
Forcael, Eric
Furet, Benoît
Gambao, Ernesto
Gan, Vincent
Gao, Maggie Y
Gao, Yuezhen
Garcia-Lopez, Nelly
Gloser, Franz-Ferdinand
Görsch, Christopher
Guidolin, Mattia

Guo, Brian
Guo, Jingjing
Hackl, Jürgen
Hall, Daniel
Hammad, Amin
Hartmann, Timo
Hosny, Fatma
Hu, Difeng
Hu, Rongbo
Hu, Yuqing
Huang, Lei
Huang, Yaojing
Ilhan Jones, Bahriye
Isaac, Shabtai
Iturralde, Kepa
Jäkel, Jan-Iwo
Jeelani, Idris
Jung, Wooyoung
Kalasapudi, Vamsi
Karamara, Merve
Kazemian, Ali
Kim, Hongjo
Kim, Hyoungkwan
Kim, Jinwoo
Kim, Seongyong
König, Markus
Lauble, Svenja
Le, Thai-Hoa
Lee, Hoonyong
Li, Mingkai
Lin, Jacob
Lin, Jing
Lin, Zhiyang
Liu, Canlong
Liu, Kexin

Reviewers

Liu, Xiaolong	Prieto, Samuel	Wang, Tao
Louis, Joseph	Quan, Guan	Wang, Xi
Lozano Galant, José	Ramon-Constanti, Amanda	Wang, Xin
Ma, JongWon	Raphael, Benny	Wei, Chialing
Ma, Jun	Rausch, Christopher	Wen, Jing
Mao, Zeyu	Rebai, Slim	Wen, Leyang
Maqsoodi, Aras	Reja, Varun	Weng, Yiwei
Martinez, Eder	Reynoso, Hamlet	Williams-Riquer, Francisco
Martinez, Jhonattan	Ryu, JuHyeong	Wong Chong, Oscar
Mengiste, Eyob	Sabek, Mohamed	Wrona, Józef
Miller, Seirgei Rosario	Saffert, Anne-Sophie	Xiao, Bo
Mohamed, Yasser	Sanchez Andrade, Benjamin	Xu, Xinghui
Mostafa, Kareem	Sánchez Rivera, Omar	Yang, Cheng-Hsuan
Muñoz La Rivera, Felipe	Sankaran, Bharath	Yang, Tao
Murguia, Danny	Sati, Ala	Yang, Xiaofei
Napps, Daniel	Schmailzl, Marc	Yantao, Yu
Ng, Ming Shan	Schneider, Oliver	Yao, Dongchi
Nornes, Stein	Senthilnathan, Shanmugaraj	Ye, Jun
Noroozinejad Farsangi, Ehsan	Shamsollahi, Dena	Ye, Yang
Núñez Varillas, Christopher	Shen, Zixin	Yeung, Timson
Olender, Margarete	Shi, Zhuoya	Yi, Wen
Osorio Gómez, Cristian	Shohet, Igal	Yin, Xianfei
Ouyang, Zhicheng	Sohn, John	Yoon, Jong Han
Paes, Daniel	Soman, Ranjith	Yu, Hongrui
Pal, Aritra	Sonkor, Muammer	Yu, Jingshuo
Pan, Mi	Soonwook, Kwon	Zhang, Cheng
Paquet, Elodie	Sun, Yuan	Zhang, Jiansong
Park, Somin	Talebi Kalaleh, Mohammad	Zhao, Charlie
Pauwels, Pieter	Tan, Tan	Zhao, Xianxiang
Pellegrino, Erika	Toca Perez, Cristina	Zhu, Aimin
Penrooz, Matias	Turkan, Yelda	Zhu, Qi
Pfister, Louis	Uribe, Adolfo	Zhu, Zhenhua
Pinheiro Santos, Pedro	Urlainis, Alon	Zhu, Zixian
Pooladvand, Shiva	Vielma, Juan	Zhuang, Dian
Poshdar, Mani	Walzer, Alexander	Zuo, Hui

Foreword

The International Association for Automation and Robotics in Construction (IAARC) and the Local Organizing Committee are pleased to present the Proceedings of the 41st International Symposium on Automation and Robotics in Construction (ISARC 2024). This event was held from June 3-5, 2024, in Lille, France, and was hosted by École Centrale de Lille. The 41st edition of ISARC marks a significant milestone in its maturity with great participation (a total of 219 papers submitted). After a rigorous peer-review process, 174 papers (81% of the submissions) were finally accepted and included in the proceedings. In total, 431 authors from 130 universities and 11 private/public research organizations and firms across 29 countries, including the Americas, Europe, Asia, Oceania, and the Middle East, submitted their work that was assessed in a two-step peer-review process, which included a rebuttal phase. The submission and review processes were supported by 23 Area Chairs who covered the eight typical technical areas of interest within IAARC, including "Robotization of Renovation in Construction," a theme proposed by the Local Organizers.

The review process involved 210 reviewers. Unlike traditional practices, submission of full papers was required without prior abstract submission, which made the revision cycles more effective. The peer review was conducted in a single-blinded manner. One hundred eighty three papers (84.7%) received two reviews, while some required additional assessments before a decision could be made: 32 papers (14.8%) had three reviews, and one paper (0.5%) underwent four reviews. This approach helped address situations where consensus was not initially reached, requiring further reviews. During the review process, the rebuttal phase has now become a part of the curation practices for the proceedings at ISARC. The standards established in this phase continue to add value for authors, enabling an ongoing process of improvement in the quality of the papers accepted for the proceedings. The peer-review process enabled the selection of high-quality submissions, incorporating both traditional scientific and short papers into the proceedings. As for presentations, the symposium featured keynote, plenary, parallel and poster sessions.

ISARC stands as the leading global conference in the field of automation and robotics in construction. To maintain its prestigious status, the IAARC Technical Committee and Area Chairs have ensured that high-quality papers are accepted. Building on the work of previous Technical Committees, who have developed editorial standards, guidelines, and a knowledge base, the current Technical Committee has introduced an automated process for generating the table of contents and proceedings. This innovation will benefit future ISARC events.

We hope you find this year's proceedings engaging and the papers included particularly stimulating. Enjoy your reading!

Vicente A. Gonzalez
Jiansong Zhang
Borja García de Soto
Ioannis Brilakis

Table of Contents

K - Keynote talks

CrackGauGAN: Semantic Layout-based Crack Image Synthesis for Automated Crack Segmentation.....	1
--	---

Honghu Chu, Weiwei Chen, Lu Deng

Safe Operations of Construction Robots on Human-Robot Collaborative Construction Sites.....	9
---	---

Marvin H. Cheng, Ci-Jyun Liang, Elena G. Dominguez

Pre-trained language model based method for building information model to building energy model transformation at metamodel level.....	17
--	----

Zhichen Wang, Mario Bergés, Burcu Akinci

P - Plenary talks

Computer Vision as Key to an Automated Concrete Production Control.....	26
---	----

Max Coenen, Maximilian Meyer, Dries Beyer, Christian Heipke, Michael Haist

Investigating the Neural Correlates of Different Levels of Situation Awareness and Work Experience.....	34
---	----

Yanfang Luo, JoonOh Seo, Sogand Hasanzadeh

Inline image-based reinforcement detection for concrete additive manufacturing processes using a convolutional neural network	42
---	----

Lukas Johann Lachmayer, Lars Dittrich, Robin Dörrie, Harald Kloft, Annika Raatz, Tobias Recker

Partial Personalization for Worker-robot Trust Prediction in the Future Construction Environment.....	49
---	----

Woei-Chyi Chang, Nestor Gonzalez Garcia, Behzad Esmaeili, Sogand Hasanzadeh

Automatic Quality Inspection of Rebar Spacing Using Vision-Based Deep Learning with RGBD Camera.....	57
--	----

Songyue Wang, Lu Deng, Jingjing Guo, Mi Liu, Ran Cao

Optimization of prefabricated component installation using a real-time evaluator (RTE) connection locating system	65
---	----

Nolan W Hayes, Bryan P Maldonado, Mengjia Tang, Peter Wang, Diana Hun

Effects of Visual Prompts in Human-machine Interface for Construction Teleoperation System.....	73
---	----

Yeon Chae, Samraat Gupta, Youngjib Ham

Zero-shot Learning-based Polygon Mask Generation for Construction Objects	81
Taegeon Kim, Minkyu Koo, Jeongho Hyeon, Hongjo Kim	
Integration of laser profiler feedback into FIM for additive manufacturing in construction	89
Simon Espinosa, Martin Slepicka, André Borrmann	
A - Automated/robotic machines, devices, and end-effectors	
Shape Control of Guide Frame to Avoid Obstacles for Tunnel Inspection System by Manual Operation	97
Fumihiro Inoue, Momoe Terata	
Component pose reconstruction using a single robotic total station for panelized building envelopes	105
Mengjia Tang, Nolan W. Hayes, Bryan P. Maldonado, Diana Hun	
Assessment of Traditional and Robotic Approaches to Interior Construction Layout: A Framework and Comparative Study	113
Catherine Caputo, Ashtarout Ammar, Ashley Johnson	
Development of automated transport system	121
Shunsuke Igarashi, Yuji Kinoshita, Taku Tani, Takayoshi Hachijo, Masahiro Indo	
Automated As-built 3D Reconstruction Using Quadraped Robot and 3D LiDAR Sensor	129
Vincent J.L. Gan, Difeng Hu, Ruoming Zhai, Yushuo Wang	
Enhancing UAV Efficiency in Bridge Inspection through BIM-based Flight Planning and Image Quality Assurance	136
Feng Wang, Yang Zou, Enrique del Rey Castillo, James B.P. Lim	
MPC-Based Proactive Swing Attenuation for Double-Pendulum Overhead Cranes....	144
Juhao Su, Shichen Sun, Ching-Wei Chang, Siwei Chang	
Reinvent reinforced concrete with robotics and 3D printing.....	152
Jean-François Caron, Nicolas Ducoulombier, Léo Demont, Victor de Bono, Romain Mesnil	
Unified framework for mixed-reality assisted situational adaptive robotic path planning enabled by 5G networks for deconstruction tasks	160
Chu Han Wu, Marit Zöcklein, Sigrid Brell-Cokcan	
Fast-Pixel-Matching Algorithm for Automated Shear Stud Welding Based on the Integration of 2D Drawings and Structured Light Cameras	168
Huiguang Wang, Lu Deng, Ran Cao, Jingjing Guo	
Implementation of a Robotic Manipulator End effector for Construction Automation ...	177
Aarón Borgo, Arturo Ruiz, Jiansong Zhang, Luis C Félix-Herrán	
Low center-of-gravity movement robotic arm with kinematic optimization algorithm for on-site construction	185

Ruiqi Jiang, Xiao Li

System integration of construction planning and robots for a joint civil engineering and robotics course..... 193

Ryosuke Yajima, Fumiya Matsushita, Keiji Nagatani, Kazumasa Ozawa

Development of Autonomous Robots for Construction..... 199

Takayoshi Hachijo, Shunsuke Igarashi, Taku Tani, Masahiro Indo

Autonomous Data Acquisition of Ground Penetrating Radar (GPR) Using LiDAR-based Mobile Robot..... 206

Amr Eldemiry, Muhammad Muddassir, Tarek Zayed

Double-Pendulum Tower Crane Trolley Trajectory Planning: A Parameter Based Time-Polynomial Optimization..... 213

Shichen Sun, Juhao Su, Yu Hin NG, Ching-Wei Chang

Advancements and Future Visions in Earthmoving Swarm Technology..... 220

Mikko Sakari Hiltunen, Rauno Kalle Heikkilä

A construction robot path planning method based on safe space and worker trajectory prediction 227

Xiaotian Ye, Hongling Guo, Ziyang Guo, Zhubang Luo

Factors Leading to Reduced Construction Productivity in Unmanned Construction.... 236

Genki Yamauchi, Takeshi Hashimoto, Mitsuru Yamada, Shinichi Yuta

A Framework and Cyber-Physical System Architecture for Cloud-Based Construction Monitoring with Autonomous Quadrupeds 243

Aras Maqsoodi, Javier Irizarry

Integrating Automation into Construction Site: A System Approach for the Brick Cutting Use Case 251

Cinzia Slongo, Vincenzo Orlando, Dietmar Siegele, Dominik T. Matt

C - Construction management techniques

Case study on the applicability of a construction site layout planning system 259

Jiyu Shin, Jongwoo Cho, Tae Wan Kim

Discrete Event Simulation to Predict Construction Equipment Emissions on a Digital Twin Platform 267

Kepeng Hong, Alexandros Apostolidis, Jochen Teizer

Performance Evaluation of Genetic Algorithm and Particle Swarm Optimization in Off-Site Construction Scheduling 275

Mizanoor Rahman, Sang Hyeok Han

Developing a novel application to digitalize and optimize construction operations using low-code technology..... 283

Eder Martinez, Louis Pfister, Felix Stauch

GPT-based Logic Reasoning for Hazard Identification in Construction Site using CCTV Data	291
Dai Quoc Tran, Yuntae Jeon, Minsoo Park, Seunghee Park	
Adopting Automation in Premanufacturing: A Two-Mode Network Analysis on Factors and Roles in Iran and North America's Construction Industry	299
Amirhossein Babaei Ravandi, Tamer El-Diraby	
Assessing the Financial Dynamics of Modular Offsite Building Projects	307
Hisham Said, Tarek Salama, Anthony Gude	
Impacts of Multi-skills on Project Productivity and Completion Time: A Building Renovation Case Study	315
Muhammad Atiq Ur Rehman, Sharfuddin Ahmed Khan, Amin Chaabane	
Integrating the AR-QR Code Approach with Positioning and Orientation Sensors to Facilitate Deploying Drawings Information on Job Sites	323
Mohsen Foroughi Sabzevar, Masoud Gheisari, L.James Lo	
Virtual Reality-based Blockchain Application for Optimized Collaborative Decisions of Modular Construction	331
Mohamed Assaf, Rafik Lemouchi, Mohamed Al-Hussein, Xinming Li	
The Role of Large Language Models for Decision Support in Fire Safety Planning	339
Dilan Durmus, Alberto Giretti, Ori Ashkenazi, Alessandro Carbonari, Shabtai Isaac	
Understanding Professional Perspectives about AI Adoption in the Construction Industry: A Survey in Germany	347
Diego Cisterna, Franz-Ferdinand Gloser, Eder Martínez, Svenja Lauble	
Enabling Construction Automation: Implementing Radio Frequency Communication Infrastructure on Construction Sites	355
Victoria Jung, Sigrid Brell-Cokcan	
State of the Art Review of Technological Advancements for Safe Tower Crane Operation	364
Avi Raj, Jochen Teizer	
A Delay Liquidated Damages (DLD) Mitigation Model based on Earned Schedule (ES) and Earned Value Management System (EVMS) concepts	374
So-Won Choi, Eul-Bum Lee	
Impact of Integrated Supply Chain Platforms on Construction Project Management ..	381
Giovanni Zenezini, Giulio Mangano, Gabriel Castelblanco	
Automated Construction Site Safety Monitoring Using Preidentified Static and Dynamic Hazard Zones	388
Kepeng Hong, Jochen Teizer	
Innovative standardized cost data structure: application on price list document for estimating public tendering	396
Jacopo Cassandro, Chiara Gatto, Antonio Farina, Claudio Mirarchi, Alberto Pavan	

A Proposed Framework to Implement Advanced Work Packaging (AWP) with the Support of Blockchain..... 404

Slim Rebai, Zoubeir Lafhaj, Olfa Hamdi, Wassim AlBalkhy, Rateb Jabbar, Hamdi Ayeche, Ahmed Moukhafi, Mohammed Chadli, Pascal Yin

Construction worker fatigue load management using IoT heart rate sensing system.. 413

LunWang Wu, HuiPing Tserng, XiuZhen Huang

Evaluation of effect of different safety training styles on mental workloads using electroencephalography..... 420

Peng Liu, Chengyi Zhang

SafeSense: Real-time Safety Alerts for Construction workers..... 428

Sivakumar K S, Vishnu Sivan, Grigary Antony

Towards Automation in 5s – Classification and Detection of Images from Construction Sites..... 436

Viral Makwana, Ganesh Devkar

Analysis of XES and OCEL Data Schemas: Towards Multidimensional Process Mining of Intertwined Construction Processes..... 444

Araham Jesus Martinez Lagunas, Mazdak Nik-Bakht

Comparative Analysis of Cognitive Agreement between Human Analysts and Generative AI in Construction Safety Risk Assessment 452

Unmesa Ray, Cristian Arteaga, Jee Woong Park

Developing Time-Cost & Storage Optimization Model for a Construction Project – A Linear Programming Approach 459

Prasanna Venkatesan Ramani, Shivaram Kandasamy, Anshul Gupta, Sugeerthi MS

Prototype development of an automated 3D shop drawing generation tool for reinforcement work 467

Kanae Miyaoka, Fumiya Matsushita, Kazumasa Ozawa

The Effect of Material Handling Strategies on Time and Labour Fatigue in Window Manufacturing 474

Omar Azakir, Fatima Alsakka, Mohomad Al-Hussein

Impact of Technology Use on Time Needed for Information Retrieval for Frontline Supervisors in the Construction Industry..... 482

Bassam Ramadan, Hala Nassereddine, Timothy Taylor, Paul Goodrum

Digital Info Screen - A Visual Management Tool for Construction Workers..... 489

Markku Kiviniemi, V. Paul C. Charlesraj, Marja Liinasuo, Pekka Siltanen, Kari Rainio

H - Human factors & human-system collaboration

Assessing the Impact of AR-Assisted Warnings on Roadway Worker Stress Under Different Workload Conditions..... 497

Fatemeh Banani Ardecani, Amit Kumar, Omidreza Shoghli

Worker-Centric Path Planning: Simulating Hazardous Energy to Control Construction Safety Using Graph Theory.....	505
Kilian Speiser, Jochen Teizer	
Experimental Evaluation of Exoskeletons for Rebar Workers Using a Realistic Controlled Test.....	513
Malcolm Dunson-Todd, Mazdak Nik-Bakht, Amin Hammad	
Digital Twin-based Instructor Support System for Excavator Training.....	521
Faridaddin Vahdatikhaki, Leon olde Scholtenhuis, Andre Doree	
Fit Islands: Designing a Multifunctional Exergame to Promote Healthy Aging for Chinese Older Adults	529
Zixin Shen, Rongbo Hu, Dong Wan, Ami Ogawa, Thomas Bock	
Use of robotics to coordinate health and safety on the construction site	537
Fabian Dornik, Niels Bartels	
The Influence of an Immersive Multisensory Virtual Reality System with Integrated Thermal and Scent Devices on Individuals' Emotional Responses in an Evacuation Experiment	546
Xiaolu Xia, Nan Li, Jin Zhang	
Envisioning Extraterrestrial Construction and the Future Construction Workforce: A Collective Perspective.....	553
Amirhosein Jafari, Yimin Zhu, Suniti Karunatilake, Jennifer Qian, Shinhee Jeong, Ali Kazemian, Andrew Webb	
Trends, Challenges, and Opportunities in Assistive and Robotic Kitchen Technologies for Aging Society: A Scoping Review	561
Rongbo Hu, Ami Ogawa, Borja García de Soto, Thomas Bock	
Guiding Visual Attention in Remote Operation: Meaning, Task and Object in Post-Disaster Scenarios	569
Xiaoxuan Zhou, Xing Su	
Conceptualizing Digital Twins in Construction Projects as Socio-Technical Systems .	577
Mohammed Abdelmegid, Algan Tezel, Carlos Osorio Sandoval, Zigeng Fang, William Collinge	
Coupled Risk Assessment of Construction Workers' Unsafe Behaviors in Human-Robot Interactions	585
Jing Lin, Qinyuan Li, Runhe Zhu, Yuhang Cai, Longhui You, Lewen Zou, Siyan Liu	
Enhancing Public Engagement in Sustainable Systems through Augmented Reality .	593
Yu-Chen Lee, Abdullah Alsuhaibani, Chih-Shen Cheng, Fernanda Leite	
Virtual Construction Safety Training System: How Does it Relate to and Affect Users' Emotions?	599
Fan Yang, Jiansong Zhang	
REBAPose -A Computer vision based Musculoskeletal Disorder Risk Assessment Framework.....	607
Sivakumar K S, Nikhil Bugalia, Benny Raphael	

Developing Joint-level Scoring Models Tailored to Whole-body Ergonomic Assessment of Construction Workers.....	615
--	-----

Zirui Li, Yantao Yu, Qiming Li

Enhancing Human-Robot Teaming in Construction through the Integration of Virtual Reality-Based Training in Human-Robot Collaboration.	623
--	-----

Adetayo Onososen, Innocent Musonda, Christopher Dzuwa

Interpretation Conflict in Helmet Recognition under Adversarial Attack.....	631
---	-----

He Wen, Simaan AbouRizk

Enhancing Decision-Making for Human-Centered Construction Robotics: A Methodological Framework	637
--	-----

Marc Schmailzl, Anne-Sophie Saffert, Merve Karamara, Thomas Linner, Friedrich Anton Eder, Simon Konrad Hoeng, Mathias Obergriesser

I - Information modeling techniques

Interpretable Machine Learning Approaches for Assessing Maximum Force in Fiber-Reinforced Composites	645
--	-----

Soheila Kookalani, Erika Parn, Ioannis Brilakis

An OpenBIM-based Framework for Lifecycle Management of Steel Reinforcement ...	653
--	-----

Mingkai Li, Yuhan Liu, Boyu Wang, Jack Cheng

A Digital Twin Based Approach to Control Overgrowth of Roadside Vegetation	661
--	-----

Varun Kumar Reja, Diana Davletshina, Mengtian Yin, Ran Wei, Quentin Felix Adam, Ioannis Brilakis, Federico Perrotta

Bridging the Annotation Gap: Innovating Sewer Defects Detection with Weakly Supervised Object Localization.....	669
---	-----

Jianyu Yin, Xianfei Yin, Yifeng Sun, Mi Pan

Extracting roof sub-components from orthophotos using deep-learning -based semantic segmentation	675
--	-----

Jiajun Li, Boan Tao, Frédéric Bosché, Chris Xiaoxuan Lu, Lyn Wilson

A Step from Virtual to Reality: Investigating the Potential of a Diffusion-Based Pipeline for Enhancing the Realism in Synthetic Construction Imagery	683
---	-----

Sina Davari, Ali TohidiFar, Daeho Kim

Integration of Construction Progress Monitoring Results using AI Image Recognition from Multiple Cameras onto a BIM	691
---	-----

Chang-Cheng Hsieh, Hung-Ming Chen, Wan-Yu Chen, Ting-Yu Wu

Transformer-based Pavement Crack Tracking with Neural-PID Controller on Vision-guided Robot	699
---	-----

Jianqi Zhang, Xu Yang, Wei Wang, Ioannis Brilakis, Hainian Wang, Ling Ding

Toward shotcrete process simulation to support robotic operation.....	707
---	-----

Mohammad Reza Yazdi Samadi, Ralf Waspe, Christian Schlette

A Framework of Integrating HBIM and GIS for Automated Fire Risk Assessment of Heritage Buildings	714
Yutong Qiao, Chi Chiu Lam, Mun On Wong, Yujin Xu	
3D reconstruction of a bridge with concrete damage classification using deep learning	722
Christopher Joseph Núñez Varillas, Marck Steewar Regalado Espinoza, Luis Mario Huaypar Acurio, Antonio Stefano Bedon Rosario, Jordan Antony Romani Chavez, Oscar Manuel Solis Garcia, Karol Maricruz Agreda Estela, Micaela Anthoaneth Cardenas Contreras	
Point Cloud-based Computer Vision Framework for Detecting Proximity of Trees to Power Distribution Lines	730
Fardin Bahreini, Amin Hammad, Mazdak Nik-Bakht	
Reinforcement Learning for Smart Mobile Factory Operation in Linear Infrastructure projects	738
Jianpeng Cao, Irfan Čustović, Ranjith Soman, Daniel Hall	
Semantic annotation of images from outdoor construction sites	745
Layan Farahat, Ehsan Rezazadeh Azar	
Structural design in the era of digital twins. A case study	753
Hector Posada, Rolando Chacón, Carlos Ramonell	
Removal of Construction Machinery Occlusion using Image Segmentation and Inpainting for Automated Progress Tracking.....	759
Ahmet Bahaddin Ersoz, Onur Pekcan	
Context-Adaptive CCTV Pan-Tilt-Zoom method for Personal Protective Equipment Detection.....	768
Seokhwan kim, Minwoo Jeong, Minkyu Koo, Taegeon Kim, Hongjo Kim	
Smart Passive Mixed Reality-Based Construction Inspection Framework	776
Boan Tao, Jiajun Li, Frédéric Bosché	
Unveiling Building Façade Deterioration: A Drone-Powered Deep Learning Approach for Seamless Tile Peeling Detection.....	784
Ngoc-Mai Nguyen, Minh-Tu Cao, Wei-Chih Wang	
Accelerating Indoor Construction Progress Monitoring with Synthetic Data-Powered Deep Learning	792
Mathis Baubriaud, Stéphane Derrode, René Chalon, Kevin Kernn	
Single-Stage Spatiotemporal Activity Recognition of Excavators: A Case Study	800
Ali Ghelmani, Ghazaleh Torabi, Amin Hammad, Chen Chen	
Development of an immersive digital twin framework to support infrastructure management: a case study of bridge asset health monitoring.....	808
Muhammad Fawad, Marek Salamak, Qian Chen, Kalman Koris	
A Graph Neural Network Approach to Conceptual Cost Estimation	815
Hao Liu, Jack C.P. Cheng, Chimay J. Anumba	

Automatic Assessment of Leanness of Construction Contract Provisions: A Proof-of-Concept	822
Ramesh Balaji, Sivakumar KS, Murali Jagannathan, Venkata Santosh Kumar Delhi	
BIM and IPA - Excerpt of an automated assessment system for an autodidactic teaching concept	830
Christian Heins, Gregor Grunwald	
Flat and Level Analysis Tool (FLAT) for real-time automated segmentation and analysis of concrete slab point clouds.....	838
Nolan W Hayes, Bryan P Maldonado, Mengjia Tang, Diana Hun	
Evaluation of Mapping Computer Vision Segmentation from Reality Capture to Schedule Activities for Construction Monitoring in the Absence of Detailed BIM	847
Juan D. Nunez-Morales, Yoonhwa Jung, Mani Golparvar-Fard	
Automating Construction Safety Inspections using Robots and Unsupervised Deep Domain Adaptation by Backpropagation	855
Vimal Velusamy Bharathi, Samuel Prieto, Borja Garcia de Soto, Jochen Teizer	
From Unstructured Data to Knowledge Graphs: An Application for Compliance Checking Problem	863
Ankan Karmakar, Chintan Patel, Venkata Santosh Kumar Delhi	
Towards Automated Physics-Based Modeling: Fusion of Construction Equipment Data for Efficient Simulation	872
Liqun Xu, Dharmaraj Veeramani, Zhenhua Zhu	
Reinforcement Learning-Enhanced Path Planning for Mobile Cranes in Dynamic Construction Environments: A Virtual Reality Simulation Approach	880
Rafik Lemouchi, Mohamed Assaf, Ahmed Bouferguene, Mohamed Al-hussein	
Automated Decision Support System Based on Quantification of Defective Tubular Steel Temporary Materials for Quality Circles	888
James Mugo Njoroge, Sejoon Lee, Kyuhyup Lee, Junsung Seol, Younghee Chang, Soonwook Kwon	
Towards autonomous shotcrete construction: semantic 3D reconstruction for concrete deposition using stereo vision and deep learning.....	896
Patrick Schmidt, Dimitrios Katsatos, Dimitrios Alexiou, Ioannis Kostavelis, Dimitrios Giakoumis, Dimitrios Tzovaras, Lazaros Nalpantidis	
Effect of Missing Data on Machine-Learning Algorithms for Real-time Safety Monitoring in Scaffolds	904
Laura Alvarez, Mahendra Ghimire, Jee Woong Park	
Digital Twin in Prefabricated Construction – Approaches, Challenges and Requirements	912
Veerakumar Rangasamy, Jyh-Bin Yang	
Exploring Digital Twin platforms across industries	920
Amin Khoshkenar, Hala Nassereddine	

Exploring Self-Supervised GPR Representation Learning for Building Rooftop Diagnostics	928
Kevin Lee, Wei-Heng Lin, Bilal Sher, Talha Javed, Sruti Madhusudhan, Chen Feng	
Underground Utility Network Completion based on Spatial Contextual Information of Ground Facilities and Utility Anchor Points using Graph Neural Networks	936
Yuxi Zhang, Hubo Cai	
Adaptive Geological Uncertainty Modeling in Excavation	944
Hannu Juola, Sara Johansson, Olof Friberg, Rauno Heikkilä	
Planning of Formwork Accessories in a BIM-based Aluminum Formwork Layout Planning System	951
Guan-Yong Xiong, Tzong-Hann Wu, Shang-Hsien Hsieh, Kuan-Yi Chen, Budy Setiawan, Wen-Tung Chang	
Development of 3D Digital-Twin System for Remote-Control Operation in Construction	958
Yoshihito Mori, Masaomi Wada	
Exploring the Potential of Reinforcement Learning in Pipe Spool Scheduling in Industrial Projects	965
Mohamed ElMenshawy, Lingzi Wu, Brian Gue, Simaan AbouRizk	
Current State and Trends of Point Cloud Segmentation in Construction Research	972
Samuel A. Prieto, Eyob T. Mengiste, Uday Menon, Borja Garcia de Soto	
Monitoring of 3D concrete printing quality through multi-view RGB-D images	980
Ahmed Magdy Ahmed Zaki, Marco Carnevale, Hermes Giberti, Christian Schlette	
Robust High-Precision LiDAR Localization in Construction Environments	988
Andrew Alexander Yarovoi, Pengyu Mo, Yong Kwon Cho	
Buildability assessment of 3D printed concrete elements through computer vision	995
Shanmugaraj Senthilnathan, Benny Raphael	
Automated Inspection Report Generation Using Multimodal Large Language Models and Set-of-Mark Prompting	1003
Hongxu Pu, Xincong Yang, Zhongqi Shi, Nan Jin	
Assisting in the identification of ergonomic risks for workers: a large vision-language model approach	1010
Chao Fan, Qiwei Mei, Xinming Li	
Knowledge Graph-based Deconstruction Planning of Building's Products	1018
Amr Allam, Mazdak Nik-Bakht	
Development of Online Course for Open Infra Built Environment Information Model	1025
Annika Kemppainen, Tanja Kolli, Rauno Heikkilä	
Integration of BIM and RFID-Sensing Technology for Automated Progress Monitoring in Modular Construction	1033
Maggie Y Gao, Chengjia Han, Yaowen Yang, Robert L.K. Tiong	

Scan-to-BIM: Unlocking current limitations through Artificial Intelligence	1040
--	------

Maxime Queruel, Stefan Bornhofen, Aymeric Histace, Laure Ducoulombier

R - Robotization of renovation in construction

Assessing the Viability of Robotic Disassembly of Building Components for Resource Recovery.....	1048
--	------

Christopher Rausch, Seungah Suh, Nikiforos Repousis, Nathan Titterington, Han Nguyen

Robotic platform for the (semi-) automated assembly of façade panels	1057
--	------

Christoph Heuer, Aaron Lentos, Zhongqian Zhao, Sigrid Brell-Cokcan

Automating the Tower Crane: Integrating the Development and Simulation of Path Planning and Trajectory Tracking of Tower Crane in ROS Framework	1065
---	------

Muhammad Muddassir, Mohamed A. A. Abdelkareem, Tarek Zayed, Zoubeir Lafhaj

Development of Robotics for Building Exterior Inspection: A Literature Review	1073
---	------

Tianxi Chen, Mi Pan, Thomas Linner, Honghao Zhong

Cable-Driven Parallel Robot (CDPR) for Panelized Envelope Retrofits: Feasible Workspace Analysis.....	1081
---	------

Yifang Liu, Rui Zhang, Nolan Hayes, Diana Hun, Bryan Maldonado

S - Sensing systems & data infrastructures

Development of a Prototype Data-system Integration Platform for As-Built Inspection	1089
---	------

Fumiya Matsushita, Kanae Miyaoka, Bumpei Miyazaki, Masamiki Matsubara, Kazumasa Ozawa

Automated Defect Inspection in Building Construction with Multi-Sensor Fusion and Deep Learning.....	1097
--	------

Juhyeon Kim, Jeehoon Kim, Yulin Lian, Hyoungkwan Kim

Towards Efficient Construction Monitoring: An Empirical Study on Action Recognition Models	1104
--	------

Sudheer Kumar Nanduri, Venkata Santosh Kumar Delhi

Preliminary Study on Enhancing Detection of Concrete Bridge Surface Spalling by Infrared Thermography.....	1115
--	------

Ching-Wei Chen, Hui-Ping Tserng, En-Wei Chiang

Variable Speed Limits Control for Smart Work Zone with Connected Vehicles	1120
---	------

Shuming Du, Saiedeh Razavi, Harith Absulsattar

VL-Con: Vision-Language Dataset for Deep Learning-based Construction Monitoring Applications.....	1128
---	------

Shun-Hsiang Hsu, Junryu Fu, Mani Golparvar-Fard

Acoustic Signal-Based Excavator Fault Detection Using Deep Learning Method	1136
--	------

Yuhan Zhou, Zhen Lin Foo, Yong Zhi Koh, Ker-Wei, Justin Yeoh

UAV Thermography for Building Energy Audit: Comparing Image Acquisition Strategies	1144
--	------

Cheng Zhang, Yang Zou, Johannes Dimyadi, Vicente A. Gonzalez

NeRF-Con : Neural Radiance Fields for Automated Construction Progress Monitoring	1152
--	------

Yuntae Jeon, Almo Senja Kulinan, Dai Quoc Tran, Minsoo Park, Seunghee Park

Identification of Factors and Metrics to Compare Vision Based Data Acquisition Devices	1160
--	------

Aakar Garg, Megha S Pradeep, Koshy Varghese

Robotic IoT-Enabled 1D Line Scanner Integration for 3D Point Cloud Data Processing	1168
--	------

Emre Ergin, Sven Stumm, Sigrid Brell-Cokcan

Instance Segmentation of Exterior Insulation Finishing System using Synthetic Datasets	1176
--	------

Mingyun Kang, Sebeen Yoon, Juho Han, Sanghyeon Na, Taehoon Kim

Virtual Construction Equipment Sensor for Determining Soil Stiffness during Compaction	1182
--	------

Maximilian Schöberl, Michael Schneider, Claus Lechner, Johannes Fottner

T - Technology management and innovation

Motives and Barriers for Offsite and Onsite Construction 3D Printing	1190
--	------

Wassim AlBalkhy, Elias Hernandez Valera, Dorra Karmaoui, Zoubeir Lafhaj, Thomas Linner, Syrine Ayadi, Reda Zerrari, Assia Boutabba

BIM Model View Definition (MVD) for disassembly planning of buildings	1198
---	------

Benjamin Sanchez Andrade, Srijeet Halder, Ranjith K. Soman, OK-Youn Yu

Customer Satisfaction in Construction Robots: A Multi-Stakeholder Perspective.....	1206
--	------

Siwei Chang, Tai Wai Kwok, Ching-Wei Chang, Yixing Yuan, Juhao Su

PESTEL Analysis of Factors Influencing the Demand and Supply of Modular Construction: Perspectives from Hong Kong.....	1214
--	------

Yidan Zhang, Wei Pan, Mi Pan

Digital Transformation, Sustainability and Construction 5.0	1222
---	------

Ming Shan Ng, Thomas Bock, Fabian Kastner, Silke Langenberg

A Safety Framework to Assess Autonomous Construction and Mining Equipment...	1230
--	------

Cynthia Brosque, Bibhrajit Halder

Human-Robot Partnership: An Overarching Consideration for Interaction and Collaboration.....	1240
--	------

Jingshuo Yu, Qian Chen, Samuel Antonio Prieto Ayllon, Borja Garcia de Soto

3D Printing vs. Traditional Construction: Cost Comparisons from Design to Waste Disposal Stages.....	1247
--	------

Svetlana Besklubova, Muhammad Huzaifa Raza, Ray Y. Zhong, Mirosław J. Skibniewski

Predicting Indicators of Design Quality for Cast-in-Place Reinforced Concrete Structures Using Machine Learning 1255

Leonardo Garcia-Bottia, Daniel Castro-Lacouture

A Generic 3D Printing Life Cycle Assessment (LCA) Framework for AEC Applications 1263

Bharadwaj R. K. Mantha, Ala Sati, Fatma Hosny, Mohamed Abdallah, Saleh Abu Dabous

Justification of Construction 4.0 maturity model with a case study of a data-driven façade company 1272

Orsolya Heidenwolf, Dr. Ildikó Szabó

A Bibliographic Multi Theme Review of Cybersecurity, AEC and 3D Printing: Learnings and Way Forward for the AEC Industry 1280

Bharadwaj R. K. Mantha, Jannatul Tashfia, Jannatul Tabassum, Saleh Abu Dabous

Analogizing in Construction and Education Research: A Case Study 1289

Ronie Navon, Yoram M Kalman

S - Services and business applications / Industry & Short Papers

Enhanced Precision in Built Environment Measurement: Integrating AprilTags Detection with Machine Learning 1295

Shengtao Tan, Aravind Srinivasaragavan, Kepa Iturralde, Christoph Holst

Training and Generating Apartment Plan Graphic Images for Commercial Uses 1299

Taesik Nam, Youngchae Kim, Sumin Chae, Youngjin Yoo, Jin Kook Lee

Research on an Open Source Physical Simulator for Autonomous Construction Machinery Development 1303

Daisuke Endo, Yosuke Matsusaka, Genki Yamauchi, Takeshi Hashimoto

Development of an Educational Package for a Construction Robots Simulator 1307

Ci-Jyun Liang, Cheng-Hsuan Yang, Liang-Ting Tsai, Bruce Alton

Collaborative R&D and Mutual Utilization of Construction Robotics in the Construction RX Consortium 1311

Hiroshi Tabai

UAV for sticking markers in the built environment 1315

Wen-Shuo Hsu, Soroush Fazeli, Kepa Iturralde, Christoph Holst

Development of Mobile Inspection Robot for Concrete Wall Surface and Strength Estimation Considering Coarse Aggregate by Small Diameter Drill 1319

Momoe Terata, Fumihiro Inoue

Analysis of openBIM Adoptions and Implementations: Global Perspectives and Canadian Industry Adoption 1323

Farzad Jalaei, Saman Davari, Yasir Sultan, Claudia Cozzitorto, Erik Poirier, Arash Shahi

Evaluating Site Scanning Methods: An Assessment System for Quantitative Comparison and Selection 1328

Maria Mortazavi, Bin Wu, Sofie Steinbrenner, Shervin Haghsheno

CrackGauGAN: Semantic Layout-based Crack Image Synthesis for Automated Crack Segmentation

Honghu Chu^{1,2}, Weiwei Chen^{1,3*}, Lu Deng²

¹ Bartlett School of Sustainable Construction, University College London, UK

² College of Civil Engineering, Hunan University, China

³ Division D Civil Engineering, Engineering Department, University of Cambridge, UK

weiwei.chen@ucl.ac.uk, honghu.chu.22@ucl.ac.uk, denglul@hnu.edu.cn

Abstract

Automated crack inspection, particularly deep learning (DL)-based crack segmentation, is crucial for the effective and efficient maintenance, repair, and operation of civil infrastructure. However, the performance of DL-based segmentation methods is often limited by the scarcity of pixel-wise labeled crack images. This paper presents CrackGauGAN, an automated crack image synthesis network that can be used to generate realistic and diverse crack image and mask pairs, which are instrumental in improving the performance of DL-based crack segmentation models. The CrackGauGAN is developed with three customized improvements based on the original GauGAN architecture. Firstly, a Criminisi-based crack image inpainting operator is introduced before the image encoder, enabling the exclusion of crack noise interference during background color feature extraction. Secondly, a background texture extraction method is proposed, assisting the SPADE-based generator in decoupling background textures as prior information. Lastly, an adaptive pseudo-augmentation strategy is introduced in the discriminator, allowing the model to be effectively trained on small-scale crack datasets. Ablation studies are conducted to prove the effectiveness of each component, and further crack image generation experiments demonstrate that the CrackGauGAN can synthesize various cracks with excellent diversity and fidelity. The CrackGauGAN-generated crack images show average improvements of over 1.97 and 7.91 in the Inception Score (IS) and Fréchet Inception Distance (FID), respectively, compared to the previously most advanced GauGAN and Pix2PixHD. As a fully automated crack image mask pair generation architecture, the CrackGauGAN can be used to provide reliable data support for the application of DL-based segmentation models in crack inspection tasks.

Keywords –

Deep Learning; Generative Adversarial Network; Crack Image Synthesis; Feature Decoupling; Crack Segmentation

1 INTRODUCTION

In recent decades, the aging of bridges and escalating traffic loads have markedly intensified the issue of structural safety [1, 2]. Cracks, as one of the most prevalent and evident indicators of structural safety [3], are particularly noteworthy. Early detection and prompt maintenance of these cracks can substantially lower maintenance expenses over the bridge's operational lifespan.

Traditionally, crack detection has been conducted through visual inspections by qualified experts, a process that is both costly and time-consuming [4, 5]. This approach struggles to keep pace with the escalating global demand for bridge inspections, particularly for long-span structures [6]. However, the advent of visual sensors in recent decades has facilitated the continuous acquisition of image data from civil infrastructure, utilizing automated robots, unmanned aerial vehicles (UAVs), camera-equipped vehicles, and fixed surveillance cameras on bridges [7]. The emergence of visual recognition models based on deep learning (DL) technologies offers the potential for accurate and efficient identification of cracks in these images, garnering significant interest from both industry and academia [8].

Recent research in bridge crack image processing has leveraged DL techniques, achieving significant advancements in crack classification, object detection, and segmentation. These techniques are increasingly recognized as the most promising solution for automating and streamlining detection tasks, potentially replacing manual labor [9-11]. While DL technologies represent the state-of-the-art in the field of pixel-level crack detection, their efficacy hinges on the availability of extensive source data for accurate network training. Limited datasets can lead to network overfitting, where the model excels on training data but exhibits markedly reduced performance in unfamiliar environments [12].

To mitigate the issue of limited training data,

researchers have explored data augmentation techniques to expand crack datasets [13, 14]. These methods fall into two categories: basic image processing and DL-based approaches [15, 16]. Basic image processing techniques, such as flipping, cropping, and rotating, generate new crack images that retain the original's semantic information [17]. However, these methods do not significantly enhance the diversity of crack samples, thus offering limited improvement in the accuracy of crack image recognition. In contrast, DL-based data augmentation algorithms, particularly Generative Adversarial Networks (GANs), can generate diverse, high-resolution images that mirror the distribution of the original dataset, thereby effectively enriching the variety of training samples [12].

Exploring GANs for synthesizing crack images with targeted domains has always been a topic worthy of discussion [18]. Numerous studies have shown that adding GAN-generated images to the original training data can make DL-based recognition models more accurate [19]. It is important to note that the generated images can be used to improve the performance of DL-based segmentation models on the condition that the images have detailed pixel-level annotations. However, carrying out the annotation process is an extremely tedious task. Therefore, to significantly enhance the convenience of using generated images, this study aims to use the GAN to generate realistic crack images while automatically obtaining their corresponding masks. To the best of the authors' knowledge, similar studies have not been reported. The main contributions to this work are as follows:

This study proposes a novel Generative Adversarial Network (GAN) architecture, named CrackGauGAN, which facilitates the generation of realistic crack images solely from semantic maps. The core innovation lies in integrating crack inpainting and texture priors into the GauGAN architecture [20], enabling the network to pre-decouple three key features in crack image generation - crack morphology, texture, and color. This innovation allows the SPADE (Spatially-Adaptive Normalization) blocks in GauGAN, originally designed for natural scene image generation, to be used for creating clear and realistic crack images. Furthermore, given the extensive image data required for parameter training due to the large number of parameters in the original GauGAN architecture, the authors implemented an adaptive data augmentation strategy in the discriminator. This approach allows for effective training of the CrackGauGAN on a limited dataset. The well-trained model is capable of automatically generating crack image and mask pairs, which can be directly used to enhance the performance of DL-based segmentation models, thereby promoting the application of DL-based crack segmentation methods in practical engineering.

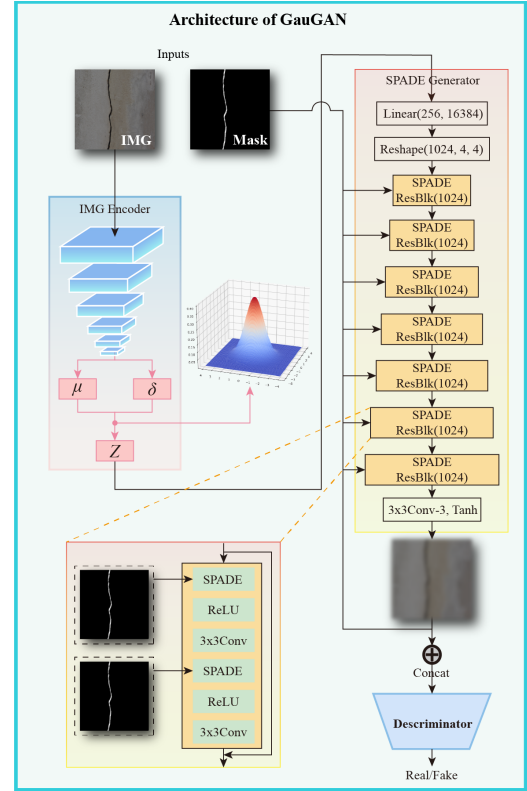


Figure 1. The overview architecture of the original GauGAN

2 METHODOLOGY

2.1 Overview of GauGAN

The CrackGauGAN model is built upon the GauGAN model [20], and it is necessary to explain the composition of the GauGAN model when introducing CrackGauGAN. The GauGAN architecture, as shown in Figure 1, mainly consists of three key components: Image Encoder, SPADE-based Generator, and Discriminator.

The Image Encoder is designed to extract the mean μ and variance δ related to the color feature distribution from the real image. Then, the extracted mean μ , variance δ , and the Gaussian distribution x would be denormalized, ultimately obtaining a random vector z that contains the color information of the real image.

The function of the SPADE-based Generator is to receive the random vector z generated in the previous step and enhance the realism of the pixels in the generated image by continuously using the semantic map.

The Discriminator is customized to process the tensor resulting from the integration of the semantic map and the generated image. It executes conditional discrimination across multiple scales, enabling the effective assessment of both global features, like background color and crack distribution, and local

features, including texture details, in the generated image. This functionality is crucial for ensuring the image's overall clarity.

Through continuous training involving a contest between the generator and discriminator, GauGAN ultimately produces realistic images that match the target distribution locations in the input semantic maps.

2.2 Revised Architecture for Crack Images

Unlike natural scene images with fixed shapes and significant semantic arrangement relationships, crack images consist only of cracks with random morphological distributions and backgrounds lacking semantic information. Consequently, traditional GauGAN, when dealing with such hard samples, tends to produce blurred and artifact-ridden images due to the difficulty in adequately decoupling the deep semantic features of cracks and backgrounds during the training phase. Additionally, the limited pixel-level annotated open-sourced crack datasets make GauGAN prone to overfitting during training. The authors address these issues effectively through three customized designs, and the revised architecture is shown in Figure 2, where the customized components are highlighted in red in the revised architecture.

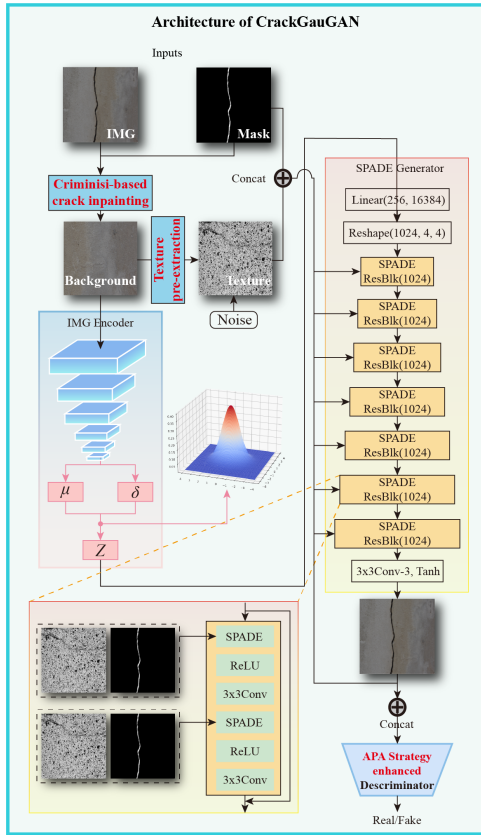


Figure 2. The overview architecture of the proposed CrackGauGAN

2.2.1 Revision 1: Criminisi-based crack inpainting

Considering that both cracks and backgrounds lack distinct semantic information that can be effectively differentiated by the image encoder, the random vector z often fails to accurately represent the color information of the image target due to feature coupling. To address this, the authors introduce a crack inpainting operation before the original image is input into the image encoder. This operation repairs the crack areas in the background image, thereby eliminating the interference of crack pixels in the extraction of the background color information vector.

To effectively repair the crack areas, it is necessary to ensure that the background texture of the repaired crack areas can transition smoothly. In this study, the Criminisi method [21] is employed. The Criminisi method, with its advantages in diffusion-based repair and texture synthesis, performs well on images with large missing areas and those composed of textures and structures. The specific implementation method is as follows:

1. Define the repair area: First, identify the crack areas in the crack image that need to be repaired from the input semantic labels;
2. Initialize priority: The algorithm then assigns priorities to the edge pixels of the area to be repaired based on structural information like pixel gradients and texture information based on the proportion of known pixels;
3. Select the source area: The algorithm searches for an undamaged area in the image that best matches the texture and structure at the current highest priority edge as the source area for repair;
4. Texture and structure replication: Copy the selected source area to the edge location with the highest priority, thereby filling a part of the repair area;
5. Update priorities and repair area: After each fill, the algorithm updates the edges of the repair area and the priorities of the corresponding pixels, and repeats steps 3 to 5 until the entire repair area is filled.

2.2.2 Revision 2: DTCWT and PSO-based background texture pre-extraction

The visual quality judgment of crack images by the naked eye primarily relies on three indicators: background color, crack distribution, and background texture. The background color and crack distribution have already been individually represented by the image encoder and the semantic map, respectively. To enable the network to effectively control all three critical indicators that determine the generation of crack images, it is necessary to add texture prior information to the network to help decouple texture information in advance.

Specifically, this study introduces an image texture feature extraction algorithm based on Dual-Tree Complex Wavelet Transform (DTCWT) and Particle

Swarm Optimization (PSO), designed to introduce texture prior information of the crack background into each SPADE block. Compared to traditional image texture feature extraction algorithms like Canny and Sobel, the texture information extraction method proposed in this study effectively leverages the efficient texture analysis capability of DTCWT and the global search advantage of PSO, thereby enhancing the accuracy and efficiency of image texture feature extraction. The implementation details of the DTCWT and PSO methods can be found in [22].

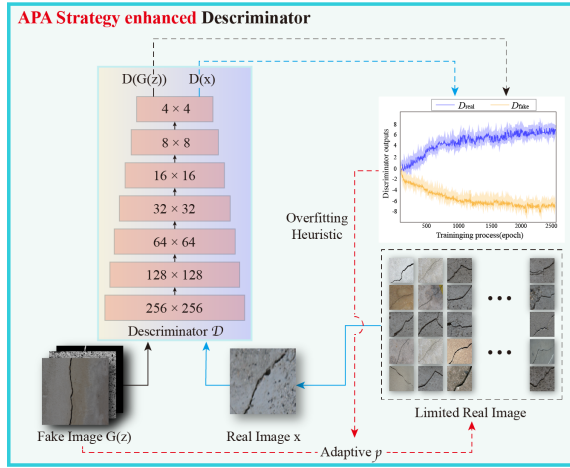


Figure 3. Workflow for adaptive pseudo-augmentation strategy in the discriminator

2.2.3 Revision 3: Adaptive pseudo-augmentation strategy in the Discriminator

To mitigate the potential overfitting issue due to the insufficiency of initial training source crack images, a new data augmentation strategy called adaptive pseudo-augmentation strategy (APA) is introduced in the discriminator, as shown in Figure 3.

Researchers have been challenged by obtaining well-trained generative models based on carefully constructed GAN frameworks, with traditional methods relying on a large number of training images to ensure the model avoids overfitting. For the GauGAN architecture, designed for generating natural scene images, open-source datasets like Pascal VOC 2012, COCO, and Cityscapes, which contain tens of thousands of pixel-level annotated images, have alleviated the data issue to some extent. However, due to the difficulties in collecting crack images and the reliance on professional personnel for annotation, there is no such large-scale open-source dataset available for cracks. Therefore, effectively training CrackGauGAN on small sample datasets becomes the problem to be addressed in this section. To this end, the authors designed a data augmentation pipeline for the limited crack training samples, called the adaptive pseudo-augmentation

Strategy, first proposed by Huang et al. [23]. It can dynamically adjust the intensity of augmentation based on the degree of overfitting in the field of medical imaging generation, without leaking augmentation patterns. Its effectiveness has been proven through the ablation study described in section 4.1.3.

3 IMPLEMENTATION DETAILS AND EXPERIMENTS

3.1 Datasets

The open-source crack segmentation dataset HRCD-282, previously established by the authors, was used as the data source for model training and evaluation. From HRCD-282, 1200 crack patch images with a resolution of 256×256 and their corresponding crack labels were cropped from the included HR crack images to serve as training data for CrackGauGAN. Additionally, to assess the quality of images generated by the model, 1200 non-crack patch images with the same resolution of 256×256 were collected from the backgrounds of the selected original HR crack images. These 1200 non-crack background images, along with crack labels of the same size, were input into the well-trained CrackGauGAN model to generate pseudo-crack images under the corresponding backgrounds.

3.2 Implementation Details

In this study, all experiments were conducted on a high-performance workstation equipped with an Intel Core i9-9820X CPU and NVIDIA RTX 3090 Ti GPU. The workstation runs on the Ubuntu 20.04 LTS operating system, providing a stable and efficient computing environment. Furthermore, to ensure the fairness of comparative experiments, all networks involved in the comparison were implemented under the PyTorch framework.

To optimize the training effectiveness of the CrackGauGAN model, the authors carefully selected the following key hyperparameters. The initial learning rate was set at 0.0001 to ensure stable gradient descent, and a learning rate decay strategy was adopted to address potential overfitting during training. Considering the limitations of hardware resources, the batch size was cautiously set to 8 to balance memory consumption and training efficiency. Additionally, the Adam optimizer was chosen, with its beta1 and beta2 parameters set to 0.5 and 0.999, respectively, to take advantage of its adaptive learning rate. The weight decay parameter was set at 1×10^{-4} to further prevent model overfitting.

In terms of the update strategy for the generator and discriminator, a 1:1 ratio was followed to ensure balanced optimization of both during training. The

weighting of the loss function was also carefully adjusted to balance the impact of adversarial loss, feature matching loss, and VGG perceptual loss, thereby optimizing the overall performance of the model. Considering the importance of crack image details and the limitations of computational resources, a training image size of 256×256 was chosen. Finally, to ensure that the model learned sufficiently and converged, the number of training epochs was set to 900. The selection of these hyperparameters was based on a comprehensive analysis of previous studies and the results of preliminary experiments, aiming to achieve the best training effect and image quality.

3.3 Evaluation Indicators

To comprehensively evaluate the model performance of CrackGauGAN, the authors conducted qualitative assessments through visualized generation results of all models and also introduced two quantitative evaluation metrics: Inception Score (IS) and Fréchet Inception Distance (FID). IS evaluates the diversity and clarity of the generated images, while FID measures the distance between the generated images and real images in the feature space.

4 EXPERIMENTS AND RESULTS ANALYSIS

4.1 Ablation Studies

4.1.1 Ablation study of the Criminisi-based crack inpainting

Table 1 quantitatively shows the ablation study for the image background inpainting. It can be found that by removing the background inpainting, the training process becomes more difficult to converge, resulting in a larger FID score and a smaller IS score, which indicates a decrease in the realism and clarity of the generated images. This is because crack images differ from natural scene images in that there is no significant color contrast difference between the foreground and background, and neither possesses a fixed topological structure. This makes it challenging for the image encoding architecture to effectively decouple their deep semantic information. As a result, the network's process of extracting background color features is easily disrupted by crack features, leading to a significant decline in the quality of color feature extraction. Experimental results confirm that repairing crack pixels in the image background effectively mitigates this issue.

4.1.2 Ablation study of the DTCWT and PSO-based texture pre-extraction

This ablation study primarily focuses on the impact

of the texture prior information intensity, extracted by the proposed DTCWT and PSO operations, on the quality of image generation. Figure 4 illustrates these comparisons across texture information intensities of 0%, 25%, 50%, 75%, and 100%. A visual evaluation of the reconstruction results in Figure 4 revealed that the model utilizing 100% texture information extraction intensity achieved the closest resemblance to the original image. Furthermore, it was observed that at texture information extraction intensities below 50%, the GAN struggled to independently decouple crack features from background texture information. This challenge is attributed to the excessive redundancy in low semantic information, a consequence of the random data distribution in crack images. Consequently, early decoupling of high-intensity texture information significantly enhances the network's proficiency in distinguishing and capturing both background semantic information and crack features, a critical factor for reconstructing high-quality crack images.

Table 1. The impact of the proposed inpainting operation on the quality of generated images and the difficulty of the model convergence

Criminisi-based crack inpainting operation	IS(↑)	FID(↓)	Iterations (epoch)
W/O	3.43	34.56	1700
W/	5.26	26.33	900

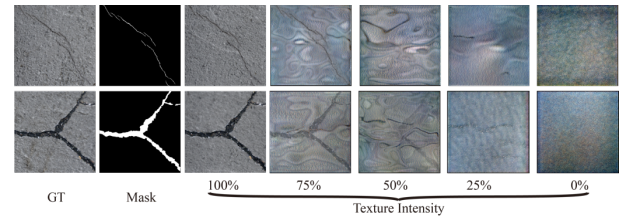


Figure 4. Visualization results of crack images generated by models with added texture prior information of varying intensities

4.1.3 Ablation study of adaptive pseudo-augmentation Strategy

Figure 5 summarizes the impact of the adaptive pseudo-augmentation (APA) strategy on both the generator and discriminator losses in the CrackGauGAN. The implementation of the APA strategy led to a notable reduction in the oscillation amplitudes of both losses during training on a small-scale dataset. This indicates that the APA strategy effectively stabilized the feature learning of both the discriminator and generator within the bounds of maximum gradient convergence. This is particularly significant for the discriminator loss, where the extremely limited training data can rapidly lead to local overfitting, thereby creating a false impression of

model convergence. The APA strategy successfully counters this issue. Analysis of the discriminator (D) loss and generator (G) loss behaviors suggests that the APA strategy not only prevents overfitting in the discriminator but also mitigates the problem of gradient vanishing, thereby ensuring the generator's continuous learning.

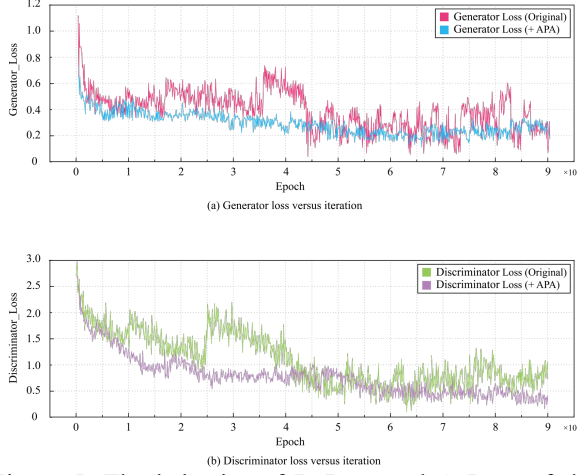


Figure 5. The behavior of D Loss and G Loss of the proposed architecture with and without APA strategy in the training phase

4.2 Performance Comparison with Current State-of-the-art Methods

To demonstrate the advancement of the method proposed in this study, it was compared with the current state-of-the-art semantic image synthesis methods: GauGAN [23] and pix2pixHD model [24]. The Ground

Truth (GT) used for evaluating the quality of the generated images originates from the HR crack images collected by the authors, as described in section 3.1. For each GT, the corresponding crack mask and the non-crack background images were used as inputs for the models. Additionally, it is important to note that to ensure fairness in comparison, all models involved in the experiment were trained using the default hyperparameters provided by their original authors.

Qualitative Results Analysis: Figure 6 illustrates some of the crack images generated by all the methods involved in the comparison. It is intuitively evident from the figure that all the compared methods can generate crack images broadly consistent with the distribution of the crack semantic map. Among these, the crack images synthesized through CrackGauGAN are closer to real crack images. Specifically, they exhibit fewer artifacts, richer details, and clearer edges in terms of visual quality. As for the Pix2PixHD, due to the lack of decoupling operations for cracks and background textures during the training phase, tends to lose background texture information in the generated images. Although GauGAN, which employs SPADE, can utilize the spatially adaptive mechanism to improve the entangled coupling between background and crack features, the limited crack training samples restrict its performance, resulting in the generated crack background textures being blurred. Further observation of the crack images generated by the CrackGauGAN reveals that different levels of noise input do not affect the clarity of the images or the distribution of cracks in the images, but only alter the distribution of background textures. This is significant for enhancing the diversity of the images.

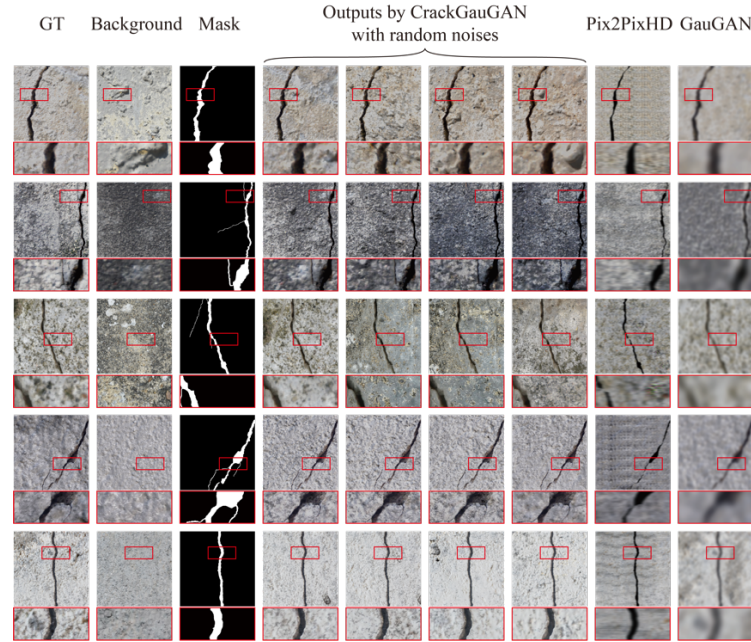


Figure 6. Qualitative comparison of CrackGauGAN with two state-of-the-art image synthesis methods

Table 2. Quantitative results of the quality of images generated by CrackGauGAN with different intensity noise added and the current state-of-the-art methods

Model	The intensity of random noise added	IS	FID
Pix2PixHD	/	4.20	31.65
GauGAN	/	4.78	29.38
CrackGauGAN	20%	6.47	22.74
	40%	6.53	22.09
	60%	6.43	22.76
	80%	6.39	23.12

Quantitative Results Analysis: Table 2 reports the quantitative results of the crack image quality generated using different methods. As can be seen from Table 2, adding different proportions of random noise has almost no impact on the evaluation metrics. This can be inferred from the qualitative analysis results shown in Figure 4, as the random noise only affects the distribution of the background texture and does not impact the clarity of the background texture or other details related to image quality. Moreover, under four types of noise inputs, the IS and FID of the images generated by GauGAN fluctuate within the range of 6.46 ± 0.07 and 22.61 ± 0.52 , respectively. This represents an average improvement of over 1.97 and 7.91 compared to the images generated by Pix2PixHD and GauGAN, respectively. This effectively confirms the advanced nature of the method proposed in this study.

5 CONCLUSION

This paper presents a novel CrackGauGAN, a semantically-driven generative adversarial network specifically tailored for crack image generation. By effectively decoupling three key features for the generator, including background color, background texture, and crack morphology, the CrackGauGAN is capable of generating realistic crack images with high fidelity and diverse data distribution types. Its performance surpasses that of the most advanced semantic layout-based GANs, with average improvements in Inception Score (IS) and Fréchet Inception Distance (FID) exceeding 1.97 and 7.91, respectively, compared to GauGAN and Pix2PixHD.

The proposed generative model can quickly generate a large number of realistic crack image datasets with pixel-level labels for in-service bridges, overcoming the challenge of insufficient training samples with similar data distribution types. In the future, this approach will be extended to create customized crack image datasets for additional bridges and train corresponding crack detection models. Such models will facilitate accurate and efficient intelligent detection methods in engineering

practice, including bridges, hydropower projects and historical buildings, enabling evidence-based infrastructure maintenance and management decisions.

Acknowledgement

This work is supported by Horizon Europe project D-HYDROFLEX (Project ID: 101122357) and Horizon Europe project INHERIT (Project ID: 101123326).

References

- [1] Zhang C, Chang C-C, Jamshidi M. Concrete Bridge Surface Damage Detection Using A Single-Stage Detector. *Computer-Aided Civil And Infrastructure Engineering*, 35(4):389-409, 2020.
- [2] Abu Dabous S, Feroz S. Condition Monitoring Of Bridges With Non-Contact Testing Technologies. *Automation In Construction*, 116:103224, 2020.
- [3] Yu Z, Shen Y, Shen C. A Real-Time Detection Approach For Bridge Cracks Based On Yolov4-FPM. *Automation In Construction*, 122:103514, 2021.
- [4] Jiang S, Zhang J. Real-Time Crack Assessment Using Deep Neural Networks With Wall-Climbing Unmanned Aerial System. *Computer-Aided Civil And Infrastructure Engineering*, 35(6):549-64, 2020.
- [5] Rodriguez-Lozano FJ, LeÓN-García F, GÁmez-Granados JC, Palomares JM, Olivares J. Benefits Of Ensemble Models In Road Pavement Cracking Classification. *Computer-Aided Civil And Infrastructure Engineering*, 35(11):1194-208, 2020.
- [6] Guo F, Qian Y, Liu J, Yu H. Pavement Crack Detection Based On Transformer Network. *Automation In Construction*, 145:104646, 2023.
- [7] Bianchi E, Abbott Amos L, Tokekar P, Hebdon M. COCO-Bridge: Structural Detail Data Set For Bridge Inspections. *Journal Of Computing In Civil Engineering*, 35(3):04021003, 2021.
- [8] Alipour M, Harris Devin K, Miller Gregory R. Robust Pixel-Level Crack Detection Using Deep Fully Convolutional Neural Networks. *Journal Of Computing In Civil Engineering*, 33(6):04019040, 2019.
- [9] Corbally R, Malekjafarian A. A Deep-Learning Framework For Classifying The Type, Location, And Severity Of Bridge Damage Using Drive-By Measurements. *Computer-Aided Civil And Infrastructure Engineering*, 37(3):1163-1188, 2023.
- [10] Yang Q, Shi W, Chen J, Lin W. Deep Convolution Neural Network-Based Transfer Learning Method For Civil Infrastructure Crack Detection. *Automation In Construction*. 116:103199, 2020.

- [11] Zhou S, Canchila C, Song W. Deep Learning-Based Crack Segmentation For Civil Infrastructure: Data Types, Architectures, And Benchmarked Performance. *Automation In Construction*, 146:104678, 2023.
- [12] Shi J, Liu W, Zhou G, Zhou Y. Autoinfo GAN: Toward A Better Image Synthesis GAN Framework For High-Fidelity Few-Shot Datasets Via NAS And Contrastive Learning. *Knowledge-Based Systems*, 276:110757, 2023.
- [13] Liu B, Zhang T, Yu Y, Miao L. A Data Generation Method With Dual Discriminators And Regularization For Surface Defect Detection Under Limited Data. *Computers In Industry*, 151:103963, 2023.
- [14] Liu Q, Zhang Q, Liu W, Chen W, Liu X, Wang X. WSDS-GAN: A Weak-Strong Dual Supervised Learning Method For Underwater Image Enhancement. *Pattern Recognition*, 143:109774, 2023.
- [15] Xu G, Yue Q, Liu X. Real-Time Monitoring Of Concrete Crack Based On Deep Learning Algorithms And Image Processing Techniques. *Advanced Engineering Informatics*, 58:102214, 2023.
- [16] Li S, Zhao X. High-Resolution Concrete Damage Image Synthesis Using Conditional Generative Adversarial Network. *Automation In Construction*, 147:104739, 2023.
- [17] Liu X, Pedersen M, Wang R. Survey Of Natural Image Enhancement Techniques: Classification, Evaluation, Challenges, And Perspectives. *Digital Signal Processing*, 127:103547, 2022.
- [18] Zhang G, Cui K, Hung T-Y, Lu S, Editors. Defect-GAN: High-Fidelity Defect Synthesis For Automated Defect Inspection. *Proceedings Of The IEEE/CVF Winter Conference On Applications Of Computer Vision*, Pages 35–43, Virtual Online, 2021.
- [19] Jin T, Ye X, Li Z. Establishment And Evaluation Of Conditional GAN-Based Image Dataset For Semantic Segmentation Of Structural Cracks. *Engineering Structures*, 285:116058, 2023.
- [20] Park T, Liu M-Y, Wang T-C, Zhu J-Y, Editors. Semantic Image Synthesis With Spatially-Adaptive Normalization. *Proceedings Of The IEEE/CVF Conference On Computer Vision And Pattern Recognition*, Pages 56–64, Long Beach, America, 2019.
- [21] Yao F. Damaged Region Filling By Improved Criminisi Image Inpainting Algorithm For Thangka. *Cluster Computing*, 22:13683-91, 2019.
- [22] Chi J, Eramian M. Enhancing Textural Differences Using Wavelet-based Texture Characteristics Morphological Component Analysis: A Preprocessing Method for Improving Image Segmentation. *Computer Vision and Image Understanding*, 158:49-61, 2017.
- [23] Huang X, Belongie S, editors. Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization. *Proceedings of the IEEE International Conference on Computer Vision*, page: 77-85, Venice, Italy, 2017.
- [24] Wang T-C, Liu M-Y, Zhu J-Y, Tao A, Kautz J, Catanzaro B, editors. High-resolution Image Synthesis and Semantic Manipulation with Conditional Gans. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, page: 121-129, Salt Lake City, America, 2018.

Safe Operations of Construction Robots on Human-Robot Collaborative Construction Sites

Marvin H. Cheng¹, Ci-Jyun Liang², and Elena G. Dominguez³

¹Division of Safety Research, National Institute for Occupational Safety and Health, USA

²Department of Civil Engineering, Stony Brook University, USA

³PILZ Automation Safety, USA

MCheng@cdc.gov, Ci-Jyun.Liang@stonybrook.edu

Abstract –

Construction robots have become essential tools on a variety of jobsites. These devices can be revolutionary tools for improving construction efficiency and reducing musculoskeletal disorders and traumatic injuries. However, this innovative technology comes with corresponding dangers and hazards if a robot is not operated properly. Construction workers can be injured by unexpected contact. Therefore, construction robots need to be operated under specific safety procedures to prevent workers from being injured. In this study, a mechanical approach was proposed to derive the dynamic models of unexpected contact during human-robot interaction. With the dynamic models, contact forces and deformations of body parts of human workers can be estimated. The estimated results can be used as reference values to help safety engineers or others to adjust the operations in different scenarios on the construction jobsite for improved safety.

Keywords –

Collaborative Robots; Safety; Mechanical Model

1 Introduction

With the rapid advances in robotics, the construction industry is beginning to be revolutionized with the help of robots designed for this labor-intensive sector [1]. Robotic devices with various functions have been deployed or studied in different applications in the construction sector. With the deployment of various robotic devices, collaborative workspaces that require human-robot interaction have become more common in the past decade. On construction sites, robots are already used to assist construction workers with labor-intensive tasks, such as bricklaying, carrying heavy materials, and demolition tasks [2-3]. Type C mobile robots have been used in construction logistics to prevent long-term musculoskeletal disorders in construction workers [4].

Various research groups have also investigated how robotic on-site additive manufacturing can speed up the construction process [5,6]. Wearable robotic devices have also been widely used to prevent occupational traumatic injuries and musculoskeletal disorders [7]. However, robotic applications deployed on construction sites remain limited due to the lack of computational power, sensory assessment, and effective human-machine interface capabilities, which are important for construction work that requires multiple steps, various tool sets, and the need to follow specific work protocols.

Modern industrial robotic devices can detect the conditions of the jobsite and communicate among each other, sharing site information in real-time. With the capabilities to efficiently sense and communicate between robots, engineers can program robots for upcoming construction jobs to actively assist workers with repetitive tasks while providing required assistance during heavy-duty manual operations. For example, masonry robots have been used to reduce potential injuries due to the need for construction workers to move heavy objects. The collaborative partnership allows construction workers to focus on the quality of the construction tasks as well. However, an open jobsite such as a construction site is often not an ideal environment for robots to have all necessary sensors to detect worker's movements and environmental changes. Environmental disturbance, noises, and an insufficient number of sensors can greatly affect the ability of the robot to detect surrounding hazards and movements of existing objects in the construction space [8]. Thus, although robotic devices can greatly assist construction workers in performing repetitive and labor-intensive tasks and prevent potential injuries, unexpected contact between robots and construction site workers can still be dangerous and even fatal [9,10].

In a human-robot collaborative environment, robots can perform repetitive and labor-intensive work while construction workers focus on planning and inspecting the results to ensure quality. However, human-robot collaboration on construction sites can be dangerous for

a number of reasons, such as the height of the job site, unstructured workspaces, and accidentally dropped items [11]. If safety regulations do not clearly dictate a specific operation, workpieces carried by existing robots could strike human workers while operating in an open environment. In this constrained environment, the robot controllers need to be able to minimize potential injuries from unexpected contact between human workers and a coexisting robotic device, or the payload they carry. Unfortunately, while safety standards for both collaborative and mobile robots have been published for the manufacturing industry, related safety standards for robotic equipment used on collaborative construction sites have not yet been developed. Therefore, the construction industry urgently needs safety guidelines for the use of robotic equipment on open construction sites.

According to recent studies of construction occupational incidents, falling of a person from heights, striking against or struck by moving objects, and struck by falling objects are ranked as the top incident types of occupational injury cases in the construction industry [10]. As an important piece of equipment on the jobsite, many different operational methods have been investigated. In these safety studies, computer vision plays an important role in detecting the presence of construction workers as well as the relative distance between construction workers and robots [12-14]. With the help of computer vision devices, robots can sense the movements of workers around them and identify the contact avoidance zones. The moving trajectory of the robot is then automatically programmed to proactively avoid potential contact with workers in the targeted zone [15]. In addition to path planning methods based on visual feedback, industry and academia have also studied task allocation methods by optimizing the individual capabilities of both robots and human workers [16]. However, most of these human-robot interaction studies focused on applications in manufacturing, healthcare, transportation, and warehouse logistics. Although the developed technologies can be applied to the human-robot collaboration in the construction industry, they are currently not well adopted in construction applications. Some tasks that are not structured enough need to rely on human reasoning in complex manipulation tasks in unstructured environments [17]. The emerging field of human-robot collaboration has significant potential applications in construction and continues to advance the state-of-the-art in defining the roles of both humans and robots in collaborative work. In this study, we focused on a common incident: injury caused by being struck by a moving object. To prevent such an incident, the transferred energy during the impact should not generate a force that yields deformation on a human body surface greater than the permissible values. With this specific condition, the moving speeds of the robot parts, or the

object the robot carries, need to be determined. A mechanical modeling approach was applied to estimate the allowed velocity of the robot at the contact point while the robot is working on its assigned task.

The second section of this manuscript discusses the safety concerns of construction robots. Safety operations of an existing safety standard are assessed whether they are applicable to construction sites. Section three presents the dynamic models for two human-robot collision scenarios. The permissible conditions are discussed in this section. The fourth section discusses the physical conditions for two possible cases of the transient responses of the collision. Proper actions are suggested based on the simulation results in this section. The final section summarizes the overall results of this study.

2 Safety Concerns of Construction Robots

2.1 Safety Standards for Collaborative Robots

Current industrial standards of collaborative robot safety, ANSI/RIA R15.06 [18] and ISO/TS 15066 [19], have identified four types of operations to control injury risks for collaborative robots to work with workers, including safety-rated monitored stop, hand guiding, speed and separation monitoring, and power and force limiting. While these types of operations are effective in protecting workers from potential injury in designated workspaces, some are not as easy to implement on construction sites. For example, a robot that assists carpenters in building a timber frame structure of a residential house needs to team up with the carpenter. In this multi-step task, the robot needs to pass tools, carry materials, slide frames, or hammer components into place [20]. Specific safety operations identified in the safety standards would need to be applied in each step to meet the current safety requirements. Implementing all safety operations can be difficult and time-consuming if the robot needs to complete all the steps of the task. Figure 1 illustrates the interaction between a construction worker and a masonry robot. In this open jobsite, it is difficult to install all the required sensors in the environment for effective safe operation.

2.2 Safety Concerns on Construction Sites

The most effective way to reduce impact injuries from moving robotic devices is to monitor the movements of surrounding construction workers and program the robot's path planning to proactively avoid unexpected contact. Contact between robots and human workers is allowed at designated locations with limited forces. However, while a robot may be able to detect the presence of human workers, it might not be able to continuously monitor the worker's movement from all

directions on an open jobsite. If a robot cannot effectively recognize the worker's movements, it is less likely that the robot can perform all the required safety operations based on the worker's body posture. Thus, there is a need to regulate energy transfer between the human worker and the robot, or the construction material carried by the robot. One important consideration is that the transferred energy caused by unexpected contact should not be greater than a threshold value [21], which shall not cause injury on the surface of different body parts or alter the position of the worker due to the contact.



Figure 1. Human-robot collaboration on the construction jobsite using a masonry robot [22].

Taking these two factors into consideration, several physical criteria need to be defined to program motions of the robot. These two criteria include the permissible deformation of the human body surface and the allowable force in the contact area of the human body. These two values are usually not the same for different human body parts. The allowed deformation of the body surface and the allowed transferred energy can be used to determine moving speed at which the robot moves on contact. In ISO/TS 15066, it is recommended that the end-effector of a robotic device should move less than 0.25 m/s when the presence of a human worker is detected [19].

3 Modeling of Impacts Between Human Workers and Robotic Devices

To develop safety requirements on collaborative construction sites, knowledge of the environment and operation of construction robots is required. The construction site is usually an open environment, so the number of sensors can be limited and not easy to install. Although the parameters of the robot, such as the moving speed of the end effector, the angular positions and velocities of joints, and the effective configuration space, can be obtained from the controller of the robot, information related to the moving speeds of surrounding workers and changes of geometric shapes of the worksite may not be readily available. Without these data being readily acquired, robotic devices need to determine

operations based only on the detectable characteristics, such as moving speeds and locations. Thus, operations need to be adjusted if the presence of human workers is detected. In this section, two contact models were investigated to determine the physical interaction of the unexpected contact between human workers and robotic devices. In this study, two cases of unexpected contact were considered:

1. Struck and pushed: The robot contacts the worker and forcibly moves the worker from his/her original position.
2. Struck and bent: The robot contacts the worker from behind, and the worker's upper body bends forward around the waist joint. The lower body of the worker remains unmoved.

To derive the dynamic models of these impact scenarios, the equations of motion of the whole body need to be derived. However, each model needs to be divided into two phases. The first phase is the compression phase, starting from the contact between the object and the human body surface, until the compression reaches the permissible deformation of the human body. The second phase is the retraction phase, where the body surface begins to recover from its permissible deformation to its original state. The contact area of the body part acts in the first phase as a mass-damper-spring system. During the retraction phase, the elastic force exists only when the interaction force between the body surface and the object is greater than zero. Once the interaction force disappears, or the object is no longer in contact with the human body, the moving object is excluded from this mass-damper-spring system, but the skin surface is still deformed due to the previous impact. The configuration of the compression and retraction phases is shown as Figure 2.

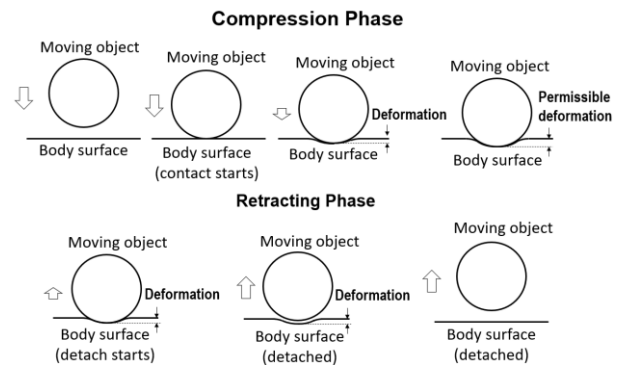


Figure 2. Compression and retraction of reaction of between skin surface and moving object due to impact.

3.1 Struck and Pushed

For the impact between a moving object and a human worker, Figure 3 demonstrates the mechanical model of such an impact. Depending on the initial velocity of the moving object carried by the robot, the dynamic response is also different. If the initial speed of the robot is below a threshold, the object bounces back after impact. If the initial velocity of the object transported by the robot is too fast, the contact surface of the worker's body is compressed to the maximum, and the worker's body moves together with the object. In this case, the object and the worker move with the same final velocity. If the object carried by the robot is heavy and moving too fast, the worker can be thrown out after being impacted. The last two situations can lead to severe worker injury and should be avoided when human workers collaborate with robotic devices in the workspace. In this study, only the first situation is discussed. Figure 3 illustrates the situation when an object carried by the robot strikes the worker from the side. In this configuration, the horizontal movement of the object transported by the robot is x_B , and the horizontal movement of the center of mass of the worker's body is x_h . The equations of motion of the worker and the moving object can be written as

$$M_B \ddot{x}_B + C_h \dot{x}_B - C_h \dot{x}_h + K_h x_B - K_h x_h = 0 \quad (1)$$

$$M_h \ddot{x}_h + C_h \dot{x}_h - C_h \dot{x}_B + K_h x_h - K_h x_B - K_h x_B + f_s = 0 \quad (2)$$

where the mass of the moving object is M_B , the mass of the human body is M_h , the stiffness of the contact surface of the body part is K_h , the corresponding viscous damping is C_h , and f_s is the static friction between the shoes and the ground. The compression of the body at the impact location Δs is

$$\Delta s = x_B - x_h \quad (3)$$

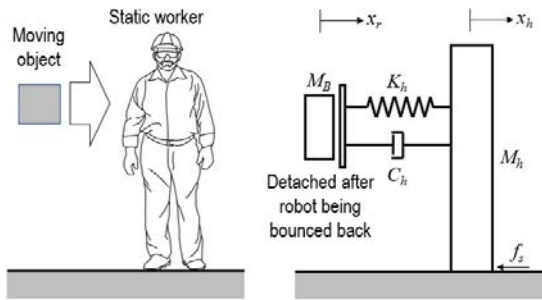


Figure 3. Mechanical model of the side impact between the static worker and the moving robot.

The contact force F_c that yields the compression on the human body is

$$F_c = K_h \Delta s_{max} \quad (4)$$

The contact force reaches its maximum when the

deformation Δs is also its maximum. The initial velocity of the moving object at the moment of contact is \dot{x}_{B0} . In the compression phase, the object carried by the robot compresses the human body surface until the relative velocity is zero. Then the retraction phase starts. Once the object is detached from the worker's body, or the surface of the worker is no longer compressed, the object stops interacting with the worker's body. Therefore, the worker is no longer subjected to the external force from the object. The velocities of the worker and the object after the impact can also be verified by conservation of momentum, which is

$$M_B \dot{x}_{B0} + M_h \dot{x}_{h0} = M_B \dot{x}_B(t) + M_h \dot{x}_h(t) \quad (5)$$

In this configuration, three possible outcomes are expected. If the moving speed of the object is low, the body surface of the worker might deform without moving the entire body. In this case, the static friction between the worker and the floor is high enough to prevent the worker from moving. The worker can also move after the impact. In this case, the object might bounce off of the worker after impact, and the friction between the worker and the floor changes from static to kinetic friction. If the object is moving fast enough, the object might push the worker and move the worker with it. Then the friction changes from static friction to kinetic friction.

3.2 Struck and Bent

To derive the dynamic model of the two phases, Figure 4 demonstrates the mechanical model of the impact. In this configuration, the horizontal movement of the moving object driven by the robot is x_B , and the angular movement of the upper body, or the waist joint, is θ_u . It is assumed the object contacts the human worker's back. An assumption has been made that the upper body center of mass is at one-half the upper body length. The horizontal movement of the center of mass of the work's upper body is x_h , which is

$$x_h = \frac{l_u}{2} \theta_u \quad (6)$$

The corresponding moving velocities and accelerations of the object carried by the robot and the affected worker are \dot{x}_B , \dot{x}_h , \ddot{x}_B , and \ddot{x}_h , individually. The range of motion of the waist joint is limited according to the physical condition of the worker, which is $\theta_{u,max}$. In the compression phase, the initial contact velocity of the moving object is \dot{x}_{B0} . When contact starts, the moving velocity and the angular velocity of the upper body of the worker are $\dot{x}_{h0} = 0$ and $\dot{\theta}_{u0} = 0$. The mass of the upper body is M_{hu} , and the length of the upper body is l_u . The lower body, including the thighs, legs, and feet are lumped as a single mass $M_{l,lumped}$. The compression of the body at the impact location Δs is $x_B - x_h$, which is

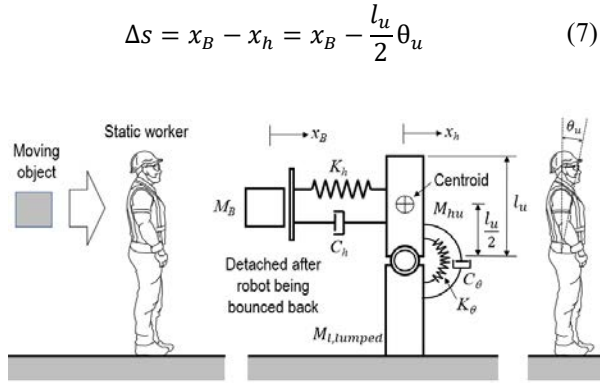


Figure 4. Mechanical model of the back impact between the static worker and the moving robot.

With this configuration, the equations of motion of both the moving object and worker's upper body in the first phase can be described as

$$M_B \ddot{x}_B + C_h \dot{x}_B - \frac{C_h l_u}{2} \dot{\theta}_u + K_h x_B - \frac{k_\theta l_u}{2} \theta_u = 0 \quad (8)$$

and

$$\begin{aligned} \frac{M_{hu} l_u^2}{4} \ddot{\theta}_u - \frac{C_h l_u}{2} \dot{x}_B + \left(C_\theta + \frac{C_h l_u^2}{4} \right) \dot{\theta}_u - \frac{k_\theta l_u}{2} x_B \\ + \left(\frac{K_h l_u^2}{4} + k_\theta \right) \theta_u = 0 \end{aligned} \quad (9)$$

where M_{hu} is the mass of the upper body, and k_θ is the stiffness of the joint between the upper and the lower body. During the whole impact process, there is no external force applied to either the robotic device or the worker. The maximal force occurs during the approaching process before the compression Δs reaches to its maximum. The contact force F_c that yields compression on the human body is

$$F_c = K_h \Delta s_{max} \quad (10)$$

The contact force reaches to its maximum when the deformation Δs is also maximum. Then the process of impact changes to the retraction phase. Once the moving object separates from the worker's body, the object stops interacting with the worker's body. The detachment occurs in the retraction phase. Though the body surface has not returned to its original position, the moving object has moved a sufficient distance and no longer touches the body surface of the worker. After the robot is separated from the worker, it moves at a constant speed. Since the interaction between the human body and the object carried by the robot no longer exists, the contact force generated by the impact also becomes zero. The human postures still change due to the dynamics induced by the stiffness and viscous properties of the human body. Thus, the equation of motion of the worker's body becomes

$$\frac{M_{hu} l_u^2}{4} \ddot{\theta}_u + C_\theta \dot{\theta}_u + k_\theta \theta_u = 0 \quad (10)$$

By modeling the dynamics of the given status, the maximum compression of the body surface and the related contact force can be derived if the initial velocity of the moving object is given. If the permissible compression on the human body is specified, an upper limit for the contact velocity of the object can also be established. Additionally, the worker cannot maintain a stationary location if the worker's center of gravity is not within the area of support, or the feet.

3.3 Allowed Pressure and Permissible Deformation at Different Body Parts

To prevent potential injuries from robot operations, unexpected contact between object carried by robot and human workers needs to be avoided. However, if the potential contact is likely to occur at a collaborative worksite, the speed at which the robot moves needs to be regulated to prevent injury. To determine the permissible moving speed of the robotic device while transporting construction materials, two parameters are required, permissible impact force and permissible deformation. In ISO/TS 15066 [19], maximum impact force and stiffness for different body parts are specified. Table 1 lists the values of these properties. The permissible deformation of individual parts can be calculated from the maximum permissible impact force and the corresponding stiffness. In ISO/TS 15066, the permissible pressure multiplier in a transient process is 2, which means the deformation can be doubled during the impact.

Table 1. Stiffness, allowable impact forces, and surface pressure regulated in ISO/TS 15066 [19].

Body part	Impact force (N)	Surface pressure (N/mm ²)	Stiffness (N/mm)
Skull/forehead	175	0.3	150
Face	90	0.2	75
Neck (sides/neck)	190	0.5	50
Neck (front/larynx)	35	0.1	--
Back/shoulders	250	0.7	35
Chest	210	0.45	25
Belly	160	0.35	10
Pelvis	250	0.75	25
Buttocks	250	0.8	--
Upper arm/elbow	190	0.5	30
Lower arm/hand	220	0.5	40
Hand/finger	180	0.6	75
Thigh/knee	250	0.8	50
Lower leg	170	0.45	60
Feet/toes/joint	160	0.45	--

4 Physical Conditions and Simulation Results

To estimate the speed limit of a robot for transporting construction materials on a worksite, a collaborative masonry robot was used to simulate the collision between a standard cored concrete masonry block and a construction worker. In this simulated scenario, a masonry robot assists construction workers in moving concrete blocks for bricklaying.

4.1 Physical Parameters of Human Body and Robotic Device

To estimate the upper limit of the moving speed of the construction materials carried by the robot, the physical conditions of the human worker and the robot need to be specified in the simulation. In the case of struck and pushed, it is assumed that the block touches the upper arm of the worker. In the case of struck and bent, the block contacts the worker's back. In both cases, two types of concrete mason units (CMUs) were used in the simulation. The dimensions of these blocks were 203.2 mm × 203.2 mm × 406.4 mm (~17 kg) and 203.2 mm × 304.8 mm × 406.4 mm (~25 kg), individually. The worker's height and weight were assumed to be 1.75 m and 90.71 kg, respectively, based on the record of average height and weight of adult males in the United States [23]. According to [25], the upper body includes 55.1% of the total male body mass [24]. The stiffness and the viscous damping of the back surface used in the simulation were 35 N/mm [19] and 100 Ns/mm [25]. The stiffness of the surface on the human's upper arm is 30 N/mm. The stiffness and viscous damping of the waist joint used in the simulation were 366 Nm/rad [26] and 60 Nm/rad [27], individually. Viscous damping has only been partially validated. This value may vary for various reasons, such as age, fatigue level, and physical condition.

According to ISO/TS 15066 (see Table 1), the constant force applied to the back of the human body should not be greater than 250 N. The multiplier of the maximum permissible force during the transient contact is 2, that is, the maximum force of impact should be less than 500 N. Assuming that the deformation of the body surface is within the linear range, the permissible deformation is defined as

$$\text{Permissible deformation} = 2 \times \frac{\text{Allowable force}}{\text{Stiffness}} \quad (10)$$

According to this value, the permissible deformation is 14.3 mm if an unexpected contact does occur on the worker's back at the construction site. The constant force applied to the upper arm of the human body should not be greater than 190 N. The multiplier of the maximum permissible force during the transient contact is 2. This means the maximum impact force should be less than 380 N

N. Correspondingly, the permissible deformation is 12.7 mm if an unexpected contact does occur on the worker's upper arm. In the case of struck and pushed, the human worker might be moved by the impact. Whether the human worker can be moved depends on the friction between the shoes and the floor, which depends on the level of the striking force. The coefficient of static friction between a rubber shoe sole and dry ceramic floor is between 0.8 to 1.2. On wet floors, this value changes to ~0.3 [28]. In this study, 0.9 was used as the static friction coefficient.

4.2 Simulated Results of Struck and Pushed

In the case of struck and pushed, the 17 kg and 25 kg CMUs collided with the human worker from the upper arm on the side of the worker's body. Figure 5 and Figure 6 demonstrate the transient responses of the two individual impacts. The interactions shown in the figures started from the initial contact between the moving object and the human worker, to the time the object detached from the human body. The worker was not moved until the contact force was greater than the static friction force. Once the worker was moved by the CMU block, the contact force existed until the block was separated from the worker. With the permissible deformation being specified as 12.7 mm, the maximum velocities of 17 kg and 25 kg blocks are 760 mm/s and 680 mm/s. The contact forces in both scenarios are around 380 N, and the worker was moved away from the original location.

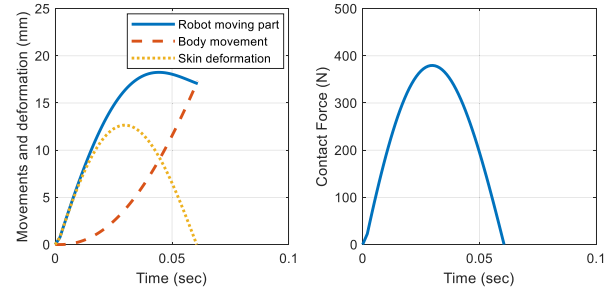


Figure 5. Struck and pushed between the worker and the CMU (17 kg) with 760 mm/s at contact.

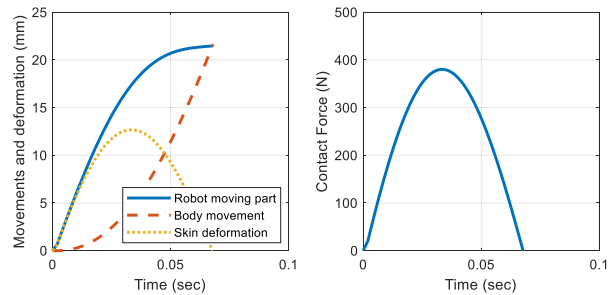


Figure 6. Struck and pushed between the worker and the CMU (25 kg) with 680 mm/s at contact.

4.3 Simulated Results of Struck and Bent

In this case, the 17 kg and 25 kg CMUs collided with the worker from behind on the worker's back. Figure 7 and Figure 8 demonstrate the transient responses of these collisions. With ~ 14 mm of the permissible deformation on the back, the maximum allowed contact velocity is 700 mm/s for a 17 kg block and 570 mm/s for a 25 kg block. In both cases, the maximum interaction force was around 490 N. In this simulation, it was assumed that the worker's foot remains at the same location without moving. All of the energy was absorbed by the viscous damping and bending of the worker's upper body.

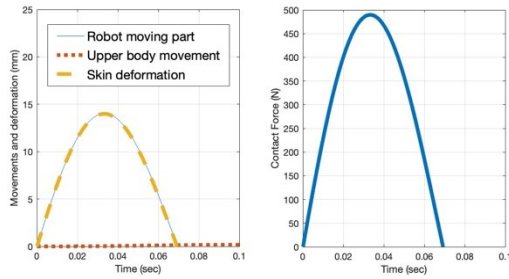


Figure 7. Struck and bent between the worker back and the CMU (17 kg) with 700 m/s at contact.

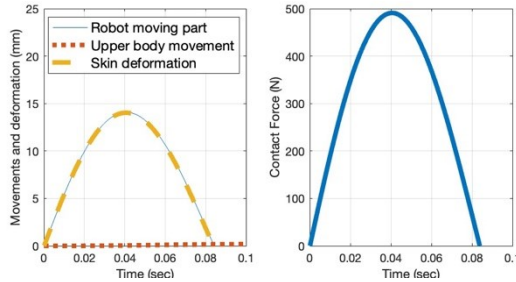


Figure 8. Struck and bent between the worker back and the CMU (25 kg) with 570 m/s at contact.

4.4 Suggested Operation

From the simulation results, the contact forces applied on the body surface and the corresponding deformation can be derived from the contact velocity \dot{x}_B . As the contact velocity increases, the chance of deformation and potential injuries both increases. The proposed model can be used to estimate the operating speeds of robotic devices while transporting materials and collaborating with construction workers on site. In this simulated environment, a male worker was working on a dry construction site. According to the simulated results, the case of struck and bent requires the payload of the robot to move at slower speeds to avoid greater contact force than the case of struck and pushed. Thus, the maximum speed of the masonry robot should be less than 700 mm/s if the robot is carrying a 17 kg block. The

maximum speed is 570 mm/s if the robot is carrying a 25 kg block. When the robot operates at the recommended speeds, it can ensure that the body surface deformation is less than the permissible value if collision does occur. Potential injury from unexpected contact or collision can thus be adequately prevented.

Table 2. Calculated deformations on body surface and contact forces applied on human body.

M_B	Struck and Bent (Back)				Struck and Pushed (Upper Arm)			
	\dot{x}_B	Δs	F_c		\dot{x}_B	Δs	F_c	
		17 kg	25 kg			17 kg	25 kg	
200	4	140	4.9	172	2.2	65	2.3	70
300	6	210	7.4	258	3.9	117	4.3	128
400	8	280	9.8	345	5.7	172	6.4	192
500	10	350	12.3	431	7.6	229	8.6	258
600	12	420	14.8	517	9.5	286	10.9	326
700	14	490	17.2	603	11.5	344	13.1	394
800	16	560	19.7	689	13.4	403	15.4	463
900	18	630	22.1	775	15.4	462	17.7	532
1000	20	700	24.6	861	17.4	521	20.1	602

\dot{x}_B (mm/s), Δs (mm), F_c (N)

5 Conclusion

This study presents a simulation approach based on equations of motion to analyze the deformation of the human body during impact while using a construction collaborative robotic device. Appropriate operating speeds can be estimated based on the proposed dynamic models of human-robot interaction. This approach can provide reference values for safely operating robots on construction sites. In the future, dynamic models for different scenarios and different operational environments could be developed, such as rainy weather and muddy ground.

References

- [1] Pan M. and Pan W. Stakeholder perceptions of the future application of construction robots for buildings in a dialectical system framework. *Journal of Management in Engineering*, 36(6): 04020080, 2020.
- [2] Melenbrink N., Werfel J., and Menges A. On-site autonomous construction robots: Towards unsupervised building. *Automation in construction*, 119: 103312, 2020.
- [3] Liang C. J., Wang X., Kamat V. R., and Menassa C. C. Human-robot collaboration in construction: classification and research trends. *Journal of Construction Engineering and Management*, 147(10): 03121006, 2021.

- [4] Follini C., Terzer M., Marcher C., Giusti A., and Matt D. T. Combining the robot operating system with building information modeling for robotic applications in construction logistics. In *Proceedings of the International Conference on Robotics in Alpe-Adria-Danube Region (RAAD)*, page 245-253, Kaiserslautern, Germany, 2020.
- [5] Lim S., Buswell R. A., Le T. T., Austin S. A., Gibb A. G., and Thorpe T. Developments in construction-scale additive manufacturing processes. *Automation in construction*, 21: 262-268, 2012.
- [6] Paolini A., Kollmannsberger S., and Rank E. Additive manufacturing in construction: A review on processes, applications, and digital planning methods. *Additive manufacturing*, 30: 100894, 2019.
- [7] Cho Y.K., Kim K., Ma S., and Ueda J. A robotic wearable exoskeleton for construction worker's safety and health. In *Proceedings of Construction Research Congress (CRC)*, page 19-28, New Orleans, LA, 2018.
- [8] Liang C. J., Lundeen K. M., McGee W., Menassa C. C., Lee S. H., and Kamat V. R. A vision-based marker-less pose estimation system for articulated construction robots. *Automation in Construction*, 104: 80-94, 2019.
- [9] Liang C. J., Kamat V. R., and Menassa C. C. Teaching robots to perform quasi-repetitive construction tasks through human demonstration. *Automation in Construction*, 120: 103370, 2020.
- [10] Shafique M. and Rafiq M. An overview of construction occupational accidents in Hong Kong: A recent trend and future perspectives. *Applied Sciences*, 9(10): 2069, 2019.
- [11] Liang C. J. and Cheng M. H. Trends in robotics research in occupational safety and health: a scientometric analysis and review. *International Journal of Environmental Research and Public Health*, 20(10): 5904, 2023.
- [12] Liu H. and Wang L. Gesture recognition for human-robot collaboration: A review. *International Journal of Industrial Ergonomics*, 68: 355-367, 2018.
- [13] Ragaglia M., Zanchettin A. M., and Rocco P. Trajectory generation algorithm for safe human-robot collaboration based on multiple depth sensor measurements. *Mechatronics*, 55: 267-281, 2018.
- [14] Scimmi L.S., Melchiorre M., Troise M., Mauro S., and Pastorelli S. A practical and effective layout for a safe human-robot collaborative assembly task. *Applied Sciences*, 11(4): 1763, 2021.
- [15] Cheng M. H. Real-time adjustment of moving trajectories for collaborative robotic devices. presented at the 2022 National Occupational Injury Research Symposium, Morgantown, WV.
- [16] Müller R., Vette M., and Geenen A. Skill-based dynamic task allocation in human-robot-cooperation with the example of welding application. *Procedia Manufacturing*, 11: 13-21, 2017.
- [17] Brosque C., Galbally E., Khatib O., and Fischer M. Human-robot collaboration in construction: Opportunities and challenges. In *Proceedings of the International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, pages 1-8, Ankara, Turkey, 2020.
- [18] American National Standard for industrial robots and robot systems – Safety requirements, ANSI Standard R15.06, 2012.
- [19] Robots and robotic devices – Collaborative robots, ISO Technical Specification 15066, 2016.
- [20] Reinhardt D., Haeusler M. H., London K., Loke L., Feng Y., de Oliveira Barata E., Firth C., Dunn K., Khean N., Fabbri A., Wozniak-O'Connor D., and Masuda R. CoBuilt 4.0: Investigating the potential of collaborative robotics for subject matter experts. *International Journal of Architectural Computing*, 18(4): 353-370, 2020.
- [21] Unfallversicherung D.G. *BG/BGIA Risk Assessment Recommendations According to Machinery Directive: Design of Workplaces with Collaborative Robots*, Institute for Occupational Safety and Health of the German Social Accident Insurance, Sankt Augustin, Germany, 2009.
- [22] Lampus, Mule Lifting Systems by Construction Robotics. Online: <https://www.lampus.com/Product-MULE-Lifting-System>, Accessed: 31/10/2023.
- [23] Fryar C. D., Carroll M. D., Gu Q., Afful J., and Ogden C. L. Anthropometric reference data for children and adults: United States, 2015–2018. *National Center for Health Statistics, State 3*, 46, Washington, D.C., 2021.
- [24] Plagenhoef S., Evans F. G., and Abdelnour T. Anatomical data for analyzing human motion. *Research Quarterly for Exercise and Sport*, 54: 169-178, 1983.
- [25] de Leva P. Adjustments to Zatsiorsky-Seluyanov's Segment Inertia Parameters. *Journal of Biomechanics*, 29(9): 1223-1230, 1996.
- [26] Farley C. T. and Morgenroth D. C. Leg stiffness primarily depends on ankle stiffness during human hopping. *Journal of Biomechanics*, 32: 267-273, 1999.
- [27] Desplantez A., Cornu C., and Goubel F. Viscous properties of human muscle during contraction. *Journal of Biomechanics*, 32: 555-562, 1999.
- [28] Mohamed M. K., Samy A. M., and Ali W. Y. Friction Coefficient of Rubber Shoe Soles Sliding Against Ceramic Flooring. *KGK*, 64(4): 44-49, 2011.

Pre-trained language model based method for building information model to building energy model transformation at metamodel level

Zhichen Wang¹ Mario Bergés¹ Burcu Akinci¹

¹Department of Civil and Environmental Engineering, Carnegie Mellon University, United States

zhichenw@andrew.cmu.edu, marioberges@cmu.edu, bakinci@cmu.edu

Abstract -

Building energy model (BEM) creation based on building information models (BIM) can save model re-creation time during building design phase. However, current BIM-to-BEM transformation is at the model level so the BEM is re-generated every time when the change happens to the BIM. Since design changes happen frequently and the generated BEM needs fine-tuning, such model regeneration is still time-consuming. Mapping rules between BIM and BEM are needed to achieve component level transformation, so that only the corresponding part in BEM instead of whole BEM are updated when changes happen to the BIM. These mapping rules can be defined explicitly using model transformation languages. However, these rule-based transformation methods have limitations in scalability. To solve this issue, this study proposes a pre-trained language model (PLM) based method to construct the mapping relationships between BIM and different types of BEM at the metamodel level. In summary, we formulate the BIM-BEM mapping as a machine translation task and solve it using PLM. For evaluation we collected and generated 35 pairs of BIM and BEM metamodels, and these metamodels are preprocessed into formatted texts that are readable by PLM. The 82% matching accuracy is achieved by proposed method which is higher than the 61% accuracy achieved by a baseline model in previous work. This paper shows the potential to utilize PLMs to facilitate the BIM to BEM transformation from BIM to a varying type of BEMs at the metamodel level. Future work will be focus on the realizing instance-level mapping and transformation.

Keywords -

Model transformation, Building information model, Building energy model, Interoperability, Natural language processing, Pre-trained language model

1 introduction

According to the report by International Energy Agency[1], the building sector accounted for 30% of global energy consumption and 27% of total energy sector emissions in 2021. Building energy models (BEMs) are utilized to predict energy consumption of the building and to improve its energy performance. However, traditional way to create BEM has some limitations: firstly, BEM modelers need to manually transform or rebuild the building geometry [2]. Secondly, the architects and BEM modelers are relatively separate in traditional BEM creation, which causes change conflicts [3, 4] and the discrepancy between architectural and thermal views [5, 6]. Therefore,

building information model (BIM) based BEM modelling is proposed by prior studies since it can help automate the BEM modelling process so that the cost and re-modelling time can be saved [7, 8, 9], for example, Bazjanac [10] identified the potential time savings of 75% for creating the geometry of small and medium buildings through the appropriate application of automated processes.

Existing studies related to BIM-to-BEM transformation [11, 9, 12, 13, 5, 14] mainly focus on how to extract information from source BIM model and convert them into the format or description required by target BEM model. This BIM-to-BEM transformation process usually happens at the model level by generating the new BEM model based on the information from the corresponding BIM model. However, considering the fact that design changes frequently happen during the design phase, such model-level transformation means that the BEM is regenerated when the designer wants to update BEM based on the updated BIM, which results in the risk of losing previous BEM fine-tuning effort and expert knowledge (Seen Figure 1). Therefore, the mapping relationship at the component level should be constructed between BIM and BEM to realize the updating of necessary part of BEM instead of regenerating the whole BEM model. BIM-to-BEM transformation and mapping rules in current studies [15, 16, 17] rely on the manual definition with expert knowledge, but there are two main limitations. Firstly, existing methods require the explicit definition of rules between each class definition in BIM and BEM, and hence the number of rules needed to be defined is enormous due to large amounts of classes defined in BIM and BEM. For example, there more than 770 entities in IFC schema [18] and more than 500 classes in Modelica buildings library [19]. Second, different types of BEM, such as EnergyPlus and Modelica, are used in practice, and the transformation rules explicitly defined are specific to a BEM, so the scalability of these methods can be limited.

To overcome these issues, the research described in this paper utilizes an example-based method to learn transformation rules from pairs of BIM and BEM. Generally speaking, the model transformation process happens at two levels: instance and metamodel. The transformation

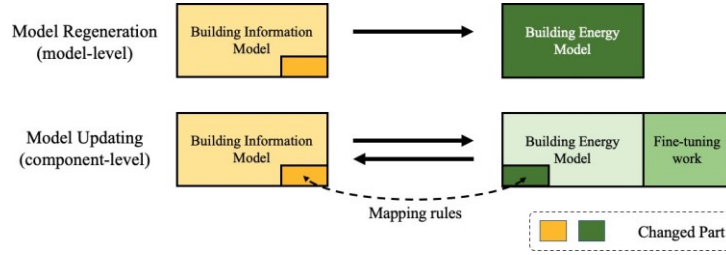


Figure 1. Model transformation at model level and component level

at the instance level focuses on the individual instantiated components while the one at the metamodel level focuses on the transformation between classes defined in the BIM and BEM models. Since the metamodels serve as the basis for instances, this paper will focus on transformations at the metamodel level. The proposed method regards the model transformation at the metamodel level as a text-to-text task. Therefore, a pretrained large language model, T5 (Text-to-text transfer transformer), is adopted and fine-tuned to assess the applicability of large language models in streamlining model transformation. The input is the metamodel of IFC4 files of building systems in JSON format, while the output is the corresponding metamodels of simulation models, including Modelica models, EnergyPlus models and CFD models, in JSON format. In the implementation, both the pretrained T5 model and baseline model from a prior study [20] are trained and evaluated based on the dataset containing 47 pairs of BIM and BEM metamodels with more than 2500 tokens. The proposed method achieves an accuracy of 84% in the token matching task on the test dataset, which is higher than the 61% accuracy achieved by the baseline model. This highlights the applicability for large language models in helping with the model transformation and improving the scalability of model transformation between BIMs and various types of BEMs at the metamodel level.

2 Related work

We review the related work in two sub-fields: (1) BIM and BEM interoperability; (2) Model transformation in software engineering.

2.1 BIM and BEM interoperability

Based on the coupling relationships between building design tools and building simulation tools, current BIM-BEM interoperability approaches can be categorized into three types: centralized, distributed, and combined [21, 22]. Centralized BIM-BEM interoperability approaches utilize a central database, a file format or a schema such as BIM and then share it with other simulation tools. For example, OpenStudio[23] uses gbXML as the

central schema, and Simergy [24] uses IFC as the central data schema for data exchange and model interoperability. Distributed approaches couple pairs of design tool and simulation tools through a middleware specific to them so that a point-to-point network is created between different design tools and simulation tools. The common example of using the distributed approach is to import the geometry model from design tools such as Revit or SketchUp to simulation tools such as EnergyPlus through a middleware. The combined approach usually packages design tools and simulation tools together so the models created by these tools communicate with each other internally, such as IESVE[25]. Compared with the other two types, centralized approaches enable high levels of customization, and usability of the central model can vary depending on the linked simulation model tools[22, 26]. Because of these advantages, existing studies have largely focused on centralized interoperability approaches [15, 27, 28]. Therefore, in this paper we focus on the centralized methods for BIM and BEM interoperability, and leverage the most popular BIM representation, namely IFC [29].

Recent studies related to transforming BIMs to different types of BEMs are shown in the Table 1. The column of “Unique Requirement” shows the necessary information or data that are required by the transformation from BIM to each type of BEM. Meanwhile, items listed in this column are not shared by all other types of BEM due to the differences between the simulation engine or assumptions adopted by each type of BEM. For instance, the CFD model requires the mesh to enable the model creation and calculation while others do not [12, 13]. The Modelica, as the object-oriented simulation, requires the topology between instantiated components[11, 9]. The reviewed literature shows that the proposed BIM-to-BEM methods in current studies are limited to specific type of BEM models. However, practical application facts that various types of BEMs are widely adopted means that the BIM to BEM transformation method should have enough scalability to facilitate the transformation process for lower cost [30, 31]. Therefore, the low scalability is one common gap existing in the current BIM-to-BEM area. [32]

Table 1. BIM-to-BEM transformation literature summary

Type	Input	Output	Unique Information Requirement	Ref.
Object-oriented Model	IFC	Modelica	Topology;	[11, 9, 33, 34, 32]
Computational Fluid Dynamics	IFC	OpenFoam	Space boundary; Mesh; Geometry representation;	[12, 13]
Equation-based Model	IFC; gbXML; IDD;	EnergyPlus	HVAC system specification; Geometry representation; Load information;	[34, 5, 30, 14, 35]

2.2 Model transformation

Model transformation originates from the software engineering discipline [36], referring to the process of converting or mapping a model from one representation to another. The model is defined as the abstraction of a system or environment, and it can be a program code, UML model, or data schema [37]. The main intended applications of model transformation include (1) Model mapping or synchronizing at the same level or different levels of abstraction [38]; (2) Model evolution or refactoring [39]. The model transformation usually consists of three components: source model, target model and transformation rule [36, 32, 40]. To enable automatic model transformation, several model transformation languages are developed to define the transformation rules. Existing model transformation languages (MTLs) can be categorized into declarative MTL [41, 42], imperative MTL [43], and hybrid MTL [44, 45].

Model transformation languages explicitly define the transformation rules for each comparable pair of classes between source model and target model. However, considering the number of classes and entities in BIM and different types of BEMs, explicit definition brings high cost and limits in scalability [20, 46].

A previous study [47] shows that it is easier for experts to show transformation examples than to express complete and consistent transformation rules, so model transformation by example (MTBE) has since been more thoroughly explored. MTBE approaches can be categorized into search-based methods and machine learning based methods. The search-based methods regard the MTBE as an optimization problem and try to find the optimal one in provided mapped pairs [48, 49]. The machine learning based methods utilize machine learning models to learn the mapping rules based on provided mapped pairs [50, 20, 51].

The search-based MTBE approaches have limitations in requiring the mapping traces or defined transformation rules between example pairs, while the machine learning based MTBE shows potential to save cost and improve scalability without need for mapping traces.

In summary, two identified issues need to be mitigated for automated updating of BEMs. The first one is that there is a need to automate the definition of transformation rules between BIM and BEM, as this process can be time-consuming and costly due to the large number of entities and classes defined in IFC and the different BEM models. The second issue is that these transformation rules are specific to the different BEMs adopted in practice, including Modelica, CFD and EnergyPlus, requiring customization for each case.

In the next section, we introduce the machine learning based approach that we propose to learn the embedding mapping rules between BIM and different types of BEM at the metamodel level, so that explicit definition of transformation rules can be avoided and the scalability of the model transformation methods can be improved.

3 Method

This paper focuses on the model transformation at the metamodel level. The definition of metamodel and instances in this paper are the class defined in the schema and the entities instantiated from the defined classes respectively. Instead of directly using the entire schema containing all classes, we adopted portions of the schema (e.g. IFC4 schema [18]) or the class definition (e.g. Modelica Buildings Library [19]) as the metamodels. Metamodels of BIM and BEM are highly structured with defined classes and references among classes, so both of them can also be converted into structured text. Therefore, given the structured and textualizable BIM and BEM metamodels as input and output respectively, the model transformation process between BIM metamodel and BEM metamodel can be regarded as the text-to-text translation task, which is a typical natural language processing (NLP) task.

Considering the significant progress achieved in pre-trained language models (PLMs) in recent years, in this paper we adopt a pretrained PLM, namely T5 (Text-to-Text Transfer Transformer) [52], as the base model to realize the transformation between metamodels of BIM and BEM. The T5 is an encoder-decoder model pre-trained on a multi-task mixture of unsupervised and supervised tasks

with the idea of reframing all NLP tasks into a unified text-to-text-format, within which the input and output are always text strings [52]. Examples of tasks are machine translation, document summarization, question answering, and sentiment analysis.

The overview of the T5-based method for model transformation can be seen in Fig.2. The input of the proposed method is the metamodel of IFC, while the output is the metamodels of different types of BEMs. All these metamodels are created by following the standard schema or class definitions that are made publicly available, and they are represented in abstract syntax trees (AST) in the JSON format. The AST of metamodels follows the binary root-children structure which is proven to be effective for code translation by Chen et al[53]. The preprocessing block converts the JSON format AST of BIM metamodel into text, which is then sent to the T5 model. After the text generation, the post-processing block converts the generated text into JSON format AST of the output BEM metamodel. The pretrained T5 model is fine-tuned based on a dataset containing pairs of IFC and BEM metamodels, so that the implicit rules are learnt.

4 Implementation and results

To implement our proposed method we first need to obtain a robust set of samples to provide as input and output to our framework. Specifically, we need a collection of metamodel pairs (BIM as input and corresponding BEMs as output). Since, to the best of our knowledge, this dataset of metamodels does not exist (or at least not publicly), we resorted to creating the dataset ourselves following standard schemas and online documentation for the model class families we used. In particular, we created 44 metamodel pairs, all stemming from model instances representing a specific building system in a commercial building at Anonymous Location. The BIM metamodels in the implementation are IFC metamodels, while the BEM metamodels are metamodels of Modelica, CFD and Energyplus models.

As mentioned in Section 6, the metamodels of BIM and BEMs are converted into text files. The metamodels are initially represented using UML because it is suitable for visualization and manual operations. Subsequently, these metamodels are manually converted into an AST format and realized as JSON files for machine readability. Finally, we used Python to process these JSON-formatted metamodels into text files for tokenization. Figure 3 shows this conversion process using an example pair of IfcBoiler class and Boiler class in EnergyPlus. The AST of metamodels are created based on the metamodels visualized in the UML format, then the JSON formatted AST of metamodels are converted into plain text following given rules. Following above procedure, we created 44 pairs of BIM

Table 2. Implementation configuration

Item	Value
Adopted model	T5-small (60 million parameters)
Baseline	LSTM encoder-decoder model [20]
Number of epochs	100
Batch size	4
Learning rate	0.002
Train dataset	1500 tokens, 30 model pairs
Test dataset	660 tokens, 7 model pairs
Validation dataset	680 tokens, 7 model pairs

and BEM metamodels (Accessed through this link) divided them with the ratio of 70%, 15%, and 15%. The resulting training dataset contains 30 model pairs with around 1500 tokens, while both the validation and test dataset contain 7 model pairs with more than 650 tokens. As for the implementation of the pretrained language model, the T5-small model from HuggingFace transformer API [54] is adopted as the base model, then it is fine-tuned based on the collected metamodel pairs. The model training process is implemented under Google Colaboratory environment with 100 training epochs. Meanwhile, as a baseline, we benchmark against a LSTM neural network model from a previous study [20]. The baseline model also comprises the encoder-decoder architecture, but the encoder and decoder are implemented with a single-layer LSTM network [20]. A summary of the implementation configuration is shown in Table 2.

The implementation results are evaluated by the accuracy of token matching on the test dataset. Given the input BIM metamodel and target BEM model type, the fine-tuned pretrained large language model generates the corresponding metamodel of BEM, then the accuracy is calculated by:

$$Acc = \frac{N_{matched}}{N_{total}}$$

where $N_{matched}$ is the number of matched token pairs, N_{total} is total the number of token pairs. As shown in the Table 3, when the pretrained T5-small model is trained after 100 epochs, the accuracy of token matching on the test dataset is around 82%, while the baseline model achieved 61% accuracy. Figure 5 shows examples of token matching results. The tokens highlighted in bold are mismatched pairs. The mismatched pairs can be due to the type of attribute value or the name of attributes belonging to the given class.

5 Discussion

As shown in the previous section, our T5 pretrained model achieved an 82% token matching accuracy in the test set, which is higher than 61% accuracy achieved by the baseline model from [20]. This result points to the promise

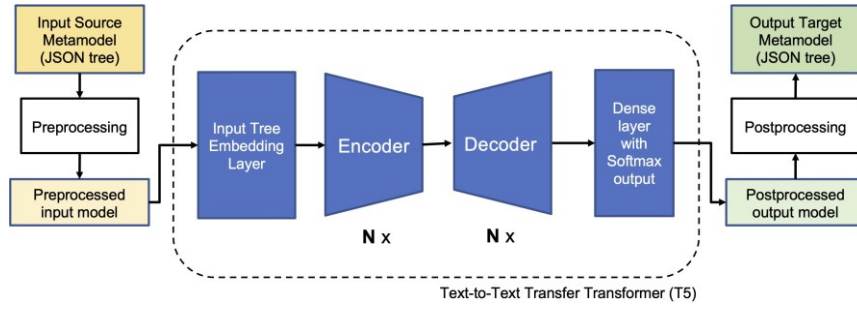


Figure 2. Proposed example-based model transformation method

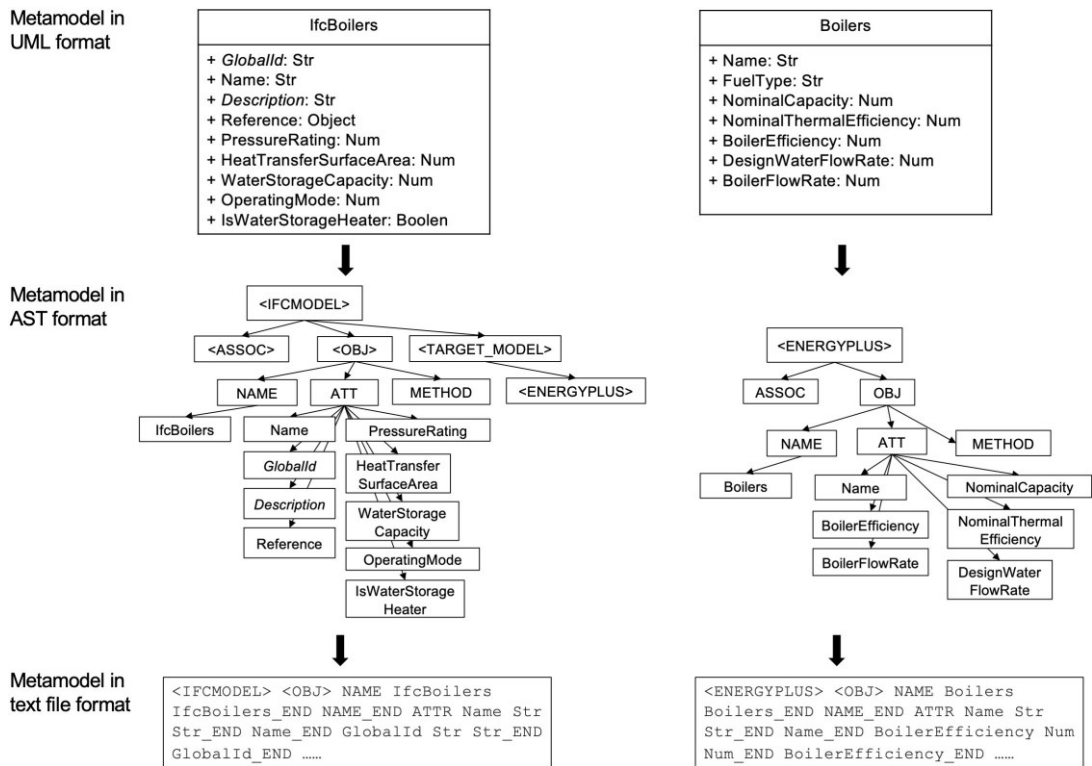


Figure 3. Workflow of converting the metamodels to text files

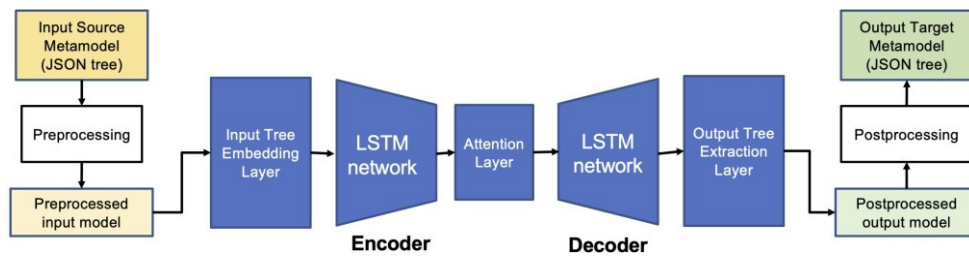


Figure 4. The workflow of LSTM based model transformation method from [20] as the baseline

Example 1	Ground Truth	<CFDModel> <OBJ> NAME CFDModel CFDModel_END NAME_END ATTR Volume Num Num_END Volume_END ThermalCapacity Num Num_END <unk>ASSOC> Composition Composition_END RoomGeometry RoomGeometry_END Wall Wall_END <unk>ASSOC>_END <CFDModel>_END
	Prediction	<CFDModel> <OBJ> NAME CFDModel CFDModel_END NAME_END ATTR Volume Num Num_END Volume_END ThermalCapacity Num Num_END <ASSOC> Composition Composition_END RoomGeometry RoomGeometry_END Wall Wall_END <ASSOC>_END <CFDModel>_END
Example 2	Ground Truth	<ENERGYPLUSModel> <OBJ> NAME ChillerHeater ChillerHeater_END NAME_END ATTR Name Str Str_END Name_END NominalCoolingCapacity Real Real_END
	Prediction	<ENERGYPLUSModel> <OBJ> NAME ChillerHeater ChillerHeater_END NAME_END ATTR Name Str Str_END Name_END NominalCoolingCapacity Num Num_END
Example 3	Ground Truth	<MODELICA> <OBJ> NAME ATT res2 Real Real_END res2_END flowCharacteristics3 Real Real_END flowCharacteristics3_END res1 Real Real_END res1_END fraK
	Prediction	<MODELICA> <OBJ> NAME ATT flowCharacteristics3 Real Real_END flowCharacteristics3_END res1 Real Real_END res1_END res2 Real Real_END res2_END fraK

Figure 5. Examples of token matching results

#Epoch	Metric	Proposed	Baseline
30	Accuracy of token matching	0.21	0.22
100	Accuracy of token matching	0.82	0.61

of using model-based MTBE methodology to automate model transformation between BIM and BEM. In particular, it shows that we may not need domain-specific expert knowledge to define every transformation rule among entities and classes in the metamodels of BIM and BEMs. Concurrently, the incorporation of metamodels from three types of BEM within the training and testing datasets highlight the capability of the trained algorithm to fulfill the practical application demand of transforming the BIM into multiple types of BEM.

That said, the results also highlight that there are shortcomings to this method. Though we did not conduct experiments to fully characterize the failure cases, we can hypothesize about the possible reasons for the mismatched pairs of tokens highlighted in Figure 5. For example, as the examples in the figure show, the errors could be attributed to the characteristics of the training dataset and the representation format of metamodels. Firstly, for the mismatched pair of “eal” and “Num” in Example 2, these tokens stand for the data type of the attribute in those classes. In the training dataset, the occurrence frequency “Num” surpasses 80%, while the “Real” accounts for less than 10% of the total number of tokens representing attribute data type. As a result, the model tends to predict the “Num” with higher probability. To mitigate this issue, diversifying the data sources can be explored to prevent the dominance of a specific data value in the dataset. Secondly, for the mismatched pairs of “res2”, “flowCharacteristics3” and “res1” in Example 3, all of them are attributes of class Valves.ThreeWayTable in Modelica Buildings library. Lastly, as mentioned in Sections 6 and 4, the tree structured metamodel is represented using a sequential

structure in plain text (AST). However, in the model prediction results, the generated “flowCharacteristics3” appears at the position of “res2” in the ground-truth. Since token matching is performed based on sequential positions, this pair is considered mismatched. One possible solution for this issue is to improve the representation format of the tree structure of metamodels in text file to avoid the impact of altered sequential positions of predicted tokens.

One significant difficulty met in the implementation of the proposed method is the acquisition and collection of data. There is a lack of open-source repositories providing metamodels of different types of BEMs, so the metamodels in the dataset were manually constructed by the authors, adhering to the online standard data schema and official documents. This limited dataset potentially exerts an impact on the performance of the fine-tuned large language model.

A further limitation of the PLM-based approach for BIM-to-BEM transformation pertains to the intricacies of instance-level transformation and mapping. This process necessitates the intricate alignment of specific attribute values of instances in BIM and BEM. However, these attribute values may exhibit randomness or be heavily influenced by contextual factors, leading to a scenario where the dataset collected for model fine-tuning does not cover all potential attribute values. For example, when considering a BIM-BEM pair that includes multiple room instances, these room instances vary in attributes such as their names, volumes, locations, and topological relationships with other instances. To address the challenge of instance-level transformation, one viable strategy involves the integration of constraints and rules that are specifically tailored to the attribute values. This approach would facilitate the transformation of instances with attribute values, from one instance from BIM to the other instance from BEM, after the instance class have been initially filtered through the metamodel transformation results derived from the PLM-based method.

6 Conclusion

Existing methods for model transformation between BIM and BEM are circumscribed in terms of cost and scalability. To mitigate these issues, this paper starts from the model transformation process at the metamodel level, and conceptualizes it as a text-to-text translation task. Subsequently, the pretrained large language model based MTBE approach is proposed to actualize the automated BIM to BEM model transformation at the metamodel level. To best of our knowledge, this paper is the first to apply the large language model based methods in the research area of BIM to BEM transformation. The implementation results envince the feasibility of the proposed method for BIM to BEM transformation. The contribution of this paper resides in (1) highlighting the applicability for large language models in helping with the model transformation and (2) improving the scalability by enabling the model transformation between BIM and various types of BEMs. To mitigate the issues of mismatched pairs observed the implementation result, the future work will focus on diversifying the data sources and improving the representation of metamodels in plain text. Meanwhile, the PLM-based method proposed in this paper is primarily concentrated on transformations at the metamodel level. Nonetheless, it encounters limitations when applied to transformations at the instance level, due to the necessity of converting specific attribute values for individual instances. Consequently, future work will be investigating methodologies for the incorporation of constraints and rules. This is aimed at facilitating the conversion of attribute values pertinent to instance-level transformation, thereby potentially enhancing the efficacy and applicability of the PLM-based approach in more complex, attribute-specific scenarios within the BIM to BEM transformation at all levels.

References

- [1] International Energy Agency. Iea building energy report, 2022. URL www.iea.org/reports/buildings.
- [2] Georgios N. Lilis, Georgios I. Giannakis, and Dimitrios V. Rovas. Automatic generation of second-level space boundary topology from IFC geometry inputs. *Automation in Construction*, 76:108–124, April 2017. ISSN 09265805. doi:10.1016/j.autcon.2016.08.044.
- [3] Vladimir Bazjanac, Tobias Maile, Cody Rose, James T O'Donnell, Elmer Morrissey, and Benjamin R Welle. AN ASSESSMENT OF THE USE OF BUILDING ENERGY PERFORMANCE SIMULATION IN EARLY DESIGN. In *Proceedings of Building Simulation*, Sydney, Australia, 2011.
- [4] Benjamin Welle, John Haymaker, and Zack Rogers. ThermalOpt: A methodology for automated BIM-based multidisciplinary thermal simulation for use in optimization environments. *Building Simulation*, 4(4):293–313, December 2011. ISSN 1996-3599, 1996-8744. doi:10.1007/s12273-011-0052-5.
- [5] Robert J Hitchcock, Justin Wong, and Hitchcock Consulting. TRANSFORMING IFC ARCHITECTURAL VIEW BIMS FOR ENERGY SIMULATION: 201.
- [6] Vladimir Bazjanac. IFC BIM-based methodology for semi-automated building energy performance simulation. Technical report, Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), Berkeley, CA (United States), 2008.
- [7] Mohamed H. Elnabawi. Building Information Modeling-Based Building Energy Modeling: Investigation of Interoperability and Simulation Results. *Frontiers in Built Environment*, 6:573971, December 2020. ISSN 2297-3362. doi:10.3389/fbuil.2020.573971.
- [8] Gabriela Bastos Porsani, Kattalin Del Valle de Lersundi, Ana Sa'nchez-Ostiz Gutie'rrez, and Carlos Ferna'ndez Bandera. Interoperability between Building Information Modelling (BIM) and Building Energy Model (BEM). *Applied Sciences*, 11(5):2167, March 2021. ISSN 2076-3417. doi:10.3390/app11052167.
- [9] WoonSeong Jeong and Kee Kim. A Performance Evaluation of the BIM-Based Object-Oriented Physical Modeling Technique for Building Thermal Simulations: A Comparative Case Study. *Sustainability*, 8(7):648, July 2016. ISSN 2071-1050. doi:10.3390/su8070648.
- [10] Vladimir Bazjanac. SPACE BOUNDARY REQUIREMENTS FOR MODELING OF BUILDING GEOMETRY FOR ENERGY AND OTHER PERFORMANCE SIMULATION.
- [11] Ando Andriamamonjy, Dirk Saelens, and Ralf Klein. An automated IFC-based workflow for building energy performance simulation with Modelica. *Automation in Construction*, 91:166–181, July 2018. ISSN 09265805. doi:10.1016/j.autcon.2018.03.019.
- [12] Minhyung Lee, Gwanyong Park, Hyangin Jang, and Changmin Kim. Development of Building CFD Model Design Process Based on BIM. *Applied Sciences*, 11(3):1252, January 2021. ISSN 2076-3417. doi:10.3390/app11031252.
- [13] Eric Fichter, Veronika Richter, Je'ro'me Frisch, and Christoph van Treeck. Automatic generation of second level space boundary geometry from IFC models. In *2021 Building Simulation Conference*, September 2021. doi:10.26868/25222708.2021.30156.
- [14] Oliver Spielhaupter. *BIM to BEM Transformation Workflows: A Case Study Comparing Different IFC-Based Approaches*. PhD thesis.
- [15] Issa J. Ramaji, John I. Messner, and Ehsan Mostavi. IFC-Based BIM-to-BEM Model Transformation. *Journal of Computing in Civil Engineering*, 34(3):04020005, May 2020. ISSN 0887-3801, 1943-5487. doi:10.1061/(ASCE)CP.1943-5487.0000880.
- [16] Jong Bum Kim, WoonSeong Jeong, Mark J Clayton, Jeff S Haberl, and Wei Yan. Developing a physical bim library for building thermal energy simulation. *Automation in construction*, 50:16–28, 2015.
- [17] Yikun Yang, Yiqun Pan, Fei Zeng, Ziran Lin, and Chenyu Li. A gbXML Reconstruction Workflow and Tool Development to Improve the Geometric Interoperability between BIM and BEM. *Buildings*, 12(2):221, February 2022. ISSN 2075-5309. doi:10.3390/buildings12020221.
- [18] buildingSmart. Industrial foundation classes (ifc) 4.0, 2020. URL https://standards.buildingsmart.org/IFC/RELEASE/IFC4_1/FINAL/HTML/link/annex-e.htm.
- [19] Lawrence Berkeley National Laboratory. Modelica buildings library, 2023. URL <https://simulationresearch.lbl.gov/modelica/>.
- [20] Loli Burguen'o, Jordi Cabot, Shuai Li, and Se'bastien Ge'rard. A generic LSTM neural network architecture to infer heterogeneous model transformations. *Software and Systems Modeling*, 21(1):139–156, February 2022. ISSN 1619-1366, 1619-1374. doi:10.1007/s10270-021-00893-y.
- [21] Goran Sibenik and Iva Kovacic. Assessment of model-based data exchange between architectural design and structural analysis. *Journal of Building Engineering*, 32:101589, November 2020. ISSN 23527102. doi:10.1016/j.jobbe.2020.101589.

- [22] Kristoffer Negendahl. Building performance simulation in the early design stage: An introduction to integrated dynamic models. *Automation in Construction*, 54:39–53, June 2015. ISSN 09265805. doi:10.1016/j.autcon.2015.03.002.
- [23] NREL U.S. Department of Energy. Openstudio, July 10, 2023. URL <https://openstudio.net/>.
- [24] Digital Alchemy. Lbnl simergy, lawrence berkeley natl. lab, 2023. URL <https://d-alchemy.com/products/simergy>.
- [25] Integrated Environmental Solutions. Integrated environmental solutions, iesve, July 10, 2023. URL <https://www.iesve.com/>.
- [26] Hang Li and Jiansong Zhang. Interoperability between BIM and BEM Using IFC. In *Computing in Civil Engineering 2021*, pages 630–637, Orlando, Florida, May 2022. American Society of Civil Engineers. ISBN 978-0-7844-8389-3. doi:10.1061/9780784483893.078.
- [27] Cody M. Rose and Vladimir Bazjanac. An algorithm to generate space boundaries for building energy simulation. *Engineering with Computers*, 31(2):271–280, April 2015. ISSN 0177-0667, 1435-5663. doi:10.1007/s00366-013-0347-5.
- [28] Huaquan Ying and Sanghoon Lee. An algorithm to facet curved walls in IFC BIM for building energy analysis. *Automation in Construction*, 103:80–103, July 2019. ISSN 09265805. doi:10.1016/j.autcon.2019.03.004.
- [29] Shu Tang, Dennis R Shelden, Charles M Eastman, Pardis Pishdad-Bozorgi, and Xinghua Gao. Bim assisted building automation system information exchange using bacnet and ifc. *Automation in Construction*, 110:103049, 2020.
- [30] Hang Li and Jiansong Zhang. Improving IFC-Based Interoperability between BIM and BEM Using Invariant Signatures of HVAC Objects. *Journal of Computing in Civil Engineering*, 37(2):04022059, March 2023. ISSN 0887-3801, 1943-5487. doi:10.1061/(ASCE)CP.1943-5487.0001063.
- [31] Hang Li, Jiansong Zhang, Soowon Chang, and Anthony Sparkling. BIM-based object mapping using invariant signatures of AEC objects. *Automation in Construction*, 145:104616, January 2023. ISSN 09265805. doi:10.1016/j.autcon.2022.104616.
- [32] Jun Cao. *SimModel Transformation Middleware for Modelica-based Building Energy Modeling and Simulation*. PhD thesis.
- [33] Matthias Thorade, Joërg Raßler, Peter Remmen, Tobias Maile, Reinhard Wimmer, Jun Cao, Moritz Lauster, Christoph Nytsch-Geusen, Dirk Müller, and Christoph van Treeck. An Open Toolchain for Generating Modelica Code from Building Information Models. In *The 11th International Modelica Conference*, pages 383–391, September 2015. doi:10.3384/ecp15118383.
- [34] James O'Donnell, Richard See, Cody Rose, Tobias Maile, Vladimir Bazjanac, and Phil Haves. SIMMODEL: A DOMAIN DATA MODEL FOR WHOLE BUILDING ENERGY SIMULATION.
- [35] Veronika Richter, Eric Fichter, Maximilian Azendorf, Je'rome Frisch, and Christoph van Treeck. Algorithms for Overcoming Geometric and Semantic Errors in the Generation of EnergyPlus Input Files based on IFC Space Boundaries.
- [36] Krzysztof Czarnecki and Simon Helsen. Feature-based survey of model transformation approaches. *IBM systems journal*, 45(3): 621–645, 2006.
- [37] Perdita Stevens. A landscape of bidirectional model transformations. *Generative and Transformational Techniques in Software Engineering II: International Summer School, GTTSE 2007, Braga, Portugal, July 2-7, 2007. Revised Papers*, pages 408–424, 2008.
- [38] Igor Ivkovic and Kostas Kontogiannis. Tracing evolution changes of software artifacts through model synchronization. In *20th IEEE International Conference on Software Maintenance, 2004. Proceedings.*, pages 252–261. IEEE, 2004.
- [39] Gerson Sunye', Damien Pollet, Yves Le Traon, and Jean-Marc Je'ze'quel. Refactoring uml models. In *International Conference on the Unified Modeling Language*, pages 134–148. Springer, 2001.
- [40] Ethan Post, Kevin Dinkel, Erich Lee, Bjorn Cole, Hongman Kim, and Bassem Nairouz. Cloud-based orchestration of a model-based power and data analysis toolchain. In *2016 IEEE Aerospace Conference*, pages 1–12, Big Sky, MT, USA, March 2016. IEEE. ISBN 978-1-4673-7676-1. doi:10.1109/AERO.2016.7500648.
- [41] Yue Cao, Yusheng Liu, Hongri Fan, and Bo Fan. SysML-based uniform behavior modeling and automated mapping of design and simulation model for complex mechatronics. *Computer-Aided Design*, 45(3):764–776, March 2013. ISSN 00104485. doi:10.1016/j.cad.2012.05.001.
- [42] Andy Schurr and Felix Klar. 15 Years of Triple Graph Grammars. In Hartmut Ehrig, Reiko Heckel, Grzegorz Rozenberg, and Gabriele Taentzer, editors, *Graph Transformations*, volume 5214, pages 411–425. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. ISBN 978-3-540-87404-1 978-3-540-87405-8. doi:10.1007/978-3-540-87405-8.28.
- [43] Joel Greenyer and Ekkart Kindler. Comparing relational model transformation technologies: Implementing Query/View/Transformation with Triple Graph Grammars. *Software & Systems Modeling*, 9(1):21–46, January 2010. ISSN 1619-1366, 1619-1374. doi:10.1007/s10270-009-0121-8.
- [44] Marcel van Amstel, Steven Bosems, Ivan Kurtev, and Lu'is Ferreira Pires. Performance in Model Transformations: Experiments with ATL and QVT. In Jordi Cabot and Eelco Visser, editors, *Theory and Practice of Model Transformations*, volume 6707, pages 198–212. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011. ISBN 978-3-642-21731-9 978-3-642-21732-6. doi:10.1007/978-3-642-21732-6.14.
- [45] Dimitrios S. Kolovos, Richard F. Paige, and Fiona A. C. Polack. The Epsilon Transformation Language. In Antonio Vallecillo, Jeff Gray, and Alfonso Pierantonio, editors, *Theory and Practice of Model Transformations*, volume 5063, pages 46–60. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. ISBN 978-3-540-69926-2 978-3-540-69927-9. doi:10.1007/978-3-540-69927-9.4.
- [46] K. Lano, S. Kolahdouz-Rahimi, and S. Fang. Model Transformation Development Using Automated Requirements Analysis, Meta-model Matching, and Transformation by Example. *ACM Transactions on Software Engineering and Methodology*, 31(2):1–71, April 2022. ISSN 1049-331X, 1557-7392. doi:10.1145/3471907.
- [47] U. Behrens, M. Flasiniski, L. Hagge, J. Jurek, and K. Ohrenberg. Recent developments of the ZEUS expert system ZEX. *IEEE Transactions on Nuclear Science*, 43(1):65, February 1996. ISSN 0018-9499, 1558-1578. doi:10.1109/23.486006.
- [48] Islem Baki and Houari Sahraoui. Multi-Step Learning and Adaptive Search for Learning Complex Model Transformations from Examples. *ACM Transactions on Software Engineering and Methodology*, 25(3):1–37, August 2016. ISSN 1049-331X, 1557-7392. doi:10.1145/2904904.
- [49] Zolta'n Balogh and Da'niel Varro'. Model transformation by example using inductive logic programming. *Software & Systems Modeling*, 8(3):347–364, July 2009. ISSN 1619-1366, 1619-1374. doi:10.1007/s10270-008-0092-1.
- [50] Loli Burgueno, Jordi Cabot, and Sebastien Gerard. An LSTM-Based Neural Network Architecture for Model Transformations. In *2019 ACM/IEEE 22nd International Conference on Model Driven Engineering Languages and Systems (MODELS)*, pages 294–299, Munich, Germany, September 2019. IEEE. ISBN 978-1-72812-536-7. doi:10.1109/MODELS.2019.00013.
- [51] Paulius Danenas and Tomas Skersys. Exploring Natural Language Processing in Model-To-Model Transformations. *IEEE Access*, 10:116942–116958, 2022. ISSN 2169-3536. doi:10.1109/ACCESS.2022.3219455.
- [52] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. Exploring the limits of transfer learning with a unified text-to-text transformer. *The Journal of Machine Learning Research*, 21(1): 5485–5551, 2020.

- [53] Xinyun Chen, Chang Liu, and Dawn Song. Tree-to-tree Neural Networks for Program Translation, October 2018.
- [54] HuggingFace. Huggingface transformer api, 2023. URL <https://huggingface.co/docs/transformers/index>.

Computer Vision as Key to an Automated Concrete Production Control

Max Coenen¹, Max Meyer², Dries Beyer¹, Christian Heipke², Michael Haist¹

¹Institute of Building Materials Science, Leibniz University Hannover, Germany

²Institute of Photogrammetry and GeoInformation, Leibniz University Hannover, Germany

[m.coenen, d.beyer, haist]@baustoff.uni-hannover.de, [meyer, heipke]@ipi.uni-hannover.de

Abstract -

The need to reduce CO₂ emissions from concrete leads to increasingly complex mix designs involving e.g. CO₂ reduced cements, recycled materials, and various chemical additives. This complexity results in a larger sensitivity of the concrete to unpredictable fluctuations in both, the base material properties and in boundary conditions such as temperature and humidity during the production process. Digital sensor systems and quality control schemes are considered as key to counteract this problem by enabling an automated production control. As contribution towards this goal, this paper investigates the research question whether Computer Vision can be used for the predictive characterisation of raw materials (here: of concrete aggregates) and of the fresh concrete quality during the mixing process. In particular, we propose the usage of imaging sensors for the observation of both, aggregate material and the flow behaviour of fresh concrete during the mixing process, and present deep learning methods for the prediction of granulometric and rheological properties from the image observations, respectively. Incorporating such systems into the concrete production process enables the facilitation of a digital control loop for ready-mixed concrete production by allowing an in-line reaction to raw material fluctuations and to deviations of the concrete from the target properties.

Keywords -

Concrete 4.0, Computer Vision, Deep Learning, image-based granulometry, image-based rheology, CNN, ViT

1 Introduction

Concrete, which is the most used (building) material worldwide, contributes approximately 8% to the global CO₂-emissions. The urgent need to reduce CO₂-emissions of the concrete production results in increasingly complex concrete mixtures: Historically, being a 3-component-material (cement, water, and aggregates), modern recipes contain a large variety of different components, including e.g. CO₂ reduced cements (with various sub-components), recycled aggregates, and various chemical additives (super-plasticiser, flow-retarder, etc.) [1]. However, the growing complexity of the concrete mixtures results in a strongly increased sensitivity of the concrete properties to unpredictable fluctuations in the raw materials and the production processes, leading to diminished robustness of the concrete, to difficulties in reaching the desired concrete quality and ultimately, to higher rejection rates. These challenges are extended by the fact that until

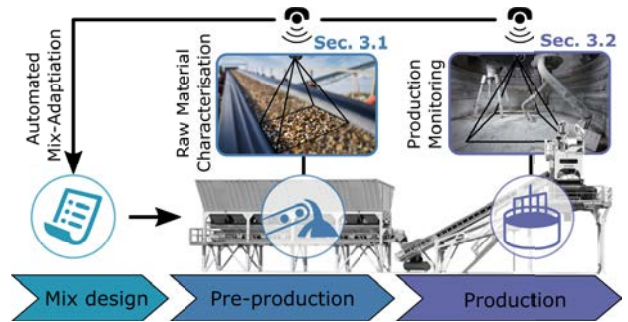


Figure 1. High-level overview on the proposed concrete production control scheme using computer vision based methods for the in-line characterisation of raw materials and fresh concrete quality.

today, the mixture design of the concrete and its production are mostly empirically driven, leading to additional difficulties due to the lack of experience with such new and complex concrete compositions. As a result, the extensive use of new, environmentally and resource-friendly building materials is hampered by the inability of controlling and guaranteeing the desired concrete properties during the production process.

In the authors' opinion, a digital mapping of the essential process steps within the concrete production chain is key for a transition towards a more sustainable high quality concrete construction industry [2]. While in many manufacturing industries, digitisation and automation have enabled a strong increase in productivity over the last decades, the developments in the construction industry stagnated during this period, resulting in the construction industry - and here, especially the concrete sector - to be one of the least digitised industries of the global economy [3]. In the context of the topic addressed by this paper, the lack of digitisation and automation is for instance reflected in the conventional, batch-based, and manual methods that are currently still used as standard for the characterisation of raw materials (e.g. manual sieving for estimating the particle size distribution of aggregates) and the quality control of the concrete production (e.g. slump flow or rheometer tests for determining fresh concrete properties). In order to overcome current limitations and to enable an

automated and improved concrete production, this paper presents computer vision based methods for an in-line monitoring of the raw materials used for concrete production as well as for an in-line quality control of the fresh concrete properties during the mixing process (cf. Fig. 1). Given real-time information about the raw material characteristics before mixing (pre-production) allows to account for deviations from the expected properties by adapting the mix design accordingly, while real-time information on the fresh concrete properties (during production) allows for taking measures to assure that the target quality is reached. In particular, to investigate the suitability of imaging systems for the characterisation of building materials, we make the following contributions in this paper:

- We propose a Vision Transformer (ViT) based method for the visual granulometry of concrete aggregates from image observations.
- We present a Convolutional Neural Network (CNN) based approach for the prediction of fresh concrete properties from images acquired during the mixing process of the concrete and mix design information.
- In the experimental part of this paper, we demonstrate and evaluate the performance of the proposed methods on challenging real-world data sets.

2 Background

2.1 Raw material characterisation

The size distribution (formally known as grading curve) of the aggregates used for concrete production significantly influences many important concrete properties including e.g. the consistency, workability, and segregation tendency of the concrete in the fresh state as well as the compressive strength and the durability in the hardened state [4]. As a consequence, the size distribution of the aggregates has to be considered during mix design (typically by considering the aggregates' *k-value*), since it affects the water demand and the amount of required super-plasticizer of the concrete composition. In practice, however, the size distribution is usually determined on a low-frequency basis by mechanical sieving of small sample batches of aggregates (a few kilograms), which are considered as representative for large amounts (a few tons). As a consequence, deviations of the actual material used for concrete production from the test sample cannot be accounted for.

A camera based sensor setup, observing the material while being transported over the conveyor belt into the mixer, and algorithms that are able to predict the size distribution from the image observations allow to consider the actual aggregate grading curves in the mix design for the production of each concrete batch. In the literature, the derivation of size distributions from images of aggregate

is often approached in a two-step procedure. In this context, the first step consists of a segmentation of individual particles, e.g. based on grayscale thresholding, edge detection, or watershed transformations in early approaches [5, 6], while modern procedures apply deep learning based methods for instance segmentation or panoptic segmentation of aggregate particles [7, 8]. In a second step, the particle size distribution is derived from the segmentation results. However, this procedure suffers from partial occlusions and an often insufficient spatial image resolution, and needs an explicit conversion from the two-dimensional segmentations to volumetric entities [5], introducing inaccuracies for the task of estimating size distributions.

In contrast to the described object-based procedure, statistical approaches avoid the explicit detection and modelling of individual instances by relying on global image statistics in order to predict the size distribution directly from the raw image. In this context, Olivier et al. [9] and Coenen et al. [10] propose to learn a CNN in order to distinguish different predefined particle size distributions. However, in this way, only a classification of a discrete set of grading curves is possible. In order to overcome this limitation, CNN based approaches for the prediction of the continuous percentiles defining the size distribution were presented in [11]. While the latter approaches used CNN-architectures, recently the application of transformer based models for vision tasks has shown great potential and encouraging results [12]. Since it was shown in the literature, that Vision Transformers (ViT) perform equal to and even outperform CNN-based methods for the task of grading curve prediction [13], we follow the recent success of transformer-based models and choose a ViT approach for the determination of concrete aggregate size distributions.

2.2 Fresh concrete characterisation

In current practice, the quality inspection of fresh concrete is mainly conducted offline, i.e. after the mixing and production process, using empirical test methods based on small batch samples taken from the concrete. One of these methods is, for example, the slump flow test [14]. This test involves applying a force to a fresh concrete sample and measuring how much it spreads. The manually measured slump flow diameter δ is an indicator of the consistency of the fresh concrete. However, the manual measurement entails a higher degree of inaccuracy due to possible human error. Furthermore, at this stage of the production process, only very limited control of the concrete properties remains possible. For this reason, an inline quality assessment during the mixing process is desirable, since it enables a real-time reaction on potential deviations from the target properties. Besides, caused by the hydration process of concrete, its properties can change significantly between the time of mixing and the time of the actual

placement of the concrete. Therefore, it is also striven for a prediction of the concrete properties at the time of placement by modelling the time-dependent behaviour of the concrete. In addition to the consistency, the fresh concrete can also be characterised by its rheological parameters, namely the plastic viscosity μ and the yield stress τ_0 of a Bingham model which is typically used to describe the non-Newtonian properties of fresh concrete. Founded on fluid dynamic laws, these properties determine the flow behaviour of the fluid under specific boundary conditions, e.g. during a mixing process. For the prediction of fresh concrete properties, two central procedures can be found in the literature. One approach is to make use of the information of the mix design on the basis of which machine learning methods are trained to predict the target parameters. In [15], the information from the mix design is additionally extended by temporal information to enable a time-dependent prediction. Although these approaches are promising, they heavily depend on the quality of the information from the mix design which, however, are often subject to larger inaccuracies, e.g. caused by unknown fluctuations in the raw material properties. The second approach is to use sensor observations of the fresh concrete, e.g. in the form of image observations, and to establish deep learning methods to predict the fresh concrete properties on the basis of the observed data. In this context, [16] used a camera setup to observe the concrete flow at the outlet of a mixing truck and a CNN-based method was proposed to learn a mapping between the sequential image data and the fresh concrete's rheological properties. However, compared to channel flow, a significantly more complex behaviour is expected for the dynamic flow of the concrete during mixing. In [17], a method for predicting the rheology of Bingham fluids from stereo-images and 3D surface reconstructions of the material during the mixing procedure was presented. In [18], a similar approach using a CNN and a long-short-term memory network has been proposed. Although these approaches are not prone to possible uncertainties in the mix design, they do not take into account the time dependency of the properties. In this paper, we extend the CNN-based method of [17] and propose a novel procedure for time-dependent fresh concrete characterisation during the mixing process based on stereo-camera observations, temporal information, and mix design. Investigations whether the deployment of ViT-based models instead of a CNN architecture can improve the performance are planned in future work.

3 Methodology

3.1 Image-based granulometry

Given a three-channel (RGB) image I depicting concrete aggregates, we aim at automatically deriving the

grading curve $G = [p_1, p_2, \dots, p_N]$. Per definition, the grading curve is a histogram in which grain size intervals (bins) are represented on the abscissa (x-axis) and the quantity proportion is shown on the ordinate (y-axis). Consequently, G is parameterised by a vector whose elements p_j correspond to the histogram percentiles of each grain size interval $j = [1 \dots N]$. In this representation, each percentile is a continuous variable with $\{p_j \in \mathbb{R} \mid 0 \leq p_j \leq 1\}$ under the constraint $\sum p_j = 1$. As a result, mapping the image I to the grading curve G corresponds to a constrained multi-regression problem of the individual percentiles p_j , in which the determined percentiles must sum up to 1. In order to tackle this problem, we make use of a Vision Transformer (ViT) based architecture [12] acting as mapping function $f : I \rightarrow G$. As is shown in Fig. 2, the ViT takes a single image as input and decomposes it into a sequence of n non-overlapping image patches $x_i \in \mathbb{R}^{h \times w}$ which are transformed into 1D tokens $z_i \in \mathbb{R}^d$ of length d using a linear projection E . The sequence of tokens $\mathbf{z}^0 \in \mathbb{R}^{(n+1) \times d}$ with

$$\mathbf{z}^0 = [z_{\text{cls}}, \mathbf{E}z_1, \mathbf{E}z_2, \dots, \mathbf{E}z_n] + \mathbf{p} \quad (1)$$

then serves as input to a transformer encoder architecture [19]. As indicated in Eq. 1, a learnable classification token z_{cls} is prepended to the sequence, whose representation at the final layer of the encoder is used as input embedding for the output layer.

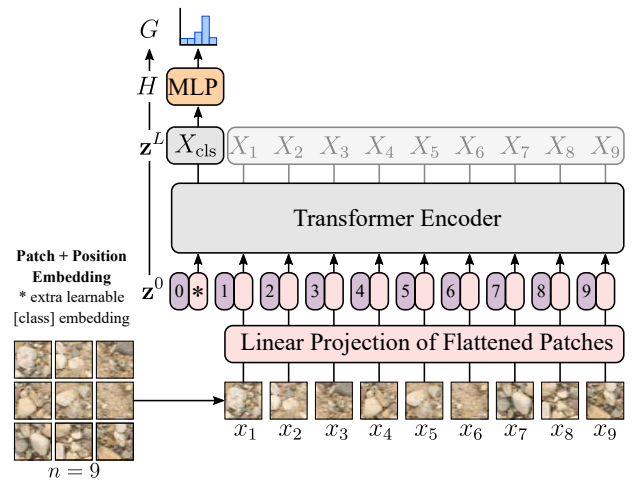


Figure 2. Structure of the ViT (adapted from [12])

Furthermore, a learnable position embedding $\mathbf{p} \in \mathbb{R}^{n \times d}$ is added to the tokens (cf. Eq. 1) in order to retain positional information throughout the permutation invariant self-attention operations of the encoder. The tokens are passed through the transformer encoder which consists of a stack of $l = 1 \dots L$ residual layers, each comprising Multi-Head Self-Attention (MSA) [19], layer normalisation (LN), and Multi-Layer Perceptron (MLP) blocks. The

output of the last layer of the transformer encoder is denoted as embedding $\mathbf{z}^L = [X_{\text{cls}}, X_1, X_2, \dots, X_n]$, where n is the total number of patch tokens and X_{cls} and X_1, \dots, X_n correspond to the embedded class token and patch tokens, respectively. Finally, a MLP head H is used on top of the transformer encoder and typically produces the prediction output based on the final encoded class token embedding X_{cls} . We make use of the softmax-function as final activation in H , returning $j = 1 \dots N$ output values representing the size distribution G which consequently comply with the constraint $\sum p_j = 1$. For training, the loss for image I can be written as

$$L_{\text{cls}} = D(H(X_{\text{cls}}), Y_I), \quad (2)$$

where $H(X_{\text{cls}})$ is the output of the final prediction head for the class embedding X_{cls} , Y_I is the reference for image I , and $D(\cdot, \cdot)$ is a distance function. In this work, we apply the Kullback-Leibler divergence D_{KL} as measure for $D(\cdot, \cdot)$ which computes the similarity between the predicted and the reference grading curves according to

$$D_{\text{KL}} = \sum_{j=1}^N p_j \cdot \log\left(\frac{p_j}{\hat{p}_j}\right), \quad (3)$$

where p_j and \hat{p}_j are the reference and the predicted percentiles of the size distribution, respectively.

3.2 Image-based fresh concrete characterisation

Until today, existing test methods for the determination of fresh concrete properties are exclusively applicable in post-production. As a result of the inability to estimate the fresh concrete properties already during the production process, an adaptation and adjustment of the concrete in case of deviations from the desired properties is not possible in current practice. To overcome this limitation, in the paper we present an image based method for the inline prediction of fresh concrete quality during the production, i.e. during the mixing process in this paper. In addition to the flow behaviour, which contains valuable information about the rheological properties, previous studies have shown that the 3D surface carries ancillary information able to support the task of rheological characterisation [17]. For this reason, we propose a stereo-camera setup observing the flow behaviour of the concrete during mixing, allowing to reconstruct the 3D surface of the concrete via dense stereo-matching and triangulation. More specifically, we compute a digital elevation model (DEM) D together with an orthophoto O for each stereo image pair. Both entities D and O are used as input to a Convolutional Neural Network (CNN) with the goal to predict a state vector $C = [\delta, \tau_0, \mu]$ representing the fresh concrete characteristics, namely the slump flow diameter δ (defining the concrete consistency), and the Bingham parameters

τ_0 and μ , describing the rheological properties. As additional input information, we introduce the composition data of the mix design m , consisting of information about the water-cement ratio, the grading curve, the paste content, the additive content and the time difference between the start of the mixing process and the image acquisition, and append it to the latent feature embedding produced by the CNN in a late-fusion manner. A high-level overview on the procedure is shown in Fig. 3. Since deeper networks tend to overfit, the CNN consists of only 7 convolutional layers. Each layer has a kernel size of 5x5 and a stride of 2, followed by a batch normalisation and a Rectified Linear Unit (ReLU) activation function. The convolutional layers are followed by three fully connected (FC) layers, each with a leaky ReLU activation function using a slope of 0.2. O and D are passed through the convolutional layers, producing the flattened feature embedding $z_{O,D}$. By concatenating $z_{O,D}$, the mixture design information m , and the time epoch Δ_t we obtain the feature vector F , which is used as input for the FC layers. Here, Δ_t represents the temporal difference between the acquisition time of the input image pair and the point in time where the consistency is to be determined. As a consequence, the target state vector C becomes time-dependent, and therefore, the FC-layers finally map the feature vector to the three time-dependent output parameters δ_{Δ_t} , τ_{0,Δ_t} and μ_{Δ_t} .

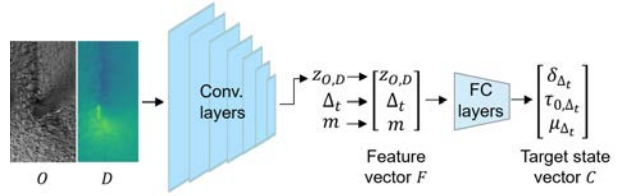


Figure 3. Architectural structure of the CNN.

For optimising the network weights ω the Mean Square Error (MSE) is used as loss function, which is iteratively minimised during training. In this context, the loss is computed for a mini-batch consisting of N samples, each associated with the state vector C containing the $k = 1 \dots K$ target parameters y^k , where $K = 3$ and $N = 32$ in this paper. For loss calculation, the squared differences between the reference values y^k and the predicted values \hat{y}^k are determined and averaged over all parameters and samples in a batch, such that

$$L_{\text{MSE}}(\omega) = \frac{1}{N \cdot K} \sum_{k=1}^K \sum_{n=1}^N (y_n^k - \hat{y}_n^k)^2. \quad (4)$$

To prevent over-fitting, weight decay is used during training. This method penalises large weights by a factor λ (in this paper $\lambda = 0.001$) and leads to an additional term in

the final loss function

$$L(\omega) = L_{\text{MSE}}(\omega) + \lambda \cdot \sum \omega^2. \quad (5)$$

4 Experiments

4.1 Image-based granulometry

Test setting: For the experimental evaluation of the proposed method for particle size distribution estimation, we make use of the publicly available *Deep Granulometry data set*¹. The data set consists of images showing concrete aggregate particles and reference data of the particle size distribution (grading curves) associated to each image. The data consists of two independent sets of images, the *Coarse Aggregate Data* (D_{coarse}) and the *Fine Aggregate Data* (D_{fine}). Both contain approximately 1700 images associated to one of 34 different particle size distributions. While the data in D_{coarse} shows aggregates with particle sizes ranging from 0.1 to 32 mm and provides reference percentiles for $N = 9$ bins with upper bounds of 0.25, 0.5, 1, 2, 4, 8, 16, 32.5, 63 [mm], the D_{fine} data contain fine material with grain sizes between 0 and 2 mm and with references for $N = 6$ bins, namely 0.063, 0.125, 0.25, 0.5, 1.0, 2.0 [mm]. We train networks for each set of images (fine and coarse) individually. Due to computational reasons we do not make use of the full image resolution in which the data is provided but we downsample the images to a size of 512x704 [px] for the D_{coarse} data, corresponding to a ground sampling distance (GSD) of 0.5 mm, and to a size of 480x480 [px] for the D_{fine} data, corresponding to a GSD of 0.1 mm. We adapt the *hybrid ViT-Base* architecture according to the definition in [12], i.e. we feed the image to a convolutional module and form the input sequence for the transformer based on the resulting feature maps (cf. [13] for details). For the architecture of the transformer encoder, we chose $L = 12$ layers, each comprising 12 multi-head-self-attention (MSA) modules. For evaluation, we split the data sets into two subsets and perform a two-fold crossvalidation. In order to reduce overfitting effects, we make use of random geometric and radiometric image augmentations during training.

Results: In order to assess the performance of the predicted particle size distributions, we compute the mean absolute error ε_p over the individual percentiles p_j of the different bins. Furthermore, for testing the grading curve predictions as a whole, we compute the average ε_H of the discrete *Hellinger distances* $D_H \in [0, 1]$, used as measure of the similarity between the reference grading curves G

and the predicted grading curves \hat{G} , with

$$D_H(G, \hat{G}) = \frac{1}{\sqrt{2}} \sqrt{\sum_j (\sqrt{p_j} - \sqrt{\hat{p}_j})^2}. \quad (6)$$

Here, values close to 0 represent highly similar distributions and values close to 1 highly different distributions. For concrete mix design, often the k-value of the aggregates, which is computed as the sum of the percentiles of the cumulative grading curve representation, and which represents a measure for the granularity of the aggregate, is used to adjust the water content of the composition. We therefore also compute the k-value from the predicted size distributions, compare it to the reference values, and report the mean absolute error $\varepsilon_{\text{k-value}}$. Tab. 1 contains the results for the described metrics on both data sets. As is visible from the table, the percentile-wise predictions lead to an average error of $\varepsilon_p = 4.21\%$ on the fine aggregate data, and to only 1.57% on the coarse aggregate data. As a consequence, also the values for the average Hellinger distance and the average errors of the k-values are comparably smaller on the D_{coarse} data.

In order to gain more detailed insights into the distribution of the respective errors, Fig. 4 shows the cumulative histogram of the absolute percentile errors (Fig. 4a) as well as a plot of the reference k-values vs. predicted k-values on the fine (Fig. 4b) and the coarse (Fig. 4c) data sets. Regarding the percentile-wise prediction errors it can be seen from Fig. 4a, that approx. 90% of all values of the fine aggregate data are predicted with an error of less than 10%, and nearly 100% of all values on the coarse aggregate data. W.r.t. the resulting k-values, the most important indicator for the concrete production, it can be seen that the predicted values show some fluctuations but, on average, are very close to the reference values, demonstrating a highly promising suitability of the proposed method for a future implementation in practice.

Table 1. Quantitative results of the image based granulometry on the two data sets D_{coarse} and D_{fine} .

	ε_p [%]	ε_H	$\varepsilon_{\text{k-value}}$
D_{fine}	4.21	0.127	0.187
D_{coarse}	1.57	0.071	0.145

4.2 Image-based fresh concrete characterisation

Test setting: For training and testing of the method proposed for the determination of fresh concrete properties, we acquired an extensive data set using a similar setup as Ponick et al. [17]. More specifically, we designed a *surrogate-mixing system* consisting of a linear

¹<https://doi.org/10.25835/61y9pei9>

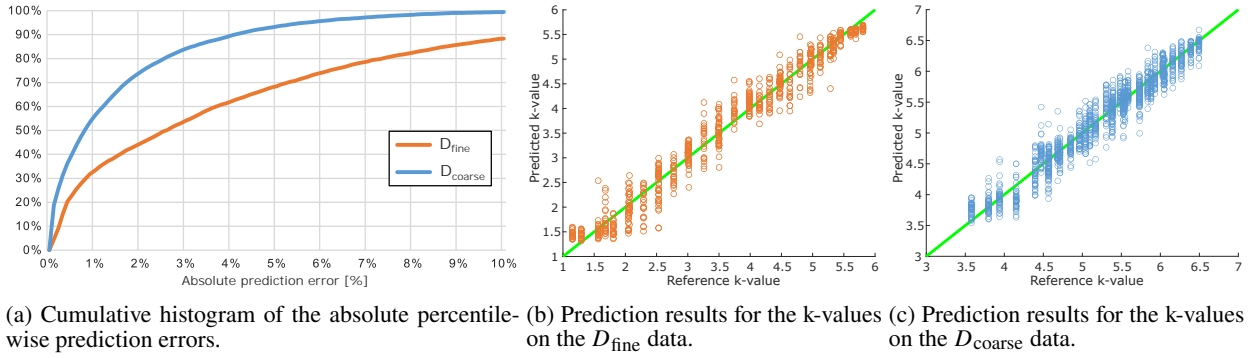


Figure 4. Performance results for the predictions of the aggregate particle size distributions.

mixing geometry in form of a channel, which is filled with fresh concrete and in which a single mixing paddle moves back and forth along a linear trajectory in order to simulate a simplified mixing process of a concrete mixer. During movement of the paddle, a stereo camera setup (2xGrasshopper 3 USB cameras with a focal length of 8 mm) is used to observe the motion behaviour of the concrete by acquiring images of size 1920x1200 [px]. Fig. 5 shows the schematic test setup.

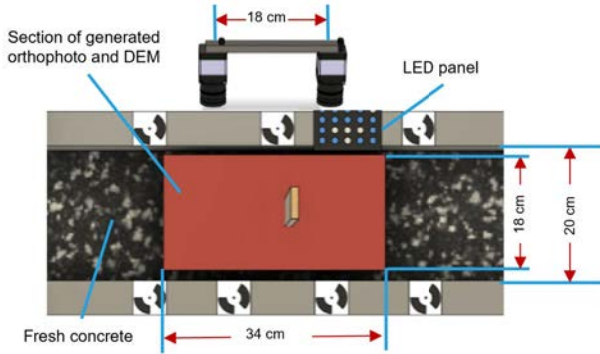


Figure 5. Top view on the *surrogate mixing system* used for generating the fresh concrete mixing data.

For the experiments in this paper, a total of 45 concretes with different mix compositions were selected. In particular, we generated different mixture designs by varying the parameters of water-cement ratio, paste content, grading curve, cement type, sand-lime powder content and additive content. In order to generate reference values for the resulting fresh concrete properties, independent measurements were conducted in coordination with the image data acquisition of the mixing behaviour. The slump flow diameter δ was determined from the slump flow test [14], yield stress τ_0 and viscosity μ were determined using an *eBT-V* rheometer from *Schleibinger*. The first slump flow test for each concrete was made directly after finishing the mixing process. The slump test and the rheometer mea-

surement are independent of each other and were repeated several times at intervals of about 30 minutes in order to account for the time-dependent properties of the concrete.

Regarding image acquisition, 14 runs were recorded for each concrete, respectively, where one run corresponds to six back-and-forth movements of the mixing paddle. In run 1-7, the paddle moved with a velocity of $200 \frac{mm}{s}$ and images were recorded with 30 frames per second (fps). In run 8-14, the velocity was set to $450 \frac{mm}{s}$ while the frame rate was increased to 60 fps. To account for effects of different mixing velocities and frame rates, we include both information in m . A number of 1300 image pairs were captured during each run. For the generation of O and D , only the image pairs in which the paddle was visible in both images were used, resulting in a total amount of approx. 314k images in the data set. Note, that due to the dynamic nature of the scene, the images must be captured strictly at the same time for a valid generation of O and D from the stereo-pairs. To verify the synchronisation of the cameras, a LED panel with 20 LEDs was used, allowing to generate time stamps for each image at millisecond intervals. By assuming that the concrete properties remain constant during the relatively short time period of one run (approx. 44 and 22 sec), we associate each image pair within one run with the time stamp of the central image pair of the run. Consequently, Δ_t represents the time difference between the associated time stamp of the image pair and the point in time of the respective reference measurement.

Training: Since the three reference values for the multi-task output of the CNN are generated by two independent measurements, we created multiple reference combinations for each set of inputs O and D during training. For this purpose, we computed all possible reference combinations for each concrete and randomly assigned all entities O and D of a concrete to one of these combinations. As a consequence, Δ_t consists of two values, one for δ and one for τ_0 and μ , respectively. Training was performed by applying a five-fold cross-validation. The 45 concretes were divided into 5 sets of 9

concretes each. First, we sorted the concretes according to the length of the first slump flow diameter δ_1 and formed 3 groups (with the 15 longest δ_1 , the 15 central δ_1 and the 15 smallest δ_1 , respectively). Subsequently, three concretes of each group were randomly assigned to one of the 5 sets to ensure a balanced distribution of the concretes. In each cross-validation step, one of the sets was used as the test set. The validation set, consisting of 5 concretes, was formed randomly from the concretes of the other sets, again by taking δ_1 into account. The remaining 31 concretes form the training set.

Results: After training, the network is evaluated on the test sets. In order to assess the performance of the network, the absolute and relative error of the predicted parameters are determined for each pair of O and D in the test set. Subsequently, we compute the mean absolute deviation ε_{abs} and the mean relative deviation ε_{rel} for each parameter. The results are shown in Tab. 2. The results improve if the predictions of the same label combination (and different inputs) in a run are averaged beforehand and then ε_{abs} and ε_{rel} are calculated (on average approx. 41 predictions were averaged). These results are shown in brackets. With 6.1% of relative error for the prediction of the slump flow diameter, the results are already in a very promising range, especially considering that the slump flow test, which is used as reference method, is a manual measure associated with a considerable uncertainty. Compared to the slump flow, the results obtained for the yield stress and the viscosity are distinctly less precise. The reason for this is probably that the rheometer measurements are actually intended for a testing directly after the mixing process, as the concrete is more fluid at this point. With increasing time and increasing viscosity, the uncertainty of the rheometer measurements increases.

Table 2. Performance metrics of the CNN for the prediction of fresh concrete properties.

	ε_{rel}	ε_{abs}
$\delta_{\Delta t}$	6.4% (6.1%)	2.8 cm (2.7 cm)
$\tau_{0,\Delta t}$	28.3% (26.9%)	53.6 Pa (50.6 Pa)
$\mu_{\Delta t}$	43.8% (42.4%)	18.5 Pa·s (17.8 Pa·s)

By using the time difference as input, the CNN also implicitly learns a time-dependent prediction model for the fresh concrete properties. This enables the prediction of the properties of the concrete at the time of placement during the mixing process (provided the time between mixing and placement is known). Fig. 6 shows examples of the temporal predictions for the slump flow diameter. On the y-axis the consistency parameters are plotted and

the x-axis represents the time difference Δt between the time of image acquisition and the time at which the target parameters are to be determined. For the reference values, the precision 2.46 cm of the slump test is given as error bar [14]. Since some reference measurements were taken before the image acquisition, the x-axis in the figures starts with a negative value. The results show that the network is able to learn the time-dependent behaviour of the concrete consistency represented by the slump flow diameter.

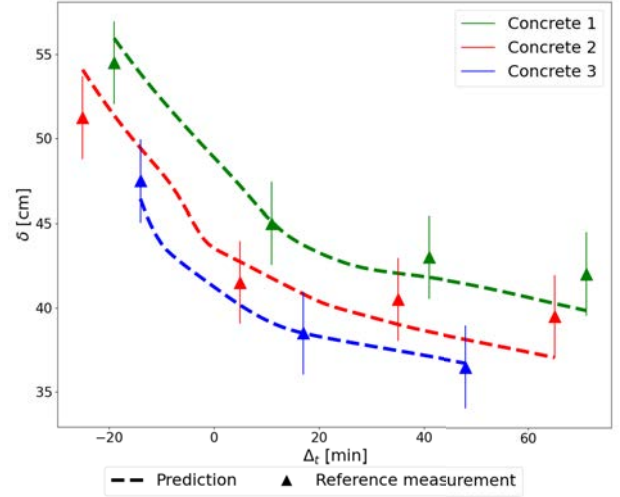


Figure 6. Three examples for the prediction of the time-dependent behaviour of the slump flow δ .

5 Conclusion

In this paper, we presented computer vision based methods for the characterisation of concrete aggregate and fresh concrete from image sequences acquired during the mixing process. The results obtained on challenging data sets demonstrate a promising performance for both, predicting the precise particle size distribution of fine and coarse aggregate particles, and deriving relevant fresh concrete properties. In the future, we aim at applying both methods under real-world conditions within a concrete production facility in order to integrate these methods into a control loop for the concrete production. In particular, strategies for an in-line adaptation of the mix design and countermeasures for a control of the fresh concrete properties during mixing will be developed on the basis of the approaches proposed in this paper.

Acknowledgements

This work was supported by the Federal Ministry of Education and Research of Germany (BMBF) as part of the research project ReCyCONTROL [No. 033R260A].

References

- [1] I. González-Taboada, B. González-Fontoboa, F. Martínez-Abella, and N. Roussel. Robustness of self-compacting recycled Concrete: Analysis of Sensitivity Parameters. *Materials and Structures*, 51(8), 2018. doi:10.1617/s11527-017-1136-1.
- [2] M. Haist, T. Schack, M. Coenen, C. Vogel, D. Beyer, and C. Heipke. Concrete 4.0 - Sustainable Concrete Construction with Digital Quality Control. In *ce/papers 6(6)*, pages 1555–1562, 2023. doi:10.1002/cepa.2965.
- [3] B. Green. *Productivity in Construction: Creating a Framework for the Industry to Thrive*. Chartered Institute of Building (CIOB), 2016.
- [4] W. Fuller and S. Thomson. The Laws of Proportioning Concrete. *Transactions of the American Society of Civil Engineers*, 59(2):67–143, 1907.
- [5] E. Hamzeloo, M. Massinaei, and N. Mehrshad. Estimation of Particle Size Distribution on an Industrial Conveyor Belt using Image Analysis and Neural Networks. *Powder Technology*, 261:185–190, 2014. doi:10.1016/j.powtec.2014.04.038.
- [6] F. C. Lira and P. Pina. Grain Size Measurement in Images of Sands. In *International Conference on Computer Vision Theory and Applications (VIS-APP)*, pages 371–374, 2006.
- [7] A. Soloy, I. Turki, M. Fournier, S. Costa, B. Peuziat, and N. Lecoq. A Deep Learning-Based Method for Quantifying and Mapping the Grain Size on Pebble Beaches. *Remote Sensing*, 12(21), 2020. doi:10.3390/rs12213659.
- [8] M. Coenen, T. Schack, D. Beyer, C. Heipke, and M. Haist. ConsInstancy: Learning Instance Representations for Semi-Supervised Panoptic Segmentation of Concrete Aggregate Particles. *Machine Vision and Applications*, 33(57), 2022. doi:10.1007/s00138-022-01313-x.
- [9] L. E. Olivier, M. G. Maritz, and I. K. Craig. Deep Convolutional Neural Network for Mill Feed Size Characterization. *IFAC-PapersOnLine*, 52(14):105–110, 2019. doi:10.1016/j.ifacol.2019.09.172.
- [10] M. Coenen, D. Beyer, C. Heipke, and M. Haist. Learning to Sieve: Prediction of Grading Curves from Images of Concrete Aggregate. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences V-2-2022*, pages 227–235, 2022. doi:10.5194/isprs-annals-V-2-2022-227-2022.
- [11] N. Lang, A. Irniger, A. Rozniak, R. Hunziker, J. D. Wegner, and K. Schindler. GRAINet: Mapping Grain Size Distributions in River Beds from UAV Images with Convolutional Neural Networks. *Hydrology and Earth System Sciences*, 25(5):2567–2597, 2021. doi:10.5194/hess-25-2567-2021.
- [12] A. Dosovitskiy, L. Beyer, A. Kolesnikov, and et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. In *International Conference on Learning Representations (ICLR)*, 2021.
- [13] M. Coenen, D. Beyer, and M. Haist. Granulometry Transformer: Image-based Granulometry of Concrete Aggregate for an Automated Concrete Production Control. In *Proceedings of the 2023 European Conference on Computing in Construction (EC3)*, 2023. doi:10.35490/EC3.2023.223.
- [14] EN 12350-5. Testing Fresh Concrete - Part 5: Flow Table Test. European Committee for Standardization, 2019.
- [15] I. Navarrete, I. La Fé-Perdomo, J. A. Ramos-Grez, and M. Lopez. Predicting the Evolution of Static Yield Stress with Time of blended Cement Paste through a Machine Learning Approach. *Construction and Building Materials*, 371, 2023. doi:10.1016/j.conbuildmat.2023.130632.
- [16] M. Coenen, C. Vogel, T. Schack, and M. Haist. Deep Concrete Flow: Deep Learning based Characterisation of fresh Concrete Properties from Open-Channel Flow using Spatio-Temporal Flow Fields. *Construction and Building Materials*, 411, 2024. doi:10.1016/j.conbuildmat.2023.134809.
- [17] A. Ponick, A. Langer, D. Beyer, M. Coenen, M. Haist, and C. Heipke. Image-Based Deep Learning for Rheology Determination of Bingham Fluids. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLIII-B2-2022*, pages 711–720, 2022. doi:10.5194/isprs-archives-XLIII-B2-2022-711-2022.
- [18] L. Yang, X. An, and S. Du. Estimating Workability of Concrete with Different Strength Grades based on Deep Learning. *Measurement*, 186, 2021. doi:10.1016/j.measurement.2021.110073.
- [19] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. Attention is All you Need. In *Advances in Neural Information Processing Systems (NIPS)*, 2017.

Investigating the Neural Correlates of Different Levels of Situation Awareness and Work Experience

Yanfang Luo¹, JoonOh Seo¹, and Sogand Hasanzadeh, A.M.ASCE²

¹Dept. of Building and Real Estate, Hong Kong Polytechnic University, Hong Kong SAR

²Lyles School of Civil Engineering, Purdue University, United States

yanfang.luo@connect.polyu.hk, joonoh.seo@polyu.edu.hk, sogandm@purdue.edu

Abstract

Maintaining good situation awareness is crucial for workers' safety on the dynamic and complex construction site, and workers with more experience may contribute to better performance in safety. However, little research has investigated the cognitive differences between experienced and novice workers regarding different levels of situation awareness (Level 1: perception, Level 2: comprehension, Level 3: projection). To address this gap, this study investigated the cognitive processes of hazard recognition behaviors among experienced and novice participants using functional near-infrared spectroscopy (fNIRS) across 12 virtual reality scenarios. The results revealed that novice participants showed higher activation in the left prefrontal cortex across all three levels of situation awareness, indicating their tendency to focus on detailed information when faced with unfamiliar environments. In contrast, more experienced workers exhibited increased activation in the right prefrontal cortex, particularly in hazard comprehension and projection (Level 2 and Level 3). This suggests that experienced participants prioritize global control mechanisms by activating the right prefrontal cortex associated with spatial awareness. These results highlight cognitive differences at different levels of situation awareness between experienced and novice participants, providing insights into following behavioral patterns and decisions. Furthermore, these findings offer a theoretical foundation for tailoring safety interventions to address the cognitive deficiencies in each level of situation awareness.

Keywords

Situation awareness, fNIRS, Work experience, Neural, Cognitive process

1 Introduction

Construction sites include various potential and active hazards, necessitating workers to maintain proper

spatial awareness to prevent accidents and injuries [1–3]. Situation awareness, encompassing three Levels as proposed by Endsley (perception, comprehension, and projection), is a pivotal factor in successfully recognizing hazards within construction environments [4,5]. Various factors, including sensitivity to stimuli, limitation in attention resources, constraints in working memory, and possession of experience in construction sites, affect situation awareness [4,6,7]. Previous literature argued that experienced workers exhibit a heightened capacity to tackle the challenges inherent in complex construction sites as they are familiar with the construction environment and better allocate their limited cognitive resources to potentially hazardous cues [8,9]. Despite the pivotal role of situation awareness in shaping safety behaviors, few studies have an in-depth understanding of the intricate cognitive mechanisms at each level of situation awareness, particularly regarding the differences between experienced and novice workers. Dzung et al. used an eye-tracker to investigate the search patterns between experienced and novice workers in site hazard identification but did not measure the situation awareness [8]. Hasanzadeh et al explored the worker's situation awareness under fall and tripping hazard conditions but did not probe into each level of situation awareness [5]. A deeper exploration of the cognitive processes underlying situation awareness may help unveil the cognitive mechanisms behind successful hazard recognition among more experienced workers [10].

Achieving higher levels of situation awareness (i.e., comprehension, and projection) represents a complex cognitive process with neural activation in different regions of the brain cortex [11]. Existing research has illustrated a strong association between the prefrontal cortex, and planning and decision-making [12]. Activation in this region, reflected by oxygen concentration, can be a reliable indicator of attaining higher levels of situation awareness [11,12]. Functional Near-Infrared Spectroscopy (fNIRS) is a non-invasive and portable device capable of measuring changes in the concentration of oxygenated and deoxyhemoglobin

(oxyHb, deoxy-Hb) in brain tissue [13]. Examining oxygen consumption in the prefrontal regions at various levels of situation awareness between novice and experienced participants may provide insights into cognitive limitations that impede higher levels of situation awareness. However, previous studies predominantly focused on the overall hazard recognition performance (i.e., recognized or not) and the corresponding neural activity, neglecting to explore the differential neural mechanism across various levels of situation awareness.

To bridge this gap, this study examined the cognitive disparities within each level of situation awareness between more experienced and novice participants, employing fNIRS technology within multiple virtual reality (VR) scenarios. These findings offer insights into potential training strategies to address cognitive deficiencies during situation awareness, ultimately contributing to reduced injury rates at construction sites.

2 Background

2.1 Effect of Work Experience on Situation Awareness

The construction site is a multifaceted and dynamic environment with diverse stimuli [14,15]. Hazard recognition abilities are essential for ensuring safety on construction sites [3]. Situation awareness, i.e., comprehensive perception, comprehension, and anticipation of environmental stimuli [4], is crucial for successful hazard recognition [16]. The attainment of higher levels of situation awareness primarily depends on developing mental models in long-term memory shaped by practical experience [17].

Hasanzadeh et al. [16] utilized eye-tracking technology to investigate hazard recognition on real construction sites, noting that levels of work experience influenced situation awareness and attentional allocation during hazard searches. More experienced workers demonstrated a greater ability to detect relevant hazard-related changes in the construction environment [18]. This was consistent with Aroke et al. [9], who found work experience and dwell time positively correlated with hazard identification. In the realm of driving, Kass et al. [19] observed that experienced drivers exhibited better performance following traffic regulations and greater situation awareness. This aligns with Wright et al. [20], which illustrated that middle-aged drivers with more experience were found to have higher levels of situation awareness and were better at anticipating hazards. Conversely, research by Zhou and her colleagues, [11] indicated that novice workers with higher sensitivity achieved a higher hazard recognition rate, coupled with a stronger electroencephalogram

signal than experienced workers in a hazard recognition task involving a series of construction photos. While many studies suggest a positive relationship between experience and situation awareness, a consensus remains unclear. Most previous literature focused on the overall hazard recognition rate rather than delving into the cognitive processes at each level of situation awareness. Therefore, exploring the intricate mechanisms of each level of situation awareness is crucial to understanding the differences between novices and experienced individuals in hazard recognition.

2.2 fNIRS Application in Construction-Related Research

The consumption of glucose and oxygen drives neuronal activity, so active nerves represent increased consumption of glucose and oxygen in local brain regions [13]. Functional Near Infrared spectroscopy (fNIRS) is a non-invasive, safe, and reliable device to measure changes in the concentration of oxygenated and deoxyhemoglobin (oxy-Hb, deoxy-Hb) in the brain [21]. This concentration change was detected by emitting NIR light from 760nm to 850 nm into the head [13]. When NIR light enters the brain, it is absorbed, scattered, or reflected. Intensity changes of scattered or reflected light can be quantified using the modified Beer-Lambert law, which is essentially an empirical description of the attenuation of light in a highly scattering medium [13,21].

fNIRS, an emerging neuroimaging technology, has not yet seen widespread application in the construction sector. While some researchers have explored its application in architectural design [22], wayfinding [23], and hazard perception [24], its application remains limited. For instance, Hu et al. [22] experimented with design brainstorming, comparing outcomes with and without real-time neurocognitive feedback of brainstorming for students. Results indicated that when feedback was provided, a higher percentage of right-hemispheric dominance, associated with increased design idea fluency, was observed. Zhu et al. [23] modeled the relationship between fNIRS features and cognitive load by extracting hemodynamic response features in the prefrontal cortex during a building wayfinding information memory test. Pooladvand et al. [24] investigated individual decision-making and cognitive demand changes under normal or stressful conditions in VR scenarios. Findings suggested that limited cognitive resources led to a failure in comprehensively processing environmental information. By analyzing pupil responses and cerebral oxyhemoglobin signals, Liao et al. [22] demonstrated that different hazard types can induce varying cognitive demands. Despite these efforts, research focusing on situation awareness has been relatively scarce. This study seeks to bridge this gap by employing fNIRS technology

to investigate the neural correlates of different levels of situation awareness in the hazard recognition process.

3 Research Method

The virtual reality (VR) experiment was designed, covering four common hazard types identified by the Occupational Safety and Health Administration (OSHA), which included: (a) Struck by an equipment: The crane was moving toward a worker, and the worker may be hit by this equipment. (b) Struck by objects: Some materials were stacked on the edge of the upper floors, and a worker passing by may be hit by the falling materials). (c) Fall to a lower level: A painting worker was standing on a ladder without fall protection and may fall from the ladder. (d) Electrocution: A worker was navigating an uneven floor cluttered with electrical wires and may be hit by electricity. These hazards were visually represented in Figure 1.

Twelve VR scenarios were developed, and 33 civil engineering participants from Purdue University were recruited to identify hazards within these scenarios. Participants with two or more years of work experience were classified as experienced participants, while those with less than six months of work experience were categorized as novice participants. Equipped with an HTC VIVE Pro Eye VR headset and Artinis fNIRS Brite II devices (Figure 2), they could freely scan the VR scenarios to identify hazards. Two regions of interest (RoIs) were selected in the fNIRS device, which were the left prefrontal cortex (LPFC) and right prefrontal cortex (RPFC), as shown in Figure 3. Both RoIs contained six transmitters and six receivers, instrumenting 14 channels covering the prefrontal cortex.

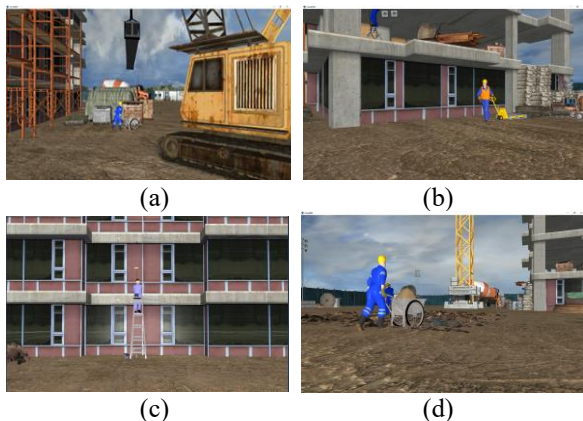


Figure 1. Hazard types in VR scenarios: (a) Struck by an equipment; (b) Struck by an object; (c) Fall to a lower level; (d) Electrocution.

After each scenario, participants were asked to answer three Situation Awareness Global Assessment Technique (SAGAT) questions by randomly freezing the

scenarios. There are a few types of situation awareness measurement techniques such as SAGAT, Situational Awareness Assessment Technique (SART) [25] and Crew Awareness Rating Scale (CARS) [26]. SART and CARS are self-rating assessments that might be limited by an individual's ability to assess their situation awareness [27]. Conversely, SAGAT measures objective reflections regarding situation awareness and is the most widely used freeze-probe technique that aims to assess different levels of situation awareness [28]. Choi et al. used SAGAT to measure workers' situation awareness while operating the forklift at the construction site, where SAGAT assisted in accurately capturing situation awareness at every level [27].

Based on their response, participants were categorized based on the level of situation awareness they have achieved [Achieved Level 1 (AL1), Achieved Level 2 (AL2), Achieved Level 3 (AL3)]. Notably, to examine the changes in the cognitive processing of participants, fNIRS data was extracted during periods when participants specifically focused on the hazardous area within the scenarios. Homer 3, a MATLAB-based toolbox, assisted in processing the fNIRS data. Physiological noise, such as cardiac activities, respiration, and blood pressure, was addressed by a bandpass filter. The modified Beer-Lambert law with different partial pathlength factor (ppf) based on individuals' age (e.g., ppf = 6,6,6) was adopted to convert the optical intensity changes into variations of oxy-Hb and deoxy-Hb. The General Linear Model (GLM) was used to analyze fNIRS data and calculate the Hemodynamic Response Function (HRF), which was computed for all participants across multiple channels.



Figure 2. The participant with a VR headset and fNIRS device.

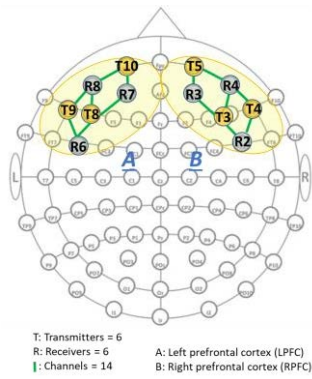


Figure 3. fNIRS layout: Region A is the Left Prefrontal Cortex (LPFC). It includes 3 Transmitters (T8, T9, T10), 3 receivers (R6, R7, R8), and 7 channels (green lines between transmitters and receivers). Region B is the Right Prefrontal Cortex (RPPC). It includes 3 Transmitters (T3, T4, T5), 3 receivers (R2, R3, R4), and 7 channels.

4 Results

Table 1 demonstrates the distribution of novice and experienced participants. 23 novice participants and 10 experienced participants were recruited for this research. The average age of novice participants was about 27 years while experienced participants exhibited an average age of around 30 years and an average work experience of 5.50 years.

Table 1 Age and work experience distribution.

Type of Participants	# of Participants	Age Mean (\pm SD)	Years of work experience Mean (\pm SD)
Novice	23	27.09 (\pm 2.56)	-
Experienced	10	30.40 (\pm 4.40)	5.50 (\pm 3.13)

Note: Most novice participants had no work experience, and a few had less than or equal to six months of work experience.

To fully investigate the hazard recognition behaviors at construction sites, all participants performed hazard searches on 12 scenarios. Consequently, the study involved 23 novice participants with a total sample size of 276 and 10 experienced participants with a total sample size of 120, as shown in Table 2. It is noteworthy that 43 samples from novice participants and 12 from experienced participants were excluded due to the absence of recorded levels of situation awareness in these assessments. Therefore, 233 samples were selected from novice participants and 108 samples were selected from experienced participants.

Table 2 Sample distribution.

Types of Participants	# of VR scenes	Total sample size	Selected	Excluded ^a
Novice participant	12	276	233	43
Experienced participant	12	120	108	12

^a means the participant did not achieve any level of situation awareness.

Figure 4 shows the percentages of participants who achieved each level of situation awareness to better understand hazard recognition behaviors across novice and experienced groups. In AL1, 84.42% of novice participants and 90% of more experienced participants were recorded, while in AL2, 71.38% of novice participants and 80% of experienced participants were observed. Ultimately, 63.41% of novice participants and 65.83% of experienced participants attained AL3. Notably, a declining trend from AL1 to AL3 was evident in both participant groups. The data indicates that irrespective of work experience, fewer participants achieved AL2 and AL3 compared to AL1. Analyzing the percentages of participants reaching each level of situation awareness across all 12 VR scenarios, it becomes apparent that more experienced participants exhibited slightly higher percentages than novice participants. However, the difference between the two groups was not substantial.

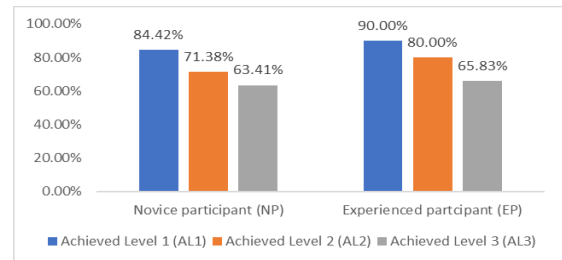


Figure 4. Percentages of participants achieved each level of situation awareness across two groups (Novice and Experienced)

Under observation of Oxy-Hb Concentration across levels of situation awareness, novice participants demonstrated a positive Oxy-Hb concentration in all situations. Specifically, the activation in the LPFC of novice participants was moderately higher than experienced participants in the perception level of situation awareness (AL1) ($p=0.06<0.1$) and significantly higher in the comprehension level of situation awareness (AL2) ($p=0.02<0.05$), as shown in Table 3. Experienced participants exhibited a negative

value of Oxy-Hb in LPFC across all levels of situation awareness. However, experienced participants did have a higher Oxy-Hb concentration value in the RPFC in the comprehension level of situation awareness AL2 (Oxy-Hb = $4.11\text{E-}07$) and in the projection level of situation awareness AL3 (Oxy-Hb = $2.41\text{E-}06$). The brain activation examples of selective experienced participants with higher activation in AL3 were shown in Figure 5. As can be seen, experienced participants exhibited lower activations during the initial stage of situation awareness (AL1). As the stimuli increased, the activations in the brain became more pronounced. Significantly, during AL3, RPFC demonstrated the highest activation compared to other levels of situation awareness in experienced participants.

Table 3 Oxy-Hb Concentration across two groups.

RoIs	Cat.	Oxy-Hb Concentration across Levels of Situation Awareness		
		Achieved Level 1 (AL1)	Achieved Level 2 (AL2)	Achieved Level 3 (AL3)
L-PFC	NP	1.65E-06	3.37E-06	3.31E-06
	EP	-2.34E-06	-1.94E-06	-1.32E-06
	NP-EP	3.99E-06	5.31E-06	4.63E-06
	<i>p</i> -value	0.06*	0.02**	0.13
R-PFC	NP	1.92E-08	4.55E-08	7.55E-08
	EP	-5.28E-07	4.11E-07	2.41E-06
	NP-EP	5.47E-07	-3.65E-07	-2.33E-06
	<i>p</i> -value	0.82	0.89	0.44

Note: RoIs: Regions of interest; LPFC: Left prefrontal cortex; RPFC: Right prefrontal cortex; NP: Novice participant; EP: Experienced participant. * $p < 0.1$; ** $p < 0.05$.

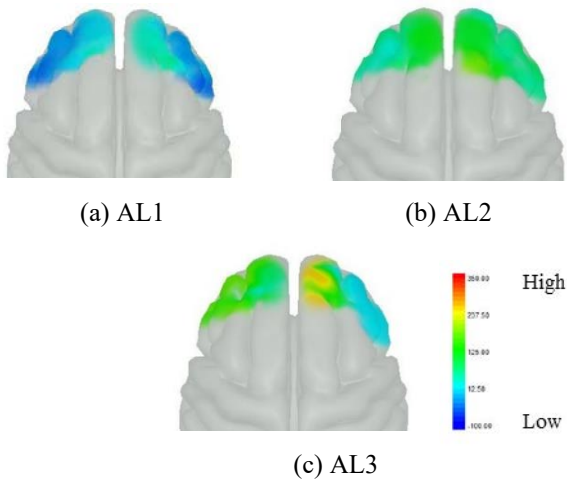


Figure 5. Brain activation of representative experienced participants in each level of situation awareness. The color red represents a higher activation level, whereas blue indicates a lower activation level.

5 Discussion

The findings revealed distinct patterns of brain activation representations associated with various situation awareness levels among novice and experienced participants. Novice participants exhibited positive neural activity in both the left and right prefrontal cortex across the three levels of situation awareness (AL1, AL2, AL3). These outcomes are consistent with Zhou et al. [11], where increased neural activation was observed in novice participants during a hazard-searching task, suggesting heightened cognitive effort in novices aiming for higher levels of situation awareness. It must be noted that higher neural activity in novices in the left prefrontal cortex compared to the right is associated with recognizing specific features of objects, particularly in terms of remembering what an object is [29,30]. Within the hazard recognition process, the left prefrontal cortex plays a more important role compared to the right prefrontal cortex for the novices [30]. Given their lower familiarity with the work environment, they tend to allocate more attention to the details and exert more cognitive efforts from Level 1 to Level 3 situation awareness. Moreover, the diminished activation in the right forehead of novices could be linked to the absence of corresponding experience or knowledge, which serve as important factors that assist in forming mental models in their long-term memory [31].

Experienced participants consistently exhibited reduced activation in the left prefrontal cortex, a region crucial for detail recognition, across all three levels of situation awareness (AL1, AL2, AL3). This suggests a reduced cerebral cortex stimulation due to the familiarity with the construction site. However, this familiarity among experienced individuals also contributes to developing intrinsic mental representations that enhance hazard recognition. The positive changes in Oxy-Hb concentration in the right prefrontal cortex among more experienced participants were observed in higher levels of situation awareness (AL2 and AL3). These may link to the activation of mental models innated in the RPFC when they achieved higher levels of situation awareness [31]. In addition, given that the right prefrontal cortex is responsible for spatial processing and awareness rather than detailed information [29–31], this suggests that experienced individuals were inclined towards employing a global searching strategy to obtain a comprehensive understanding of the construction site. It is noteworthy that, despite these distinctions, the percentages of achieving each level of situation awareness were similar between novices and experienced individuals in this study. This implies that experienced individuals may underestimate risks due to familiarity, whereas novices exhibit a more cautious approach.

This paper provided a new perspective to investigate hazard recognition behaviors by comparing different

situation awareness levels. Employing neuroimage techniques helped gain insights into the cognitive mechanism at different situation awareness levels among individuals with various experience levels. Findings revealed that novices had different neural reactions across three levels of situation awareness compared to the experienced, calling for customized training to address the specific weaknesses among novice workers. In addition, the phenomenon of diminished neural activations in experienced participants warrants careful consideration and attention. Injuries may occur when experienced workers underestimate or do not adequately acknowledge potential hazards. More effective safety training can be developed based on recognizing the deficiencies in brain activity during hazard recognition, aiming for enhanced hazard identification and overall safety performance.

While this paper offers valuable insights into the interplay between neural activation across various levels of situation awareness, some limitations should be noted. First, a few types of hazards have been considered in this study, and only one main hazard exists in each scenario, limiting the generalizability of the findings. Second, although the experienced group comprised over 100 samples, the participant count was limited to ten individuals, future research can consider recruiting more experienced participants. Third, it's noteworthy that the more experienced participants with five years of work experience were also students in the civil engineering domain, potentially limiting the generalizability of the findings to real-world construction scenarios. Future studies could benefit from a larger and more diverse worker pool. Last, the experiment was conducted with VR scenarios, replicating it in actual dynamic construction sites using portable fNIRS devices can enhance ecological validity.

6 Conclusion

Maintaining situation awareness is pivotal for worker safety in the dynamic and hazardous environments of construction sites. A detailed examination of various levels of situation awareness between novice and experienced individuals unveils insights into the cognitive neural correlates affecting the attainment of higher awareness levels. The findings indicated that novice participants exhibit higher neural activations in the left prefrontal cortex, which is linked to detailed object information, particularly in Level 1 perception and Level 2 comprehension. In contrast, more experienced participants showed increased activations in the right prefrontal cortex, associated with global comprehension and mental representation, particularly in achieving Level 2 comprehension and Level 3 projection. These results underscore distinct cognitive patterns of hazard

recognition in novice and experienced participants facing the same hazardous scenarios. Greater activation in the left prefrontal cortex among novices is mainly due to their lower familiarity with the dynamics of construction sites, which may demand more attention to identify the related hazards in the scene. Simultaneously, the lack of mental representation in the right prefrontal cortex weakens hazard recognition ability, leading to lower neuron activations. These results highlight the potential of fNIRS data to differentiate the status of situation awareness between novice and experienced individuals, offering valuable insights for tailoring training strategies to address these cognitive deficiencies.

References

- [1] Abbas, A., Seo, J., Ahn, S., Luo, Y., and Wyllie, M.J., Lee, G., Billingham, M. How Immersive Virtual Reality Safety Training System Features Impact Learning Outcomes: An Experimental Study of Forklift Training. *Journal of Management in Engineering*, 39(1), 2023.
- [2] Chang, W.-C., Borowiak, A., and Hasanzadeh, S. The Importance of Situational Awareness in Future Construction Work: Toward the Effects of Faulty Robots, Trust, and Time Pressure. In *Proceedings of the Computing in Civil Engineering*, pages 821–828, Corvallis, United States, 2023.
- [3] Luo, Y., Seo, J., and Hasanzadeh, S. Using Eye-Tracking to Examine the Cognitive Process of Hazard Recognition in Construction VR Scenarios., in: *Computing in Civil Engineering*, In *Proceedings of the Computing in Civil Engineering*, pages 541–548, Corvallis, United States, 2023.
- [4] Endsley, M.R. Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37(1): 32–64, (1995).
- [5] Hasanzadeh, S., Esmacili, B., and Dodd, M.D. Examining the Relationship between Construction Workers' Visual Attention and Situation Awareness under Fall and Tripping Hazard Conditions: Using Mobile Eye Tracking. *Journal of Construction Engineering and Management*, 144(7), 2018.
- [6] Endsley, M.R., and Garland, D.J. Theoretical underpinnings of situation awareness: A critical review. *Situation Awareness Analysis and Measurement*, 1(1): 3–21, 2000.
- [7] Endsley, M.R. Situation awareness misconceptions and misunderstandings. *Journal of Cognitive Engineering and Decision Making*, 9(1): 4–32, 2015.

- [8] Dzung, R.-J., Lin, C.-T., and Fang, Y.-C. Using eye-tracker to compare search patterns between experienced and novice workers for site hazard identification. *Safety Science*, 82(): 56–67, 2016.
- [9] Aroke, O., Esmacili, B., Hasanzadeh, S., Dodd, M.D., and Brock, R. The Role of Work Experience on Hazard Identification: Assessing the Mediating Effect of Inattention Under Fall-Hazard Conditions. In *Proceedings of the Construction Research Congress*, pages 509–519, Tempe, United States, 2020.
- [10] Luo, Y., Seo, J., Hasanzadeh, S., and Abbas, A., Exploring the Schema of Attention and Search Strategy in Different Levels of Situation Awareness under Fall Hazard Conditions. In *Proceedings of the Construction Research Congress*, pages 802–811, Des Moines, United States, 2024.
- [11] Zhou, X., Liao, P.C., and Xu, Q. Reinvestigation of the Psychological Mechanisms of Construction Experience on Hazard Recognition Performance. *Human Factors*, 66(1): 221–233, 2022.
- [12] Coutlee, C.G., and Huettel, S.A. The functional neuroanatomy of decision making: Prefrontal control of thought and action. *Brain Research*, 1428(): 3–12, 2012.
- [13] Bunce, S.C., Meltem, I., Kurtulus, I., Banu, O., and Kambiz, P. Functional near-infrared spectroscopy. *IEEE Engineering in Medicine and Biology Magazine*, 25(4): 54–62, 2006.
- [14] Luo, Y., Ahn, S., Abbas, A., Seo, J.O., Cha, S.H., and Kim, J.I. Investigating the impact of scenario and interaction fidelity on training experience when designing immersive virtual reality-based construction safety training. *Developments in the Built Environment*, 16(), 2023.
- [15] Chang, W.-C., and Hasanzadeh, S. Toward a Framework for Trust Building between Humans and Robots in the Construction Industry: A Systematic Review of Current Research and Future Directions. *Journal of Computing in Civil Engineering*, 38(3), 2024.
- [16] Hasanzadeh, S., Esmacili, B., and Dodd, M.D. Measuring Construction Workers' Real-Time Situation Awareness Using Mobile Eye-Tracking. In *Proceedings of the Construction Research Congress*, pages 2894–2904, San Juan, Puerto Rico, 2016.
- [17] Krems, J.F., and Baumann, M.R.K. Driving and Situation Awareness: A Cognitive Model of Memory-Update Processes, In *Proceedings of the Human Centered Design: First International Conference*, pages 986–994, San Diego, United States, 2009.
- [18] Solomon, T., Hasanzadeh, S., Esmacili, B., and Dodd, M.D. Impact of Change Blindness on Worker Hazard Identification at Jobsites. *Journal of Management in Engineering*, 37(4), 2021.
- [19] Kass, S.J., Cole, K.S., and Stanny, C.J. Effects of distraction and experience on situation awareness and simulated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 10(4): 321–329, 2007.
- [20] Wright, T.J., Samuel, S., Borowsky, A., Zilberstein, S., and Fisher, D.L. Experienced drivers are quicker to achieve situation awareness than inexperienced drivers in situations of transfer of control within a level 3 autonomous environment, In *Proceedings of the Human Factors and Ergonomics Society*, pages 270–273, Washington, D.C., United States, 2016.
- [21] Pinti, P., Tachtsidis, I., Hamilton, A., Hirsch, J., Aichelburg, C., Gilbert, S., and Burgess, P.W. The present and future use of functional near-infrared spectroscopy (fNIRS) for cognitive neuroscience. *Annals of the New York Academy of Sciences*, 1464(1): 5–29, 2020.
- [22] Hu, M., Shealy, T., Milovanovic, J., and Gero, J. Neurocognitive feedback: a prospective approach to sustain idea generation during design brainstorming. *International Journal of Design Creativity and Innovation*, 10(1): 31–50, 2022.
- [23] Zhu, Q., Shi, Y., and Du, J. Wayfinding Information Cognitive Load Classification Based on Functional Near-Infrared Spectroscopy. *Journal of Computing in Civil Engineering*, 35(5), 2021.
- [24] Pooladvand, S., and Hasanzadeh, S. Neurophysiological evaluation of workers' decision dynamics under time pressure and increased mental demand. *Automation in Construction*, 141(), 2022.
- [25] Taylor, R.M. Situational awareness rating technique (SART): The development of a tool for aircrew systems design, In *Situational Awareness*, pages 111–128, Routledge, 2017.
- [26] McGuinness, B., and Foy, L. A subjective measure of SA: The Crew Awareness Rating Scale (CARS), In *Proceedings of the First Human Performance, Situation Awareness, and Automation Conference, Savannah*, pages 286–291, Georgia, United States, 2000.
- [27] Choi, M., Ahn, S., and Seo, J.O. VR-Based investigation of forklift operator situation awareness for preventing collision accidents. *Accident Analysis and Prevention*, 136(), 2020.
- [28] Salmon, P., Stanton, N., Walker, G., and Green, D. Situation awareness measurement: A review

- of applicability for C4i environments. *Applied Ergonomics*, 37(2): 225–238, 2006.
- [29] Wilson, F.A., Séamas PÓ, S., and Patricia S., G.-R. Dissociation of object and spatial processing domains in primate. *Science*, 260(5116): 1955–1958, 1993.
- [30] Zhou, X., Hu, Y., Liao, P.C., and Zhang, D. Hazard differentiation embedded in the brain: A near-infrared spectroscopy-based study. *Automation in Construction*, 122(), 2021.
- [31] Biswal, B.B., Mennes, M., Zuo, X.N., Gohel, S., Kelly, C., Smith, S.M., Beckmann, C.F., Adelstein, J.S., Buckner, R.L., Colcombe, S., Dogonowski, A.M., Ernst, M., Fair, D., Hampson, M., Hoptman, M.J., Hyde, J.S., Kiviniemi, V.J., Kötter, R., Li, S.J., Lin, C.P., Lowe, M.J., Mackay, C., Madden, D.J., Madsen, K.H., Margulies, D.S., Mayberg, H.S., McMahon, K., Monk, C.S., Mostofsky, S.H., Nagel, B.J., Pekar, J.J., Peltier, S.J., Petersen, S.E., Riedl, V., Rombouts, S.A.R.B., Rypma, B., Schlaggar, B.L., Schmidt, S., Seidler, R.D., Siegle, G.J., Sorg, C., Teng, G.J., Veijola, J., Villringer, A., Walter, M., Wang, L., Weng, X.C., Whitfield-Gabrieli, S., Williamson, P., Windischberger, C., Zang, Y.F., Zhang, H.Y., Castellanos, F.X., and Milham, M.P. Toward discovery science of human brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 107(10): 4734–4739, 2010.

Inline image-based reinforcement detection for concrete additive manufacturing processes using a convolutional neural network

Lukas Lachmayer¹, Lars Dittrich, Tobias Recker¹, Robin Dörrie², Harald Kloft² and Annika Raatz¹

¹Institute of Assembly Technology and Robotics, Leibniz University Hannover, Germany

²Institute of Structural Design, Technische Universität Braunschweig, Germany

lachmayer@match.uni-hannover.de

Abstract –

Within the scope of additive manufacturing of structural concrete components, the integration of reinforcement provides an inevitable opportunity to enhance the load bearing capacity of the components. Besides the rebar integration itself, ensuring as-planned concrete cover is key to achieve a stable and long-term legally permissible integration. The thickness of the as-built concrete cover however is unpredictably altered during printing by the varying material behaviour of the printed concrete. In addition, the lack of opportunities to anchor reinforcement elements before printing can lead to a displacement of reinforcement during printing. In this publication, we present an approach for determining the position of reinforcement elements within additively manufactured components without post-process measurement steps. During the printing process, RGB images and depth camera data are recorded by a camera mounted to the print head. Subsequently, a neural network is employed to distinguish between reinforcement structures and the deposited material within the coloured image. By overlaying the colour image data with the depth information a 3D point cloud is generated, within which the reinforcement is marked.

Keywords –

Additive Manufacturing; Process Control; Image Processing; Neural Network; Printing Robot

Initial reinforcement strategies for additive manufacturing are presented by Kloft et al. [3]. Aside from the mere integration during the printing process, an additional challenge lies in ensuring and providing proof of accurate positioning of the reinforcement [4]. Evidence is required not only to comply with and fulfil legal guidelines but, also for planning potential post-processing steps and guaranteeing component stability. There are two challenges in providing this evidence. At first, additional process steps should be avoided, so it is advisable to opt for inline data recording during printing. Second, additive manufacturing processes based on concrete generally present a contaminated environment, so robust data acquisition should be emphasized in this case.

In this publication, we present an approach for the fully automatic determination of the reinforcement position within additively manufactured components. Therefore, we record and evaluate image data during two large-scale 3D printing processes with a sensing unit to observe the process and to identify the integrated reinforcement. To distinguish between printed concrete and the integrated reinforcement within the recorded images and data, we propose and prove the applicability of convolutional neural networks (CNN). In section 2, we first describe the state of the art regarding reinforcement integration and detection. Section 3 explains our methodological approach, while section 4 outlines the description of the experimental investigations. Section 5 addresses the results, and section 6 covers the summary and outlook.

1 Introduction

In the context of concrete component manufacturing, the integration of reinforcement is essential for the generation of tensile strengths [1]. To achieve comparable loadbearing capabilities with additively manufactured components, reinforcement elements must also be integrated during 3D printing processes [2].

2 Related Work

Within this section, we first address the strategies for reinforcement integration during concrete 3D printing processes. This allows us to derive boundary conditions for the data acquisition, e.g. environmental influences such as the degree of air pollution influencing the

captured data quality, the required data, and image resolution to detect the reinforcement or the expected colour contrast between material and reinforcement. Afterwards, we evaluate state of art sensing approaches for their suitability and efficiency for the reinforcement determination task during 3D printing processes. Finally, we present research on potential methodological approaches and challenges to evaluate the recorded data.

2.1 Reinforcement Integration and boundary conditions for sensing

According to Pacillo et al., [5], most additive manufacturing processes in construction can be categorized into extrusion-based, binder jetting, or powder bed-based methods. Most of these processes occur without significant dust emission, leaving no special constraints to consider for integrating sensors in most cases. However, Shotcrete 3D Printing (SC3DP), developed at ITE TU Braunschweig [6], belonging to the extrusion-based processes, generates a considerable amount of contamination due to the concrete jetting. This additional challenge prompts us to utilize SC3DP as the most demanding printing process for reinforcement detection. Thus, we aim to develop a sensing approach that works even in such demanding conditions. The process is illustrated exemplarily in Fig. 1.



Figure 1. Shotcrete 3D Printing (SC3DP) at the DBFL of the ITE TU Braunschweig including short reinforcement bar insertion, as proposed by Dörrie et al. [7], and a 2D laser profiler running ahead of the printing nozzle. The goal of our paper is the detection of the depicted type of reinforcement elements.

In addition to the influences arising from the additive manufacturing process itself, the various options for reinforcement integration have different impacts on the to be-chosen measurement approach. Classen et al. [8] provide a brief overview of currently developed reinforcement strategies for additive manufacturing processes. All strategies can be divided according to the usage of flexible reinforcement e.g. textile fibres [9], or stiff reinforcement such as pre-manufactured steel meshes [10], short steel rebar pieces [11] or welding-while-printing utilizing wire-and-arc-additive-manufacturing (WAAM). In all cases, the diameter of an individual reinforcement bar is in the range of a few millimetres, allowing this to be considered as the desired sensor resolution.

2.2 Inline Sensing

For the examination of the position and condition of reinforcement bars within a conventionally cast component, various measurement methods exist [12]. In addition to X-ray imaging [13], for example, ultrasound [14] provides a different approach. These common methods share the commonality that only small areas are captured, and the examinations are carried out on the finished component.

However, in the context of additive manufacturing, there is the option to acquire data during the layer-by-layer construction process, eliminating the need for an additional measurement step. For inline measurement and recording during additive manufacturing processes, generally, three types of sensors are used. Wolfs et al. [15] utilized a 1D time-of-flight sensor to measure the nozzle-to-strand distance while printing. Yet, purely one-dimensional information is insufficient for the detection of (3D) reinforcement elements. An early method for 2D data recording is described by Doumanidis et al. [16], using a line laser and a CCD camera to record 2D profiles of the printed strand. An advanced setting is described by Xiong et al. [17], who use a similar CCD camera aside from a welding setup to generate image data of the process. Lindemann et al. [6] fundamentally transferred such 2D approaches to concrete additive manufacturing. Fig. 2 shows 2D profile data, recorded during SC3DP. The presence of slight elevations within each profile suggests the potential existence of reinforcement structures, see Fig. 1, at these points. However, it is important to note that a definitive identification is not achievable.

More meaningful data can be generated through the use of coloured images and three-dimensional measurements. A 3D data-gathering approach is presented by Kazemian et al. [18]. The authors utilized a nozzle-mounted camera to surveil a concrete additive manufacturing process gaining 3D images at each layer. The images are evaluated to detect over- or under-

extrusion for layer width adjustment. Since 3D data, especially in combination with colour images, provides the highest information content for detecting reinforcing elements, we choose a colour and depth camera as the most suitable sensor technology.

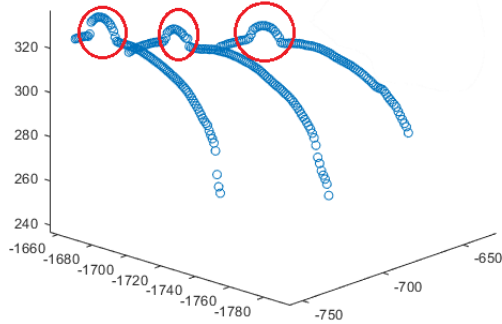


Figure 2. Three 2D profiles recorded during SC3DP at ITE TU Braunschweig. The slight elevations, marked in red in the data, could be part of reinforcing elements, but an explicit identification is not possible.

2.3 Data Evaluation

In the context of data analysis, there are fundamentally two approaches available for image data sets. On the one hand, mathematical methods such as Canny edge detection [19] or thresholding can be implemented. On the other hand, machine learning such as convolutional neural networks is suitable for the examination of image data recorded during additive manufacturing [20]. Previous related work has not yet conclusively shown which method is more effective. We intend to investigate this topic in this work. The following chapter takes a closer look at the individual methods.

3 Methodology

From the state of the art, we derive the necessity to determine the position of reinforcement elements within additively manufactured components. For this purpose, we implement a robust measurement technology which provides meaningful data through colour and depth images. The data is evaluated through mathematical methods and neural networks. The results provide insights into the more suitable approach. Therefore, we have designed a four-stage algorithm, including the following steps:

1. Depth and color image recording by running a camera ahead of the printing process.
2. Identification of rebar pixels within the colored image.
3. Extracting the correlated rebar pixels from the depth image to gain 3D information on rebar positioning.
4. Transforming the depth image into the robot workspace to generate an as-built model. Use a colour contrast to mark rebar elements.

To implement this approach, we first developed a module for the acquisition of inline image data.

3.1 Data acquisition

For inline data acquisition in the context of large-scale additive manufacturing processes based on concrete, we developed a dedicated multi sensor-measuring system. As shown in Fig. 3, this consists of an Intel RealSense D435 for colour and depth image recording. This data is used, as proposed previously, for reinforcement detection. A Raspberry Pi 4 is added to the system for data processing, a Wi-Fi module for data transmission and communication with the printing robot, and batteries ensure the power supply. All of these components are attached to a 2Dlaser profiler. The data of the profiler is utilized in parallel for online control of the printed strand width and height. The whole system runs ahead of the printing nozzle and is rotated by a stepper motor. The 2D-laser profiler and our online strand control approach are described in detail in our previous work [21].

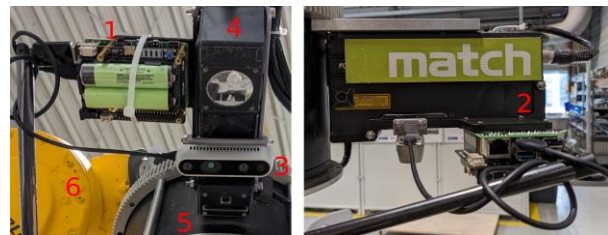


Figure 3. Sensor-module for image processing during 3D printing. 1) power supply 2) Raspberry Pi 4 for data processing 3) Intel RealSense D435 camera 4) Keyence 2D-laser profiler for additional online control 5) Slip ring and gearing for 360° endless rotation 6) printing robot

3.2 Selection of a data processing approach

To compare the two data evaluation approaches, mathematical methods and machine learning, and further develop the more beneficial one in the context of more extensive research, a test image was captured. Fig. 4 shows the test image which contains a printed strand of PU foam.



Figure 4. 3D printed PU-foam strand with inserted short reinforcement bars marked by red circles.

PU foam is used as a substitute material instead of concrete and was already determined to be a suitable substitute for concrete printing process investigations in previous process-related experiments. Short reinforcement bars (diameter 15mm, length 150mm) were inserted into the foam during printing. The strand dimensions are approximately 10 cm in width and 2 cm in height, closely resembling the concrete-based SC3DP process shown in Fig. 2. Printing and data collection are performed at the additive manufacturing test bench at the Institute of Assembly Technology and Robotics. The test rig is shown in Fig. 5. In this figure, the light colour of the PU-foam is visible. The PU-foam is occasionally darkened during the printing process using spray paint to reduce the colour contrast, resulting in the colour depicted in Fig. 4.

In an initial approach, we attempt to identify the reinforcing elements by detecting circular and angular characteristics using the mathematical method [19]. The first step of identifying the reinforcement bars using the mathematical approach is to detect the edges using Canny Edge Detection. In the second step a Hough transformation is performed, to search for circles and lines. Fig. 6 shows the subsequent algorithms' outputs. As a result, a multitude of arbitrary circles is recognized, making further development seem impractical.

A likewise unpromising result is observed when implementing thresholding methods for edge identification. Depending on the limit settings, either too many or no edges are detected. Consequently, we proceed to use convolutional neural networks (CNN) in further development. In particular, we train and utilize YOLOv8 [22].

4 Experimental data and CNN training

For the CNN training and validation, three different types of images are captured. The types are grouped as



Figure 5. Foam printing test bench at Institute of Assembly Technology and Robotics. 1) printing robot 2) printing head including nozzle and slip ring 3) material supply 4) sensor-module 5) printed column with colouring and rebar integration

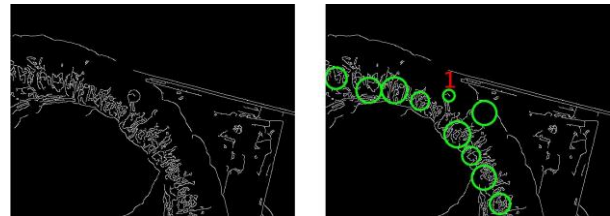


Figure 6. Canny edge detection and Hough transformation applied to a process image of PU-foam printing including short rebar insertion. Only the circle labelled with one correlates to a real rebar.

as follows:

1. sharp: The first set of image data was captured statically above the PU-foam column shown in Fig.
2. blurry: The second set of image data was captured during the printing of the PU-foam column. Due to the printing motion, this data contains motion blur.
3. Concrete: The last set contains image data captured at TU Braunschweig during a Shotcrete 3D printing process as shown in Fig. 1, used for validation only.

The advantage of using statically recorded data, as within the data set "sharp", of existing objects is that no

active printing process is needed. Therefore, an arbitrary number of training data can be generated without material wastage. As shown in Fig. 7, motion blur appears during the printing process. Such images are only part of the data set "blurry". Two identical YOLOv8 CNNs were trained and validated to determine whether the training set "sharp" without motion blur is suitable. The training of both networks ran for 150 cycles, each with 640 pre-labelled images. For validation, 59 images were used. Each image set contains 31 images with short reinforcement bars. The validation datasets for both networks include 50% of blurred images, as these occur during the printing process. The detection accuracy of the network trained with blurred data is ultimately 66%. The network without blurred data is close to 0%. Training data without blur is therefore insufficient.

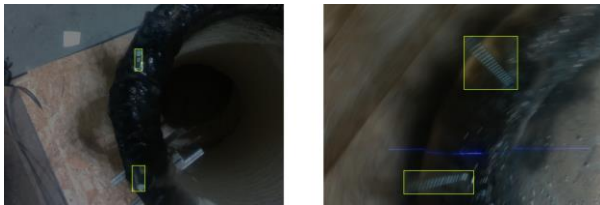


Figure 7. Image recorded with a static robot after printing (left), image recorded during the printing process (right)

5 Results

After training the networks and validating the net trained with blurred data, testing is carried out within the scope of the PU-foam printing processes. Fig. 8 depicts the top layer of a second manufactured column with incorporated reinforcement



Figure 8. Printed and coloured PU-foam column for evaluation. The rebar was added to the last layer and the camera was moved along the printing trajectory over the surface for one full circle. Movement speed was set to 100mm/min

Images were captured and evaluated with a frequency of 0.5 Hz using the trained network on the Raspberry PI 4.

After identification of the rebar pixels, labelling is transferred to the depth image, and the depth image is transformed into the robot's workspace. Fig. 9 shows the assembled and transformed data from all recorded and labelled images.

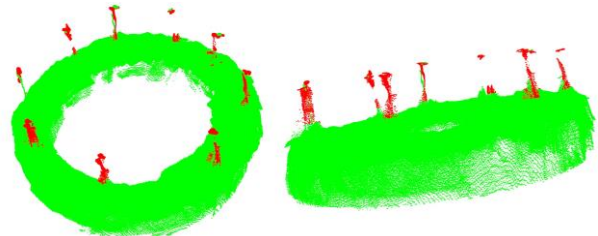


Figure 9. Depth camera data combined to a single point cloud by using the transformation into the robot base coordinate system. Data points identified as rebar are marked in red.

Within the image, the data points recognized as reinforcement are been marked in red colour. It is evident that all reinforcing bars were detected in at least one image, and thus, there exists a portion of red data points for each bar. Furthermore, it can be seen that there are green data points between the red ones, indicating that in some images, the reinforcing bar was not recognized. This correlates with the 66% accuracy achieved in section 4. Such low accuracy, however, is tolerable since multiple images are recorded and evaluated per reinforcement bar. The quantity of available images depends on three factors: the feed rate, the image size, and the data processing speed. In our applications, one reinforcement bar is visible within three images each. With an individual image accuracy of 66%, the overall probability of not recognizing a rod is as low as 4%. After successful identification during the experiments conducted with PU foam, the sensor module was used during the SC3DP. Fig. 10 shows the images marked by the same neural network.



Figure 10. Images recorded during SC3DP evaluated by the trained CNN to label the inserted short rebars.

Basic recognition of the reinforcing bars is possible without training the network with additional data from the concrete process. However, it is evident that the detection accuracy further decreases. This results in clear visible bars (see fig. 10 bottom left) being no longer recognized by the net.

6 Conclusion & Outlook

Inline identification of reinforcement during additive manufacturing processes provides an efficient means to obtain as-built data of a manufactured component without an additional process step. Especially for the verification of, for example, concrete cover or for planning subsequent post-processing, these data provide a solid foundation. In the course of our work, we first developed a module for recording such data. Furthermore, an approach for systematic evaluation was presented, tested and validated in two different production processes. While the presented approach represents a functional methodology for the identification and localization of reinforcement elements during printing, the following basic statements can be derived. Typical for the use of neural networks, and this also applies to the proposed use case, is the relevance of good training data. It turns out that the mere use of post-printing process images is not sufficient to achieve adequate classification during the printing process. The experimental investigations also show that mediocre classification quality can be partially compensated for by high-frequency recording and evaluation of the image data. Such oversampling is limited by the capabilities of the Raspberry Pi 4 in our case. A limitation in the use of the methodology also results from the need for a wireless network for data transmission, which may not always be available for onsite construction. The system is also not yet dust and waterproof in its current state. In the context of future research, we aim to further investigate the accuracy of the recorded data to determine the applicability for the generation of quality inspection documentation. Additionally, we need to evaluate additional network architectures regarding their resulting detection accuracy to improve robustness.

7 Acknowledgement

The authors gratefully acknowledge the funding by the Deutsche Forschungsgemeinschaft (DFG – German Research Foundation) – Project no. 414265976. The authors would like to thank the DFG for the support within the SFB/Transregio 277 – Additive manufacturing in construction. (Subproject B04)

References

- [1] Renato S. O. and Zuccarello F. A. An experimental study on the tensile strength of steel fiber reinforced concrete. *Composites Part B: Engineering*, 41(3):246–255, 2010.
doi:<https://doi.org/10.1016/j.compositesb.2009.12.003>.
- [2] Bos F., Wolfs R., Ahmed Z. and Salet T. Additive manufacturing of concrete in construction: potentials and challenges of 3D concrete printing. *Virtual and physical prototyping*, 11(3):209–225, 2016.
doi:<https://doi.org/10.1080/17452759.2016.1209867>.
- [3] Kloft H., Empelmann M, Hack N, Herrmann E. and Lowke D. Reinforcement strategies for 3d-concrete-printing. *Civil Engineering Design*, 2(4):131–139, 2020.
doi:<https://doi.org/10.1002/cend.202000022>.
- [4] Adam Neville. Concrete cover to reinforcement or cover-up? *CONCRETE INTERNATIONAL-DETROIT*-, 20:25–30, 1998.
- [5] Arcangelo P. G., Ranocchiai G., Loccarini F. and Fagone M. Additive manufacturing in construction: A review on technologies, processes, materials, and their applications of 3D and 4D printing. *Material Design & Processing Communications*, 3(5):e253, 2021.
doi:<https://doi.org/10.1002/mdp2.253>.
- [6] Lindemann H., Gerbers R., Ibrahim S., Dietrich F., Herrmann E., Dröder K., Raatz A. and Kloft H. Development of a shotcrete-3d-printing (sc3dp) technology for additive manufacturing of reinforced freeform concrete structures. In *First RILEM International Conference on Concrete and Digital Fabrication– Digital Concrete 2018*, pages 287–298. Springer, 2019.
doi:https://doi.org/10.1007/978-3-319-995199_27.
- [7] Dörrie R., Freund N., Herrmann E., Baghdadi A., Mai I., Galli F., David M., Dröder K., Lowke D. and Kloft H. Automated force-flow-oriented reinforcement integration for shotcrete 3d printing. *Automation in Construction*, 155:105075, 2023.
doi:<https://doi.org/10.1016/j.autcon.2023.105075>.
- [8] Classen M., Ungermann J. and Sharma. R. Additive manufacturing of reinforced concrete—development of a 3D printing technology for cementitious composites with metallic reinforcement. *Applied sciences*, 10(11):3791, 2020.
doi:<https://doi.org/10.3390/app10113791>.
- [9] Gantner S., Rothe T. N., Hühne C. and Hack. N. Reinforcement strategies for additive manufacturing in construction based on dynamic fibre winding: Concepts and initial case studies. In *Open Conference Proceedings*, volume 1, pages 45–59, 2022.
doi:<https://doi.org/10.52825/ocp.v1i.78>.
- [10] Marchment T. and Sanjayan. J. Mesh reinforcing

- method for 3D concrete printing. *Automation in Construction*, 109:102992, 2020.
doi:<https://doi.org/10.1016/j.autcon.2019.102992>.
- [11] Freund N., Dressler I. and Lowke. D. Studying the bond properties of vertical integrated short reinforcement in the shotcrete 3d printing process. In *Second RILEM International Conference on Concrete and Digital Fabrication: Digital Concrete 2020 2*, pages 612–621. Springer, 2020.
doi:https://doi.org/10.1007/978-3-030-49916-7_62.
- [12] Song H.-W. and Saraswathy V. Corrosion monitoring of reinforced concrete structures *Int. J. Electrochem. Sci.*, 2(1):1–28, 2007.
doi:[https://doi.org/10.1016/S1452-3981\(23\)170490](https://doi.org/10.1016/S1452-3981(23)170490).
- [13] Pei C., Wu W. and Ueaska M. Image enhancement for on-site x-ray nondestructive inspection of reinforced concrete structures. *Journal of X-Ray Science and Technology*, 24(6):797–805, 2016.
doi:10.3233/XST-160588.
- [14] Li D., Zhang S., Yang W. and Zhang. W. Corrosion monitoring and evaluation of reinforced concrete structures utilizing the ultrasonic guided wave technique. *International Journal of Distributed Sensor Networks*, 10(2):827130, 2014.
doi:<https://doi.org/10.1155/2014/827130>.
- [15] Wolfs R., Bos F. P., Van Strien E. C.F. and Salet T. A real-time height measurement and feedback system for 3D concrete printing. In *High Tech Concrete: Where Technology and Engineering Meet: Proceedings of the 2017 fib Symposium, held in Maastricht, The Netherlands, June12-14, 2017*, pages2474–2483.Springer, 2018.
doi:https://doi.org/10.1007/978-3-319-594712_282.
- [16] Domanidis C. and Kwak Y.-M. Geometry modelling and control by infrared and laser sensing in thermal manufacturing with material deposition. *J. Manuf. Sci. Eng.*, 123(1):45–52, 2001.
doi:<https://doi.org/10.1115/1.1344898>.
- [17] Xiong J. and Zhang. G. Adaptive control of deposited height in gmaw-based layer additive manufacturing. *Journal of Materials Processing Technology*, 214(4):962–968, 2014. ISSN 0924-0136.
doi:<https://doi.org/10.1016/j.jmatprotec.2013.11.014>.
- [18] Kazemian A., Yuan X., Davtalab O. and Khoshnevis B. Computer vision for real-time extrusion quality monitoring and control in robotic construction. *Automation in Construction*, 101:92–98, 2019.
doi:<https://doi.org/10.1016/j.autcon.2019.01.022>.
- [19] Heini M., Schmitt F. K. and Hausotte T. In-situ contour detection for additive manufactured workpieces. *Procedia CIRP*, 74:664–668, 2018.
doi:<https://doi.org/10.1016/j.procir.2018.08.051>.
- [20] Valizadeh M. and Wolff S. J. Convolutional neural network applications in additive manufacturing: A review. *Advances in Industrial and Manufacturing Engineering*, 4:100072, 2022.
doi:<https://doi.org/10.1016/j.aime.2022.100072>.
- [21] Lachmayer L., Müller N., Herlyn T. and Raatz A. Volume flow-based process control for robotic additive manufacturing processes in construction. In *2023 IEEE 19th International Conference on Automation Science and Engineering (CASE)*, pages 1–6. IEEE, 2023.
doi:10.1109/CASE56687.2023.10260620.
- [22] Jocher G., Chaurasia A. and Qiu. J. Ultralytics yolov8, 2023.
URL <https://github.com/ultralytics/ultralytics>.

Partial Personalization for Worker-robot Trust Prediction in the Future Construction Environment

Woei-Chyi Chang¹ Nestor F. Gonzalez Garcia² Behzad Esmaceli^{1,3} Sogand Hasanzadeh^{1*}

¹Lyles School of Civil Engineering, Purdue University, USA.

²Department of Systems Engineering and Computing, Universidad de los Andes, Columbia.

³School of Industrial Engineering, Purdue University, USA.

* Corresponding author

chang803@purdue.edu, nf.gonzalez@uniandes.edu.co, besmaei@purdue.edu, sogandm@purdue.edu

Abstract

Establishing proper trust between human workers and robots is crucial for ensuring safe and effective human-robot interaction in various industries, including construction. An accurate trust prediction facilitates timely feedback and interventions, helping workers calibrate their trust levels. While machine-learning modeling personalization (i.e., tailoring models to individual characteristics) has garnered attention in the literature, the conventional approach of developing a personalized model for each individual is impractical in labor-intensive industries like construction. Such an approach compromises efficiency and leads to an accuracy-efficiency tradeoff. To address this gap, this study aims to investigate the tradeoff inherent in model personalization and identify a cost-effective solution to enhance trust prediction accuracy without compromising efficiency. The results suggested that a partial model personalization method can effectively balance this tradeoff. Moreover, the proposed feature-based partial personalization approach enables a cost-effective trust prediction model development for the construction industry, demonstrating its broader applicability to other worker-related predictions in other settings. This study provides insights into the strategies to improve trust prediction accuracy while maintaining the efficiency of model development by considering the distinctiveness of the future construction industry.

Keywords –

Worker-robot trust; future construction; feature-based partial personalization; psychophysiology.

1 Introduction

As robots become more prevalent in the construction industry to improve automation in construction, how workers build their trust in robots during the interaction

has drawn increasing interest [1]. Trust has been identified as an essential element in any successful relationship, and its importance should be further highlighted in such dynamic and hazard-rich workplaces as construction sites to ensure occupational safety [2–6]. Because robots' perfect performance cannot be guaranteed to date in construction, an appropriate level of trust (neither excessive nor inadequate) represents a prerequisite to a secure and effective worker-robot interaction. Trust has been discerned as a dynamic concept where workers continuously update their trust levels based on human-related (e.g., gender), robot-related (e.g., transparency), and workplace-related factors (e.g., time pressure) [4,7]. To understand varying human trust and acknowledge the implicit nature of trust, there has been a growing interest in using real-time psychophysiological responses rather than self-report subjective measures [8].

The literature has identified the latent safety issue in the construction domain accompanied by workers' inappropriate trust levels in robots [9,10]. For example, in the study investigating workers' situational awareness of robots, Chang and his colleagues found that scheduling pressure in construction projects provoked workers' overtrust in a faulty robot, leading to their ignorance of the robot as a dynamic hazard [2]. To address the trust dynamics and the challenges of appropriate trust-building on the job site, a recent review study suggested developing trust prediction models trained by psychophysiological responses to better monitor and understand workers' real-time trust in robots [4]. Such predictive models are envisioned to facilitate early feedback and interventions to reduce accidents and enhance the safety of worker-robot interactions.

Given the pivotal role of prediction accuracy, researchers have deployed various techniques across domains, e.g., data augmentation (i.e., increase the diversity of a training dataset) [11], ensemble methods (i.e., combining the predictions of multiple individual models) [12], and personalization (i.e., tailor a model to

individual characteristics) [13]. Personalization, aiming to tailor the predictive model to the specific preferences, characteristics, or behavior of an individual user, has been prevalently deployed in sensor-based human activity recognition [13]. Considering worker variability in construction, personalizing a trust prediction model for individuals holds the potential to enhance accuracy. However, discussions in the literature also highlight several efficiency and privacy-related concerns associated with model personalization [14,15]. These tradeoffs might potentially impede the acceptance and deployment of personalized models. There is a paucity of research regarding the accuracy-efficiency tradeoff associated with model personalization, especially within labor-intensive sectors like construction. Given the criticality of trust in worker-robot interaction, this study aims to (1) explore the trust prediction model personalization in the future construction industry and (2) provide insights into cost-effective personalization strategies for worker-robot trust.

2 Background

2.1 The Review on Human-robot Trust Prediction

As robots become seamlessly integrated with workplaces across industries, the understanding and prediction of human trust have created considerable value for ensuring the efficiency and safety of human-robot interaction. In the existing literature, the fusion of machine learning (ML) with psychophysiological measurements has emerged as a powerful tool for trust prediction [16–19]. For example, Ayoub and his colleagues designed a simulation task where participants needed to drive an autonomous vehicle in a simulator while collecting their heart rate, eye-tracking, and galvanic skin response (GSR) during the experiment [16]. Multiple ML models (e.g., Decision Tree and Naïve Bayesian) were trained with the collected data to predict their trust levels in the vehicle. In the study investigating users' trust in an AI assistant, the authors developed a predictive model using an electroencephalography (EEG) sensor [18]. The proposed models in the literature have attained over 70% accuracy of trust prediction in a static setting where the participants did not necessarily exhibit physical movements. However, when dealing with dynamic and physically demanding work environments, the inherent challenge arises characterized by inevitable worker movements. These movements may inadvertently introduce motion artifacts and signal noises to psychophysiological data in such dynamic and physically demanding work environments, thereby impacting prediction accuracy. To tackle this challenge, recent research has implemented trust prediction strategies by (i)

leveraging deep learning (DL) techniques, (ii) incorporating diverse types of psychophysiological responses as training datasets, and (iii) employing DL auto-encoders to automatically extract important features and remove noises from the raw data [20]. The results showed an accuracy level comparable to what the existing literature attains in static environments. Model personalization may hold significant potential in tailoring models to individuals' specific demands to further enhance the current prediction performance [21].

2.2 Model Personalization

Model personalization refers to customizing models to individual preferences, characteristics, or behaviors [22]. However, the conventional model takes a generalized approach to train a one-size-fits-all model for all individuals. The personalization approach primarily aims to improve the model's performance by considering each user's unique features and patterns, thereby providing more accurate predictions. While model personalization has been advocated as an effective strategy for enhancing accuracy, previous studies have underscored latent privacy and efficiency issues associated with this approach [14,16]. For instance, in the study examining the relationship between personalization and privacy, the findings indicated that personalized model creation would trigger users' privacy concerns [15]. Moreover, there is an intuitive inference that the development of personalized models requires a significant investment of time and computational resources [17]. Notably, the construction industry presents a unique challenge because of the substantial workforce and frequent changes. This characteristic of the construction industry might lead to the inefficiency of personalizing models for individual workers. To address the accuracy-efficiency tradeoff in model personalization, researchers have proposed "partial model personalization" as a potential solution [14]. Partial personalization uses specific features (e.g., layers and parameters) for individuals to develop a tailored model for a particular group [23,24]. In the construction context, workers can be categorized into groups based on their specific features, accommodating the variability among workers, and alleviating the demands of training personalized models. Therefore, the partial model personalization facilitates achieving a balanced tradeoff between accuracy and efficiency in construction. In summary, this research strives to investigate an effective model personalization strategy for optimizing trust prediction in future construction.

3 Methodology

3.1 Experimental Design and Procedure

This study developed a human-robot collaborative bricklaying experiment in which workers must complete the bricklaying tasks with a bricklaying co-bot (i.e., MULE) and different types of drones. While MULE was designed to automatically lift and drop heavy concrete blocks for workers, human interventions were still needed (i.e., applying mortar and moving MULE to correct positions). Moreover, drones assisted with (i) monitoring the environment for safety, (ii) delivering new materials to workers in an elevated platform for efficiency, and (iii) inspecting workers' behaviors to examine work progress.

This study designed two modules to manipulate workers' trust levels in drones: (i) Baseline and (ii) Error modules. The Baseline module refers to the scenario where all types of drones exhibit error-free performance. On the contrary, the Error module includes various drone-related system failures and errors (e.g., workers were struck by drones) to assess how the workers' trust in drones might change. Noteworthy, workers might decrease their trust in the Baseline module due to personal preferences even though the drones performed flawlessly. Similarly, workers might increase their trust in the Error module because they did not identify or perceive the drone failures as not risky [8].

Participants were presented with an introduction to the experiment and the functionalities of the bricklaying

co-bot and drones. Training was provided to familiarize participants with the designated bricklaying task. Participants were equipped with a HTC VIVE headset, two controllers, three motion trackers, and psychophysiological wearable sensors (e.g., a Functional Near-Infrared Spectroscopy (fNIRS) and Empatica E4 wristband) and were asked to finish the Baseline and Error modules. Furthermore, A widely used 5-point Likert-scale trust questionnaire [25] was administered to collect their self-report trust levels in drones before and after each module, denoted by t_i (*initial trust before the Baseline module*), t_b (*trust after completing the Baseline module*), and t_e (*trust after completing the Error module*). All the procedures were approved by the Purdue Institutional Review Board (IRB). Figure 1 illustrates the experimental design and procedure of this study.

3.2 Participants

Eighty-nine participants (60 males and 29 females) were recruited to participate in this experiment. All the participants were from the departments of Civil Engineering and Construction Engineering and Management majors at Purdue University, representing the next generation of the workforce. The age range of participants was between 19 and 36 years ($M = 22.54$, $STD = 3.32$), with 64% of them having more than one year of experience in the construction industry.

3.3 Apparatus

The selected VR device was the HTC Vive Pro Eye

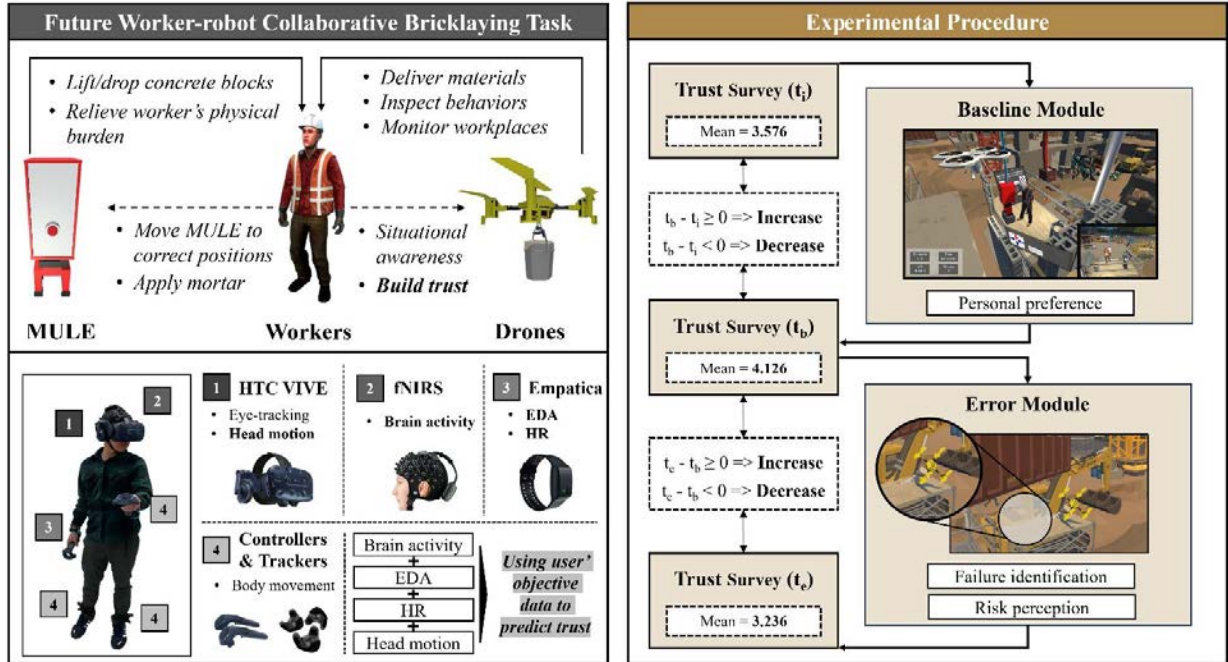


Figure 1. The experimental design and procedure of the bricklaying task.

with a refresh rate of 90 Hz and a field of view of 110° (manufactured by HTC Corporation, Taoyuan, Taiwan). The experiment was run on an Alienware PC with an AMD Ryzen 9 5950X 16-Core processor and an NVIDIA GeForce RTX 3090 graphics card. To capture participants' psychophysiological responses, (i) Empatica E4 wristband (manufactured by Empatica, Boston, United States) was to collect electrodermal activity (EDA) with a sampling frequency of 64 Hz and heart rate (HR) with a sampling frequency of 1 Hz and (ii) fNIRS (Brite 23, Artinis, Netherlands) was used to collect brain activation and cognitive processing with a sampling frequency of 10 Hz.

3.4 Data Collection and Pre-Processing

Various types of objective data were collected from participants during the experiment and used to train a trust prediction model. The data included (i) fNIRS, (ii) EDA, (iii) HR, and (iv) head motion captured by the headset. Due to the heterogeneous data sources, pre-processing approaches were used to ensure consistent sampling frequencies and compliance with model training requirements. Specifically, raw fNIRS signals were processed by Homer3 packages to output the hemodynamic response function (HRF). Then, the HRF data was divided into segments by considering a 10-second time window due to sequential dependencies and hemodynamic delayed activation [26,27]. Each 10s segment was converted into a 2D image because this transformation offers an enhanced representation of the high spatial resolution in fNIRS data. Furthermore, this study conducted max-min normalization, re-sampling, and 10s time-window segmentation for the EDA, HR, and head motion data. Figure 2 presents a schematic overview of the pre-processing in this study.

Participants were requested to subjectively report their trust levels in drones three times (i.e., $t_{i-\text{Mean}}$: 3.576; $t_{b-\text{Mean}}$: 4.126; $t_{e-\text{Mean}}$: 3.236). While the data indicated the trust increased during the Baseline module and decreased during the Error module, this labeling method would not consider the variance among participants (e.g., some participants decreased their trust in Baseline module). As highlighted above, the drone system failures did not necessarily lower human trust, and its perfection did not

guarantee increased trust. To address this issue, this study deployed customized labeling for each individual (i.e., the module in which each participant increased their trust was labeled as “increase” and the same logic for the “decrease”), as shown in Figure 1.

3.5 Model Development

The model development comprises two phases (Figure 2): (i) feature extraction and (ii) trust prediction. Compared to the traditional approach of manual feature extraction, this study considered automated extraction by applying the autoencoder (AE) technique. This technique has been proven to effectively reduce the dimension of the data with high representativeness of the extracted features. Due to the dissimilarity between the image (i.e., fNIRS) and time series (i.e., EDA, HR, and head motion) data, two types of AEs were implemented in this study. Specifically, convolution neural network (CNN)-based AE was developed for the fNIRS data, while the essential features of time-series data were extracted by long short-term memory (LSTM)-based AEs. All extracted features were then aggregated as the final trust prediction model input, which is a CNN model, to predict whether workers increase or decrease participants' trust during the interaction with robots in the future construction environment. Due to the page limit here, the details of the model structures and parameters were discussed in [20], and the personalization results and discussions are covered in this paper.

3.6 Partial Model Personalization

Multiple trust prediction models were developed for different groups of participants to investigate the performance of partial model personalization. Unlike full personalization, where an individual model is trained with one participant's data, this study employed a randomized grouping approach to achieve partial model personalization. The randomization process was repeated ten times, which refers to 10-fold cross-validation, to mitigate the potential bias introduced by the one-time randomized grouping. Each repetition involved the random distribution of participants into groups of 1 (generalization model), 2, 4, 8, and 16. Subsequently,

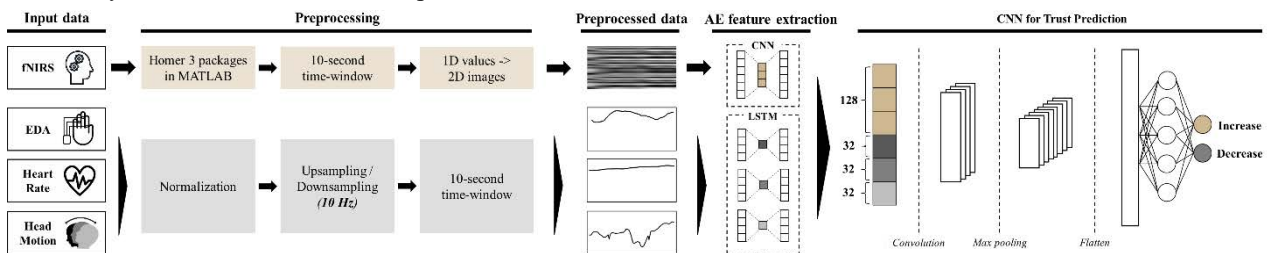


Figure 2. A schematic overview of the data preprocessing approaches and model development.

models were independently trained for each group within each repetition. The results were then averaged across the ten replications to provide a more robust and unbiased assessment of the model's performance. Lastly, the accuracy and efficiency values were normalized between 0 and 1 to examine the balance within the accuracy-efficiency tradeoff.

4 Results

Table 1 shows the average accuracy, training time, and efficiency of models across different levels of personalization. The accuracy of training and testing data from the generalized trust prediction model (i.e., only training one generalized model with no personalization) indicated 79.3% and 71.8%, respectively. As can be seen in Table 1, the testing accuracy increased as participants were divided into more groups (#1: 71.8%; #2: 71.3%; #4: 72.7%; #8: 77.1%; #16: 77.5%). However, higher level of personalization led to an increase in training time ((#1: 183.57s; #2: 181.67s; #4: 185.89s; #8: 192.26s; #16: 216.40s) and a decrease in efficiency (#1: 5.44e-3; #2: 5.50e-3; #4: 5.38e-3; #8: 5.20e-3; #16: 4.62e-3). These results demonstrate a trade-off between accuracy and efficiency within model personalization.

Table 1 Overview of the changes in accuracy and efficiency in partial model personalization.

# of groups	Testing accuracy	Training time	Efficiency (1/time)
1	71.8%	183.57s	5.44e-3
2	71.3%	181.67s	5.50e-3
4	72.7%	185.89s	5.38e-3
8	77.1%	192.26s	5.20e-3
16	77.5%	216.40s	4.62e-3

Figure 3 provides a visual representation of changes in the accuracy and efficiency variations based on a normalized scale. The two lines intersect approximately when the participants were divided into seven groups. These findings suggest an intermediate level of model personalization to better balance the model's accuracy and efficiency.

5 Discussion

An accurate understanding and prediction of workers' trust are paramount for ensuring the safety of human-robot interaction in such dynamic and hazardous workplaces as construction environments [28]. Although an acceptable prediction accuracy has been achieved by combining varied sources of workers' objective data and employing the AE technique, the various individual characteristics of workers present an opportunity for further enhancement, particularly through the integration

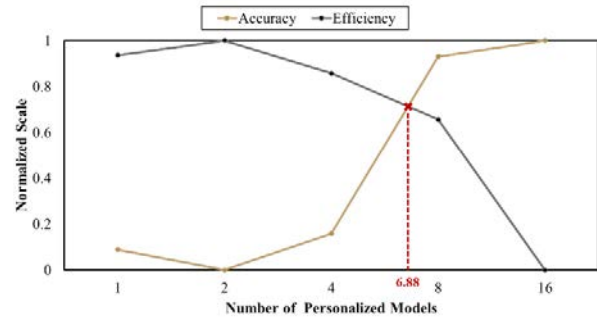


Figure 3. A graphical representation of the trade-off between model accuracy and efficiency.

of model personalization. The results of this study align with the construction literature, suggesting the effectiveness of model personalization to improve accuracy. For example, in the analysis focusing on cognitive load classification using fNIRS responses, Zhu and his colleagues found that the fully personalized models outperformed the generalized model [21]. However, the previous literature has not extensively addressed the efficiency issue, which is presumed to be very critical in the construction industry. This suggests an avenue for research in this paper to explore the efficiency implications of model personalization in the construction context.

The findings showed a decline in efficiency accompanied by a higher degree of personalization, which poses challenges in a labor-intensive industry where tasks require the presence of many workers on jobsites. Even though model personalization enhances trust prediction accuracy, full model personalization is impractical due to the complexity and time-consuming nature of training a personalized model for each worker, coupled with the dynamic crew-based nature of the industry. In recent research that aimed to develop DL models to recognize construction workers' postures in manual construction tasks, Zhao and Obonyo argued that the full model personalization is infeasible and proposed an improved model (i.e., integrating one CNN layer with two LSTM layers) with high generalizability to accommodate the variation among workers [29]. However, their proposed model was limited to the posture-related data of four participants, highlighting the need for an alternative approach to tackle the accuracy-efficiency tradeoff.

According to the results of this study, the balance between accuracy and efficiency was achieved when training seven partially personalized models to accurately predict workers' trust without compromising efficiency. That is, grouping workers into a limited number of groups and training a model for each group are suggested to achieve a better balance of efficiency and accuracy. While employing a randomized grouping method to categorize workers, this study proposes using feature-

based model personalization as an optimal grouping strategy to enhance cost-effectiveness. Specifically, this feature-based approach necessitates an initial grouping based on workers' unique characteristics affecting trust-building. For example, prior studies have suggested the effects of human factors on trust, such as gender [30,31], personality [32,33], experience [34,35], and self-confidence [36] which can be considered as features for grouping. Understanding how worker-related characteristics will impact their trust in robots is a prerequisite to implementing this approach. This proposed featured-based model personalization could enhance the generalizability, scalability, and inclusivity of the proposed model because new workers can be classified into one of the existing groups without allocating extra resources to train another model [37]. More importantly, this approach extends its applicability beyond trust prediction to other worker-related classifications and predictions in construction. Ultimately, human-centered construction can be established in the foreseeable future [38].

While this study offers considerable value to the body of knowledge, there are some limitations worth mentioning. First, the recruited participants in this experiment represent the next generation of the construction workforce, whereas it is worthwhile to replicate the study with the current workforce and incorporate a larger group of participants with varied backgrounds. Second, the literature has mentioned the privacy concern associated with the full level of model personalization. While this issue is assumed to be mitigated by deploying the proposed feature-based personalization approach, future research is recommended to explore the accuracy-privacy trade-off within model personalization.

6 Conclusion

While existing literature underscores the benefit of model personalization in enhancing accuracy, its inherent inefficiency poses challenges, particularly within the construction context. Therefore, this study addresses this challenge by identifying the nuanced balance within the accuracy-efficiency tradeoff of model personalization and proposing a cost-effective (i.e., achieving high accuracy without consuming excessive time by regarding time as cost) solution to predict workers' trust. The proposed feature-based partial personalization approach addresses the unique labor-intensive and crew-based nature of the construction sector. This approach suggests classifying workers into groups based on a specific influential trust feature and training partial personalized models for each group. The proposed approach ensures an accurate prediction of models while maintaining the efficiency of the training process.

This research contributes to the body of knowledge by (i) introducing the concept of partial model personalization to the construction industry, recognizing the inefficiency of full personalization, (ii) navigating the balance within the accuracy-efficiency tradeoff, and (iii) proposing a cost-effective feature-based personalization strategy for conducting trust-related and other worker-related predictions in the construction domain. This research calls for further research initiatives to refine and expand the application of cost-effective model personalization strategies to foster safer and more efficient human-robot collaborations in construction environments.

Acknowledgment

This work was supported by the National Science Foundation Future Work-Human Technology Frontier (FW-HTF) (2310210 and 2128970). Any opinions, findings, and conclusions in this material are those of the authors and do not reflect the views of the National Science Foundation.

References

- [1] Chang WC, Ryan SM, Hasanzadeh S, Esmaili B. Attributing responsibility for performance failure on worker-robot trust in construction collaborative tasks. In: *European Conference on Computing in Construction.*; 2023. doi:10.35490/EC3.2023.205
- [2] Chang WC, Borowiak A, Hasanzadeh S. The Importance of Situational Awareness in Future Construction Work: Toward the Effects of Faulty Robot, Trust, and Time Pressure. In: *ASCE International Conference on Computing in Civil Engineering.* ; 2023.
- [3] Gheisari M, Esmaili B. Unmanned Aerial Systems (UAS) for Construction Safety Applications. In: *Construction Research Congress 2016.* American Society of Civil Engineers; 2016:2642-2650. doi:10.1061/9780784479827.263
- [4] Chang WC, Hasanzadeh S. Towards a Framework for Trust-Building between Humans and Robots in the Construction Industry: A Systematic Review on Current Research and Future Directions. *Journal of Computing in Civil Engineering.* 2024;38(3). doi:https://doi.org/10.1061/JCCEE5.CPENG-5656
- [5] Luo Y, Seo J, Hasanzadeh S. Using Eye-Tracking to Examine the Cognitive Process of Hazard Recognition in Construction VR Scenarios. In: *Computing in Civil Engineering*

2023. American Society of Civil Engineers; 2024:541-548. doi:10.1061/9780784485248.065
- [6] Luo Y, Ahn S, Abbas A, Seo J, Cha SH, Kim JI. Investigating the impact of scenario and interaction fidelity on training experience when designing immersive virtual reality-based construction safety training. *Developments in the Built Environment*. 2023;16:100223. doi:10.1016/j.dibe.2023.100223
- [7] Khavas ZR. A Review on Trust in Human-Robot Interaction. *arXiv preprint*. Published online May 20, 2021. <http://arxiv.org/abs/2105.10045>
- [8] Chang WC, Esmaceli B, Hasanzadeh S. Understanding Worker-Robot Trust in Future Construction: The Impacts of System Failures and Risk Perception. *Autom Constr*. Published online 2024.
- [9] Jeelani I, Gheisari M. Safety challenges of UAV integration in construction: Conceptual analysis and future research roadmap. *Saf Sci*. 2021;144:105473. doi:10.1016/j.ssci.2021.105473
- [10] You S, Kim JH, Lee SH, Kamat V, Robert LP. Enhancing perceived safety in human-robot collaborative construction using immersive virtual environments. *Autom Constr*. 2018;96:161-170. doi:10.1016/j.autcon.2018.09.008
- [11] Shorten C, Khoshgoftaar TM. A survey on Image Data Augmentation for Deep Learning. *J Big Data*. 2019;6(1):60. doi:10.1186/s40537-019-0197-0
- [12] Sagi O, Rokach L. Ensemble learning: A survey. *WIREs Data Mining and Knowledge Discovery*. 2018;8(4). doi:10.1002/widm.1249
- [13] Ferrari A, Micucci D, Mobilio M, Napoletano P. Deep learning and model personalization in sensor-based human activity recognition. *J Reliab Intell Environ*. 2023;9(1):27-39. doi:10.1007/s40860-021-00167-w
- [14] Pillutla K, Malik K, Mohamed AR, Rabbat M, Sanjabi M, Xiao L. Federated Learning with Partial Model Personalization. In: Chaudhuri K, Jegelka S, Song L, Szepesvari C, Niu G, Sabato S, eds. *Proceedings of the 39th International Conference on Machine Learning*. Vol 162. Proceedings of Machine Learning Research. PMLR; 2022:17716-17758. <https://proceedings.mlr.press/v162/pillutla22a.html>
- [15] Toch E, Wang Y, Cranor LF. Personalization and privacy: a survey of privacy risks and remedies in personalization-based systems. *User Model User-adapt Interact*. 2012;22(1-2):203-220. doi:10.1007/s11257-011-9110-z
- [16] Ayoub J, Avetisian L, Yang XJ, Zhou F. Real-Time Trust Prediction in Conditionally Automated Driving Using Physiological Measures. *IEEE Transactions on Intelligent Transportation Systems*. Published online 2023:1-9. doi:10.1109/TITS.2023.3295783
- [17] Akash K, Hu WL, Jain N, Reid T. A classification model for sensing human trust in machines using EEG and GSR. *ACM Trans Interact Intell Syst*. 2018;8(4). doi:10.1145/3132743
- [18] Ajenaghughrure IBen, Sousa SDC, Lamas D. Measuring Trust with Psychophysiological Signals: A Systematic Mapping Study of Approaches Used. *Multimodal Technologies and Interaction*. 2020;4(3):63. doi:10.3390/mti4030063
- [19] Hu WL, Akash K, Jain N, Reid T. Real-Time Sensing of Trust in Human-Machine Interactions. In: *IFAC-PapersOnLine*. Vol 49. Elsevier B.V.; 2016:48-53. doi:10.1016/j.ifacol.2016.12.188
- [20] Woei-Chyi Chang, Nestor F. Gonzalez Garcia, Sogand Hasanzadeh. Deep Learning-based Prediction of Human-robot Trust Dynamics in the Future Construction Using Neuro-psychophysiological Measurements. *Journal of Management in Engineering*. Published online 2024.
- [21] Zhu Q, Shi Y, Du J. Wayfinding Information Cognitive Load Classification Based on Functional Near-Infrared Spectroscopy. *Journal of Computing in Civil Engineering*. 2021;35(5). doi:10.1061/(ASCE)CP.1943-5487.0000984
- [22] Lin CY, Marculescu R. Model Personalization for Human Activity Recognition. In: *2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. IEEE; 2020:1-7. doi:10.1109/PerComWorkshops48775.2020.9156229
- [23] Pereira JA, Matuszyk P, Krieter S, Spiliopoulou M, Saake G. A feature-based personalized recommender system for product-line configuration. In: *Proceedings of the 2016 ACM SIGPLAN International Conference on Generative Programming: Concepts and Experiences*. ACM; 2016:120-131. doi:10.1145/2993236.2993249
- [24] Dave D, DeSalvo DJ, Haridas B, et al. Feature-Based Machine Learning Model for Real-Time Hypoglycemia Prediction. *J Diabetes Sci Technol*. 2021;15(4):842-855. doi:10.1177/1932296820922622
- [25] MUIR BM. Trust in automation: Part I. Theoretical issues in the study of trust and human

- intervention in automated systems. *Ergonomics*. 1994;37(11):1905-1922. doi:10.1080/00140139408964957
- [26] Quaresima V, Ferrari M. Functional Near-Infrared Spectroscopy (fNIRS) for Assessing Cerebral Cortex Function During Human Behavior in Natural/Social Situations: A Concise Review. *Organ Res Methods*. 2019;22(1):46-68. doi:10.1177/1094428116658959
- [27] Abdalmalak A, Milej D, Yip LCM, et al. Assessing Time-Resolved fNIRS for Brain-Computer Interface Applications of Mental Communication. *Front Neurosci*. 2020;14. doi:10.3389/fnins.2020.00105
- [28] Luo Y, Seo J, Hasanzadeh S, Abbas A. Exploring the Schema of Attention and Search Strategy in Different Levels of Situation Awareness under Fall Hazard Conditions. In: *Construction Research Congress 2024*. American Society of Civil Engineers; 2024:802-811. doi:10.1061/9780784485293.080
- [29] Zhao J, Obonyo E. Convolutional long short-term memory model for recognizing construction workers' postures from wearable inertial measurement units. *Advanced Engineering Informatics*. 2020;46:101177. doi:10.1016/j.aei.2020.101177
- [30] Ghazali AS, Ham J, Barakova EI, Markopoulos P. Effects of robot facial characteristics and gender in persuasive human-robot interaction. *Frontiers Robotics AI*. 2018;5(JUN). doi:10.3389/frobt.2018.00073
- [31] Law T, Chita-Tegmark M, Scheutz M. The Interplay Between Emotional Intelligence, Trust, and Gender in Human-Robot Interaction. *Int J Soc Robot*. 2021;13(2):297-309. doi:10.1007/s12369-020-00624-1
- [32] Salem M, Lakatos G, Amirabdollahian F, Dautenhahn K. Would You Trust a (Faulty) Robot?: Effects of Error, Task Type and Personality on Human-Robot Cooperation and Trust. In: *ACM/IEEE International Conference on Human-Robot Interaction*. Vol 2015-March. IEEE Computer Society; 2015:141-148. doi:10.1145/2696454.2696497
- [33] Zhou J, Luo S, Chen F. Effects of personality traits on user trust in human-machine collaborations. *Journal on Multimodal User Interfaces*. 2020;14(4):387-400. doi:10.1007/s12193-020-00329-9
- [34] Clare AS, Cummings ML, Repenning NP. Influencing trust for human-automation collaborative scheduling of multiple unmanned vehicles. *Hum Factors*. 2015;57(7):1208-1218. doi:10.1177/0018720815587803
- [35] Sanders TL, MacArthur K, Volante W, et al. Trust and prior experience in human-robot interaction. In: *Proceedings of the Human Factors and Ergonomics Society*. Vol 2017-October. Human Factors and Ergonomics Society Inc.; 2017:1809-1813. doi:10.1177/1541931213601934
- [36] Sanchez J, Rogers WA, Fisk AD, Rovira E. Understanding reliance on automation: Effects of error type, error distribution, age and experience. *Theor Issues Ergon Sci*. 2014;15(2):134-160. doi:10.1080/1463922X.2011.611269
- [37] Chang WC, Ismael Becerra J, Karalunas SL, Esmaeili B, Yu LF, Hasanzadeh S. Pioneering Research on a Neurodiverse ADHD Workforce in the Future Construction Industry. In: *Construction Research Congress 2024*. American Society of Civil Engineers; 2024:283-292. doi:10.1061/9780784485293.029
- [38] Chae Y, Gupta S, Ham Y. Design and Evaluation of Human-Centered Visualization Interfaces in Construction Teleoperation. In: *Construction Research Congress 2024*. American Society of Civil Engineers; 2024:109-118. doi:10.1061/9780784485262.012

Automatic Quality Inspection of Rebar Spacing Using Vision-Based Deep Learning with RGBD Camera

Songyue Wang¹, Lu Deng^{1,2,*}, Jingjing Guo^{1,2}, Mi Liu¹, and Ran Cao^{1,2}

¹College of Civil Engineering, Hunan University, China

²Hunan Provincial Key Laboratory for Damage Diagnosis for Engineering Structures, Hunan University, China

*Corresponding Author

songyue@hnu.edu.cn, *denglu@hnu.edu.cn, guojingjing@hnu.edu.cn, liumi@hnu.edu.cn, rcao@hnu.edu.cn

Abstract –

Non-compliant rebar placement in reinforced concrete structures directly impacts the strength and durability of the overall structure. Current inspection methods heavily rely on manual techniques, introducing subjectivity and potential errors. This paper proposes a method for automatic quality inspection of rebar spacing using vision-based deep learning combined with RGBD camera. The method consists of three modules: (1) The image preprocessing module applies point cloud plane fitting techniques to eliminate the interference of the background rebar layer and rotate the image of the current rebar layer; (2) The recognition and localization module employs YOLOV8 keypoint detection algorithm to obtain pixel coordinates of rebar crosspoints, which are then transformed into spatial coordinates in the camera coordinate system; (3) The automatic inspection module proposes an automatic spacing measurement method based on the pixel coordinates of rebar crosspoints, enabling the inspection of all rebar spacings in the image and evaluating their compliance. Experimental results demonstrate the robustness of the pixel segmentation method, showcasing its applicability for compliance inspection with an average error of 2.65mm. The study concludes by suggesting potential directions for future research.

Keywords –

Keypoint detection; RGBD camera; Rebar spacing; Distance detection; Quality inspection

1 Introduction

Rebar is a vital structural element in reinforced concrete structures, providing support and ensuring proper load distribution. The correct placement between rebars is vital for the overall quality, strength, and durability of the structure [1]. Building codes in many countries, such as China, specify allowable deviation for

rebar placement, typically $\pm 20\text{mm}$ for bidirectional rebar mesh [2], as shown in Figure 1. Currently, quality inspection for rebar placement is mainly focused on the concealed engineering acceptance stage before concrete casting. However, inadequate quality control during construction process can lead to potential issues such as rework, construction delays, and even catastrophic consequences [3]. Considering the sequential layering construction process involved in Multi-layered bidirectional rebar, monitoring rebar spacing should also adopt a layer-by-layer approach. Therefore, strengthening rebar quality inspection during the construction process is crucial to minimize errors and enhance structural safety.

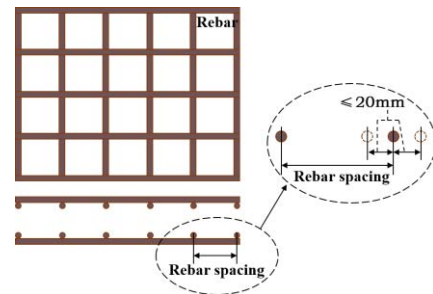


Figure 1. Illustration of the allowable deviation for rebar spacing

Traditionally, rebar placement inspection relies on manual methods involving experienced inspectors, which are cumbersome, time-consuming, and error-prone [4]. Recent advancements in computer vision technology have provided opportunities for automating dimensional quality inspection of reinforced concrete components. Current research has mainly focused on prefabricated components (e.g., precast panels and beams) [5][6][7],[8], with some exploring its application to rebar inspection [9][10],[11]. However, the reliance on 3D computer vision and point cloud models presents challenges such as complexity and limited automation. Therefore, there is an urgent need for a rapid, accurate

method to automatically inspect the dimensional quality of rebar.

The image-based measurement method enables real-time measurements [12], critical for monitoring rebar spacing layer by layer in the construction process. Existing algorithms struggle with interference from background rebar layers when inspecting the current layer. Kardovskyi et al. [13] employed Mask Region-Based Convolutional Neural Network (Mask R-CNN) for single-layer rebar instance segmentation but did not extend it to double-layer bidirectional rebar. An et al. [14] proposed a method using images and a laser rangefinder but faced challenges in layer attribution for bidirectional rebar. Jin et al. [15] introduced a neural network-based method but required consistent camera distance and angle for effective depth filtering, posing limitations in practical applications.

To address the aforementioned problems, this paper proposes a novel method for automatic quality inspection of rebar spacing using vision-based deep learning combined with RGBD camera. The key innovations are as follows: (1) Introducing a method to filter the background rebar layer using point cloud plane fitting, addressing interference in the current layer. (2) Presenting an automatic spacing measurement method based on the pixel coordinates of the rebar crosspoint for inspecting all rebar spacing in the image. (3) Proposing an intelligent end-to-end method for rebar spacing inspection. The feasibility and reliability of the proposed method are validated in physical double-layer bidirectional rebar structures.

2 Theoretical method

This paper presents a method for automatic quality inspection of rebar spacing using keypoint detection algorithm combined with RGBD camera. As shown in Figure 2, the automatic rebar spacing inspection consists of three modules: the image preprocessing, recognition and localization, and automatic inspection module. A detailed explanation is as follows.

2.1 Image preprocessing module

Obtaining the rebar pixels is crucial for spacing inspection. However, because of the feature similarity existing in double-layer bidirectional rebar, the background rebar layer could greatly influence the inspection of rebar spacing in the top layer. To extract the pixels belonging to the top layer, this paper employed a point cloud plane fitting technique based on depth information. Firstly, the RGB and depth images, captured by using the RGBD camera, are transformed into point clouds by incorporating the camera intrinsics, and then the passthrough filter is applied to denoise the point clouds, retaining the region of interest containing the rebar points. Subsequently, the denoised point clouds undergo RANSAC plane fitting algorithm, wherein various planes are segmented to obtain multiple point cloud planes along with their respective plane equations $Ax + By + Cz + D = 0$. Optimal plane parameters are determined through the minimization of aggregate vertical distances from each rebar layer's data points to the plane. To identify the point cloud plane containing the foremost rebar layer, considering that the coordinate reference system of the point cloud has its origin at the camera's position, the distance from the origin (0,0,0) to each point cloud plane is calculated, as shown in Equation (1). The plane with the shortest distance is identified as the point cloud plane of the current rebar layer.

$$dis = \frac{|D|}{\sqrt{A^2 + B^2 + C^2}} \quad (1)$$

To achieve automatic spacing inspection in Section 2.3, it is essential to ensure that the rebar in the image is approximately parallel to the x and y axes. For images that do not meet this requirement, this paper employs a rotation-based processing approach, which involves rotating the point cloud plane of the current rebar layer with the rotation center at the coordinate origin (0,0,0). This paper determines the rotation angle by taking the minimum angle between each rebar and the xz plane in the current layer. Utilizing the RANSAC algorithm, the

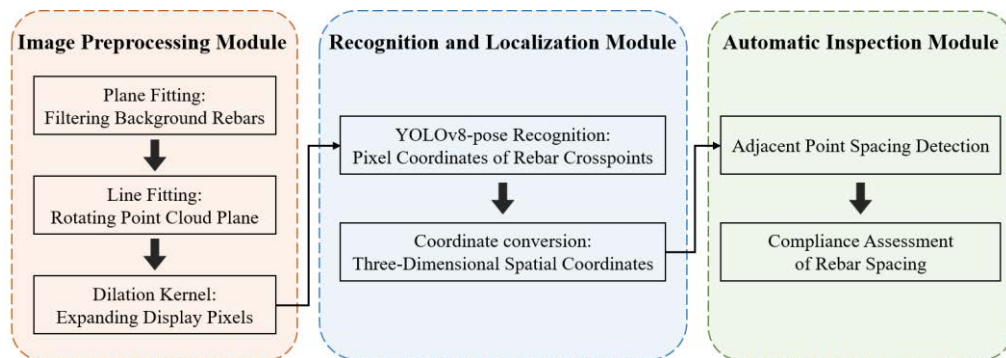


Figure 2. Flowchart of proposed formwork

spatial line fitting is applied to the point cloud plane of the current rebar layer, resulting in multiple rebar lines and their corresponding equations. Subsequently, the angle θ_i between each rebar line and the xz plane is calculated, as shown in Equation (2). By comparing these angles, the minimum value θ_{\min} is identified as the rotation angle. Correspondingly, the direction vector \vec{A}_{\min} of the rebar line associated with θ_{\min} is obtained.

$$\sin \theta_i = |\cos \phi_i| = \left| \frac{\vec{A}_i \cdot \vec{u}}{|\vec{A}_i| \cdot |\vec{u}|} \right| \quad (2)$$

where ϕ_i is the angle between the direction vector of the rebar and the normal vector $\vec{u} = (0,1,0)$ of the xz plane, \vec{A}_i is the directional vector of the rebar line defined by $y = \vec{A}_i x + \vec{B}_i$, \vec{B}_i is the 3D intercept of the rebar line.

$\vec{A}_{\min xz}$ is the projection of \vec{A}_{\min} on the xz plane. The rotation direction of the point cloud is determined by the z-axis component of the cross product $\vec{A}_{\min} \times \vec{A}_{\min xz}$. If it is greater than zero, the rotation direction is counterclockwise; otherwise, the rotation direction is clockwise. Finally, the rotated point cloud plane is mapped into an image of the current rebar layer. After processing, the recognition of rebar crosspoints on this image effectively avoids interference from other rebar layers.

2.2 Recognition and localization module

2.2.1 Introduction to the YOLOV8 framework

YOLOv8 from Ultralytics [16], the latest iteration of the YOLO-based object detection algorithm [17] series, boasts advanced capabilities encompassing object detection, instance segmentation, keypoint detection, tracking, and classification. By replacing the detection head of YOLOv8 with a pose detection head, it can be repurposed for keypoint detection. In contrast to existing keypoint detection algorithms, YOLOv8-pose eliminates the need for post-processing steps inherent in bottom-up algorithms to group detected keypoints into a target object, and also circumvents the top-down approach of first employing object detection algorithms to find bounding boxes and then conducting keypoint detection within each box. YOLOv8-pose introduces an end-to-end training method that associates all keypoints of the detected objects with bounding boxes. This model simultaneously learns the tasks of object detection and keypoint detection, employing a joint loss function to account for the correlation between the two tasks and shared features.

The YOLOv8-pose model comprises the following key components: input, backbone network, neck network,

and detection head, as shown in Figure 3. After undergoing the preprocessing outlined in Section 2.1, the images are scaled and subjected to data augmentation techniques such as mosaic and mixup before being input into the network. In the backbone, the C2f module based on cross stage partial (CSP) [18] and the SPPF module based on spatial pyramid pooling (SPP) [19] are used to extract image features to eliminate redundant operations and expedite feature fusion. The neck network combines the Feature Pyramid Network (FPN) [20] and the Path Aggregation Network (PAN) [21] structure, facilitating bidirectional fusion of low-level and high-level features, thereby improving the model detection performance across different scales. YOLOv8-pose uses a decoupled detection head to calculate the loss of localization and classification for bounding boxes and keypoints through four parallel branches of convolution.

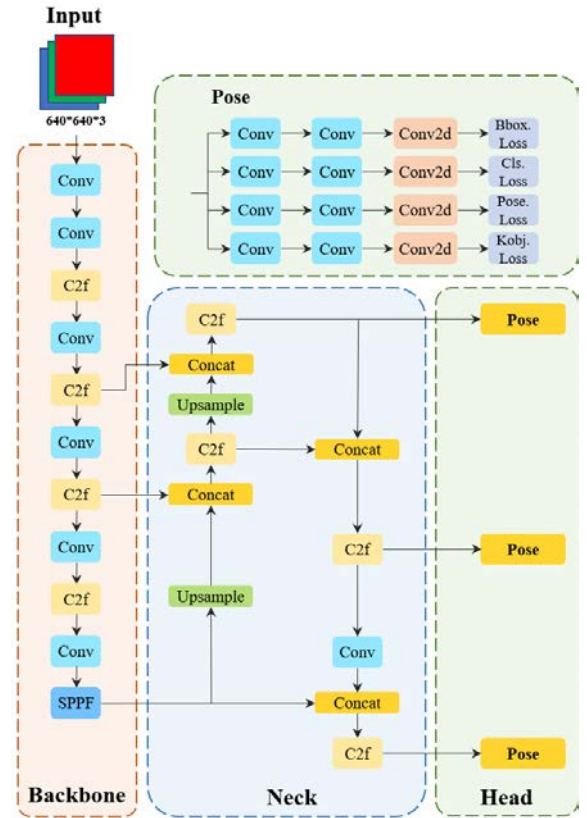


Figure 3. The network architecture of YOLOv8-pose

2.2.2 Keypoint spatial localization

The pixel coordinates of rebar crosspoints in the image can be predicted through the above keypoint detection algorithm. By integrating depth information and camera intrinsic parameters, these pixel coordinates can be transformed into spatial coordinates in the camera coordinate system.

Depth information can be obtained from the depth image. Camera intrinsic parameters are the inherent properties of a camera, including focal lengths (f_x and f_y), principal point coordinates (c_x , c_y), etc. The camera intrinsic matrix K is shown in Equation (3).

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

For any given point, the transformation relationship between the camera coordinate system and the pixel coordinate system is described by Equation (4).

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = K \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} \quad (4)$$

where (u, v) is the coordinate in the pixel coordinate system, and (X_c, Y_c, Z_c) is the coordinate in the camera coordinate system.

2.3 Automatic inspection module

Through the aforementioned steps, the 3D spatial coordinates (X_c, Y_c, Z_c) and pixel coordinates (u, v) of all rebar crosspoints in the image have been obtained. This paper proposes a method for automatic spacing measurement based on the pixel coordinates of rebar crosspoints, aiming to obtain the rebar spacing values.

This paper transforms the problem of measuring the rebar spacing into measuring the distance between two rebar crosspoints. Only the distance between adjacent crosspoints accurately reflects the spacing of adjacent rebars, as shown in Figure 4. Therefore, this paper defines the adjacent points for a given point as the four points closest to it in the up, down, left, and right directions. In terms of code implementation, two points are considered adjacent if they satisfy the following conditions: (1) one of the pixel coordinates of the two points is considered extremely close, which is equal to or less than a small threshold; (2) the other pixel coordinate is the minimum among all points satisfying the first condition. It is noted that a point is only connected to adjacent points on its right and above, excluding those on its left and below. This design ensures that there is no re-

petition in distance measurements while traversing all points.

By measuring the distances between all adjacent rebar crosspoints, all rebar spacing in the image can be obtained. Finally, based on the allowable deviation, assess whether these comply with the specified standards.

3 Experiment setup

The proposed method was tested on a double-layer bidirectional rebar cage with 8mm diameter rebar, as shown in Figure 5. The rebar cage, fabricated by spot welding at crosspoints, had dimensions of $2\text{m} \times 0.2\text{m} \times 0.9\text{m}$. The designed rebar spacing was $100\text{mm} \times 10$ and $200\text{mm} \times 5$. In the experiment, data collection was performed using the Intel D435i camera device with the resolution of 848×480 . The camera intrinsic parameters were specified as follows: $f_x = 606.946$, $f_y = 607.077$, and the principal point coordinates $(c_x, c_y) = (418.495, 250.889)$. The shooting distance from the camera lens to the rebar plane was approximately 300mm. Both RGB images and aligned depth images were acquired for analysis. Then the RGB images were depth-filtered and rotated to obtain images containing only the current layer of rebar pixels. However, due to the challenging task of accurately capturing the depth information of small-diameter rebar with the camera, there were discontinuous areas in the segmentation effect, as shown in Figure 6. To expand the display of rebar pixels in the current layer, considering the linear nature of double-layer bidirectional rebar, this paper employed a row-column pixel display method. Specifically, if the number of rebar pixels in a row or column exceeded a certain threshold, all pixels in that row or column were displayed. The improved segmentation result is shown in Figure 6. But it is important to note that this method is primarily designed to address the poor performance of the Intel D435i camera in segmenting small-diameter rebar.

This study utilized a total of 124 images as the dataset, all of which were captured under indoor laboratory conditions. Considering the small size and simplicity of the dataset, it was advisable to increase the proportion of the validation set. The dataset was split into a training set (65%) and a validation set (35%) to ensure accurate and

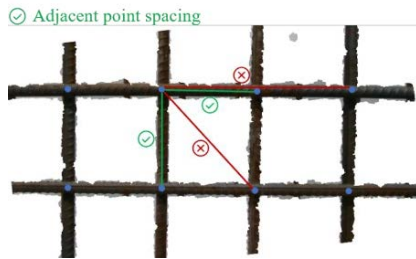


Figure 4. Schematic diagram of adjacent point spacing

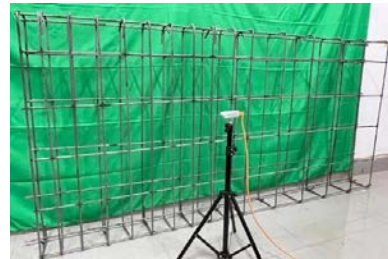


Figure 5. The double-layer bidirectional rebar cage

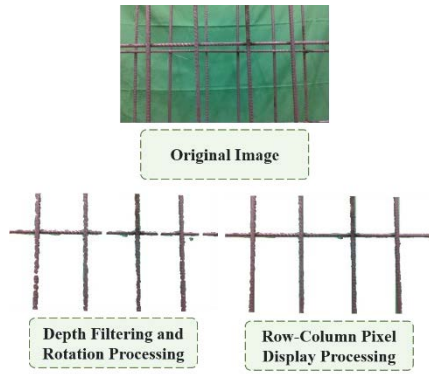


Figure 6. Schematic of the row-column pixel display method

robust prediction of model performance. 81 images were randomly selected for the training set, while the remaining 43 images constituted the validation set. Annotations were conducted on preprocessed RGB images, manually marked using the LabelMe tool to generate JSON files containing information such as image dimensions, object category names, coordinates of the four vertices of the bounding box, and coordinates of keypoints. Since the measurement of spacing is critically dependent on the accurate localization of pixels at rebar crosspoints, manual annotations points should ideally be positioned at the central point of intersection areas between two rebars to the greatest extent possible. Finally, the label files were converted from the JSON format to the TXT format suitable for YOLO training.

The YOLOv8-pose network was trained on Python 3.9 and PyTorch 2.0 environment. The training was conducted on the Windows 10 operating system with hardware specifications including an Intel(R) Core (TM)

i5-10400F CPU @ 2.90GHz and NVIDIA GeForce RTX 3060 GPU. The network training utilized the SGD optimizer with the following settings: learning rate of 0.01, 100 training epochs, and batch size of 8. Furthermore, the images fed into the network were randomly cropped and resized to the uniform size of 640×640 pixels. Data augmentation techniques were employed to enhance the training process.

4 Experimental results and discussion

4.1 Experimental result analysis

First, the effects of image depth filtering and rotation were analyzed. Figure 7 shows the rebar segmentation effects with different camera distances and rebar diameters. To explore the impact of different distances on segmentation effects, the row-column pixel display method was not applied in this case. From Figure 7, it can be observed that at longer distances, both the 8mm and 12mm diameter rebar exhibit pixel fracture or discontinuity. However, for thicker diameter rebar, the distance threshold at which this phenomenon occurs is larger and less frequent. The results indicate that, with increasing distance, the segmentation effect of a individual rebar gradually diminishes; At the same distance, thicker diameter rebar exhibit better segmentation effect compared to their thinner counterparts. This may be attributed to the difficulty of the camera depth sensor in accurately capturing long-distant and smaller-diameter rebar. Additionally, some background pixel points were retained after segmentation, possibly due to the interference of depth information by factors such as ambient noise and lighting variations.

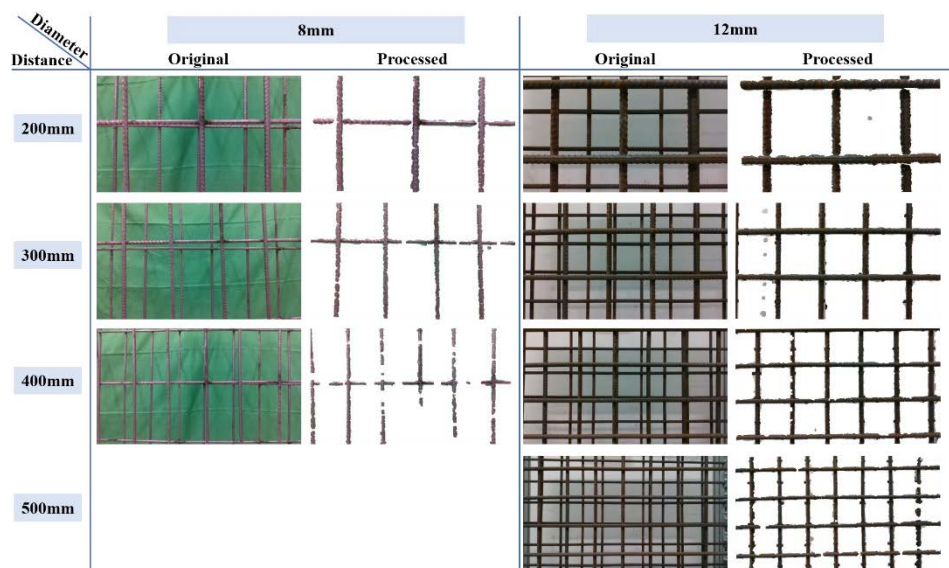


Figure 7. Segmentation effects for different distances and diameters

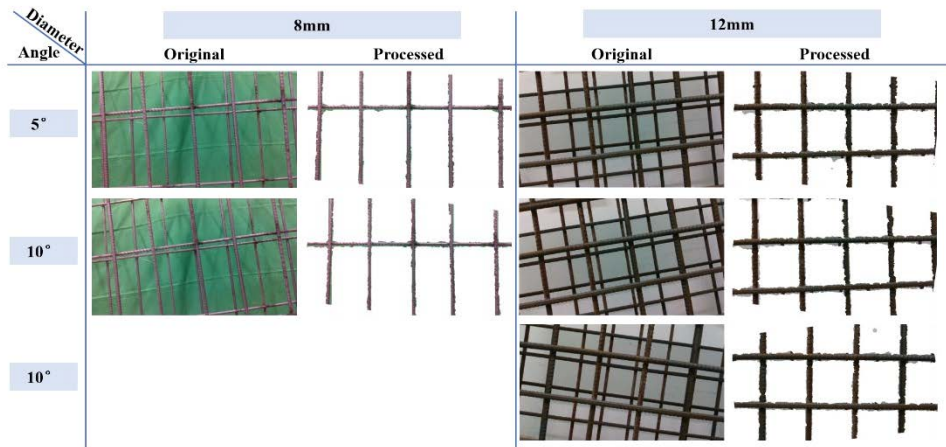


Figure 8. Segmentation effects for different rotation angles

Figure 8 shows the segmentation effects of the rebar at various angles of rotation in the camera plane, with the camera positioned 300mm away from the rebar plane. Here, the row-column pixel display method were employed for the 8mm rebar to better show the effects of plane rotation. It can be observed from Figure 8 that, for different rotation angles and diameters of rebars, the rebars were roughly parallel to the x and y axes after processing, which achieved considerable effects. However, pixel segmentation based on depth information exhibits some background point clouds at the rebar boundary. This may lead to errors in the process of fitting the rebar lines, consequently affecting the accurate calculation of the rotation angle. Therefore, having a clear rebar boundary is crucial for obtaining precise rotation angle.

Subsequently, the rebar crosspoint recognition and localization as well as the automatic spacing inspection methods were validated. After the images were processed by depth filtering and rotation, the trained YOLOv8-pose model was utilized for the prediction of rebar crosspoints, and their pixel coordinates were converted to 3D spatial coordinates relative to the camera coordinate system. Finally, the automatic spacing inspection module was utilized for rebar spacing inspection, with specific results shown in Figure 9. The errors between the rebar spacing inspection results and the manual measurement results are shown in Table 1, with an average error of 2.65 mm. Experimental results demonstrate the proposed method had high accuracy for rebar spacing inspection. By integrating these three modules, an end-to-end inspection process from image to result was achieved.

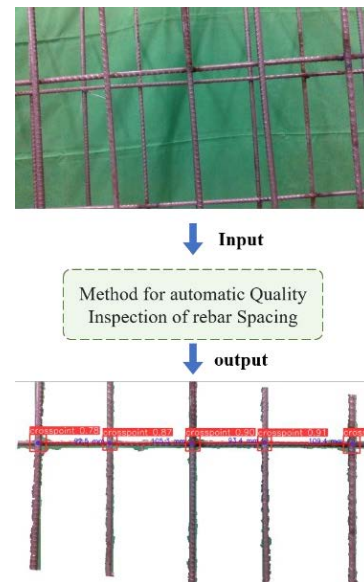


Figure 9. The method for inspecting rebar spacing

Table 1 Predicted rebar spacing and real results, and their errors in the image above

	1	2	3	4
Prediction (mm)	92.5	105.3	93.4	109.4
Real result (mm)	95	102	92	106
Error (mm)	2.5	3.3	1.4	3.4

4.2 Limitations

Rebar pixels segmentation based on depth information performs poorly in capturing small-diameter rebar. This is attributed to the challenge of accurately capturing depth information for small-diameter rebar through sensors. Additionally, factors such as ambient noise and lighting variations also introduce interference in the depth image, resulting in a few background pixels

being retained and making it difficult to clearly define the rebar pixel boundary.

In rebar spacing inspection, the limitation in the camera field of view result in each capture covering only a local region of the rebar layer. The challenge lies in effectively integrating and processing these local images.

5 Conclusion

In this paper, a method for automatic detection of rebar spacing quality is proposed, which uses a keypoint detection algorithm combined with RGBD camera. This method comprises three consecutive modules: (1) an image preprocessing module that addresses the challenge of recognizing double-layer bidirectional rebar in the image by filtering the background rebar layer; (2) a module for the recognition and localization of rebar crosspoints, obtaining their spatial coordinates; (3) an automatic inspection module enabling rebar spacing measurement in the image. The integration of these modules allows for an end-to-end inspection from image to result. Experiment testing on rebar cages in the laboratory demonstrates the efficacy of the image preprocessing module in segmenting and rotating the current rebar layer. Importantly, this module exhibits robustness without constraints related to camera distance and rotation angle. Furthermore, the average error of rebar spacing inspection is 2.65mm, which can be used for compliance inspection.

In future research, the proposed method will be validated in more complex scenarios, including construction sites and prefabrication plants. Additionally, advanced camera equipment and onboard platform will be utilized for autonomous image capture and inspection. For acceptance inspection, we will further refine its methodologies.

6 Acknowledgement

This work was supported by the National Natural Science Foundation of China (No. 52278177, No. 52308312, No. 52108136) and the National Key Research and Development Program of China (No. 2023YFC3806804).

References

- [1] Yuan X, Moreu F, Hojati M. Cost-effective inspection of rebar spacing and clearance using RGB-D sensors. *Sustainability*, 13(22): 12509, 2021.
- [2] Ministry of Construction of P.R. China. Code for acceptance of constructional quality of concrete structures: GB50204-2015. 2015.
- [3] Hwang B G, Thomas S R, Haas C T, et al. Measuring the impact of rework on construction cost performance. *Journal of construction engineering and management*, 135(3): 187-198, 2009.
- [4] Ma Z, Liu Y, Li J. Review on automated quality inspection of precast concrete components. *Automation in Construction*, 150: 104828, 2023.
- [5] Kim M K, Cheng J C, Sohn H, et al. A framework for dimensional and surface quality assessment of precast concrete elements using BIM and 3D laser scanning. *Automation in Construction*, 49: 225-238, 2015.
- [6] Kim M K, Wang Q, Park J W, et al. Automated dimensional quality assurance of full-scale precast concrete elements using laser scanning and BIM. *Automation in construction*, 72: 102-114, 2016.
- [7] Kim M K, Sohn H, Chang C C. Automated dimensional quality assessment of precast concrete panels using terrestrial laser scanning. *Automation in Construction*, 45: 163-177, 2014.
- [8] Yoon S, Wang Q, Sohn H. Optimal placement of precast bridge deck slabs with respect to precast girders using 3D laser scanning. *Automation in construction*, 86: 81-98, 2018.
- [9] Liu S L, Ma Z L. Automatic checking algorithm for the number and spacing of reinforcing bars based on point cloud. *Journal of Architecture and Civil Engineering*, 39(4): 90-99, 2022.
- [10] Wang Q, Cheng J C, Sohn H. Automated estimation of reinforced precast concrete rebar positions using colored laser scan data. *Computer - Aided Civil and Infrastructure Engineering*, 32(9): 787-802, 2017.
- [11] Yuan X, Smith A, Sarlo R, et al. Automatic evaluation of rebar spacing using LiDAR data. *Automation in Construction*, 131: 103890, 2021.
- [12] Lin F. Research on measurement technology of structural construction process based on machine vision [D], Southeast University, 2022.
- [13] Kardovskiy Y, Moon S. Artificial intelligence quality inspection of steel bars installation by integrating mask R-CNN and stereo vision. *Automation in Construction*, 130: 103850, 2021.
- [14] An M, Kang D S. The distance measurement based on corner detection for rebar spacing in engineering images. *The Journal of Supercomputing*, 78(10): 12380-12393, 2022.
- [15] Jin J, Zhang W, Li F, et al. Robotic binding of rebar based on active perception and planning. *Automation in Construction*, 132: 103939, 2021.
- [16] Jocher G. Ultralytics YOLOv8. 2023. On-line: <https://github.com/ultralytics/ultralytics>.
- [17] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection. *In proceedings of the IEEE conference on computer*

- vision and pattern recognition*, pages 779-788. 2016.
- [18] Wang C-Y, Liao H-Y M, Wu Y-H, et al. CSPNet: A new backbone that can enhance learning capability of CNN. *In proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pages 390-391. 2020.
 - [19] He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, 37(9): 1904-1916, 2015.
 - [20] Lin T-Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection. *In proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2117-2125. 2017.
 - [21] Li H, Xiong P, An J, et al. Pyramid attention network for semantic segmentation. *arXiv preprint arXiv:180510180*, 2018.

Optimization of prefabricated component installation using a real-time evaluator (RTE) connection locating system

Nolan W. Hayes¹, Bryan P. Maldonado¹, Mengjia Tang¹, Peter Wang¹, and Diana Hun¹

¹Buildings and Transportation Science Division, Oak Ridge National Laboratory, United States of America *

hayesnw@ornl.gov, maldonadopbp@ornl.gov, tangm@ornl.gov, wangpl@ornl.gov, hunde@ornl.gov

Abstract -

Prefabrication promises to industrialize the construction industry. By constructing elements within a manufacturing environment, producers can better control quality and maximize production efficiency. Since the major adoption of prefabrication, a wide variety of prefabricated components have been produced for varying applications such as new construction and exterior wall retrofits. While the production processes of these prefabricated components have seen much innovation, the installation process has remained relatively unchanged for decades. To innovate the installation process with modern technologies, a real-time evaluator (RTE) has been developed to reduce the installation cost of prefabricated components by reducing installation time, decreasing rework, and improving accuracy. The RTE uses developed software solutions with off-the-shelf hardware to assist erectors in completing an installation by measuring the real-time positions of connections and prefabricated components, providing installation guidance through a graphical user interface, and monitoring the accumulated installation errors. An overview of the RTE and proposed workflow is presented. A connection locating system that guides users in expediting the installation of connections is introduced. Laboratory experiments were conducted to determine the accuracy improvement and time savings of the RTE in installing connections for prefabricated components. RTE enabled a time saving of up to 37% compared to traditional connection installation methods using handheld measurement tools.

Keywords -

prefabrication, installation, real-time, automation, accuracy, time

1 Introduction

Prefabrication promises to industrialize the construction industry [1]. Prefabrication is the process of manufacturing and preassembling several building components, modules, and elements before their shipment and installation on a construction site [2]. By constructing elements within a manufacturing environment, producers can better control quality and maximize production efficiency [3].

Modern surveying technologies include robotic total stations that expedite and improve the accuracy of building and land surveying. Additionally, 3D laser scanners can produce point cloud data for the development of 3D models using reality capture. However, these tools are most commonly used to assess the as-built conditions of the construction after prefabricated component installation. For example, many works have focused on the automated generation of as-built 3D models using laser scanning [4, 5, 6]. However, these methods are often utilized immediately after placement of components is completed. A better use of the technology is the active monitoring of the quality of construction and using this information to assist installers while the building is being constructed such that errors can be compensated in real-time. By compensating errors in real-time, total installation time can be reduced.

The construction industry would benefit from an installation tool that utilizes modern surveying technologies to provide corrective guidance in real-time to reduce installation time, decrease rework, and reduce cost. In other industries, laser trackers are commonly used to precisely inform and assess the manufacturing process. For example, in the aviation industry, three laser trackers measuring targets attached to a wing have been used to generate real-time guidance during attachment of the wing to the aircraft body [7]. Similar techniques using machine vision are often used for robotic assembly of cars [8]. A similar technique, used in construction applications, could enable real-time evaluation of components during placement, providing guidance for users such that errors are addressed.

To innovate the installation process with modern technologies in the construction industry, a real-time evaluator (RTE) was developed to reduce the installation cost of

*Notice: This manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<https://www.energy.gov/doe-public-access-plan>).

prefabricated components by reducing installation time, decreasing rework, and improving accuracy. The function of the RTE is illustrated in Figure 1. The RTE uses developed software solutions with off-the-shelf hardware to assist erectors in completing a prefabrication component installation by measuring the real-time positions of connections and prefabricated components, providing installation guidance through a graphical user interface, and monitoring the accumulated installation errors. An overview of the RTE and proposed workflow to expedite prefabrication on-site is presented. This paper will introduce a connection locating system to guide and expedite the installation of connections. Potential time-savings and accuracy improvements are detailed through laboratory experiments.

2 Methodology

The RTE has been developed, first and foremost, as a tool to expedite the installation of overclad panel retrofits. An overclad panel retrofit consists of adding an additional layer of insulation to the exterior surfaces of an existing building [9]. These overclad panels commonly include finish, insulation, a mechanical frame, and connections to the supporting structure. The panels are prefabricated to the exact dimensions of the existing building and require tight installation tolerances because the panels serve as the new air and water barriers for the building. As a result, the figures, descriptions, and experiments will focus on illustrating and validating the RTE in a retrofit application. However, even though the RTE has been designed specifically for retrofits, the tool can be equally applied to any form of prefabrication to enable faster installation.

2.1 Overview

The real-time evaluator consists of four components. These components are illustrated in Figure 2. Each component is described in detail.

First, a digital twin is necessary for the operation of the RTE. A digital twin is a digital model of the actual building as it will be constructed [10]. Many research efforts have focused on the automated development of digital twins for construction applications [5]. As part of the digital twin, the locations and connections for each component must be modeled in space. The RTE imports the digital twin to determine the geometry of the building and installation including the goal locations of components and connections. While a necessary part of the RTE, the automated generation of the digital twin is not the focus of this paper. For the RTE, an automated method of generating the digital twin for existing buildings has been developed and detailed in a previous paper by the authors [11]. This automated digital twin generation enables the rapid mea-

surement and digitization of an existing building so that a panel retrofit can be quickly designed. For new construction, 3D models of the construction are often generated as part of the design process. For building retrofits, an as-built digital twin must be generated as an initial step. After procurement of the digital twin, the designer selects panel sizes and connection locations relative to the existing building. Whether the digital twin is automatically generated or developed as part of the construction design process, the goal install locations of each connection and component must be described within the model. The goal locations of connections and components are later used to generate installation commands.

An autonomous tracking system locates and tracks panels as they are being installed. The tracking system utilizes a robotic multi-station, operating as a laser tracker, to cycle through and locate multiple retroreflector targets (prisms) attached to the prefabricated component. Communication and control software modules direct the robotic multi-station to perform automated procedures and actively monitor the position of the component. Details on the automated tracking system and algorithms to track multiple prisms in sequence were detailed in a previous paper by the authors [12]. For the panel retrofit application, the autonomous tracking system collects the real-time position and orientations of the panel during installation.

A positioning assistant system generates installation commands for the user through a graphical interface. The system uses the data stored in the digital twin and collected by the autonomous tracking system to compare sets of goal locations to actual locations. From this comparison, required translational movements and reorientation angles can be derived to inform the user of the required installation procedures. The algorithms to reconstruct the pose of the prefabrication components during installation from the actual location to the goal location are detailed in a separate paper [13]. For the panel retrofit application, the pose reconstruction algorithms extract the translation and angles required to reposition the panel from the actual position to the goal position.

The connection locating system, which is the main focus of this paper, assists users in installing connections to the supporting structure to hold the prefabricated components. For panel retrofits, each panel must be attached to the existing building structure to support the weight of the panel. These connectors are specifically designed as part of the retrofit and are often preinstalled on the existing building before panel installation. The details of the connection locating system are further described.

2.2 Connection locating system

The connection locating system uses off-the-shelf hardware and novel algorithms to expedite the installation

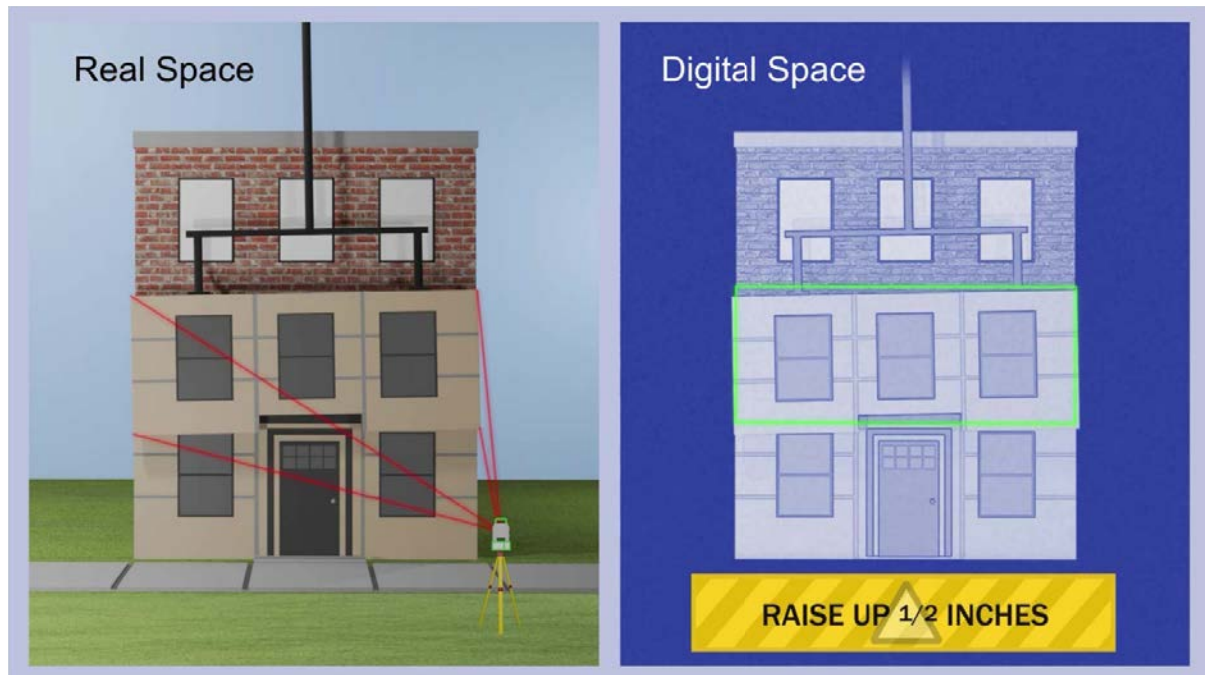


Figure 1. A real-time evaluator to expedite prefabricated component installation. It provides real-time guidance on installing prefabricated components.

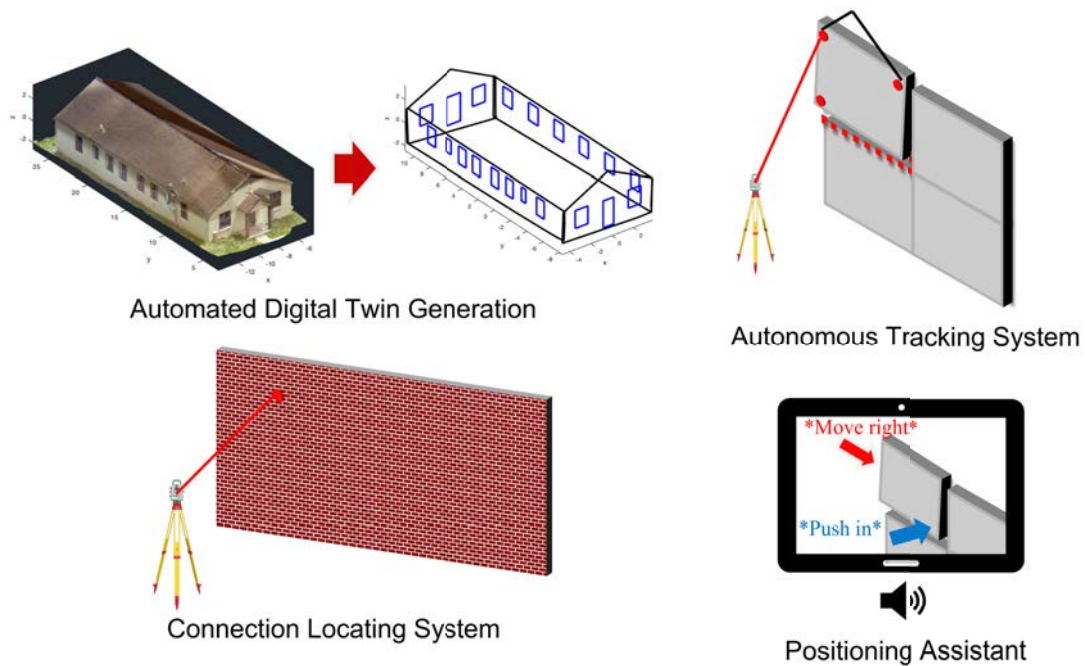


Figure 2. Components of the real-time evaluator.

of connections for prefabricated components. A robotic multi-station is controlled using communication protocols to measure the locations of reflectors attached to the phys-

ical connection hardware. The user installs connections to the supporting structure, either manually or with guiding assistance from the system. The physical location of

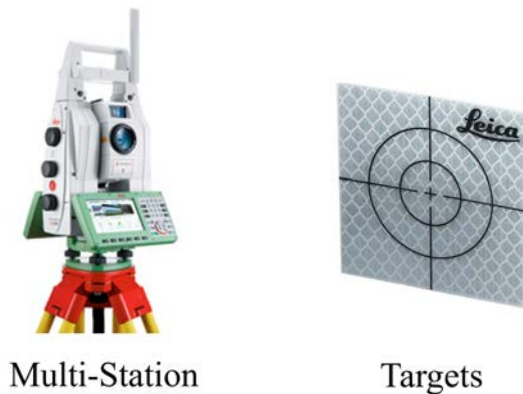


Figure 3. Hardware of the RTE connection locating system.

each connection is measured and stored during installation in the digital twin and compared to the design (goal) location of the connection. Real-time installation errors are calculated and reported to the user. The system will automatically raise flags when connections have an installation error beyond the user-specified tolerance. The user can access the reported errors to determine the corrective action required to correctly relocate the connection. The hardware, software, and processes to achieve these system features are further discussed.

2.3 Hardware and software

The hardware required by the RTE to perform connection locating functions is off-the-shelf and readily available. Additional hardware is required to perform the other functions of the RTE; however, because the focus of this paper is the connection locating system, only the hardware necessary for that system is detailed. Figure 3 shows the required hardware. A robotic multi-station is needed to track and measure the position of reflector targets. In this research, a Leica MS60 was used as the multi-station. Tape reflector targets are attached to the physical connection hardware. For each connection, the algorithms send commands to the MS60 to aim at the expected location of connection, search for the tape reflector, and measure the actual location of the reflector target attached to the connection hardware. In this research, a Leica reflector tape, GZM29, was used.

The software for the RTE was written in Python. Several modules are needed to handle the functions of the system. The communication methods of multi-stations vary by manufacturer. Leica multi-stations utilize an ex-

ternal communication protocol called GeoCom. GeoCom consists of sending and receiving string commands and feedback between the software and hardware. A communication module within the software handles the writing and receiving of string commands to the hardware. For example, the command `%R1Q,9027:Hz,V` requests the multi-station to turn to a position indicated by the horizontal angle, Hz, and vertical angle, V. Upon successful completion, the multi-station will respond with `%R1P,0,0:RC` indicating that the requested command was completed with no issues. All basic functions are performed as single request/response queries. A controller module handles the high-level functions of the instructions to the multi-station. For example, the task of measuring a tape reflector consists of a series of low-level functions including a turn to the reflector, a search for the reflector, a locking-in on the reflector, a measure of the reflector, and an output of the measured location to save for future reference. The controller module allows the user to access a single high-level function to perform a series of low-level functions for specific purposes. A calculator module performs geometry calculations such as determining horizontal and vertical angles of a specified point in 3D space, x, y, z , from the origin (location of multi-station). A digital twin module stores all of the design and actual locations of the building, control points, connections, and prefabricated components.

2.4 Implementation

To reduce installation time, the connection locating system removes the need for manual measurement of connection locations using hand tools. Two modes of operation have been developed for the connection locating system. The first mode of operation is guided installation. Guided installation utilizes the visible red laser of the multi-station along with the software implementation to point to the locations where connections should be installed based on the digital twin. The user then places each physical connector with a tape reflector aligned with the red laser as shown in Figure 4. After each connection is fastened, the user instructs the RTE to measure the actual position of the connection. The actual position is compared to the design location and an installation error is calculated. Upon comparison to a user-specified tolerance, the RTE will instruct the user in real-time if the connection has been incorrectly installed and, if so, how to correct the installation.

A second mode of operation assumes that connections have already been installed. The second mode of operation is evaluation. In this mode, the RTE will turn to the design location of each connection stored in the digital twin, radially search for the nearest reflector, measure the position of the reflector, and output the installation errors. The user will be notified of any errors beyond the spec-

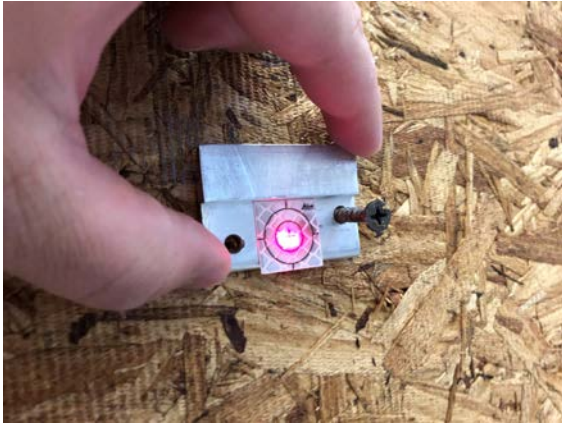


Figure 4. Guided installation of connection.

ified tolerance and given corrective actions to rectify the problems.

The first mode of operation proposes to significantly reduce installation time by removing the need to use hand tools to measure the installation positions of connectors. Instead of using traditional hand tools such as tape measures, bubble levels, or laser levels, the RTE simply points a laser to the location where each connector should be installed. Removing the need for hand tools will also improve safety since these tools will no longer need to be carried up and down ladders. To verify that the connection locating system of the RTE reduces installation time, a lab-scale demonstration experiment was conducted.

3 Lab-scale demonstration

3.1 Lab-scale mockup installation

A lab-scale demonstration experiment was conducted within the Maximum Building Energy Efficiency Research Laboratory at Oak Ridge National Laboratory. A 1/3rd scale mockup of a three-story overclad panel retrofit was constructed. A digital twin of the mockup is illustrated in Figure 5. A metal frame (1.829 m wide by 3.124 m high) was constructed to simulate a three-story building at approximately 1/3rd scale. A working surface of an oriented strand board was attached to the metal frame to simulate the building facade upon which connections would be attached. Six simulated overclad panels and a simulated crane were manufactured; however, only the connections used to secure the panels to the working surface were used in this study. For each overclad panel, two connections near the top corners allow the panels to be erected onto the working surface. Each connector is a pair of metal brackets. One bracket is attached to the building facade working surface at the indicated position. The mating bracket is attached to the back side of the panel such that two sets allow the panel to hang. While these connections are not

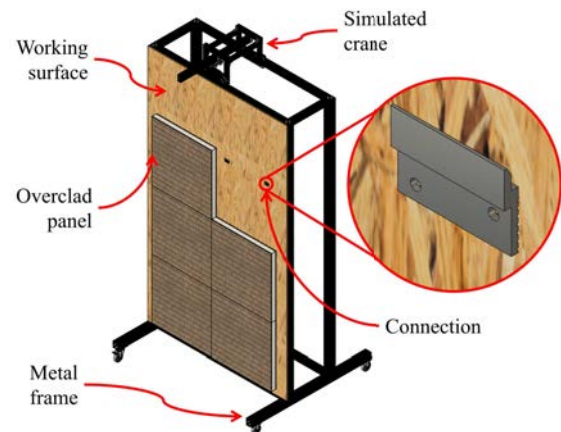


Figure 5. Digital twin of the lab demonstration.

the same as traditional connection methods for overclad panels, the basic function is similar and the installation method is largely the same. The digital twin was exported to the RTE software so that the digital design location of each connection was stored. A total of 12 connections existed in the mockup.

3.2 Experiment

Two sets of experiments were conducted to quantify the possible installation time savings of the RTE connection locating system. The 1/3rd mockup building was set up approximately 8 meters away from the Leica MS60 as shown in Figure 6. The wheels of the mockup frame were locked to ensure that the frame was stationary during all tests. The MS60 was connected via Bluetooth connection to a personal laptop computer running the RTE software. Two experiments were conducted to determine the amount of time required to install the connections for the mockup overclad panel retrofit. A tape reflector was placed on each physical connector to be attached to the simulated building surface. Each connector was numbered according to the installation position to eliminate the variability of tape placement between experiments. For each experiment, the installation process was recorded to determine the time required for each step.

In the first experiment, the user installed the connections using manual methods. A drawing was generated from the digital twin indicating the exact design locations of all connections. Dimensions were generated for each connection on the working surface plane. Handheld tools, including a tape measure and bubble level, were used to mark connection installation points on the working surface. First, the elevation of two connection points was marked by measuring with a tape measure from the bottom of the working surface. From each of these marks, a level line of the elevation was marked using a bubble level. The exact po-



Figure 6. Setup of the connection locating system experiment.

sitions of connections were then marked along the level elevation line by measuring from the sides of the worked surface. The connections were installed at each marked location using a screwdriver to install the fasteners. Each connection required the installation of two fasteners. This process was repeated for each set of two connections. The highest set of four connections was only accessible using a ladder. The installation of connections was performed at a normal pace. After installing all connections, the MS60 was used to measure the exact positions of all connectors. Installation errors were calculated by comparing the actual position of connections to the design locations in the digital twin.

In the second experiment, the user installed the connections using the guided installation from the connection locating system of the RTE. The MS60 was instructed to activate the red visible laser and point to the design location for each connection, iteratively. The user installed the physical connector by aligning the red laser with the tape reflector and driving fasteners using a screwdriver. The user did not use any hand tools to measure the actual locations of connections. After installation of each connection, the user indicated completed installation from the laptop computer, and the RTE measured the actual installed location of the connector. The installation of connections was performed at a normal pace. Installation errors were calculated by comparing the actual position of connections to the design locations in the digital twin after each installation.

In both experiments, the same single user completed the installation processes. The manual process was completed first, and the guided installation process was completed second. Both experiments were completed in immediate succession. The second experiment, being entirely guided by the RTE, is unaffected by user affinity with the installation process. As a result, as the user gained knowledge and experience of the installation process, the time required to

Table 1. Installation times for 12 connections.

Method	Total Time (MM:SS)
Manual	13:51
Guided (RTE)	8:43

complete manual installation could be reduced while the guided installation would remain relatively unchanged.

The total installation times for each method are shown in Table 1. The manual method refers to the installation of connections using handheld tools. The guided method refers to the installation of connections using the connection locating system of the RTE. It is important to note that the time reported for the manual method includes only the installation process of the connections. The reported installation time for the manual method does not include the time required to manually measure the real positions of connection for the purpose of error calculations. However, for the reported time of the guided installation method, the time required to automatically measure the real position of each connection is included. The guided installation of the RTE connection locating system reduced installation time by nearly 37%. By removing the need for hand tools, the installation time savings are significant.

Installation errors for each connection according to the installation method are shown in Figure 2. For each method, installation error in each dimension is reported. For reference, the x -axis is left (-) and right (+) across the face of the working surface. The y -axis is towards (+) and away (-) from the face of the working surface. The z -axis is elevation. The installation error corresponds to the Euclidean distance calculated as $\|\text{error}\|_2$.

In general, the installation errors between manual methods and guided methods are very similar. The guided method did consistently reduce error slightly. In both cases, errors in the x -axis (left/right) and z -axis (up/down) were small in comparison to the y -axis (away/towards). Additionally, connections 9 through 12 were located at elevations that required the use of a ladder to install. For both installation methods, the largest errors were present in these connections at the highest elevation. For installation requiring a ladder, the quality of installation was lower.

Minimum, maximum, and mean installation errors for each method are shown in Table 3. Guided installation using the RTE connection locating system reduced errors although the degree of reduction in this case was negligible by most standards. For most cases, a peak accuracy of installation of 3 mm is needed due to the installation tolerances of components. However, for panel retrofits, a lower installation tolerance may be required since the panels serve as the new air and water barrier.

The maximum error is reduced by a significant pro-

Table 2. Installation error for each of the 12 connections.

Connection	Errors (mm)							
	Manual				Guided (RTE)			
	x	y	z	$\ error\ _2$	x	y	z	$\ error\ _2$
1	-0.21	0.93	-0.05	0.96	-0.10	0.43	-0.14	0.46
2	-0.07	0.50	0.05	0.50	0.04	0.61	0.22	0.65
3	0.06	1.13	-0.18	1.15	-0.15	0.85	-0.01	0.87
4	0.05	1.00	-0.12	1.01	0.16	0.80	0.00	0.82
5	-0.11	1.15	-0.01	1.16	-0.15	0.65	-0.11	0.68
6	0.07	1.22	-0.06	1.22	0.07	0.81	0.23	0.84
7	0.11	0.78	0.27	0.83	0.32	0.88	-0.03	0.94
8	0.13	1.07	0.24	1.11	0.13	0.88	-0.05	0.89
9	0.16	0.59	-0.17	0.63	0.32	0.27	0.01	0.42
10	-0.27	1.67	-0.92	1.92	-0.23	1.65	-0.80	1.85
11	-0.21	2.48	-0.12	2.50	0.39	1.42	0.05	1.48
12	0.17	1.32	-0.01	1.33	-0.04	1.12	0.07	1.13

Table 3. Installation errors of the 12 connections.

Method	Error (mm)			Std
	Min	Max	Mean	
Manual	0.50	2.50	1.19	0.0005
Guided (RTE)	0.42	1.85	0.92	0.0004

portion by the guided installation compared to manual methods. This reduction in maximum installation error could make the difference in an actual installation such that cumulative errors do not propagate into issues. Because installation errors are cumulative, individual errors as small as 1 mm can aggregate into several centimeters during the entire installation process of a large area multi-story building. Therefore, it is still important to minimize installation errors to mitigate the risk of error propagation.

3.3 Limitations

Time savings of the guided installation of the RTE have been demonstrated in lab-scale experiments. However, time savings will vary widely between users, building types, and applications. To further investigate the time savings potential of the guided installation, multiple installers will be tested as part of the experiment to determine an average expected potential amongst all types of users. Additionally, as the size of the working surface increases, the difficulty and expertise required to perform manual installation also increases because handheld measurement tools become more complex and unwieldy. For this reason, it is expected that potential time savings will be larger for full-scale implementation.

The guided installation process requires that the multi-station have a line of sight to the reflector targets. Users must be aware of blocking the line of sight between multi-station and targets during the installation process. Avoid-

ing this issue may be more difficult for multi-story buildings that require specialized equipment such as cherry pickers or manlifts.

4 Conclusions and next steps

A connection locating system for a Real-Time Evaluator (RTE) was developed to expedite and optimize the installation of connections for prefabricated components. The connection locating system uses off-the-shelf hardware and novel software to guide the installation of connections for prefabricated components. The connection locating system was tested in laboratory-scale experiments to determine the installation time-saving potential and accuracy improvements compared to manual installation using common hand tools. Results of the experiments conclude that the connection locating system can reduce connection installation time by 37% while maintaining better installation accuracy compared to manual installation.

In future work, the RTE will be tested to determine time-saving potential and accuracy improvements of overlaid panel retrofit installation. Panel installation using the RTE will be tested at laboratory-scale and full-scale in a real overlaid panel retrofit.

5 Acknowledgements

This manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access

Plan (<https://www.energy.gov/doe-public-access-plan>).

This research was supported by the DOE Office of Energy Efficiency and Renewable Energy (EERE), Building Technologies Office, under the guidance of Sven Mumme, and used resources at the Building Technologies Research and Integration Center (BTRIC), a DOE-EERE User Facility at Oak Ridge National Laboratory.

References

- [1] Filipe Barbosa, Jonathan Woetzel, Jan Mischke, Maria Joao Ribeiro, Mukund Sridhar, Matthew Parsons, Nick Bertram, and Stephanie Brown. Reinventing construction through a productivity revolution. Technical report, McKinsey Global Institute, 2017.
- [2] Chris Goodier and Alistair Gibb. Future opportunities for offsite in the uk. *Construction Management and Economics*, 25:585–595, 06 2007.
- [3] Yuan Chang, Xiaodong Li, Eric Masanet, Lixiao Zhang, Zhiye Huang, and Robert Ries. Unlocking the green opportunity for prefabricated buildings and construction in china. *Resources, Conservation and Recycling*, 139:259–261, 2018.
- [4] Pingbo Tang, Daniel Huber, Burcu Akinci, Robert Lipman, and Alan Lytle. Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques. *Automation in Construction*, 19(7):829–843, 2010.
- [5] Qian Wang and Min-Koo Kim. Applications of 3d point cloud data in the construction industry: A fifteen-year review from 2004 to 2018. *Advanced engineering informatics*, 39:306–319, 2019.
- [6] Aravinda S. Rao, Marko Radanovic, Yuguang Liu, Songbo Hu, Yihai Fang, Kourosh Khoshelham, Marimuthu Palaniswami, and Tuan Ngo. Real-time monitoring of construction sites: Sensors, methods, and applications. *Automation in Construction*, 136: 104099, 2022.
- [7] New River Kinematics. Using Real-Time, 6D Object Tracking to Assemble Large Aerospace Components. Online: <https://www.kinematics.com/imagesweb/DeltaIV-2.pdf>. Accessed: 11/03/2024.
- [8] Alexandra Papadaki and Maria Pateraki. 6d object localization in car-assembly industrial environment. *Journal of Imaging*, 9(3), 2023.
- [9] Mikael Salonvaara, Antonio Aldykiewicz Jr, Diana E. Hun, and Andre Omer Desjarlais. Cost assessment of building envelope retrofits. In *2020 ACEEE Summer Study on Energy Efficiency in Buildings - Pacific Grove, California, United States of America*, August 2020.
- [10] Calin Boje, Annie Guerriero, Sylvain Kubicki, and Yacine Rezgui. Towards a semantic construction digital twin: Directions for future research. *Automation in Construction*, 114:103179, 2020.
- [11] Bryan P. Maldonado, Nolan W. Hayes, and Diana Hun. Automatic point Cloud Building Envelope Segmentation (Auto-CuBES) using Machine Learning. In *Proceedings of the 40th International Symposium on Automation and Robotics in Construction*, pages 48–55, Chennai, India, 2023.
- [12] Nolan W. Hayes, Bryan P. Maldonado, Diana Hun, and Peter Wang. Automated tracking of prefabricated components for a real-time evaluator to optimize and automate installation. In *Proceedings of the 40th International Symposium on Automation and Robotics in Construction*, pages 192–199, Chennai, India, 2023.
- [13] Mengjia Tang, Nolan W. Hayes, Bryan P. Maldonado, and Diana Hun. Component pose reconstruction using a single robotic total station for panelized building envelopes. In *Proceedings of the 41st International Symposium on Automation and Robotics in Construction*, Lille, France, 2024.

Effects of Visual Prompts in Human-machine Interface for Construction Teleoperation System

Yeon Chae¹, Samraat Gupta² and Youngjib Ham³

¹Department of Construction Science, Texas AandM University, USA

²Department of Computer Science and Engineering, Texas AandM University, USA

³History Maker Homes Endowed Associate Professor, Department of Construction Science Texas AandM University, USA

yeonchae62@tamu.edu, samraatg@tamu.edu, yham@tamu.edu

Abstract

In construction teleoperation, particularly in disaster restoration, delicate manipulation of heavy machinery is crucial, based on a thorough understanding of the surroundings. Current practices have utilized multiple viewpoints to facilitate a thorough understanding of the site's 3D spatial layout. However, challenges might arise as visual cues within the surroundings could create distractions for teleoperators. Drawing from visual search theory and Gibson's perception theory, exploring visual prompts in teleoperation interface could enhance performance by directing attention to key visual cues, reducing cognitive workload. Nonetheless, the evaluation of different visual prompts from human factors perspectives has been underexplored. Addressing challenges of potential distraction from multiple viewpoints and inappropriate visual prompts, this study emphasizes the necessity of exploring different visual prompts to most effectively guide operators' attention in given work environments. The experiment, designed with low and high visual cue environments, focuses on debris removal and extended 3D Fitts' law tasks, evaluating spatial awareness and depth perception during teleoperation. The experiments were conducted with participants in construction-related fields with industrial experience. Performance measurements and subjective ratings with open-ended discussion was conducted. The findings show visual prompts' effects on distraction and visibility conditions concerning task-oriented difficulty levels in teleoperation. The experimental results can inform the optimal design of visual prompts in human-machine interface for teleoperation for complicated construction environments, highlighting the importance of the considerations of environments and task characteristics.

Keywords –

Human-Machine Interface, Construction Teleoperation

1 Introduction

In the aftermath of disasters and demolition, hazardous conditions within sites create barriers to physical human access [1]. In such cases, teleoperation technology has been adopted, enabling workers to operate heavy machinery from a secure external location, minimizing on-site risks [2]. Despite this advancement, such jobsites present persistent challenges in teleoperation that directly hinder productivity, especially in constrained movement areas where teleoperators require heightened spatial awareness [3]. To alleviate these challenges, incorporating multiple viewpoints to offer additional information seems promising, yet the potential overflow of information might overwhelm teleoperators [4,5]. Previous studies have highlighted the efficacy of visual prompts in guiding teleoperators to conduct critical tasks with multiple screen settings [4]. However, the exclusive efficacy of visual prompts in teleoperation and their alignment with the naturalness principle remains unexplored, along with an incomplete understanding of how environmental settings influence the effects of visual prompts in teleoperation.

This paper aims to explore varied types of visual prompts, by differentiating the forms of conveying spatial information and the level of granularity in indicating distances to the target objects, to uncover their effectiveness in teleoperation. Examining these conditions offers refined insights into how different visual prompts impact teleoperators' performance, and task difficulty and efficiency. This study can inform how to optimize visual interfaces for teleoperation in different task environments and address potential challenges of visual prompts in certain environment settings. Through the investigation of varying impacts of different visual

prompts on teleoperation across diverse task settings, this study addresses critical gaps in understanding and contributes to the visual interface design in teleoperation systems, especially heavy machinery. Overall, this study can provide practical implications for enhancing teleoperator experience, refining task outcomes, and advancing human-robot interaction.

2 Backgrounds

2.1 Visual Prompts in Human-Machine Interfaces for Teleportation

The visual interface for human-machine teleoperation heavily relies on understanding selective attention through visual search theory and Gibson's Theory of Perception [6]. Teleoperators engaged in articulate manipulation tasks necessitate sustained attention, making the design of a robust visual interface crucial [5]. Visual prompts, guided by top-down and bottom-up influences, affecting the adjustment of bottom-up approach by the salient features and top-down approach by controlling themselves how to utilize the given visual information based on their goal, also play a crucial role in human-robot interaction, directing attention toward essential elements [4]. The design of these prompts should reduce involuntary attention shifts by carefully considering their reliability, following Gibson's ecological approach, which highlights the importance of perceiving environmental cues that support the robot's actions [7]. Understanding selective attention mechanisms and incorporating visual search theory principles can markedly enhance visual prompts' design, optimizing teleoperator performance and improving human-robot interaction during teleoperation [4].

Aided visual prompts in robotic teleoperation encompasses a range of types, addressing specific environmental and task needs, facilitating tailored prompts that enhance operators' situational awareness [8]. The selection of optimal visual stimuli can significantly influence operators' decision-making and overall task performance, aligning with the specific teleoperation requirements [9]. In the realm of teleoperation, visual prompts categorize into directional and regional types [9], each conveying distinct information. Direct visual prompts, particularly regional ones, aim to provide 3D spatial understanding, either directly or indirectly, crucial in scenarios where desktop displays offer 2D vision. These prompts aim to engage voluntary attention, improve goal-directed task execution, and reduce cognitive demands, significantly impacting teleoperation performance [4]. Strategic implementation of these visual prompts, emphasizing relevant features in alignment with top-down approaches, optimizes operators' attentional resources, ultimately

enhancing teleoperation efficiency [4,10].

2.2 Teleoperation Task Environments with Varying Level of Visual Cues

Level of visual cues could be varied by the extent to which information is provided in terms of depth perception [12]. Based on the hierarchy of the visual cues [13], there is a property of visual cues that could assist the depth perception. Based on the experiment conducted [14], low level of visual cues provides minimum spatial cues (including relative size of the objects, occlusion, linear perspective) whereas high level of visual cues provides maximum spatial cues (including distance to horizon with familiar visual cues, cast shadows).

Depth perception in varied visual environments, characterized by cue availability, poses significant challenges in tasks, especially when working with low visual cues lacking familiar landmarks [15]. Evaluating cue reliability becomes critical as the inclusion of unreliable cues can hinder rather than aiding task performance [16]. Ensuring cue accuracy and reliability is crucial to prevent misinterpretations in scenarios where depth perception uncertainties are amplified due to the absence of adequate visual cues [16]. However, in the context of teleoperation, these challenges might be exaggerated as the task itself has to be manipulated in the complicated environments which require a delicate maneuver, engaged by continuous cognitive engagement of teleoperator. Moreover, as given that the teleoperator is provided by the visual interface, indirectly experience the scene through it, there's a room to augment that interface, by varying the level of visual cues given to the teleoperator with a thorough consideration of the task characteristics and environments. Therefore, integrating multiple visual cues into the interface aims to enhance performance and usability, necessitating a keen understanding of cue reliability's impact on decision-making and task execution [10].

3 Modeling

The overview of this study is depicted in Figure 1. We build upon the two models with different visual cues, a low visual cue environment as a baseline setting and the demolition sites as a high visual cue environment.

3.1 Task Environment with Low-Level Visual Cues

The virtual space featured a clear blue-sky area strategically designed to minimize distractions for participants [17]. In our study, the low-level visual cue environments were built upon the Fitts' law task [17],

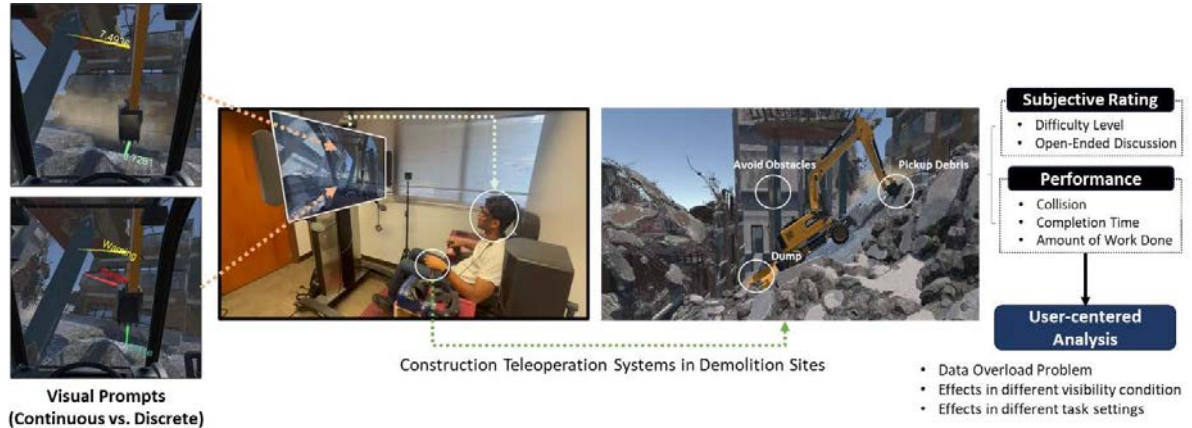


Figure 1. Overview of this study to answer “How can we most effectively map spatial data to enhance situational awareness?” by exploring diverse representations of visual prompts and task attributes in construction teleoperation.

emphasizing visual cues limited to target, destination, and a horizontal line, serving as a key perceptual reference [13]. These cues, influencing depth perception, encompass occlusion, relative size, distance to the horizon, and linear perspective [13]. Given the pivotal role of target size and distance in task complexity [17], the target consistently commences midway between center and destination. Each trial situates the destination around the target with identical 2D distances but random forward or backward depths in 3D, maintaining uniformity in depth radius.

The design units of objects and camera length aim for comprehensive scene understanding across varied camera conditions. The scenario introduces the destination in horizontal, vertical, and diagonal locations, each presented twice, totaling six occurrences. In each of these locations, both forward and backward depths are assigned randomly per trial, although the sequence is randomized, ensuring consistent difficulty across manipulation trials.

3.2 Task Environment with High-Level Visual Cues

In this experiment, the demolition site is designed as an environment with high-level visual cues. This included a demolished building and an excavator. The excavator mirrored real-world behavior, and the trajectory of excavator is controlled for the safety measures to prevent tipping on slopes during the manipulation. Camera placements were tailored for cabin-installed position. In this environment, unsound structures are near excavators, which requires careful operation. Dust effects and sound cues when collision occurs as collision effects are designed to emphasize immersion. As illustrated in Figure 2(a), debris are occluded, invisible from the cabin view. Following the task design, the environment is designed to restrict the excavator’s trajectory with the closely located obstacles (Figure 2(b)), with the dumping area (Figure 2(c)).

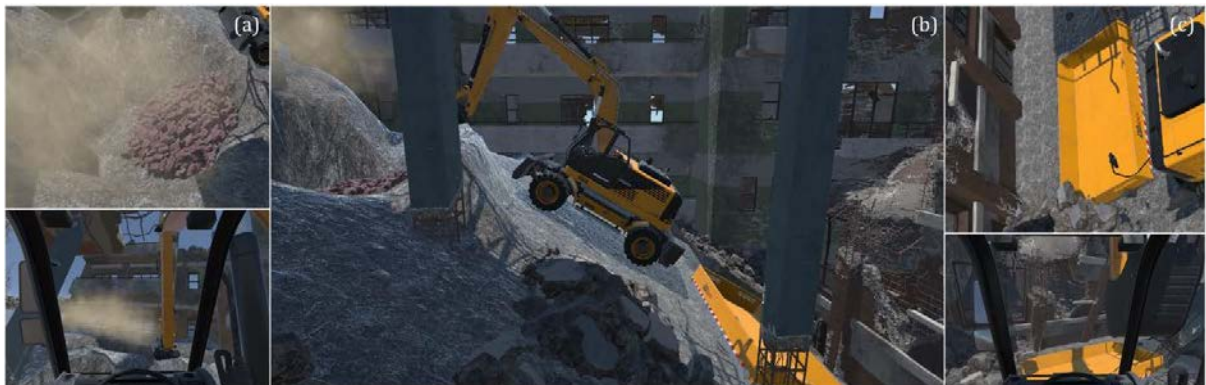


Figure 2. Environment setting – (a): Debris, (b): Overall View of the Scene, (c): Dumping Area.



Figure 3. Visual prompts design in demolition sites – (a): Continuous representation, (b): Discrete

3.3 Visual Prompts Design

Basically, visual prompts can serve to enhance comprehension of 3D spatial information in remote construction sites [4]. In our experiments, components used in the model are selectively adopted from previous studies [4,10] and customized to the given experimental scenarios and task settings. The primary aim of these visual prompts is to augment understanding, especially regarding depth perception within a 2D display representing 3D spatial information. These prompts encompass line interpolation, distance, and color coding, assisting in removing debris and monitoring progress within occluded areas. Categorized as work facilitation and obstacle avoidance prompts [4], these prompts aim to support situational awareness for the task including the debris removal and obstacle avoidance.

For “work facilitation prompts” aiding debris removal task, visual cues include the productivity monitor bar conveying the information of the number of debris being completed in picking up and dumping task, green lines guiding the nearest debris mass center, and proximity-indicating lines (Figure 3). These lines adapt color and thickness based on environmental settings,

providing intuitive cues for obstacle proximity. In “obstacle avoidance prompts”, lines indicate the closest vertical points of obstacles and excavator components, varying in width to represent depth. Continuous prompts display numerical distance values and color gradients indicating proximity (Figure 3(a)). Discrete prompts feature warning signs triggered by distance thresholds, with colors indicating danger levels (Figure 3(b)). In discrete visual prompts, the distinction between “warning” and “danger” levels is determined by proximity criteria [4]. Similarly, work facilitation prompts categorize proximity as “close” or “very close,” based on the distance relative to the stick’s length. Contrastingly, continuous visual prompts use numerical values in feet, detailed to four decimal points, to indicate proximity. When the participants operate the excavator, the continuous visual prompt changes occur at a frequency of 60 fps. However, the discrete visual prompt remains unchanged while a part of the machine is physically within the predefined range. In the baseline scene, visual prompts primarily reflect features from demolition sites, focusing on work facilitation without collision avoidance requirements. These prompts include varied line attributes indicating depth in both discrete and

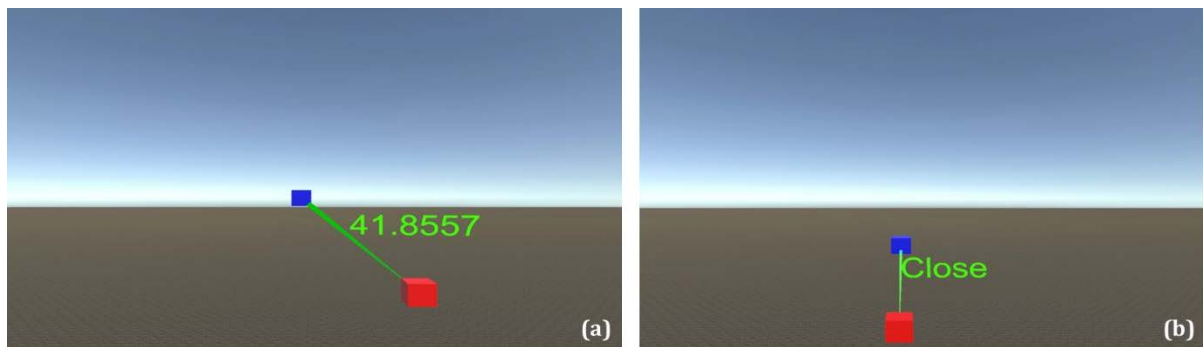


Figure 4. Visual prompts design in low level visual cues environment – (a): Continuous representation, (b): Discrete representation.

continuous settings (Figure 4).

In our experiments, we explore how different visual prompts work in teleoperation for different working environments and look into their effects from human factors perspectives and how to most effectively map spatial data to enhance situational awareness.

4 Experiments

4.1 Task Design

In the environment with the high-level visual cues, the task focuses on assessing spatial awareness and depth perception in debris removal scenarios like disaster restoration or demolition sites. Participants maneuver an excavator to pick up debris, avoid obstacles, dump debris, and return two trials within a six-minute time limit, assessing productivity and safety.

In the environment with the low level of visual cues, the extended 3D Fitts' law task evaluates 3D spatial awareness by manipulating objects in horizontal, vertical, and diagonal directions. This task requires six sequential movements of the target cube (indicating as blue) toward the destination where the red cube is located. Destination disappears once each hit is completed and pops up in a random position. The task involves six hits, each with a random position, with two locations set diagonally, vertically, or horizontally. Fixed target sizes and distances provide consistent difficulty levels across movements.

4.2 Performance Metrics

In debris removal task, task performance was measured as an objective measure in terms of amount of work done, time and collision occurrences [2]. For extending 3D Fitts' law task, performance was measured based on time, errors, and efficiency in navigating the movements in a virtual space [17]. Considering the task characteristics, if the participant mistakenly manipulates

the joystick resulting in deviating from the target destination, it is counted as an error in this task. This extension of the 3D Fitts' law task delves into the challenges of navigating a given environment in X, Y, and Z directions, considering factors like movement time, error rates, and directional complexities. Understanding these challenges aids in enhancing 3D interfaces for teleoperation, especially concerning different movement difficulties and their impact on task performance.

4.3 Effects of Visual Prompts in Different Task Environment Settings

The experiments started with a brief introduction of the study, followed by the task description detailed in section 4.1. Participants then completed tasks, including the extended 3D Fitts' law task and debris removal task in demolition sites, in a random order. Ten undergraduate students majoring in construction-related fields with industrial experience from Texas AandM University participated. Objective measures were collected during the experiments, along with subjective ratings on task difficulty [4] via a comprehensive questionnaire and open-ended discussions.

Table 1. Pre-survey results for the participants ($N=10$)

Variables	Scale	#
Age	19~21	8
	22~24	2
Gender	Male	8
	Female	2
VR Experience	Very unfamiliar	2
	Unfamiliar	3
	Somewhat familiar	3
	Familiar	1
	Very familiar	1

For the extended 3D Fitts' law task, the completion

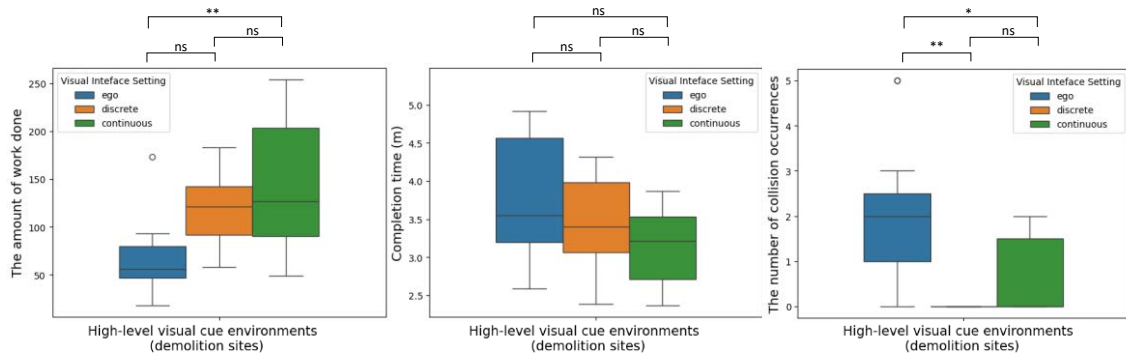


Figure 5. Results from debris removal task in demolition sites with high-level visual cues - (a): Amount of work done, (b): Consumed time, (c): Collision occurrences.

time for each trial (totaling six completion times for a single task) is documented, alongside the coordinates of the blue box (target) and red box (destination). These coordinates provide essential data for tracking the real-time distance between the target and destination throughout the manipulation process. For demolition sites, completion time, amount of work done, collision occurrences are used as performance metrics.

5 Results and Discussion

With the results, a comparative examination of group disparities is performed through One-way ANOVA (ANalysis Of VAriance) with post-hoc Tukey HSD (Honestly Significant Difference) Test.

5.1 Effects of Visual Prompts in Different Task Environment Settings

Figure 5 shows the performance in demolition sites, as an objective measure in terms of amount of work done, time and collision occurrences. Overall, visual prompts enhance the work performance across all measures compared to the counterpart with non-visual prompts. Interestingly, in terms of collision, comparing the average value, when visual prompts are provided in a discrete way, collision reduced, which is significant compared to the counterpart with non-visual prompts. Especially for continuous visual prompts, based on the open-ended discussion following the experiment, participants replied that they felt exhausted in tracking the continuously changing number, finding it challenging to gauge its proximity and its potential impact on collision avoidance. This lack of clarity made planning maneuvers with heavy machinery difficult, resulting in the occurrence of collisions. Based on the given information (i.e., number), it is hard to plan the maneuver of heavy machinery, since the number itself was not intuitive to inform the control based on that.

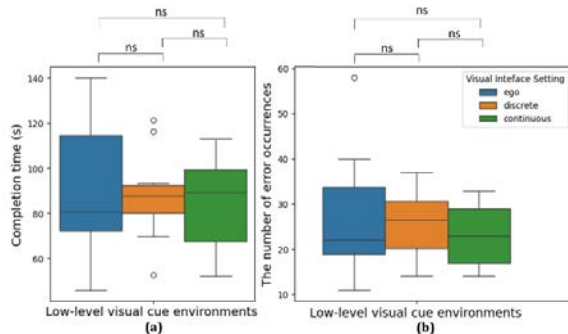


Figure 6. Results in extended 3D Fitts' law task - (a): Consumed time, (b): Error.

Contrarily, when visual prompts are provided with discrete representation, participants reported that their

focus was easily directed by the visual prompts, leading to a significant reduction in collision. These varied patterns are also aligned with their interviews, demonstrating that continuous visual prompts in the environments with high-level visual cues could potentially serve as a distraction, impacting teleoperators' attention.

On the contrary, in the low-level visual cues in Figure 6, the average value of the consumed time with visual prompt has increased compared to non-visual prompt conditions. In the environments with low level of visual cues, some participants experienced improved performance with visual prompts, while others encountered hindrances in their manipulative tasks, resulting non-significant difference within different settings ($p > 0.05$). This variance underscores the complexity of how visual prompts influence teleoperators' task execution in low-cue settings, where the impact on performance is not uniformly beneficial and could vary based on individual perceptions and adaptability to visual cues. Interestingly, despite these varied objective outcomes, participants reported lower task difficulty levels when visual prompts were available. Comparing the average number of errors occurrences between two different conditions of visual prompts, continuous visual prompts condition shows lower error incurred, compared to the discrete visual prompts. This result could be originated from the task requirements, as the task requires continuous monitoring of comparing the position of the object being manipulated in relation to its intended destination. Overall, the results highlight the criticality of the visual representation following the naturalness principle based on the consideration of task characteristics within the context of environments.

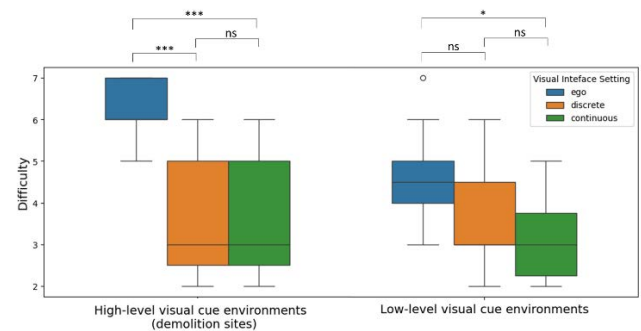


Figure 7. Difficulty level of task for debris removal task and extended 3D Fitts' law task.

Based on the level of visual cues, the results, as illustrated in Figure 7, indicate a significant disparity in perceived task difficulty among participants. In low-level visual cue environments, participants reported feeling less difficulty when engaged in the extended 3D Fitts' law task compared to the high-level visual cue environment where the task involved debris removal.

This suggests that the complexity of the task itself, coupled with the amount of visual information available, plays a crucial role in how participants perceive and experience task difficulty.

These results based on their subjective difficulty level of the task could underscore the necessity of tailoring the interface and visual prompt design based on the environmental context. In low-level visual cue settings, where tasks might be less visually demanding, simpler, and more direct visual prompts could suffice. Conversely, in high-level visual cue environments where tasks are inherently more complex due to the visual richness, visual prompts should offer more comprehensive and easily interpretable information to assist users effectively. The results could highlight that visual prompt designs should adapt to the environmental setting, aligning with task complexities and varying levels of visual cues to optimize user performance and minimize task difficulty.

5.2 Effects of Visibility Conditions and Task-Specific Difficulty Levels

Participants replied that highlighted challenges were linked to both discrete and continuous signals with the visibility condition of the working area, particularly in discerning object proximity during their task. In our scene, during the picking up task, the target object is occluded, whereas during the obstacle avoidance task, the obstacle, the target to avoid, is visible, not occluded. Specifically, discerning the proximity of objects, notably when occluded, posed significant difficulties for teleoperators.

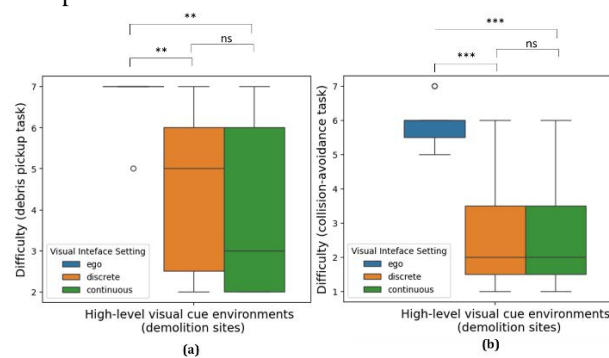


Figure 8. Difficulty level in debris removal task – (a): Picking up, (b): Collision avoidance.

During the debris picking up task, with the condition of discrete representation of visual prompts, defining the distinctions between being "close" or "very close" became ambiguous during manipulations. The sign was too vague to grasp the relative proximity between the target object (debris) and the end effector of the excavator (bucket). Contrarily, the condition of continuous representation of visual prompts, offered

comprehensive proximity details, teleoperators encountered issues in ascertaining movement orientation beforehand, requiring them excessive manipulation to check which direction they should input.

The results indicate the effect of visibility conditions to task difficulty, considering work environments as a key factor which could directly affect the teleoperator's performance and task complexity, as shown in Figure 8. In these high cue environments, the abundance of data could overwhelm operators, underscoring the crucial need for intuitive data presentation. For instance, discrete signals are typically employed to keep workers alert, while continuous representations offer accurate depth perception but can be overwhelming. Interestingly, teleoperation tasks in the high-level visual cue environments were more challenging than the low-level visual cue environments, with low visual cues posing challenges due to the absence of visual cues to infer depth more accurately. These challenges encompass diverse factors such as joystick manipulation, task intricacies, and the varying levels of visual cues.

This study endeavors to unravel the complex interactions among task performance, environmental context, and the impact of visual prompts, aiming to uncover the challenges involved. This investigation illuminates the varied challenges faced by teleoperators concerning different types of visual prompts within different levels of visual cue environments.

Conclusion

Visual interface in human-machine teleoperation is crucial particularly in addressing selective attention through visual search theory and Gibson's Theory of Perception. This paper demonstrates the crucial role of customized visual interfaces in enhancing human-robot interaction by guiding operators' attention toward essential task elements while reducing involuntary attention shifts. The investigation into depth perception complexities in varied visual environments emphasizes the necessity of assessing cue reliability to prevent performance hindrances. The comparative analysis between non-visual and visual prompt settings demonstrates the varied impacts of different visual prompts on teleoperators' performance, cognitive load, and task efficiency across diverse environments and tasks. This highlights the criticality of visual representation formats in influencing performance, with varying the environments with different level of visual cues where the effects of visual prompts on task execution vary based on individual perceptions. Moreover, this study emphasizes the need for further exploration of visual prompts' effects concerning diverse environmental and task settings, shedding light on the intricate interaction between teleoperator performance,

task characteristics, and environmental factors. The investigation provides insights into the dynamics between visibility conditions and task-oriented difficulty levels, presenting challenges and implications for future studies in optimizing visual interfaces for teleoperation, especially in contexts where heavy machinery is used.

References

- [1] Shigematsu K., Tsubouchi T. and Sarata S. Tip-over prevention system based on motion prediction for teleoperated excavator. *Advanced Robotics*, 35, 1438–1449, 2021.
- [2] Liu D., Kim J. and Ham Y. Multi-user immersive environment for excavator teleoperation in construction. *Automation in Construction*, 156, 105143, 2023.
- [3] Kamezaki M., Yang J., Iwata H. and Sugano S. Visibility Enhancement using Autonomous Multicamera Controls with Situational Role Assignment for Teleoperated Work Machines. *Journal of Field Robotics*, 33, 802–824, 2016.
- [4] Kamezaki M., Yang J., Sato R., Iwata H. and Sugano S. A situational understanding enhancer based on augmented visual prompts for teleoperation using a multi-monitor system. *Automation in Construction*, 131, 103893, 2021.
- [5] Motohashi S., Qiao Z., Moteki T. and Iwata H. Analysis of the Effect of the Operator's Spatial Cognition on Their Planning Skills in Unmanned Construction. In *2023 IEEE/SICE International Symposium on System Integration (SII)* (pp. 1–7), 2023.
- [6] Müller H. and Krummenacher J. Visual search and selective attention. *Visual Cognition*, 14, 389–410, 2006.
- [7] Hong Z., Zhang Q., Su X. and Zhang H. Effect of virtual annotation on performance of construction equipment teleoperation under adverse visual conditions. *Automation in Construction*, 118, 103296, 2020.
- [8] Li X., Fan J. and Su X. Influence of Spatial Ability on Virtual Annotation Response in Construction Equipment Teleoperation. In J. Li, W. Lu, Y. Peng, H. Yuan, and D. Wang (Eds.), *Proceedings of the 27th International Symposium on Advancement of Construction Management and Real Estate* (pp. 776–785). Springer Nature, 2023.
- [9] Livatino S., Guastella D., Muscato G., Rinaldi V., Cantelli L., Melita C., Caniglia A., Mazza R. and Padula G. Intuitive robot teleoperation through multi-sensor informed mixed reality visual aids. *IEEE Access*, 9, 25795-25808, 2021.
- [10] Arevalo Arboleda S., Rücker F., Dierks T. and Gerken J. Assisting Manipulation and Grasping in Robot Teleoperation with Augmented Reality Visual Cues. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–14). Association for Computing Machinery, 2021.
- [11] Wang M., Wong P., Luo H., Kumar S., Delhi V. and Cheng J. Predicting Safety Hazards Among Construction Workers and Equipment Using Computer Vision and Deep Learning Techniques. In *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction* (Vol. 36, pp. 399-406). IAARC Publications, 2019.
- [12] Burge J., Fowlkes C. and Banks M. Natural-scene statistics predict how the figure-ground cue of convexity affects human depth perception. *Journal of Neuroscience*, 30(21), 7269-7280, 2010.
- [13] Bogdanova R., Boulanger P. and Zheng B. Depth Perception of Surgeons in Minimally Invasive Surgery. *Surgical innovation*, 23(5), 515-524, 2016.
- [14] Park H., Faghihi N., Dixit M., Vaid J. and McNamara A. Judgments of Object Size and Distance across Different Virtual Reality Environments: A Preliminary Study. *Applied Sciences*, 11, 11510, 2021.
- [15] Newman P. and McNamara T. Integration of visual landmark cues in spatial memory. *Psychological Research*, 86, 1636–1654, 2022.
- [16] Jacobs R. What determines visual cue reliability? *Trends in Cognitive Sciences*, 6, 345–350, 2002.
- [17] Clark L., Bhagat A. and Riggs S. Extending Fitts' law in three-dimensional virtual environments with current low-cost virtual reality technology. *International Journal of Human-Computer Studies*, 139, 102413, 2020.

Zero-shot Learning-based Polygon Mask Generation for Construction Objects

Taegeon Kim¹, Minkyu Koo¹, Jeongho Hyeon¹, and Hongjo Kim¹

¹Department of Civil & Environmental, Yonsei University, South Korea

E-mail: ktg9655@yonsei.ac.kr, kmk0119804@yonsei.ac.kr, hyeon9404@yonsei.ac.kr, hongjo@yonsei.ac.kr (Corresponding author)

Abstract

For construction site monitoring, the use of segmentation-based computer vision technology has been proposed. In such environments, the main technical challenge is the generation of data for training the segmentation model. The training data for a segmentation model involves polygon annotation of objects within an image, which is a time-consuming task. To address this issue, this study proposes a new approach that uses the YOLOv8 object detection model to predict bounding box labels and inputs these into a Segment Anything Model (SAM) to automatically generate polygon label data. The performance of the YOLOv8 model exceeded 80%, and the automatic generation of polygon labels through SAM resulted in an IoU range of 55-86%, producing high-quality mask label data. This approach significantly reduces the time, labor, and cost associated with the labeling process.

Keywords – Polygon label generation, Instance segmentation, Zero-shot learning

1 Introduction

The construction industry is recognized globally as one of the most hazardous sectors, with a high incidence of injuries and fatalities [1]. Global statistics indicate that the fatality and injury rates in the construction industry are three and two times higher, respectively, than the average for other industries [2].

Faced with this high rate of accidents, the construction industry is progressively adopting advanced digital technologies such as Digital Twins (DT), Building Information Modeling (BIM), Artificial Intelligence (AI), the Internet of Things (IoT), and Smart Vision (SV) to improve efficiency, productivity, accuracy, and safety[3]. The introduction of these technologies represents a continuous effort to transition from traditional industrial practices and manufacturing methods to autonomous smart systems [4], and the construction industry has innovated its work processes through the digitalization of project management processes, gradually improving competitiveness [5].

Traditional methods of construction site monitoring often involve manual inspections and assessments, which are not only time-consuming but also prone to human error[6]. The application of computer vision technologies for site monitoring has emerged as a pivotal tool [7,8], significantly enhancing safety and progress management. Computer Vision-based systems offer a more efficient, accurate, and real-time alternative to traditional monitoring methods[9]. However, the efficacy of such systems is heavily reliant on the quality of the training data used to develop them, particularly in the context of object segmentation models [10].

The process of generating polygon annotations for the training of segmentation models has historically been a labor-intensive and time-consuming task [6]. These annotations are crucial for teaching models to accurately identify and segment various objects in a construction setting, such as equipment and personnel. The challenge is further compounded when adapting deep learning models to new domains, a process that traditionally requires extensive manual data annotation[6,11].

Recent advancements in deep learning have seen the exploration of few-shot learning and domain adaptation methods [12,13]. These techniques aim to reduce the reliance on large volumes of manually annotated training data when adapting models to new domains. However, the performance of segmentation models trained using these methods remains suboptimal. There is a notable lack of research focused on the preparation of polygon annotations for construction objects, despite the potential benefits they offer, such as monitoring personal protective equipment (PPE) compliance.

Addressing this gap, this research proposes an innovative method for the automatic generation of polygon masks, which serve as training data for segmentation models. This method leverages the capabilities of an instance segmentation model, utilizing detection results (bounding boxes) as inputs to fully automate the polygon mask generation process. Specifically, the study employs the "You Only Look Once version 8" (YOLOv8) [14] for predicting bounding boxes of construction objects. These bounding boxes are then used as prompts for the Segment Anything Model (SAM)[15], which generates the polygon masks.

To validate the effectiveness of the proposed method, experiments were conducted using the Moving Objects in Construction Sites (MOCS) dataset [16]. This dataset encompasses a comprehensive range of construction objects including workers, various types of vehicles, and equipment. The experimental design involved training the model with 19,404 images and testing it with a separate set of 4,000 images. The results indicate that the polygon masks generated by the method are comparable to those produced through manual human annotation, with an Intersection over Union (IoU) deviation ranging from 10.8% to 34.6%.

This research makes significant contributions to the field of computer vision in construction engineering. Firstly, it presents a method to automatically generate polygon annotations without the need for human involvement, thus streamlining the training process for segmentation models. Secondly, it demonstrates the quality and viability of automatically generated segmentation masks derived from detection results.

2 Related work

In the field of computer vision, the development of deep learning models heavily relies on the quality and quantity of training data. Particularly, segmentation models require pixel-level labeling data, a process that consumes approximately ten times more resources in terms of labor and cost compared to bounding box labeling for object detection models [17]. In this experience, the preparation of polygon masks takes ten to twenty times more efforts than the preparation of bounding box labels. This challenge becomes more complex in dynamic environments such as construction sites, where the constant movement of equipment and machinery necessitates diverse data collection. To overcome these challenges, the construction domain has conducted active research in various ways such as synthetic data generation, zero-shot or few-shot learning, and domain adaptation.

2.1 Synthetic data generation

Synthetic data has emerged as a key solution to alleviate the time and labor burdens associated with preparing data for model training. There has been research utilizing synthetic data for visual data analysis in infrastructure management, automating the data collection process to address labor-intensive and time-consuming issues [18]. Additionally, studies have been conducted on acquiring scaffolding point cloud data through Mobile Laser Scanning (MLS) and utilizing it for training data in construction sites [19].

2.2 One-shot or few-shot learning

One of the key challenges in the advancement of

segmentation model, especially supervised learning models, is effectively recognizing object with limited training data. Many models require a substantial amount of training data for each class, but obtaining sufficient data for certain classes can often be challenging. To address this issue, new learning paradigms such as few-shot and one-shot learning methods [20,21] have been proposed. These approaches utilize knowledge from instances of various classes to enable effective learning even with a small number of instances, aiming to overcome additional challenges such as classifying instances of classes that have not been encountered before [22].

2.3 Domain Adaptation

Domain adaptation addresses performance degradation due to differences in data distribution between source and target domains. Enhancing model performance by adding data from the target domain similar to the source domain has been explored. Particularly, self-training techniques, which involve using a trained model to predict labels on unlabeled data and utilizing it as training data, have been researched to improve the generalization capabilities of models [6].

2.4 Knowledge gap of previous studies

Despite the advancements in these technologies, the need for human annotation remains a crucial element in the training process of segmentation models, especially when the model is applied to a new target domain. Human annotation ensures the quality and accuracy of data, playing a vital role in reflecting the complexity and diversity of construction environments in the data. Based on this context, this study proposes a new methodology that can automatically generate training data, emulating the characteristics of human-annotated data.

The methodology developed in this research utilizes SAM to automatically generate high-quality training data comparable to human annotation. This presents an opportunity to effectively enhance the performance of deep learning models, particularly in complex and dynamic environments like construction sites. The automated data generation process can significantly replace the time-consuming and costly tasks performed by human annotators, providing rich training data in a faster and more cost-efficient manner.

3 Proposed Method

3.1 Overview of the proposed method

To automatically generate polygon masks of construction objects, the proposed method, as shown in Fig. 1, is divided into 2 steps, as follows:

- 1) Training the YOLOv8 object detection model using the MOCS training dataset to predict bounding boxes on the test dataset.
- 2) Polygon mask generation using the Segment Anything Model (SAM)

3.1.1 Object detection model training and bounding box prediction

You Only Look Once version 8 (YOLOv8) [14,23] is the latest model that enables fast and accurate object detection, similar to the human visual system. This model performs the process of classifying objects within an image and determining their location information through a single inference. Among the YOLO series, YOLOv8 has established itself as the preferred architecture in applications requiring fast inference speeds by providing the highest mAP performance and inference speed on the Microsoft Common Object in Context (MS COCO) dataset [24].

Despite these capabilities, performance degradation occurs due to visual differences between the training domain and the intended application domain (target domain), which is more pronounced in complex environments like construction sites. Therefore, to maximize the model's performance for a specific domain, it is essential to optimize or retrain the model's weights for the target domain data. In this study, the YOLOv8 model pre-trained on the COCO dataset was retrained on the MOCS dataset. The re-trained YOLOv8 model was used to predict bounding boxes of target construction object classes (see Fig.2 listing the target object classes).

3.1.2 Automated polygon mask generation with predicted bounding box

This study utilizes the Segment Anything Model (SAM), an instance segmentation model, which serves as a foundational model in the field of computer vision, aiming for the universality like that of ChatGPT. This model was trained on the 1.1 billion SA-1B dataset and possesses the capability to perform segmentation of various objects through simple prompt input [25]. SAM can process various forms of prompts, including masks, bounding boxes, points, and text, enabling the automatic generation of polygon label data [15,17]. The proposed method leverages SAM to efficiently generate high-quality polygon label data within construction site environments. Specifically, the bounding box prediction information for construction site objects predicted by the YOLOv8 model is used as input prompts, generating accurate polygon label data. This methodology automates the generation of polygon mask data in complex and dynamic construction site environments by integrating prompt-based systems with the latest computer vision model, SAM, replacing labor-intensive data labeling tasks in the target

domain.

4 Experiments

4.1 Experimental Settings

4.1.1 Computer & Datasets

The experiments were conducted on systems equipped with 4 NVIDIA GeForce RTX 4090 GPUs running on Ubuntu 20.04 operating system. The Moving Objects in Construction Sites (MOCS) dataset was utilized, comprising 19,404 training images and 4,000 validation images. All target classes (Worker, Static crane, Hanging head, Crane, Roller, Bulldozer, Excavator, Truck, Loader, Pump truck, Concrete mixer, Pile driving, Other vehicle) were used. This data was used to train a YOLOv8 object detection model, which was then utilized to predict bounding boxes on the MOCS validation dataset composed of 4,000 images. The MOCS dataset includes publicly available images for training, validation, and testing; however, bounding boxes and polygon annotations are only provided for the training and validation datasets, which limits the use of the test dataset. Therefore, the performance of the fully trained model was evaluated using the validation dataset. The predicted bounding boxes and original images were inputted into SAM to generate polygon mask label data, and the IoU of these extracted mask labels was calculated. To compare the accuracy of the generated mask label data, ground truth bounding boxes and original images were inputted into SAM to calculate the IoU, and finally, the differences between the two results were compared.

4.1.2 Model selection & hyperparameters

In this experiment, the largest 'x' model of the YOLOv8, pre-trained on the COCO dataset, was used. Although this model has a large number of parameters, resulting in longer training times, it was selected for its outstanding performance on the COCO dataset. Additionally, a threshold of 0.5 was set for the bounding boxes predicted by the detection model, and the extracted bounding boxes were used as prompt values for SAM.

4.1.3 Model training

The hyperparameters set for YOLOv8 training included an image size of 1280*720 HD, 300 epochs, a batch size of 16, and a learning rate of 0.01. The epochs were adjusted to 300 to correspond with the quality of the data. Additionally, data augmentation techniques such as Mixup and Copy-Paste were incorporated into the training process.

5 Experimental Results

5.1 Detections results of YOLOv8

As shown in Table 1, the mAP results, which are the primary performance metrics for the YOLOv8 detection model trained on the MOCS dataset. The target class that

exhibited the highest performance was 'Excavator' with mAP of 94.2%, while the lowest performing target class was 'Crane' with mAP of 77.7%. Despite a considerable number of instances in the Worker class as shown Fig. 2, which could have led to concerns about model bias, the model still achieved high performance.

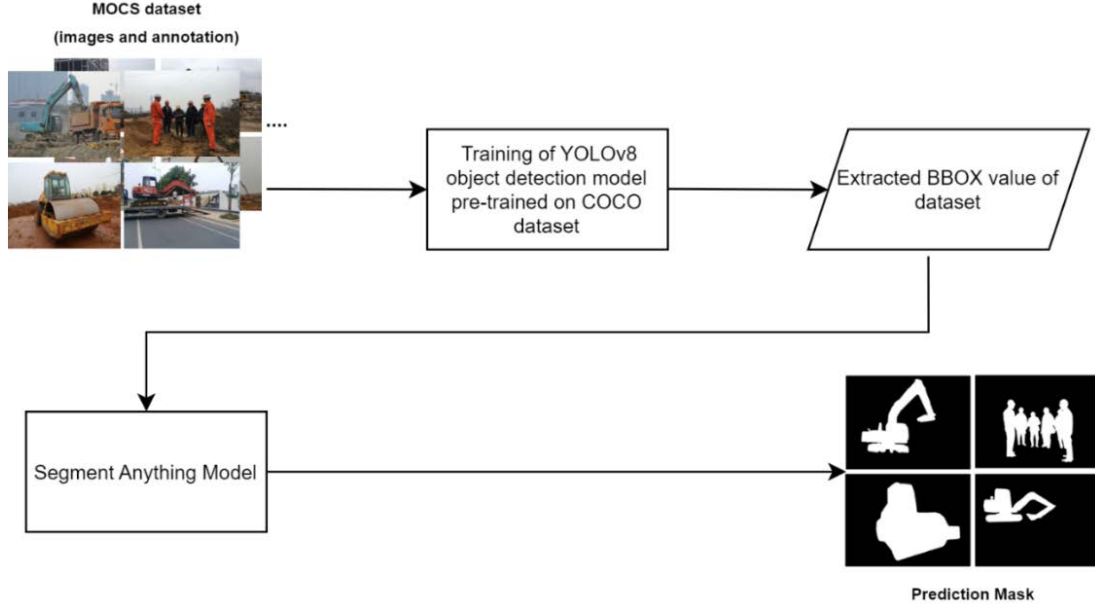


Figure 1. Overview of proposed method

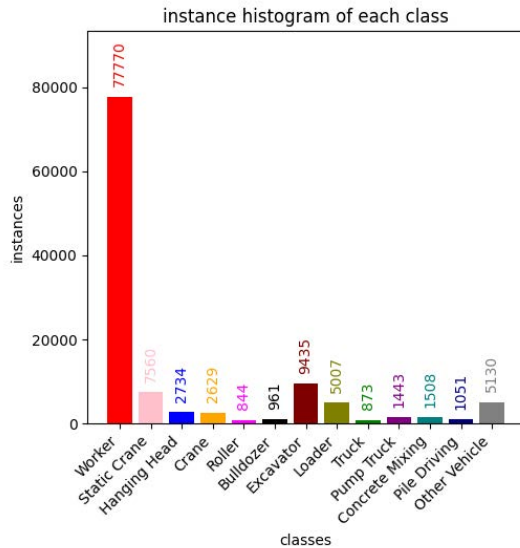


Figure 2. Number of instances each class in the MOCS dataset

5.2 SAM Result

Table 2 shows the results of the Segment Anything Model. The IoU values for masks generated by inputting predicted bounding boxes and original images into SAM were calculated and compared with those of masks created using actual ground truth bounding boxes. The most substantial discrepancy was observed in the 'Concrete Mixer' class with approximately 21%, and the smallest discrepancy was in the 'Bulldozer' class with about 3%. These figures provide crucial information on how accurately SAM classifies and labels objects across various classes.

These results indicate that SAM can generate polygon labels with relatively higher accuracy for certain classes of objects, while exhibiting lower accuracy for others. This suggests a critical interplay between the performance of the SAM algorithm and characteristics of objects such as their features, shapes, sizes, and colors. Understanding this interaction is vital in improving the algorithm of SAM and enhancing its detection and labeling performance for specific classes.

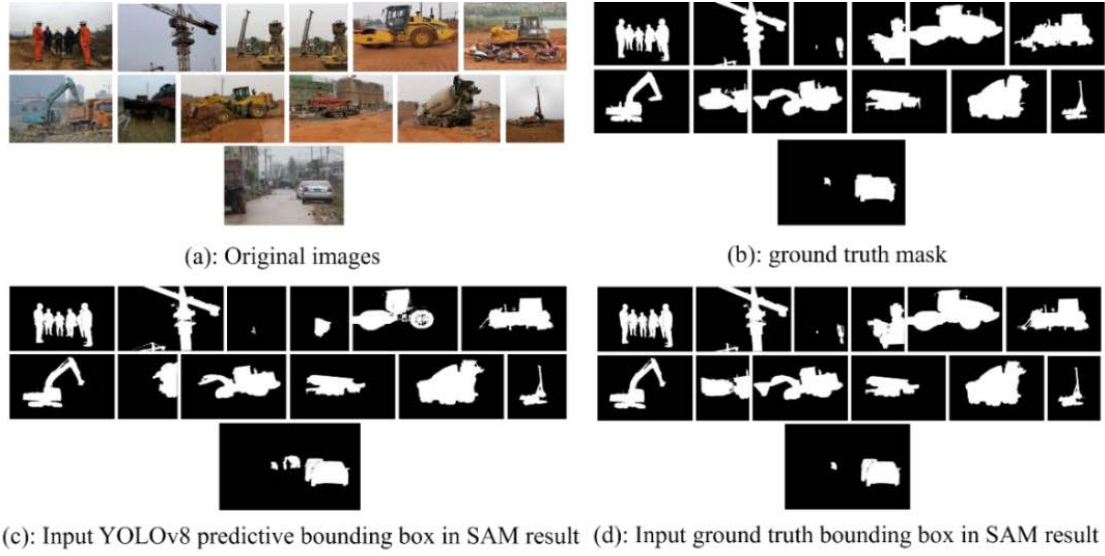


Figure 3. Original images and generated polygon labels

6 Discussion

Performance of YOLOv8 to generate bounding box prompts for SAM: In this study, the performance of the YOLOv8 object detection model was evaluated based on the mean Average Precision@50 (mAP@50) metric, showed an impressive result of 85.8% in the average mAP across all classes. Additionally, the mAP performance for each class was observed to range between 77% and 94%, further emphasizing the model's excellence (see Table 2). According to the study by Xuehui et al. [16], when other models such as YOLOv3 and Faster R-CNN were applied to the MOCS test dataset, they showed relatively low performances with mAP at 9.99% and 66.59%, respectively. However, the direct comparison between the previous methods [16] and our method is not possible as the test datasets were different (This study used the MOCS validation dataset because the ground truth of the MOCS test dataset is not publicly available).

Quality of generated polygon annotations: Furthermore, the Intersection over Union (IOU) of polygon mask labels generated by inputting predicted bounding boxes to SAM was measured between 55% to 86% among the target classes. In contrast, when actual ground truth bounding boxes were input into SAM, the IOU was evaluated to be between 65% and 89%. There was a performance difference of 5% to 22% between polygon label data generated by these two methods.

In the study by Chern et al. [26], a dataset having a single target class per image was used, and the model's Pseudo Label (P.L), Refined Pseudo Label (Refined P.L), and Feature Pyramid Networks (FPN) performance for target classes Background, Dump truck, Excavator, Mixer truck, Roller ranged between 48.48% and 78.76%.

This demonstrated that the IOU performance of mask labels generated by SAM for all classes was higher.

These results illustrate the competitiveness of the proposed system to generate target domain annotations in a fully automated manner, reducing the labor and cost involved in bounding box labeling. These results offer the possibility of automatic training data generation in complex environments like construction sites.

As shown in Fig. 3, instances were identified where ground truth labeling for certain objects was inaccurately applied. In some cases, polygon annotations generated by SAM were more accurate than the original annotations. These factors are considered to have contributed to the observed decline in overall performance.

6.1 Limitation & Future study

Several key limitations were identified in the process of automatically generating polygon mask label data. Firstly, it was revealed that the detection performance for small objects with a low number of pixels within the image was low. If small objects' bounding boxes were not accurately detected, SAM was not able to generate polygon masks for those objects. Secondly, SAM experienced difficulties in generating accurate mask labels in areas where the object's features are similar to the background. For example, as shown in Fig. 4, the generated polygon annotations by SAM are more accurate compared to the annotations by humans. This was particularly pronounced in situations involving objects partially obscured by other objects. While SAM generates masks without considering the obscured parts, ground truth data includes the obscured portions of objects in the polygon masks. Since the exact shape of objects hidden behind obstructions cannot be known, the predicted polygons by SAM are more accurate than the original masks in these

respects. Additionally, some ground truth annotations were represented differently from the actual objects' appearance. These factors are deemed to have negatively impacted the overall performance of SAM.

7 Conclusion

This study presents the automated polygon annotation generation method using YOLOv8 for an object detector and SAM for an annotation generator. The experimental results showed promising results, with quality annotations comparable to the ground truth, as shown in Figure 3. These results provide important implications for the accuracy and reliability of generating polygon label data in complex environments like construction sites. Therefore, future research should focus on improving the performance of the object detection model and

optimizing the SAM model. Through this, more accurate and reliable generation of polygon label data is expected, contributing to the expansion of the application scope in the fields of deep learning and computer vision.

Acknowledgment

This research was conducted by the support of the “2023 Yonsei University Future-Leading Research Initiative (No. 2023-22-0114)” and the “National R&D Project for Smart Construction Technology (No. RS-2020-KA156488)” funded by the Korea Agency for Infrastructure Technology Advancement under the Ministry of Land, Infrastructure and Transport, and managed by the Korea Expressway Corporation.

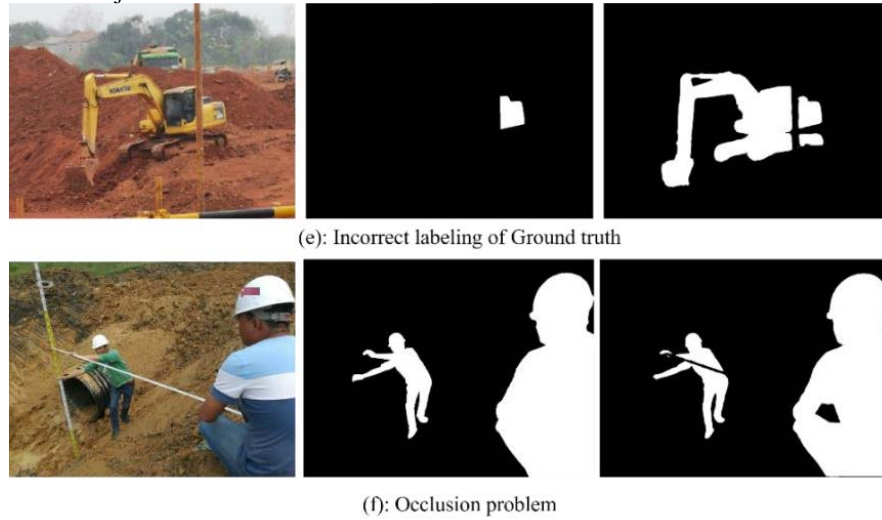


Figure 4. Original image (Left), Ground truth mask (Center), SAM predicted mask (Right)

Table 1. mAP performance for each class of object detection

Other Model	Target Class												
	Worker	Static crane	Hanging head	Crane	Roller	Bulldozer	Excavator	Truck	Loader	Pump truck	Concrete mixer	Pile driving	Other vehicle
Yolov8 object detection model (Ours)	91.3	80.7	80.5	77.7	90.5	91.3	94.2	87.0	83.3	86.8	89.5	80.8	80.9

Table 2. SAM result on MOCS dataset

Segment Anything	Target Class
------------------	--------------

Model		Worker	Static crane	Hanging head	Crane	Roller	Bulldozer	Excavator	Truck	Loader	Pump truck	Concrete mixer	Pile driving	Other vehicle
mIoU	Predicted BB OX	78.5	65.6	73.0	68.2	86.7	83.2	76.3	80.5	83.2	66.3	65.9	55.8	59.9
	GT BB OX	83.1	79.7	78.6	77.1	89.2	86.3	79.8	85.9	86.8	75.7	86.6	65.4	77.4

References

- [1] M. Abbas, B.E. Mneymneh, H. Khoury, Assessing on-site construction personnel hazard perception in a Middle Eastern developing country: An interactive graphical approach, *Safety Science* 103 (2018) 183–196. <https://doi.org/10.1016/j.ssci.2017.10.026>.
- [2] S. Teran, H. Blecker, K. Scruggs, J. García Hernández, B. Rahke, Promoting adoption of fall prevention measures among Latino workers and residential contractors: Formative research findings, *American Journal of Industrial Medicine* 58 (2015) 870–879. <https://doi.org/10.1002/ajim.22480>.
- [3] F. Craveiro, J.P. Duarte, H. Bartolo, P.J. Bartolo, Additive manufacturing as an enabling technology for digital construction: A perspective on Construction 4.0, *Automation in Construction* 103 (2019) 251–267. <https://doi.org/10.1016/j.autcon.2019.03.011>.
- [4] S.K. Baduge, S. Thilakarathna, J.S. Perera, M. Arashpour, P. Sharafi, B. Teodosio, A. Shringi, P. Mendis, Artificial intelligence and smart vision for building and construction 4.0: Machine and deep learning methods and applications, *Automation in Construction* 141 (2022) 104440. <https://doi.org/10.1016/j.autcon.2022.104440>.
- [5] Economies | Free Full-Text | Industry 4.0 for the Construction Industry: Review of Management Perspective, (n.d.). <https://www.mdpi.com/2227-7099/7/3/68> (accessed March 13, 2024).
- [6] Y. Hong, W.-C. Chern, T.V. Nguyen, H. Cai, H. Kim, Semi-supervised domain adaptation for segmentation models on different monitoring settings, *Automation in Construction* 149 (2023) 104773. <https://doi.org/10.1016/j.autcon.2023.104773>.
- [7] R. Bai, M. Wang, Z. Zhang, J. Lu, F. Shen, Automated Construction Site Monitoring Based on Improved YOLOv8-seg Instance Segmentation Algorithm, *IEEE Access* 11 (2023) 139082–139096. <https://doi.org/10.1109/ACCESS.2023.3340895>.
- [8] Z. Wang, Y. Zhang, K.M. Mosalam, Y. Gao, S.-L. Huang, Deep semantic segmentation for visual understanding on construction sites, *Computer-Aided Civil and Infrastructure Engineering* 37 (2022) 145–162. <https://doi.org/10.1111/mice.12701>.
- [9] J. Cho, K. Kim, Detection of moving objects in multi-complex environments using selective attention networks (SANet), *Automation in Construction* 155 (2023) 105066. <https://doi.org/10.1016/j.autcon.2023.105066>.
- [10] G. Fenza, M. Gallo, V. Loia, F. Orciuoli, E. Herrera-Viedma, Data set quality in Machine Learning: Consistency measure based on Group Decision Making, *Applied Soft Computing* 106 (2021) 107366. <https://doi.org/10.1016/j.asoc.2021.107366>.
- [11] N.A. Koohbanani, B. Unnikrishnan, S.A. Khurram, P. Krishnaswamy, N. Rajpoot, Self-Path: Self-Supervision for Classification of Pathology Images With Limited Annotations, *IEEE Transactions on Medical Imaging* 40 (2021) 2845–2856. <https://doi.org/10.1109/TMI.2021.3056023>.
- [12] Y. Gao, H. Li, W. Fu, Few-shot learning for image-based bridge damage detection, *Engineering Applications of Artificial Intelligence* 126 (2023) 107078. <https://doi.org/10.1016/j.engappai.2023.107078>.
- [13] H. Wang, M. Xu, B. Ni, W. Zhang, Learning to Combine: Knowledge Aggregation for Multi-Source Domain Adaptation, (2020). <https://doi.org/10.48550/arXiv.2007.08801>.
- [14] G. Jocher, A. Chaurasia, J. Qiu, YOLO by Ultralytics, (2023). <https://github.com/ultralytics/ultralytics> (accessed December 28, 2023).
- [15] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A.C. Berg, W.-Y. Lo, P. Dollár, R. Girshick, Segment Anything, (2023). <http://arxiv.org/abs/2304.02643> (accessed July 27, 2023).
- [16] A. Xuehui, Z. Li, L. Zuguang, W. Chengzhi, L. Pengfei, L. Zhiwei, Dataset and benchmark for detecting moving objects in construction sites, *Automation in Construction* 122 (2021) 103482. <https://doi.org/10.1016/j.autcon.2020.103482>.
- [17] M. Gröger, V. Borisov, G. Kasneci, BoxShrink: From Bounding Boxes to Segmentation Masks, (2022). <http://arxiv.org/abs/2208.03142> (accessed December 27, 2023).
- [18] H. Murtaza, M. Ahmed, N.F. Khan, G. Murtaza, S. Zafar, A. Bano, Synthetic data generation: State of the art in health care domain, *Computer Science Review* 48 (2023) 100546. <https://doi.org/10.1016/j.cosrev.2023.100546>.
- [19] J. Kim, J. Kim, Y. Kim, H. Kim, 3D reconstruction of large-scale scaffolds with synthetic data generation and an upsampling adversarial network, *Automation in*

- Construction 156 (2023) 105108.
<https://doi.org/10.1016/j.autcon.2023.105108>.
- [20] Li Fei-Fei, R. Fergus, P. Perona, One-shot learning of object categories, *IEEE Trans. Pattern Anal. Machine Intell.* 28 (2006) 594–611.
<https://doi.org/10.1109/TPAMI.2006.79>.
- [21] S. Ravi, H. Larochelle, Optimization as a Model for Few-Shot Learning, in: 2016. <https://openreview.net/forum?id=rJY0-Kc1l> (accessed December 28, 2023).
- [22] W. Wang, V.W. Zheng, H. Yu, C. Miao, A Survey of Zero-Shot Learning: Settings, Methods, and Applications, *ACM Trans. Intell. Syst. Technol.* 10 (2019) 1–37. <https://doi.org/10.1145/3293318>.
- [23] D. Reis, J. Kupec, J. Hong, A. Daoudi, Real-Time Flying Object Detection with YOLOv8, (2023). <http://arxiv.org/abs/2305.09972> (accessed December 27, 2023).
- [24] T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C.L. Zitnick, P. Dollár, Microsoft COCO: Common Objects in Context, (2015). <http://arxiv.org/abs/1405.0312> (accessed February 13, 2024).
- [25] C. Zhang, L. Liu, Y. Cui, G. Huang, W. Lin, Y. Yang, Y. Hu, A Comprehensive Survey on Segment Anything Model for Vision and Beyond, (2023). <http://arxiv.org/abs/2305.08196> (accessed December 28, 2023).
- [26] W.-C. Chern, T. Kim, T.V. Nguyen, V.K. Asari, H. Kim, Self-supervised sub-category exploration for Pseudo label generation, *Automation in Construction* 151 (2023) 104862.
<https://doi.org/10.1016/j.autcon.2023.104862>.

Integration of laser profiler feedback into FIM for additive manufacturing in construction

Simon Leon Espinosa¹, Martin Slepicka¹ and André Borrman¹

¹Chair of Computational Modeling and Simulation, Technical University of Munich, Germany

simon.espinosa@tum.de, martin.slepicka@tum.de

Abstract -

This research delves into the utilization of laser profiling technology within the realm of Three-Dimensional (3D) printing in the construction industry. The primary emphasis is on achieving a seamless integration of Fabrication Information Modeling (FIM), Additive Manufacturing (AM), and Digital Twinning (DT). The study involves the capture of precise measurements at various stages and storing them as Point Cloud (PC) data. Moreover, the objective is to transform operational planning and enhance real-time “as-built” data generation by intricately mapping shape and height deviations and storing the corresponding data in the FIM model.

Keywords -

BIM; FIM; RTDE; Cyber-Physical-Systems; Digital Twin

1 Introduction

In comparison to other industry sectors, the construction industry has traditionally faced challenges in improving productivity. However, recent advancements in key technologies are poised to address this longstanding issue. Building Information Modeling (BIM), Additive Manufacturing (AM), and Digital Twinning have already made their mark, offering transformative possibilities. In this context, Fabrication Information Modeling (FIM) [1] was introduced as a bridge between digital design and automated manufacturing in order to further expand the level of automation in the construction industry and thus contribute to the fourth industrial revolution, commonly referred to as “Industry 4.0”. [2]

Each of these technologies plays a unique role in reshaping the construction landscape. BIM is a digital design methodology for creating and using digital building models that are intended to represent the physical and functional properties of a building as accurately as possible [3]. AM helps to automate manufacturing on-site and off-site and enables the production of components with complex geometry and improved material efficiency. As an intermediary layer between BIM and AM, FIM manages fabrication information, facilitating a seamless transition of 3D information into discrete two-dimensional (2D) data for layer-by-layer manufacturing [1]. Digital Twinning (DT) ensures real-time synchronization between physical objects and digital models [4].

Industry 4.0 is characterized by increasing digitization, smart manufacturing, and customization, driving the development of the construction industry from planning to construction, focusing on quality control through sensors and thorough planning. As technology integration increases, DT and AM methods are employed to update digital models and automate construction processes. Both technologies reduce the need for manual labor by automating repetitive or controllable tasks and thus enhance productivity [2]. However, along with all the advantages of AM methods, there are also limitations. One of the problems with AM is the lower reliability compared to traditional manufacturing methods, especially in terms of deviations between the as-printed and as-designed state, if the AM process is not adequately controlled by feedback [5]. In this context, this study focuses on enabling digital twinning methods for AM, aiming to provide real-time monitoring for improved decision-making.

The data structure of FIM models was designed based on the BIM exchange data format Industry Foundation Classes (IFC) and can contain various model representations side by side. This means that a FIM model can include both the as-designed data extracted from the BIM model, fabrication information, and later sensor data (as-printed data). Based on such a data foundation, with a system for real-time data exchange [6], and with the sensor integration proposed in this study, it is feasible that FIM can be used for cyber-physical systems (CPS) [7].

In the following, the technologies on which this work is based and their current status are first described in section 2. Subsequently, the proposed methodology is detailed in section 3. Finally, information on this methodology’s impact on the industry is provided in section 5.

2 Background

The study presented in this article examines how sensor feedback can be integrated into FIM-based additive manufacturing. To this end, an insight into the FIM methodology and the underlying data structure is first provided. After this, the optical sensor system used in this work is described.

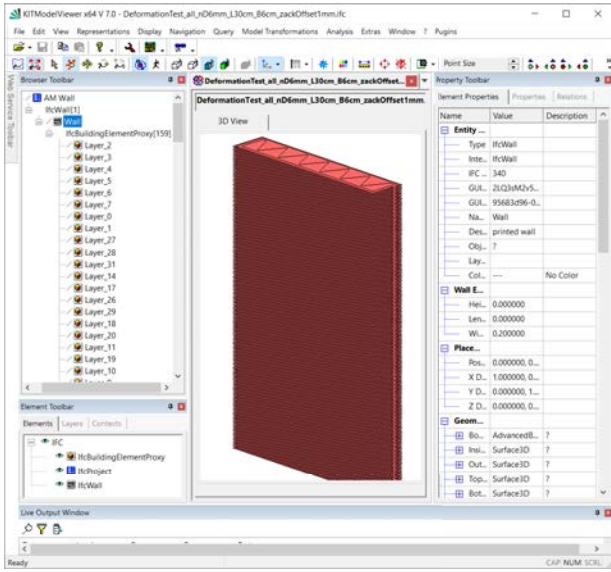


Figure 1. Visualization of the FIM-IFC data.

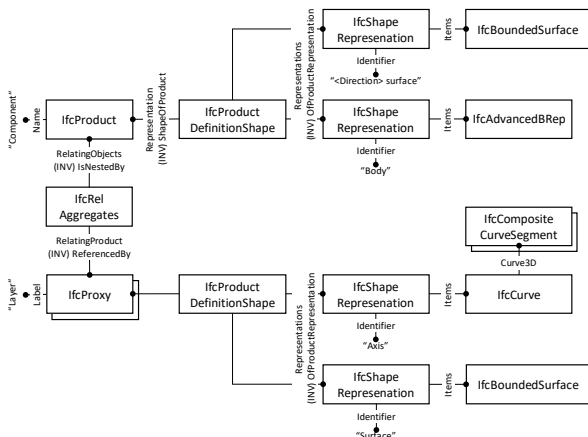


Figure 2. Express-G diagram of the FIM data structure, after [1].

2.1 Fabrication Information Modeling (FIM)

According to Slepicka et al. [1], FIM is an intermediate layer between digital design and automated manufacturing. The FIM framework implemented by Slepicka et al. [1] can be used to derive manufacturing information from digital models, such as a BIM model, which can be used directly in manufacturing. For this purpose, the tool path for creating the object is generated from the extracted information of the digital model and enriched with manufacturing data.

With the FIM Framework, components can be extracted from a BIM model and automatically enriched with manufacturing information using a pattern-based algorithm. The 3D geometry data, if not already given as boundary

representation (B-Rep), is converted into a B-Rep model, and the corresponding boundary surfaces are topologically mapped. A pattern-based path planning algorithm is then executed to generate a C2-continuous tool path using the boundary surfaces and slicing planes. Additionally, to avoid rapid tool acceleration, e.g., in tight curves, a velocity profile is generated depending on the curvature of the path. [1]

The data generated with FIM, i.e., B-Rep, slicing planes, path, and speed profile, are formatted in the BIM exchange data format, the Industry Foundation Classes (IFC). The different data is attached to the same component as different representations (IfcShapeRepresentation). The component to be printed (e.g., IfcWall) is represented as an assembly (IfcRelAggregates) of layers (IfcBuildingElementProxy), each of which is described by its slicing surface (IfcSurface) and the corresponding tool path (IfcCompositeCurve) (see figs. 1 and 2) [1].

The data structure just described, in particular the ability to access the individual surfaces of the component, has several functions. As described above, the boundary and layer surfaces are used for path planning. However, the corresponding surfaces can also be used as an aid for simulation purposes [8]. Moreover, the subject of this study is the possibility of using these surfaces as a reference for “as-printed” measurements.

2.2 Geometric data acquisition

There are various optical sensor systems for 3D geometry acquisition, e.g., active and passive stereo cameras, terrestrial LiDAR, structured light sensors, and laser profilers. This study focuses on laser profiling, a technology in which a laser line is projected onto an object to be measured, which is captured using a camera via triangulation. As in the study by Chen et al. [9], in which this technology was used for quality assurance in an AM process, a laser profiler is also attached to the robotic arm carrying out the AM process in this study in order to be able to capture the component geometry during the printing process.

The sensor works by projecting a laser line on the object to be measured and capturing the reflection to create a 2D profile, as illustrated in fig. 3. The object is scanned, and a 2D version and its surroundings are obtained. It is important to note that only the reflected data can be captured; therefore, some parts of the measured object may be occluded and not visible to the scanner. For this reason, the scanning path and orientation are crucial for the results.

This type of sensor is commonly used in production lines, where the sensor is mounted stationary, and the objects to be measured are moved under the sensor on a conveyor belt. In this way, individual line measurements

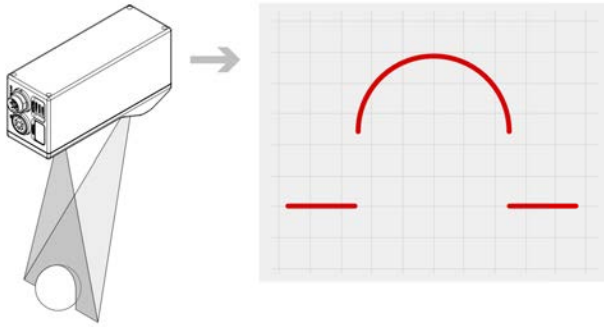


Figure 3. Functional principle of the laser profiler (left) and the corresponding measurement result (right) [10].

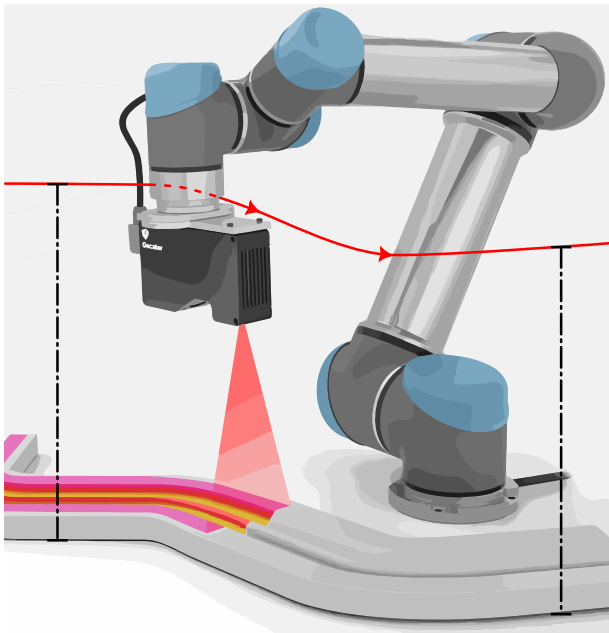


Figure 4. Robot mounted laser profiler on a path at a constant distance to the base surface.

can be combined to provide a surface measurement. However, as Chen et al. [9] show, a laser profiler can also be mounted on a robot to carry out surface measurements. A necessary prerequisite for accurate measurement is that the robot guiding the sensor moves as uniformly as possible over the object to be measured, as stated by our laser profiler documentation [10], see fig. 4.

If the distance between the sensor and the ground is known, the sensor output can be calibrated so that the height of the measured object can be returned in relation to the ground. This allows a matrix with height values to be created from several line profiles, i.e., a 2.5D representation of the object. Within the laser line, there is

a resolution depending on the distance to the object, and from one laser line to another laser line, there is a resolution depending on the speed of movement of the sensor (if robot-mounted).

3 Method

As previously indicated, the hypothesis in this study is that laser profiling can be integrated into the FIM methodology (cf. section 2.1) for “as-printed” data feedback. The idea is that the manufacturing information available in a FIM model can be used to generate robot trajectories for a scanning process with a laser profiler (cf. section 3.1). The data captured along these trajectories needs to be post processed by removing unwanted data (cf. section 3.2) and transforming it into manageable packages (cf. section 3.3), which can be fed back into the FIM model (cf. section 3.4). In both cases, the boundary and layer surfaces stored in the FIM model are used for this purpose (cf. figs. 1 and 2).

3.1 Scan planning

As already indicated in section 2.2, the sensor must be moved at a constant speed along a robot trajectory at a constant distance from a reference plane so that a scanning process with a laser profiler can deliver usable results (cf. fig. 4). To create such a robot trajectory, the corresponding surface geometry of the part to be measured can be taken from the FIM model. If, for example, a printed layer is to be checked, the corresponding layer surface can be used. When measuring the overall geometry, on the other hand, the individual surfaces of the B-Rep representation can be used. Using the profile width of the sensor, the desired surface can be divided into suitable segments from which the central isocurve can be derived and shifted to a constant distance from the surface using an offset. It is important to note that the offset curves created in this way must be extended beyond the corresponding area at both ends. A robot needs some time to accelerate to a constant speed, and the quality of the measurement with the laser profiler is impaired during acceleration. The extended curves create a buffer zone to compensate for this. Additionally, this extended area for measurement enables the detection of material outside the designed range (over-extrusion).

3.2 Data filtering

As previously described in section 3.1, the paths for the scanning process must be longer than the actual object geometry. Furthermore, if a surface had to be recorded in several segments, several data sets must be combined into one. This means that a few measured lines have to be removed, and lines have to be trimmed.

In addition, invalid measured values must also be removed. A measured value is considered invalid if it lies within the range of another measurement (e.g., in the previous layer) or far outside the expected range (background). The decision as to whether a measured value is considered invalid is determined in this study via a threshold value. By defining a region of interest (ROI) that incorporates the designed geometry, including a margin of error, unwanted data points can be removed from the captured data. Using the FIM data, it is possible to define this ROI in the form of a volume in space; every measured value outside this volume can be excluded. Additionally, this method simplifies assigning the measured values to the correct “as-designed” surface.

3.3 Deviation maps

The sensor data must be correlated with the geometric information in the FIM model to enable data feedback later. If the sensor is operated as described in section 3.1, deviations or heights are measured by the sensor in relation to the reference plane used, i.e., to the “as-designed” surface that was used to generate the robot trajectory. As described in section 2.2, the sensor output is then a matrix with the height values at the equally distributed measurement positions. In this respect, the measured data can also be interpreted in relation to surface coordinates (UV coordinates). Effectively, only the individual matrix positions of the measured values need to be mapped to the UV coordinates of the reference surface, i.e., instead of assuming a fixed Cartesian distance between matrix cells, a parameter distance can also be considered (cf. section 3.4).

If the measured height value is then coded in a color value, a heat map (or deviation map) can be derived (cf. figs. 9 and 10). The heat map is designed to show the deviation in two different scales for a more intuitive visualization. A cold scale visualizes the over-extruded sections (height values larger than zero), while the under-extruded sections (height values smaller than zero) are visualized by a warm scale.

Similar to the height deviation map, which visualizes the deviation of the measured values in the normal direction to the reference surface, a second deviation map can also be created for each of the layer surfaces, which shows the deviations from the “as-designed” printing path. These are then deviations tangential to the reference surface. These deviations are determined by filtering the height values of the sensor matrix using binary filtering and the “as-designed” position, which is determined using the printing path in the FIM model. If the “as-designed” and “as-printed” data are superimposed, different areas can be color-coded depending on whether too many, too few, or exactly the right data points were detected in the “as-printed” data set. This approach enables a detailed

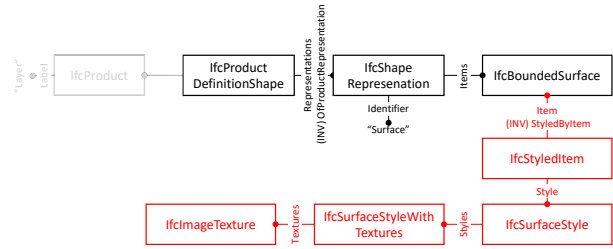


Figure 5. Data structure for attaching a texture to a geometry.

analysis, distinguishing between instances where the material is present in unintended locations and situations where the material is absent from anticipated positions.

3.4 Data feedback

Both deviation maps described in the previous section are already created in such a way that the data is referenced to the surface geometries stored in the FIM model. Accordingly, the maps can be attached to the data model as a texture without further revision, as already indicated above. The heat map textures can be attached to the corresponding surface as IfcStyledItem in the FIM model. The necessary additions to the data model are illustrated in fig. 5 (additions in red).

4 Experimental validation

In order to validate the methodology described in section 3, various tests were carried out on a small-scale setup. In the following, the setup is first described (cf. section 4.1), and then the necessary implementations are explained (cf. section 4.2). Finally, the results obtained are presented section 4.3.

The experimental part of this research was conducted in the scope of the master’s thesis by Espinosa [11].

4.1 Setup

To validate the methodology proposed in this study, a clay 3D printing manufacturing setup was used. In extrusion-based 3D printing, clay behaves similarly to concrete and is well suited for conducting preliminary laboratory tests. Figure 6 shows this setup. It consists of two Universal Robots arm robots, a UR10e and a UR5e, the former equipped with a Stoneflower3D clay extruder and the latter with a Gocator 2330 laser profiler (Resolution Z: 0.006 mm, Clearance Distance: 90 mm, Measurement Range: 80 mm, Field of View: 47–85 mm). Both robots are securely positioned on a “level” table (cf. section 4.3). Additionally, the UR5e (in the backside of fig. 6) is placed on a longitudinal axis for extended reach.



Figure 6. Full setup at home position. UR10e robotic arm outfitted with a stoneflower3D clay extruder on the right. And on the left a UR5e on a SlideKit linear axis outfitted with a Gocator 2330 laser profiler.

Each robot undergoes calibration, and a transformation matrix is computed to establish a correlation between their individual positions. This matrix enables a cross-reference of their respective coordinate systems, which is crucial for aligning the “as-built” structure with the “as-designed” model. Beyond alignment, this transformation matrix plays a pivotal role in precisely situating the “as-built” model on the table, ensuring accurate positioning.

The alignment of these two data sets is facilitated by the utilization of a Real-Time Data Exchange (RTDE) module by Slepicka et al. [6] in both UR robots, enabling precise positioning for their respective tasks. This synchronization ensures that the position of each robot is consistently known, not only facilitating alignment but also serving as a valuable resource for the calibration of measurements throughout the entire process.

The Gocator sensor is securely attached to the robot, and the scanning process is executed by moving the sensor from the start to the finish point using the URCap provided by the sensor’s manufacturer, LMI [10, pp. 1038–1063]. The connection between the Gocator and the UR robot is established through the Dynamic Host Configuration Protocol (DHCP).

Continuous robot movement ensures that the Gocator sensor moves at a speed of 60 mm/s during scans. To ensure consistent scanning performance, the robot must reach maximum speed before reaching the scanning area. This particular requirement enables accurate and reliable data capture throughout the scanning process.

4.2 Implementation

The computational tools for the Lab setup are implemented using Python and consist of two main classes to control the information flow. One class effectively manages the data import from the Gocator sensor (“as-printed” data) and all the coordinate information of it. The other class manages the data flow from and to the FIM model, in-

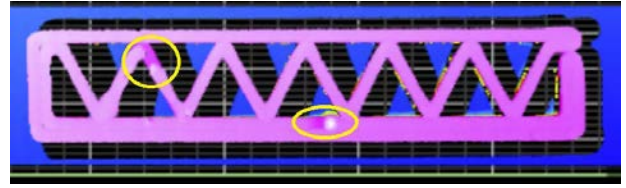


Figure 7. Imported raw sensor data of a layer scan (Layer 14). Even before post-processing, ROIs can be identified (marked by the yellow circles). Visible are a defect (left) and the transition from the previous printing layer (right).

cluding the designed path information, layer information, desired height, and scale. In addition, multiple utilities were implemented to process and communicate sensor and FIM data.

The current implementation has no tools to automate the scanning process and the data flow. During the 3D printing process, the scanning has to be triggered via a manual command while the printing process has to be paused. The Sensor then stores the measured raw data in a CSV file, which can then be utilized with the Python tool introduced in the previous paragraph.

The raw sensor data should first be organized to utilize the scan data properly. For this, it is necessary to first reformat the data into arrays and filter out undesired data points. Via a `fromCSVtoCloud` function the data in the CSV file is read and converted to a list of point coordinates, essentially a point cloud (cf. fig. 7). In this format, the data can easily be filtered as described in section 3.2 by performing inside-outside tests for each point with the specified ROI.

The filtered sensor data can then be processed by the Python tool into the deviation maps by comparing the sensor data with the FIM data. All the required processing for creating deviation maps was performed by utilizing image processing tools of the *OpenCV* module for Python. Application examples of the implementation described above can be seen in the next section.

4.3 Results

To demonstrate the proposed method, a straight wall segment was printed in model scale, as illustrated in fig. 8, and scanned with a Gocator laser profiler at certain stages of the process. The resulting data was then processed with the tools described in the previous section. During the manufacturing process of the model wall, the process was paused after finishing 7 layers at a time to perform a layer scan. After completing the manufacturing process, another scan was performed to capture the front surface of the wall.

The object shown in fig. 8 was printed until the clay

cartridge was empty with the purpose of generating large deviations from the “as-designed” data. The final layer of this print provided an exciting testing ground for the geometry-capturing method. The following will show processed layer scans of the 14th and 21st layers (height and shape deviation) and a lateral scan of the completed wall (only height deviation).

Height deviation maps

First, the height deviation map of the 14th layer of the example wall is shown, then that of the 21st layer, and finally, the deviation map of the lateral scan:

In fig. 9, many details about the printing process are already visible. The blue color on the upper left indicates over-extrusion, and the red on the lower right indicates under-extrusion. However, there seems to be an axis on which the correct layer height has been printed. This observation can be explained by the fact that the UR10e (printing robot) is slightly tilted in relation to the table. This slight tilt ($< 0.5^\circ$) results from imperfections in the lab table and could be accounted for by calibrating the tool position of one robot with respect to the base coordinate frame of the other. This positioning issue was revealed by the experiments but left uncalibrated to illustrate the system’s sensitivity.

Other areas of interest are the two sections colored in deep red. One of these sections (lower part of the image) represents the area in which the printer transitions to the next layer, while the other section (upper left) shows a filament tear caused by an air bubble in the material feed. However, the threshold filter was not set accurately in the filament tear section, as this section should be blank (no information on the lower layers should be available).

In fig. 10, it is clearly visible that the layer has not been printed entirely, as large parts of the expected shape are missing. These values have been removed by the threshold filter, as they belong to the layer below. However, similar to layer 14 (cf. fig. 9), due to the tilt, not all the values that should belong to the previous layer were removed (dark red colored part in the upper left of fig. 10).

Figure 11 shows a representative part of the deviation map derived from the lateral scan. It was trimmed for visualization purposes. As in the previous figures, defects are clearly visible; Deep red spots indicate filament tears.



Figure 8. Completed clay wall.

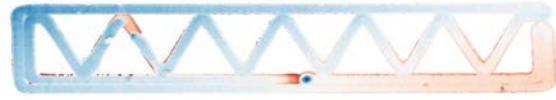


Figure 9. Height deviation map of the 14th layer during the printing process.



Figure 10. Height deviation map of the 21st layer (the last layer) shortly after the printing process.

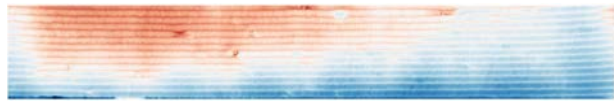


Figure 11. Trimmed deviation map of the lateral scan after completion of the printing process.

In addition, it is visible that the lower layers seem to be over-extruded. However, this excess material could be explained by the material behavior: The more material is deposited on top, the more weight is carried by the lower layers, which compresses the lower layers and pushes the material outwards.

Shape deviation maps

Same as before, for the height deviation maps, the shape deviation is illustrated for the 14th and 21st layers.



Figure 12. Shape deviation map of the 14th layer during the printing process.

Figure 13 illustrates the shape deviation of the 14th layer, derived from a scan captured during a pause in the manufacturing process. In green, the “as-printed” geometry matches with the “as-designed” model (neglecting any height deviations); in red, too much material and in blue, insufficient material was deposited. As this deviation map was derived from the height deviation map (cf. fig. 9), the issue with the unfiltered data, as mentioned before, persists, and thus the filament tear is not recognized in this deviation map.

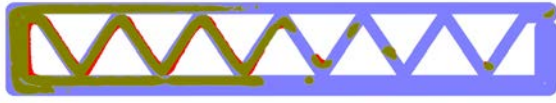


Figure 13. Shape deviation map of the 21st layer during the printing process.

Figure 13 depicts the shape deviation of the 21st layer. Similar to before, it is clearly visible that more than half of the layer is missing. But as with the shape deviation map of layer 14 (previous figure), there is the same issue with the wrongly unfiltered data that belongs to the layer below.

4.4 Discussion

A significant difficulty in data post-processing is the adjustment of the filter. Employing an overly aggressive filter risks losing crucial information, while a too-lenient filter introduces noise or retains undesired data, as observed in fig. 9. In either scenario, the accuracy of the results is compromised, incorporating inaccurate data. Given the sensor's high resolution, substantial noise is not anticipated; instead, the focus is on filtering out scanned information from undesirable positions, e.g., when performing layer scans and data from the previous layer is retained. In layer scans, the deviation map of the previous layer can be utilized to filter the captured point cloud of the current layer in order to remove points that correspond to the previous layer. In the example shown, this was not possible because not every layer was scanned.

As indicated in the previous section, the results showed that the positioning of the two robots with respect to each other did not perfectly match. However, this was not a problem in developing the proposed method; instead, it underlined its usefulness. In future experiments, when the accuracy of the "as-printed" geometry data is important, care must be taken to ensure the system is correctly calibrated.

Additionally, it's important to note that some applications of this study may not be seamlessly extrapolated to a larger scale. A notable example is the handling of layer information, where, in the current study, all layer data is encapsulated within a single scan. In the context of a larger model, the process necessitates multiple scans, subsequently requiring an alignment of various scans.

5 Conclusion

One of the selling points of AM in construction is that it can manufacture components resource-efficiently. However, as mentioned in section 1, AM methods can be unreliable and fail due to incorrect parameter settings, operating

errors, or environmental influences. To ensure error-free, high-quality, and resource-efficient 3D printing, it is almost imperative that the "as-printed" and "as-designed" states are continuously compared with each other.

As illustrated in section 4.3, even with a poorly calibrated system, a lot of helpful information can be extracted from the "as-printed" data to react early to impending problems. Although the system shown is currently not fully automated, it is still possible to recognize the possibilities this system can offer if all components are fully automated. With "as-designed" and "as-printed" data closely intertwined in the FIM data structure, data accessibility for automated quality control will not be an issue. The component's geometry, semantics, fabrication information, and captured "as-printed" data can be stored side-by-side with FIM.

Since the results are stored in textures, it is possible to store several recorded states of the corresponding component and visualize them independently so that different information can be stored and displayed in the same place. Visually embedding post-processed scan results within the FIM model provides a comprehensive and intuitive representation of the manufactured component, enhancing the interpretability of any measurements performed on it and enabling the storage of the manufacturing history of the object. Overall, this method will make it possible to develop the FIM methodology into a cyber-physical system.

Acknowledgments

The authors gratefully acknowledge the financial support of the Deutsche Forschungsgemeinschaft (DFG) in the frame of the Transregio 277 "Additive Manufacturing in Construction" (Project 414265976) and the Priority Program 2187 "Adaptive modularized constructions made in a flux" (Project 423969184).

References

- [1] Martin Slepicka, Simon Vilgertshofer, and André Borrmann. Fabrication Information Modeling: interfacing building information modeling with digital fabrication. *Construction Robotics*, 6(2):87–99, 2022. ISSN 2509-8780. doi:10.1007/s41693-022-00075-2. URL <https://doi.org/10.1007/s41693-022-00075-2>.
- [2] Orsolya Nagy, Ilona Papp, and Roland Zsolt Szabó. Construction 4.0 organisational level challenges and solutions. *Sustainability*, 13(21):12321, 2021.
- [3] André Borrmann, Markus König, Christian Koch, and Jakob Beetz. Building information modeling:

Why? what? how? In *Building information modeling – Technology Foundations and Industry Practice*, pages 1–24. Springer, 2018.

- [4] Rafael Sacks, Ioannis Brilakis, Ergo Pikas, Haiyan Sally Xie, and Mark Girolami. Construction with digital twin information systems. *Data-Centric Engineering*, 1:e14, 2020.
- [5] Yanzhou Fu, Austin Downey, Lang Yuan, Avery Pratt, and Yunusa Balogun. In situ monitoring for fused filament fabrication process: A review. *Additive Manufacturing*, 38:101749, 2021.
- [6] M. Slepicka, J. Helou, and A. Borrmann. Real-time data exchange (rtde) robot control integration for fabrication information modeling. In *Proceedings of the 40th ISARC*, Jul 2023. doi:<https://doi.org/10.22260/ISARC2023/0016>.
- [7] Yang Liu, Yu Peng, Bailing Wang, Sirui Yao, and Zihé Liu. Review on cyber-physical systems. *IEEE/CAA Journal of Automatica Sinica*, 4(1):27–40, 2017.
- [8] O. Oztoprak, M. Slepicka, A.L. Aninger, S. Kollmannsberger, E. Rank, and A. Borrmann. Linking fabrication information modeling and finite cell method for simulating the behavior of additively manufactured building components. In *EG-ICE Conference*, London, United Kingdom, 2023.
- [9] Lequn Chen, Xiling Yao, Peng Xu, Seung Moon, and Guijun Bi. Rapid surface defect identification for additive manufacturing with in-situ point cloud processing and machine learning. *Virtual and Physical Prototyping*, 16, 10 2020. doi:10.1080/17452759.2020.1832695.
- [10] LMI. Gocator line profile sensors – user manual. On-line: https://d3ejaay6gq5z4s.cloudfront.net/manuals/gocator/gocator-4.1/pdf/15159-4.1.4.12_MANUAL_User_Gocator-2300-2880-Series.pdf, Accessed: 25/12/2023.
- [11] S. Espinosa. Integration of laser profiler feedback into fim-based additive manufacturing in construction. Master’s thesis, Technische Universität München, Sep 2023.

Shape Control of Guide Frame to Avoid Obstacles for Tunnel Inspection System by Manual Operation

Fumihiko Inoue¹, Momoe Terata², Satoru Nakamura³

¹Shonan Institute of Technology, Japan

² Graduate School of Shonan Institute of Technology, Japan

³Tokyu Construction, Japan

inoue@mech.shonan-it.ac.jp, 23T1502@sit.shonan-it.ac.jp, nakamura.satoru@tokyu-cnst.co.jp

Abstract

Guide frames developed for tunnel wall inspection are designed to pass through obstacles in tunnels because the frame shape can be changed. However, in emergencies or when it is better to change the frame shape on the spot depending on the shape and arrangement of obstacles, a simple method to control the shape of the frame was required. Therefore, a shape change method using a force control technique was devised to easily change the shape of the guide frame by manual operation. The guide frame is an arch structure with a linearly connected Variable Geometry Truss (VGT) consisting of an extended actuator and a hinge. A manual control using open-loop method was devised in which the frame shape changes in response to the force acting on the puller, including the force sensor installed on the underside of the guide frame. By changing the position of the puller and the range of action of the actuator, the axial force acting on the guide frame and the shape change were confirmed through experiments and analysis. This control method is a simple way to manually deform the guide frame and can be applied to a variety of shapes.

Keywords –

Tunnel Inspection System; VGT Guide Frame; Shape control; Manual operation; Force Control

1 Introduction

A large number of the road-related infrastructure structures such as highways, bridges, and tunnels built mainly in urban areas during the high-growth period of the 1980s have exceeded their useful lives, and it is time for large-scale repair, renewal, and reconstruction. Since the tunnel is difficult to modify after completion, regular inspections have grasped the progress of deterioration, and maintenance and renewal based on future forecasts are being maintained. In particular, the collapse of highway tunnels and inspection tunnels that occurred in Japan in the past has led to an urgent



Figure 1 View of inspection system in an actual tunnel

need for inspections of aging tunnels.

With the support of the SIP program "Maintenance Management Robot," which promotes automation of inspection and maintenance of tunnel inner walls, this research developed a system to inspect and diagnose tunnel inner walls without being restricted by traffic regulations. Figure 1 shows the inspection system installed in an actual tunnel. The inspection robot moves on a variable guide frame along the inner wall, but as the inspection trolley travels through the tunnel, the frame may come into contact with obstacles in the tunnel (smoke exhaust devices, lights, traffic signs, road signs, etc.) [1]. The guide frame consists of a series of connected variable shape trusses (VGTs), and the mechanism allows the inspection robot to avoid obstacles by changing the shape of the guide frame [2].

However, the shape of the guide frame cannot be easily changed because it consists of an arch structure supported at both ends and automatically controlled by the actuator via a motor. Therefore, it became necessary to consider the shape of the guide frame and its operation method, since the shape of the frame may be easily changed by manual operation in an emergency or depending on obstacles [3].

In this study, a method of changing the shape of the guide frame using force control technology was devised with the aim of easily manipulating the shape

of the guide frame through manual operation. When the operator applies force to the puller in any direction, information from the force sensor is transmitted to each actuator to change the length and speed of the actuator. On the other hand, since the overall shape of the guide frame also changes according to the length of the actuator specified by the puller, the end position misalignment was prevented by inverse analysis of the unspecified actuator length. Ensure that the guide frame does not move. This ensures that the end positions of the guide frame remain virtually unchanged even if the shape of the guide frame changes. This result was verified based on model experiments and guide frame analysis.

2 Overview of Developed Tunnel inner Wall Inspection System

The developed tunnel inspection system is intended to replace the human visual and percussion inspection work with the application of robot technology, and is equipped with an inspection device that automatically detects cracks and floaters in the lining concrete. An overview of the tunnel inspection system is shown in Figure 2. By using elevated work platform equipped with gantry-type overall frame and inspection guide frame, work time can be reduced and labour can be saved depending on the conditions in the tunnel. In addition, inspection robot positions and inspection results can be acquired with high accuracy, and inspection reports can be efficiently prepared using

the acquired inspection data. [4], [5]

Features of this inspection system are as follows.

- (1) Flexible guide frame: The shape of the guide frame can be freely changed to suit the tunnel shape and equipment. The shape and position of obstacles can be measured in advance by 3D laser measurement [2].
- (2) Traveling protective frame: This frame runs inside the tunnel and protects third parties from falling concrete pieces during inspection.[1]
- (3) Crack inspection equipment: Capable of highly accurate detection by simultaneously acquiring images and unevenness to identify differences such as cracks and stains [6].
- (4) Percussion testing equipment: Automatically determines the location of deteriorated areas by using AI (artificial intelligence) from the sound of concrete being struck, and records the location of deteriorated areas and displays them on a screen [7], [8], [9].
- (5) Repair Expert System: Based on the inspection results, the system proposes the optimal repair method that meets the LCC (Life Circle Cost).

3 Manual Operation Experiment of Guide Frame and Control System

3.1 Page Image of Shape Change of Guide Frame due to Manual Operation

In order to establish a method for easily deforming the shape of the inspection guide frame, model experiments were conducted using manual operation

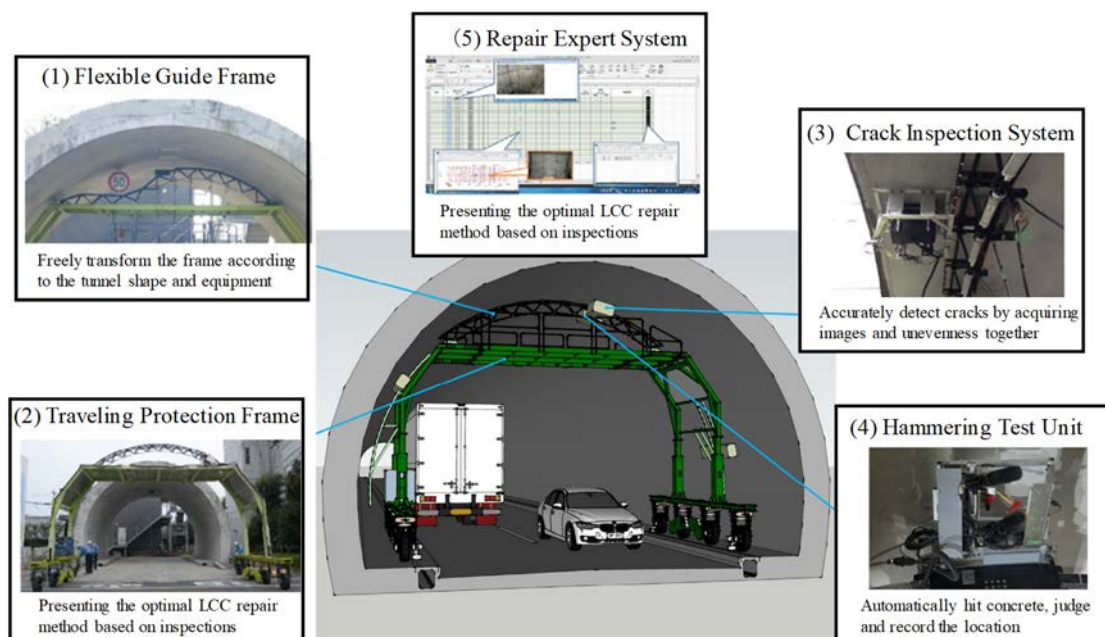


Figure 2 Overview of the tunnel inner wall inspection system and the introduced new technologies

and force control. An image of manual manipulation of the guide frame is shown in Figure 3. The guide frame consists of multiple VGTs, which are supported at both ends of the frame, so that a partial shape change affects the entire frame. The operator applies force F to the puller at the desired position, and the information from the force sensor is transmitted to each actuator to change the length and speed [10].

3.2 Flow of Force Control and Force Sensor

To verify the shape change of the guide frame by manual operation, a model of the guide frame consisting of actuators and force sensors was fabricated. Figure 4 shows the flow of manual control and the basic model of the guide frame. By operating (pushing or pulling) the manual handle installed on the upper side of the guide frame, an external force was applied to the force sensor. The external force was converted into a voltage V , and the actuator was activated when the absolute value of the voltage, $|V|$, was within the control range (minimum value $V_{Min.}$ and maximum value $V_{Max.}$). For safety reasons, the actuator was deactivated otherwise ($F = 0$ or F is greater than the set force). The actuator of the guide frame was extended or retracted according to the direction of the manual handle, and its speed u was controlled as $u = k |V|$ using the proportionality constant k . This control flow was continuously executed so that the guide frame was gradually changed by manual operation.

3.3 Force Control Experiment by Manual Operation

In order to apply manual force control to the actual guide frame, as a basic experiment, a simple force control experiment combining multiple VGTs was carried out.

Figure 5 shows a test of the operation of a guide frame with two actuators. From the initial state (Figure 5 (a)-1: left frame, (b)-1: right frame), when a pushing force was applied to the tip of the left frame, the two actuators contracted and the guide frame changed to a large curved shape. (Figure 5 (a)-2). On the other hand, when a pulling force was applied to the tip of the right frame, the two actuators extended and the guide frame assumed a large convex shape (Figure 5 (b)-2). It can be seen that as the number of actuators that change increases, the shape change of the frame also increases.

Thus, the shape varies greatly depending on the puller force and the number of actuators acting on it. It would be possible to experimentally change the

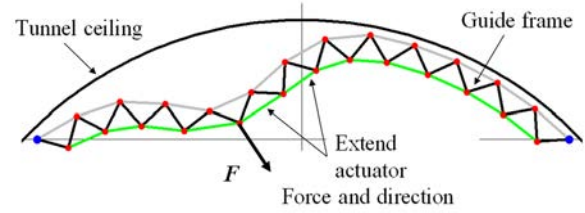


Figure 3 Image of manual operation of the guide frame by using pulling handle

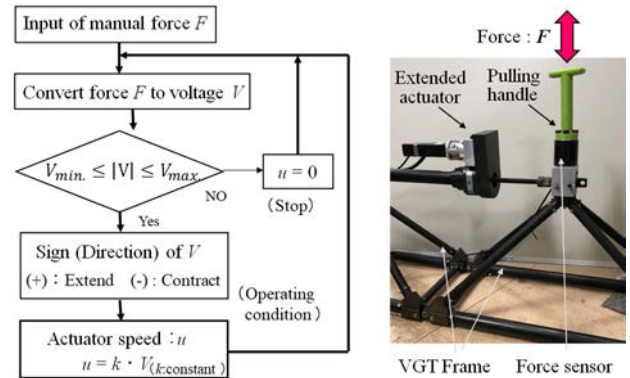


Figure 4 Flow of forced control and basic model of guide frame consisting of actuator and force sensor

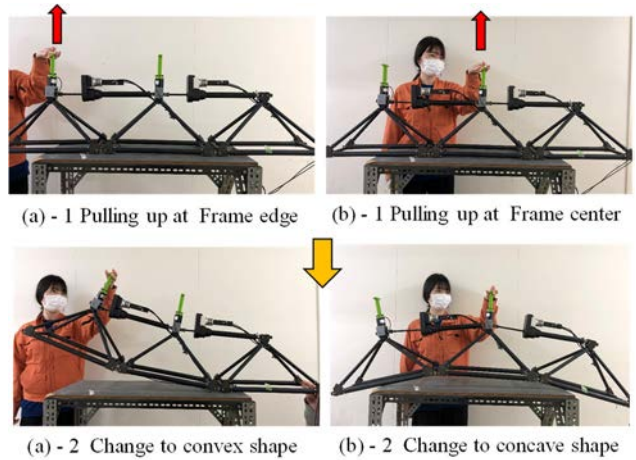


Figure 5 Shape change of two VGT actuators controlled by manual operation

guide frame to the desired shape, since the frame allows various choices of puller position and force.

In the next chapter, the shape changes of a guide frame using force control are analysed by simulation, and confirm the differences depending on the position of the puller, the number of actuators, the magnitude of the force, and the duration of the force.

4 Shape Analysis of Guide Frame Imitating Manual Operation Using Force Control Operation

Guide frames used in the actual field consist of multiple VGT mechanisms, so it is necessary to select the puller position and actuator according to the shape change. In this chapter, the guide frame changes that occur when a downward force acts on a part of the frame are analysed. The forces inside the actuator were also analysed using the finite element method.

4.1 Shape Analysis of Guide Frame

4.1.1 Basic Structure of Guide Frame

The proposed guide frame was constructed on several VGT (Variable Geometry Truss) compositions. This VGT was very simple truss structure composed of extended actuator, fixed members and hinges, as showing in Figure 6. By controlling the lengths of the extendable actuator, it is possible to create various truss shapes.

The original shape of the guide frame was an arch structure, supported at both ends by protective frames. Although the guide frame can be flexibly modified according to the VGT shape, its shape is limited by the area above the tunnel wall and the traffic space [11],[12], [13].

4.1.2 Motion Equation and Inverse Analysis of Guide Frame

As analysing the arch structure composed of VGT, the whole of guide frame was assumed to be a cantilever structure [3], [14], [15]. Analysis model of a guide frame is shown in Figure 7. The frame can replace a robot manipulator combining two fixed-length members in series. When the supported edge of frame was $q(x_0, y_0)$ at A, the top $q(x, y)$ of the x, y co-ordinates of the frame combined with n ($n \geq 2$) VGT sets was given by Eq. (1) and Eq.(2) using each hinge angle θ_j ($j=1, 2, 3, \dots, n$), and the relationship between actuator length x_j and hinge angle θ_j is shown in Eq. (3) based on the geometry of the VGT mechanism.

$$q(y, n) = l_0 \cdot \sum_{k=1}^n \sin \left\{ \sum_{j=1}^k \theta_j \right\} \quad (1)$$

$$q(x, n) = l_0 \cdot \sum_{k=1}^n \cos \left\{ \sum_{j=1}^k \theta_j \right\} \quad (2)$$

$$\theta_j = \pi - \cos^{-1} \left\{ 1 - \frac{x_j^2}{2l_0^2} \right\} \quad (3)$$

Where, l_0 was indicated as the length of fixed member of the frame. To transform the shape of arch frame, some hinge positions on the frame only had to change in proportion to target shape. However, for an intended frame shape fixed by Eq. (1) and (2), it was quite difficult to solve these equations analytically and to decide the

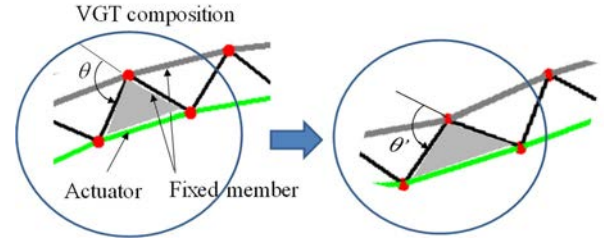


Figure 6 Basic Structure of guide frame compose of VGTs and shape changes

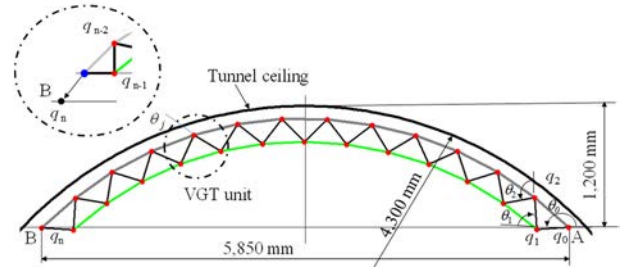


Figure 7 Analysis model of a guide frame of VGT

angle because the frame was a very highly redundant. In this case, inverse kinematics analysis was applied [3].

Considering the instantaneous change of the entire frame, the relationship between velocity $\dot{q}_n(x_n, y_n)$ and angular velocity $\dot{\theta}_n$ at each hinge position are expressed by Eq. (4).

$$\dot{q}_n(x_n, y_n) = J \cdot \dot{\theta}_n \quad (4)$$

Where, J indicates inverse Jacobean Matrix ($2 \times n$). However, J^{-1} isn't necessarily decided because J is not a regular system in $n > 2$. In this case, the suitable matrix such a pseudo inverse matrix $J^\#$ was generally induced instead of J^{-1} shown by Eq. (5), and angle velocity $\dot{\theta}_n$ of VGTs in the structure are indicated by Eq. (6) [16], [17].

$$J^\# = J^T \cdot (J \cdot J^T)^{-1} \quad (5)$$

$$\dot{\theta}_n = J^\# \cdot \dot{q}_n(x_n, y_n) \quad (6)$$

Here, when a part of the guide frame is deformed, the shape of the entire arch is also changed, so the arch tip position q_n does not match the original tip B. In order to return the tip to its original position B, the angles of the other VGTs, excluding the angle of the VTG specified by manual operation, are inversely analysed to determine the angle of each VGT such that the tip positions match. In this case, the pseudo-inverse matrix of Eq. (5) is used (i.e., the least-squares method is used) to ensure that the overall angle of movement is the smallest, which corresponds to a smaller modified error. By repeating this calculation, it is possible to change the shape of the guide frame so that it is supported at both ends and avoids obstacles.

4.2 Shape Analysis of Guide Frame Imitating Manual Operation using Force Control

4.2.1 Shape Change of Guide Frame Assuming Force Control

The displacement of the actuator when a force (tension or compression) was applied to a part of the guide frame by manual operation followed the control flow shown in Figure 4. The overall shape change of the guide frame was analysed using the inverse kinematics presented in section 4.1.

The parameters used in the analysis of the guide frame were the values of the actual model (total dead weight: approx. 650 N, vertical initial load of each VGF: 7.5 N (outer) and 26 N (inner)) and the load during manual operation ($F=50$ N). These values were relatively smaller than the load normally given by the operator. Furthermore, the force acting on each member by its own weight was calculated in advance, and then the manual load was included in the analysis, and the difference between them was taken to obtain the load for manual operation only.

4.2.2 Avoidance of Two Different Obstacles by Manual Operation

The shape changes of the guide frame avoiding two different obstacles due to manual operation were

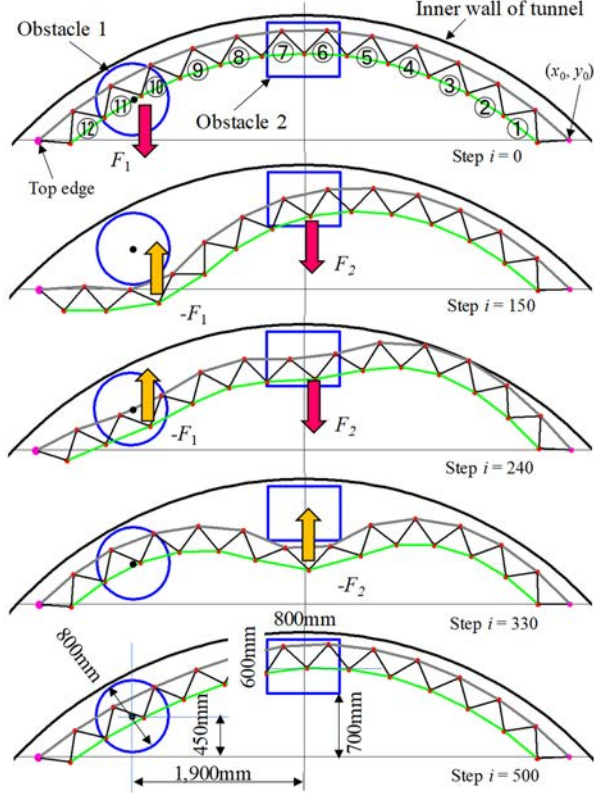


Figure 8 Shape change of the guide frame avoiding two different obstacles by manual operation

analysed. As the analytical method, first, two actuators were moved at a speed of 2 mm/s for one step in 0.1 second against the position to the puller, and the corresponding actuator lengths and angles between VGTs were calculated. The angle between the remaining VGTs for each step was then determined by inverse analysis to determine the overall shape. This procedure was repeated to determine the shape for each step, and the calculation was performed until the guide frame avoided the obstacle. Along the way, the shape of the frame was varied by changing the position of the puller.

Figure 8 shows the shape change in avoiding the obstacles. Since the obstacles were relatively small, two actuators of CASE 1 were controlled. First, to avoid the circular obstacle of diameter 800mm on the side, a pulling force F_1 was applied between ⑩, ⑪ to deform the guide frame, and avoidance was achieved in 150 time steps (15 seconds). Next, a push force $-F_1$ was applied simultaneously between ⑩ and ⑪ and a pull force F_2 between ⑥ and ⑦. In this case, the guide frame changed in a complex shape and avoidance of the central square obstacle of 800mm length *600mm width was achieved in 330 time steps (33 seconds). Finally, a pushing force $-F_2$ was applied between ⑥ and ⑦, resulting in the guide frame shape almost identical to the initial shape.

The length of each actuator and the magnitude of the

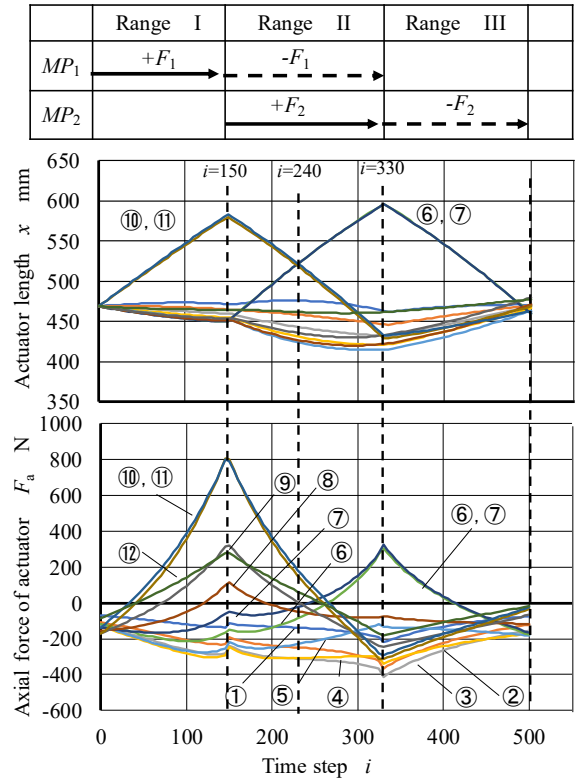


Figure 9 Relationship of axial force and actuator length to time steps when the guide frame avoiding two different obstacles

axial force is shown in Figure 9 for the change in shape of the guide frame shown in Figure 8. The control actuators are ⑩ and ⑪ in the first half and ⑥ and ⑦ in the second half, showing that they extend and retract significantly in an almost linear manner. The other actuators are the lengths controlled by inverse analysis, and the lengths of extension and contraction are small changes. The maximum axial force of about 800 N is exerted on actuators ⑩ and ⑪ in the shape that avoids the circular obstacle, and the peak axial force of 300 N is exerted on actuators ⑥ and ⑦ in the square-shaped obstacle in the center. On the other hand, the length of the actuator controlled by the inverse analysis is shortened from the initial length, with a maximum compressive force of 400 N acting on it. The allowable compressive force is about 7000 N (the same as the buckling limit force), indicating that it is within sufficient strength.

Thus, the shape of the guide frame can be changed by the position of the manual operation and the number of actuators, and by combining these cases and types, a variety of shapes can be manually controlled.

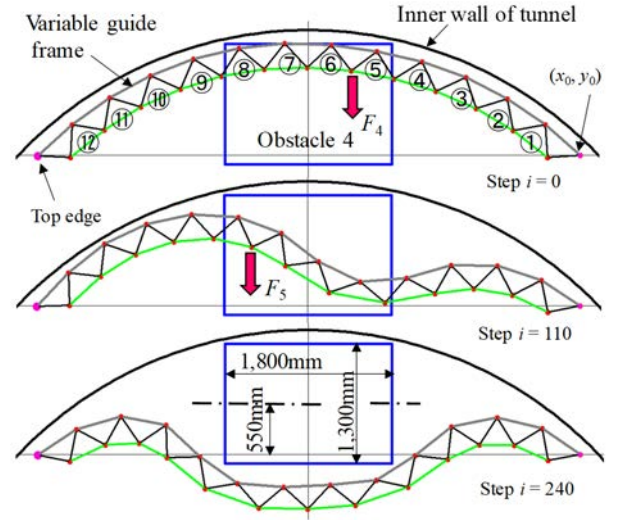
4.2.3 Avoidance of Large Square-shaped Obstacles and Multiple Obstacles by Manual Operation

To deal with the square-shaped obstacle of 800mm length * 600mm width shown in Figure 10-(a), the puller was first placed at the second position on the right side from the center and operated F_4 manually, causing the center of the guide frame to bend greatly and stop after 110 time steps (11 sec.). The manual operation was then performed by replacing the puller F_5 at the second left position from the center. As above, the center of the guide frame was greatly curved to avoid the square obstacle.

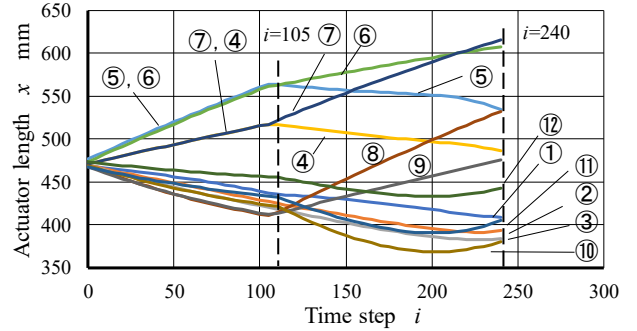
As showing in Figure 10-(b), the lengths of the controlled actuators (first half: ④, ⑤, ⑥, ⑦ Second half: ⑥, ⑦, ⑧, ⑨) increased linearly, while the lengths of the other actuators shrank by control through inverse analysis. In the second half, the length of each actuator is different because the length of the actuator in the first half is added to the length of the actuator in the second half, but the length of each actuator is balanced as a whole. Overall, the avoidance was achieved in 240 time steps (24 sec.). By pulling down around the left and right points slightly off center of the guide frame, the area around the center of the guide frame could be widened significantly. As described in 4.3.2, it was confirmed that the load was within the allowable load range of each actuator.

4.3 Determination of Final Shape of Guide Frame analyzed is by Spline Function

In order for the guide frame to avoid obstacles, it is necessary to predict the final shape of the guide frame avoiding obstacles. In this section, the shape of the guide



(a) Shape changes of guide frame to avoid obstacle 4



(b) Change of actuator length to avoid obstacle 4

Figure 10 Analysis of Shape changes and actuator of guide frame to avoid obstacle 4 by manual operation

frame was analyzed using a Spline interpolation method and in consideration of this method's effectiveness [2].

4.3.1 Shape Change of Guide Frame Assuming Force Control

To the initial guide frame shape, the overall shapes required to allow avoiding obstacles were mathematically combined using the spline interpolation function. Spline interpolation is a method of combining arbitrary shapes with polynomials to form a continuous shape. Assuming that a function interpolates the section (x_j, x_{j+1}) , the piecewise polynomial $S_j(x)$ is expressed through the Eq. (7).

$$S_k(x) = a_k(x - x_k)^3 + b_k(x - x_k)^2 + c_k(x - x_k) + d_k \quad (k = 1, 2, 3, \dots, n-1) \quad (7)$$

In order for this cubic equation to be a smooth curve, it is assumed that the value of the first derivative of and $S_k(x)$ for x_j is equal, and that the value of the

second derivative of $S_k(x)$ and $S_{k-1}(x)$ for x_j is also equal. Further, the value of the second derivative at the start point x_0 and the end point x_n are assumed to be 0. By applying the above conditions to each equation for several interpolation points, the each coefficient of function a_k, b_k, c_k, d_k were calculated, and the function by the spline interpolation was determined. Here, each coefficient was determined as follows;

$$a_k = \frac{\ddot{S}_k(x_{k+1}) - \ddot{S}_k(x_k)}{6(x_{k+1} - x_k)} \quad (8)$$

$$b_k = S_k(x_k)/2 \quad (9)$$

$$c_k = \frac{y_{k+1} - y_k}{(x_{k+1} - x_k)} - \frac{(x_{k+1} - x_k)\{2\ddot{S}_k(x_k) + \ddot{S}_k(x_{k+1})\}}{6} \quad (10)$$

$$d_k = y_k \quad (11)$$

4.3.2 Estimation of Guide Frame Shape Avoiding Obstacles

Using the frame analysis in 4.3.1, the guide frame shape estimation method for avoiding obstacles is shown below.

(1) Several interpolation points (including both end points) that avoid obstacles are selected for the shape of the guide frame, and the constants of the piecewise function of Eq. (7) were calculated.

(2) The curves of all piecewise functions are connected and the upper chord shape of the guide frame checked. At this time, the upper chord length of the transformed guide frame is different from the original upper chord length.

(3) Next, VGT elements are arranged continuously from the starting point according to the upper chord shape of the guide frame. However, the top of the guide frame is not located at the original fixed end.

(4) Therefore, the angle of each VGT is calculated by inverse analysis so that the tip matches the original fixed end, and the final shape is estimated. As described above, the change in the angle of each VGT is small, so there is little effect on the final shape.

By constructing an easy system that allowed the selection of representative points by touching the tablet screen, we were able to visually show the positions to avoid obstacles. Several examples are shown below

4.3.3 Estimation of Guide Frame Shape Avoiding Side Obstacles

The shape of a circular obstacle suspended from the left side of the tunnel ceiling overlapped the guide frame as shown in Figure 11-(a). We selected six optimum interpolation points to avoid obstacles and analyzed the

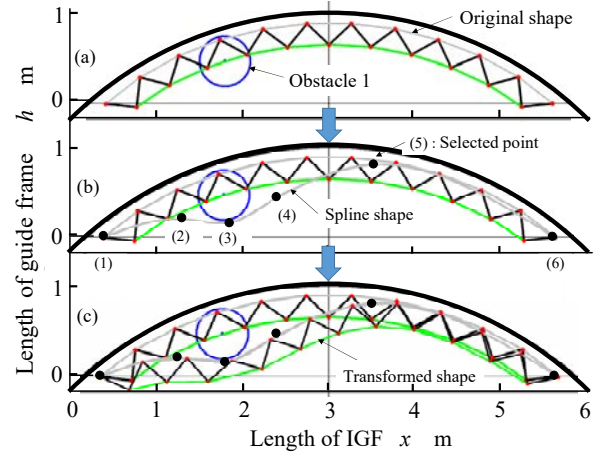


Figure 11 Shape decision of spline function for obstacle

Table 1 Each constant value of spline function for obstacle 1

$S_k(x)$	a_k	b_k	c_k	d_k
1	-0.022	0	0.47	0
2	0.086	-0.593	-0.071	0.264
3	-0.062	0.821	0.353	0.189
4	-0.003	-0.201	0.568	0.528
5	0.005	-0.293	-0.011	0.876



Figure 12 Shape control of guide frame avoiding obstacles in model tunnel

shape of the upper chord using spline interpolation. As shown in Figure 11-(b), the obtained spline curve is a smooth shape that continuously connects five regions sandwiching selected points, and is connected at the start and end points while avoiding obstacles. The shape of the guide frame was suitable for fitting because there is no discontinuity or large curvature. Each constant value analyzed by Eq. (7) is shown in Table 1.

The guide frame was reproduced by configuring the guide frame to conform to the shape of the obtained upper chord member as shown in Figure 11-(c). Although the total length of the guide frame of the new shape is slightly different from the original shape, the shape is optimized by proportional distribution. Since there is almost no significant change in shape, the obstacle can be sufficiently evaded.

Based on the above analysis results, the shape of the actual guide frame was observed in Figure 11. It was verified experimentally that the shape of the guide frame analyzed by spline interpolation.

5 Conclusion

The guide frame developed for tunnel wall inspection is a structure that is automatically controlled by a motorized telescopic actuator, so its shape cannot be easily changed. In emergencies or when there are obstacles in the tunnel, it may be easier to change the shape by human operation. We studied a method of changing the shape of the guide frame with the aim of easily changing the shape of the guide frame through multiple human operations.

The following is a description of the implementation and results.

- (1) A basic model was fabricated by attaching a puller and a force sensor to two pairs of VGTs that make up the guide frame. From
- (2) In order to manually manipulate the shape of the guide frame, the guide frame was replaced with a robot manipulator consisting of a VGT, and kinematics analysis and FEM structural analysis of an arch structure with two supported points were conducted.
- (3) As the example of typical manual operation, the direction and time of force applied to the obstacles and the position of operation were shown, and shape analysis was conducted to avoid the two obstacles.
- (4) Analysis results of manual operation in the case of a large obstacle in a tunnel or multiple obstacles in the same plane were shown, and it was confirmed that the shape of the guide frame can be adequately controlled to avoid the obstacles.
- (5) In order for the guide frame to avoid obstacles, it is possible to predict the final shape of the guide frame avoiding obstacles using Spline Function.

In the future, we would like to confirm the effectiveness of force control by manual operation using actual guide frames and aim for its practical application.

Acknowledgements

Finally, the author thanks all who supported the development of shape control of guide frame to avoid the obstacles by manual operation.

References:

- [1] S. Nakamura, A. Yamashita, F. Inoue, et al, "Inspection Test of a Tunnel with an Inspection Vehicle for Tunnel Lining concrete", *J. of Robotics and Mechatronics*, Vol.31, No.6, pages 762-771, 2019.
- [2] F. Inoue, S. Kwon, S. Nakamura, "Shape Adaptation of the Inspection Guide Frame in Tunnels to Avoid Obstacles Detected by Laser Range Finder", *J. of Robotics and Mechatronics*, Vol.31, No.6, pages 752-761, 2019.
- [3] Inoue F, "A Study on Adaptive Structure Applying Variable Geometry Truss (Mechanism of Movable Arch Roof with External Panel)", *J. of Robotics and Mechatronics*, Vol. 21, No.2, pages 172-178, 2009.
- [4] Fujii H, Yamashita A and Asama H, "Defect Detection with Estimation of Material Condition Using Ensemble Learning for Hammering Test", *Proc. of the 2016 IEEE International Conference on Robotics and Automation*, pages 847-854, Stockholm (Sweden), 2016.
- [5] F. Inoue and S. Kwon et al, "Study on Shape Determination Method and Operation Experiment of Inspection Guide Frame to Avoid Obstacle in Tunnel", *Proc. of the IASS Annual Symposium 2019*, pages 1-8, Barcelona, Spain, 2019.
- [6] D. Inoue, T. Ueno and S. Nakamura, "Robotic Inspection Tests of Tunnel Lining Concrete with Crack Lihrt-section Device on Variable Guide Frame", *Proc. of 35th Int. Symp. On Automation and Robotics in Construction*, pages 524-529, 2018.
- [7] J. Im, H. Fujii, A. Yamashita and H. Asama, "Multi-modal Diagnostic Method for Detection of Concrete Crack Direction Using Light-section Method and Hammering Test", *Proc. of 14th Int. Conf. on Ubiquitous Robots and Ambient Intelligence (URAI)*, pages 922-927, 2017.
- [8] H. Fujii, A. Yamashita and H. Asama, "Improvement of Environment Adaptivity of Defect Detector for Hammering Test Using Boosting Algorithm", *IEEE/RSJ Int. Conf. on Intelligent Robots and System*, ThFT14, 2015.
- [9] J. Kasahara, H. Fujii, A. Yamashita and H. Asama, "Unsupervised learning approach to automation of hammering test using topological information", *ROBOMECH Journal*, Vol. 4, No.13, pages 1-10, 2017.
- [10] F. Inoue, M. Terata, et al, "Manual Deployment and Shape Determining Method of Guide Frame for Tunnel Inspection by Force Control", *Proc. of the IASS Annual Symposium 2023*, pages 1-10, Melbourne, Australia, 2023.
- [11] Inoue F, "Shape Control of Variable Guide Frame for Tunnel Wall Inspection", *Proc. of 34th ISARC 2017*, pages 675-682, Taipei, Taiwan, 2017.
- [12] Inoue F, Sasaki T, Huang X and Hashimoto H, "A Study on Position Measurement System Using Laser Range Finder and Its Application for Construction Work", *Journal of Robotics and Mechatronics*, Vol. 26, No.1, pages 226-234, 2012.
- [13] Tamura H, Sasaki T, Hashimoto H and Inoue F, "Circle Fitting Based Position Measurement System Using Laser Range Finder in Construction Fields", *Proc. of IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems*, pages 209-214, 2010.
- [14] Koryo Miura, "Variable Geometry Truss Concept (Part I: Design and Operation of a Deployable Truss Structure)", *The Institute of Space and Astronautical Science Report*, No.614, 1984.
- [15] Koryo Miura, "Variable Geometry Truss Concept (Part II: Review on Special Application of Variable Geometry Truss)", *The Institute of Space and Astronautical Science Report*, No.614, 1984.
- [16] Chirikjian, "Binary paradigm for robotic manipulators", *ICRA*, pages 3063-3069, 1994.
- [17] Chirikjian, "Kinematic synthesis of mechanisms and robotic manipulators with binary actuators", *ASME DE-72*, pages 161-167, 1994.

Component pose reconstruction using a single robotic total station for panelized building envelopes

Mengjia Tang¹, Nolan W. Hayes¹, Bryan P. Maldonado¹, and Diana Hun¹

¹Buildings and Transportation Science Division, Oak Ridge National Laboratory, United States of America *

tangm@ornl.gov, hayesnw@ornl.gov, maldonadopbp@ornl.gov, hunde@ornl.gov

Abstract -

The construction industry can benefit greatly from increased automation in construction processes in terms of precision, time, and labor costs. This is especially true for prefabricated construction for either new buildings or envelope retrofits. Here, prefabricated components are manufactured offsite according to design specifications and installed on-site. Accurate information on the component's position and orientation (pose) is needed to achieve this. A tool named "Real-Time Evaluator" (RTE), designed to autonomously track prefabricated components as they are being installed and provide real-time installation instructions, is currently under development using a single robotic total station. This is challenging since at least three points are needed to determine the pose of an object in space. To achieve this goal with a single robotic total station, two key algorithms were developed: (1) the "resection" algorithm aligns the digital twin with the physical twin regardless of the location of the total station, and (2) the "transformation" algorithm gives instructions of translations and rotations to installers to achieve the desired installation pose. The algorithms were evaluated experimentally on a lab-scale demonstration. Results show that the resection algorithm achieved an average error of < 3.1 mm, while the transformation algorithm predicted the rotation angle along a single axis with an error of < 0.5°.

Keywords -

Prefabricated construction; Automation; Resection; Rotation; Accuracy

1 Introduction

Building construction requires a high degree of accuracy between as-designed and as-built structures, which is challenging given the long project duration, heavy reliance on workers' skills and experience, and complicated onsite

environments (e.g., crane operators not being able to see the load all the time). For prefabricated construction, not meeting the required tolerances during the installation of prefabricated components often requires expensive corrections. Therefore, a tool is needed to actively check the quality of construction during the installation of prefabricated components for real-time installation instructions.

Rapid advances in surveying technologies present a great potential for automating construction processes. Robotic total stations can autonomously turn their telescopes and aim at targets by following commands from software, enabling fast and precise measurement and tracking with little user input. Three-dimensional (3D) laser scanners can produce point clouds to provide reality capture, including as-built information during or after construction [1]. While commercially available, these instruments need a streamlined procedure to interact with (1) data in the digital space, (2) the building structure and prefabricated components in the physical space, and (3) the users or construction workers.

The Real-Time Evaluator (RTE) is a tool that uses a single robotic total station to gather the data needed to compare the actual position of a prefabricated component against the as-designed position in a digital twin during installation in order to provide real-time feedback to installers [2]. Two major tasks were identified during the development of the RTE. The first task is to align the 3D model of a building (i.e., digital twin) with the physical twin. When the total station is set up to obtain an initial scan of the building, track the installation process of prefabricated components, or obtain an as-built scan of the building, it will most likely have different positions and orientations in each session. Moreover, the digital twin may be obtained from a 3D laser scanner, building information modeling (BIM) software, or computer-aided design (CAD) software foreign to the total station [3]. Thus, discrepancies can exist between the coordinate system of the digital twin and the coordinate system of the total station in its current setup, which prevents the RTE from giving correct real-time installation instructions. A resection function is available in software such as Trimble® Access™ and Leica® Captivate™ to determine the coordinates of the total station based on the Helmert trans-

*Notice: This manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<https://www.energy.gov/doe-public-access-plan>).

formation [4]. However, the available third-party software does not use the digital twin as input to adjust the coordinate system of the total station accordingly. Therefore, it is necessary to develop a resection algorithm that uses the digital twin in an arbitrary coordinate system as input to achieve accurate RTE installation commands.

The second task involves generating optimal installation instructions to move a prefabricated component from its current pose to the intended (goal) pose. While prior research has focused on pose estimation of the crane and load [5] or deviations of the current pose from the as-designed pose [6], less attention has been directed to reconstructing a component pose by tracking targets attached to it. The installation time of prefabricated components could be notably reduced by explicitly specifying the steps to move them directly to the goal pose, eliminating the need for trial-and-error placement. Such steps need to consider six degrees of freedom in the movement of a rigid body, i.e., the translations along the three Cartesian coordinate axes and the rotations around such axes. The trajectory is not unique and needs to be tested and optimized based on various factors including the initial and final poses of the component, the center of rotation, rotation axes, and the connectors on the building structure.

While the conversion between coordinate systems and the transformation of rigid bodies are not new mathematical problems, they are new tasks for automating the prefabricated construction because the procedures, requirements, and available tools need to be considered as constraints. To address the above two tasks, this paper develops a resection algorithm to align the digital twin with the physical building within measurement error. This paper also develops a transformation algorithm that applies a rigid-body transformation to a component in order to achieve a goal pose with a series of translations and rotations around three axes. Accuracy of the algorithms are obtained from lab-scale experiments.

2 Methodology

This section first outlines the algorithms for resection and rigid-body transformation. Subsequently, it describes the lab-scale experimental setup used to demonstrate and evaluate these two algorithms.

The two algorithms, in essence, solve the same problem. The resection algorithm computes the rotation and translation of points from one coordinate system to another. The transformation algorithm computes the rotation and translation of points when they move in the same coordinate system. In both algorithms, the method used to calculate the rotation matrix between two sets of at least three points is the Kabsch or the Kabsch-Umeyama algorithm [7]. It uses the singular value decomposition (SVD) to find the rotation matrix that best aligns the two paired sets of points.

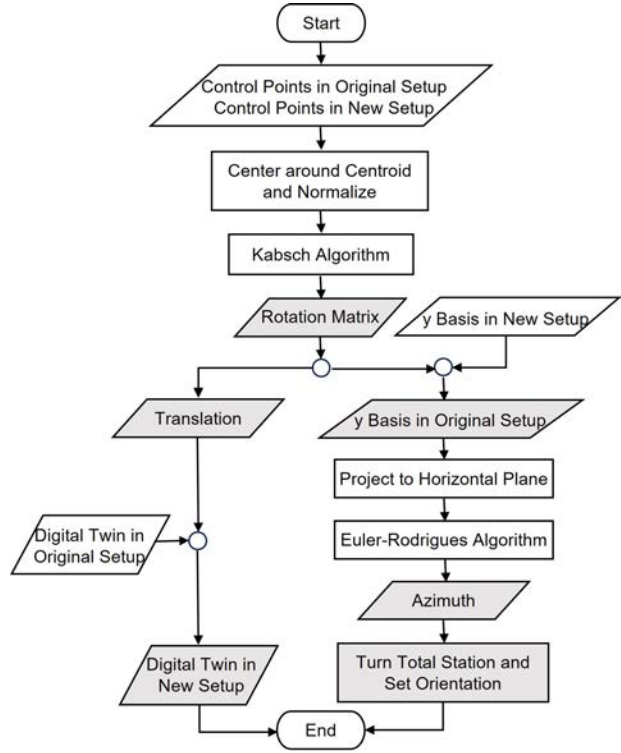


Figure 1. Workflow chart for Algorithm 1: resection.

It was implemented through the “align_vectors” function in the SciPy library in Python [8]. When the number of points is two, the Euler–Rodrigues formula was used to align the vector formed by the two points [9] and was fulfilled in this paper with a user-defined function.

2.1 Resection

Let CP_{old} be the set of three or more control points measured in the original (old) coordinate system. Let CP_{new} be the corresponding set of control points measured in the new coordinate system. Finally, let DT_{old} be the set of all the points in the digital twin. The algorithm that adjusts DT_{old} and total station orientation to align the digital twin with the physical twin using CP_{old} and CP_{new} is described in Algorithm 1 and shown in Figure 1. It has the following four steps.

First, the translation vector t and rotation matrix R are calculated such that, for any point in the new coordinate system P_{new} , the corresponding point in the original coordinate system is calculated as $P_{old} = t + RP_{new}$. This determination is based on the measurements of control points by the total station at both the original and new setups. The measured coordinates of the control points are re-centered around their centroid in the original (G_{old}) and new (G_{new}) coordinate system, respectively. The re-centered control

Algorithm 1: resection: align digital twin with physical twin**Data:** $CP_{old}, CP_{new}, DT_{old}$ **Result:** $t, Az, DT_{adjusted}$ **// Calculate rotation matrix and translation vector** $G_{old} \leftarrow \text{centroid}(CP_{old})$ $G_{new} \leftarrow \text{centroid}(CP_{new})$ $cp_{old} \leftarrow \text{normalize}(CP_{old} - G_{old})$ $cp_{new} \leftarrow \text{normalize}(CP_{new} - G_{new})$ $R \leftarrow \text{Kabsch}(\text{align } cp_{new} \text{ to } cp_{old})$ $t \leftarrow G_{old} - RG_{new}$ **// Calculate azimuth** $j_{old} \leftarrow Rj_{new}$ $\text{proj}_{xy}(j_{old}) \leftarrow \text{project_xy}(j_{old})$ $R_{xy} \leftarrow \text{Euler-Rodrigues}(\text{align } j_{new} \text{ to } \text{proj}_{xy}(j_{old}))$ $Az \leftarrow z \text{ component of factor}(R_{xy}) \text{ in degrees}$ **if** $Az < 0$ **then**| $Az \leftarrow 360 + Az$ **end****// Total station commands** $\text{total_station.turn_to_angle}(Az \times \frac{\pi}{180})$ $\text{total_station.set_orientation}()$ **// Adjust digital twin** $DT_{adjusted} = DT_{old} - t$ **return** $DT_{adjusted}$

points are normalized to unit vectors (cp_{old} and cp_{new}). Such vectors are then used as inputs to the Kabsch algorithm to obtain the rotation matrix R . The translation vector is calculated as the difference between the centroid in the original coordinate system and the one measured in the new coordinate system and rotated by R as follows: $t = G_{old} - RG_{new}$.

Then, the rotation matrix is used to calculate the azimuthal angle (Az) for the total station to turn to the same north (y-axis) as in the original coordinate system. Note that only the horizontal angle is needed since the total station is set to level before each use, which results in the x- and y-axes always forming a horizontal plane within machine tolerance. Let $j_{new} = [0 \ 1 \ 0]$ be the y basis vector with respect to the new coordinate system. Then, the y basis vector of the original coordinate system is obtained as follows: $j_{old} = Rj_{new}$. The resulting vector is then projected to the horizontal xy plane: $\text{proj}_{xy}(j_{old})$, which is used to compute the rotation matrix of y-axis on the xy plane using the Euler–Rodrigues formula. The re-

sulting rotation matrix (R_{xy}) is factorized into a sequence of Euler angles around the y-, x-, and z-axes using the user-defined function [10], which always leads to zero angles around the y- and x-axes.

The angle around the z-axis (Az) is sent to the total station to command it to turn to Az and reset north. Note that this angle is the counterclockwise rotation of the original coordinate system to the new one, so the new frame should be turned by the same angle clockwise to align with the original north.

Finally, the original digital twin points are adjusted by the translation. Now the points in the digital space ($DT_{adjusted}$) have the same coordinates as measured by the total station in its current position and orientation.

2.2 Rigid-body transformation

The movement of solid prefabricated components is considered a rigid-body transformation. By rigidly attaching prisms as reflectors at a known distance from the external corners of the component, the pose of the component can be determined with a total station by measuring the location of at least three prisms. Let $P_{curr}^n(t)$ be the set of current coordinates of n prisms at time t and P_{goal}^n be the set of corresponding prisms at the goal pose, for $n \geq 3$. An automated workflow was developed to achieve the targeted component pose by transforming $P_{curr}^n(t)$ into P_{goal}^n , as described by Algorithm 2 and shown in Figure 2.

The center of rotation is important for determining the correct translation. The center of rotation can be any point that moves in the same way as the component, e.g., the center, one corner, or the bearing point of the component. In the local coordinate system of the component, and in the case where there are $n = 3$ prisms, the basis $\mathcal{B}(P^n)$ can be created using the three prisms, i.e., two vectors formed by three prisms and one vector normal to the plane formed by the three prisms. The center of rotation is then calculated as a linear combination of the three basis vectors as follows: $G = \mathcal{B}(P^n)g$. By doing this, the coefficients g of the linear combination in the component coordinate system are independent of the component pose and the global coordinate system set by the total station, in other words, g is constant and independent of time. The coefficients can be determined by referring to the component design drawing or conducting an initial measurement of the three prisms and the center of rotation using the total station. Then the coefficients g can be used to compute the coordinates of the center of rotation using $P_{curr}^n(t)$ and P_{goal}^n , which is the first step in Algorithm 2.

Next, the rotation matrix $R(t)$ is computed with the Kabsch algorithm after the measured prisms at the current and goal poses are re-centered around their respective center of rotation and then normalized. The translation vector

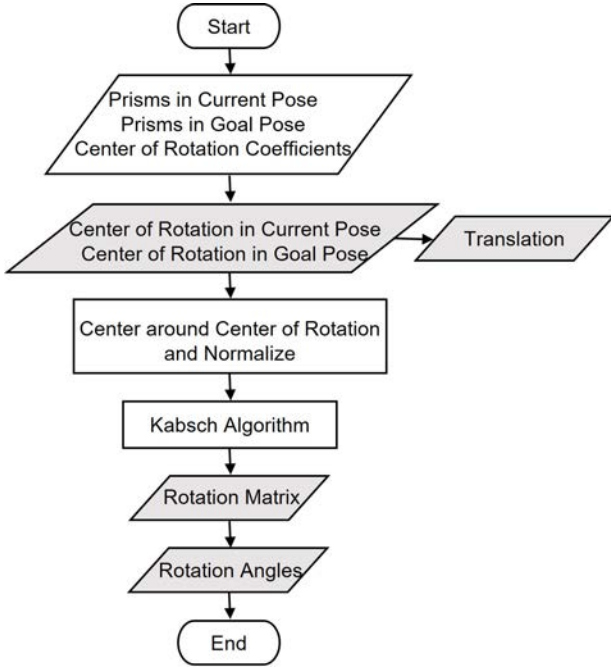


Figure 2. Workflow chart for Algorithm 2: transformation.

$T(t)$ is calculated as the difference between the centers of rotation of the goal and current poses.

Finally, the Euler angles around the x , y , and z - axes in a Cartesian coordinate system set by the total station are computed from the rotation matrix $R(t)$. The order of the rotation affects the Euler angles, so it needs to be specified. In Algorithm 2, the rotation matrix is factorized into three Euler angles in the order of y -axis, x -axis, and z -axis, similar to Algorithm 1 [10]. Therefore, the rotation is expressed as a composition of sequential extrinsic rotations about axes y , x , and z with angles $\alpha(t)$, $\beta(t)$, and $\gamma(t)$, respectively.

2.3 Lab-scale experiments

The algorithms for resection (Algorithm 1) and transformation (Algorithm 2) were tested using the Leica Nova MS60 Multi-Station on a lab-scale setup. The shortest measuring distance between the multi-station and reflector target is 1.5 m, while the longest distance varies from 200 m to 10000 m depending on the reflector type and atmospheric conditions. The measuring accuracy can be decreased by haze, sunlight, heat shimmer, beam interruptions, and moving objects within the beam path. The multi-station should be set up in such a way that the beam can reach the reflective part of the target.

The algorithms were written in Python and interfaced with the multi-station using a Bluetooth connection and

Algorithm 2: transformation: convert current pose to goal pose

Data:

$P_{curr}^n(t)$, P_{goal}^n , g

Result:

$T(t)$, $\alpha(t)$, $\beta(t)$, $\gamma(t)$

// Calculate the center of rotation at current and goal poses

$G_{curr}(t) \leftarrow \mathcal{B}(P_{curr}^n(t))g$

$G_{goal} \leftarrow \mathcal{B}(P_{goal}^n)g$

// Calculate rotation matrix and translation vector

$p_{curr}(t) \leftarrow \text{normalize}(P_{curr}^n(t) - G_{curr}(t))$

$p_{goal} \leftarrow \text{normalize}(P_{goal}^n - G_{goal})$

$R(t) \leftarrow \text{Kabsch}(\text{align } p_{curr}(t) \text{ to } p_{goal})$

$T(t) \leftarrow G_{goal} - G_{curr}(t)$

// Find rotation angles

$\alpha(t), \beta(t), \gamma(t) \leftarrow \text{factor}(R(t))$

return $T(t)$, $\alpha(t)$, $\beta(t)$, $\gamma(t)$

Leica GeoCOM commands. While the multi-station was used instead of a robotic total station, the algorithms should be able to be applied on robotic total stations as long as the communication between the software and total station can be established because the scanning capability of a multi-station is not needed for executing the algorithms. The following two sections describe the test procedures and evaluation metrics for the resection and transformation algorithms, respectively.

2.3.1 Experiment on resection

The resection algorithm (Algorithm 1) was tested using a rectangular mock-up wall (1.8 m wide by 3.1 m high), a pole with one Leica 360° full-size prism (Model GRZ122, 2.0 mm pointing accuracy) installed on the top, along with the multi-station (Figure 3). The mock-up wall was kept stationary during all tests to simulate a physical building. Four Leica reflective tapes (Model GZM31) were placed close to the four corners of the mock-up wall as control points (CP). Two positions were used to set up the multi-station: Position A was at a distance of ~8 m away from the mock-up wall on the right, and Position B was at a distance of ~5 m from the mock-up wall on the left. The pole with the prism was used to determine the exact translation vector (t) when the multi-station was moved from one position to the other and was set to a 1.5 m target height for all tests.

A total of four tests were conducted. One baseline test was performed without moving the multi-station to establish the baseline error of the resection algorithm due

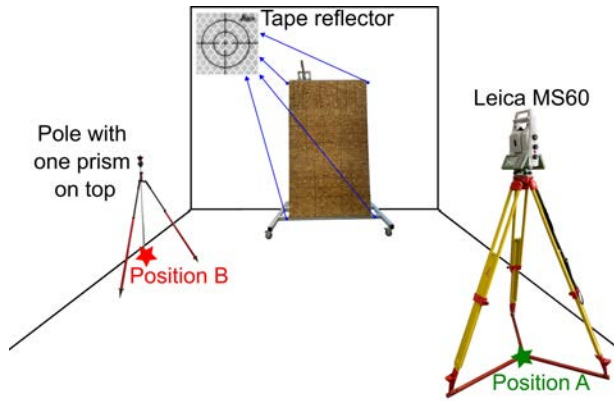


Figure 3. Experimental setup for testing resection with Algorithm 1.

to instrument accuracy. In the remaining three tests, the multi-station was moved from Position A to Position B or vice versa. During each test, the following steps were performed. For ease of explanation, the initial setup in Figure 3 is used as an example.

1. Set up the multi-station at one position (e.g., Position A in Figure 3).
2. Set up the pole at the other position (e.g., Position B in Figure 3).
3. Measure the coordinates of four reflective tapes to obtain the control points (CP_{old}).
4. Measure the coordinates of the prism on the pole.
5. Move the multi-station to Position B (except in the baseline test).
6. Reset the orientation of the multi-station randomly (except in the baseline test).
7. Measure the height of the multi-station.
8. Remeasure the coordinates of the four reflective tapes to obtain the control points in the new coordinate system (CP_{new}).
9. Run Algorithm 1 to determine translation t and change in orientation Δz . Adjust the data stored in the original coordinate system to obtain $CP_{adjusted}$ and reset the orientation of the multi-station to Δz .
10. Remeasure the coordinates of the four reflective tapes ($CP_{measured}$).

Two different measurement modes of the multi-station — “Auto” and “Manual” — were used in the tests to measure the coordinates of the reflective tapes. In the “Auto” mode, the laser of the multi-station was pointed close to the reflective tapes, then the multi-station searched for it and measured automatically. In the “Manual” mode, a human manually aimed the laser at the target and commanded the multi-station to measure and record the coordinates of the target. By varying the measurement mode, the error for

each measurement mode can be determined and compared.

The true translation of the multi-station was determined in Steps 4 and 7 above. By measuring the prism on the pole, the translations in the x - and y -axes (i.e., the two axes in the horizontal plane) relative to the original setup of the multi-station were directly obtained. To calculate the true translation in the z -axis (i.e., the vertical axis), the z -coordinate of the prism on the pole measured in Step 4 was first subtracted from the height of the prism (1.5 m), which was then subtracted from the height of the multi-station measured in Step 7 after the multi-station was moved.

To assess the accuracy of the resection algorithm, two metrics were used. The first one is the absolute difference between the true and calculated translation along three axes. The second one is the Euclidean distance between $CP_{measured}$ in Step 10 and $CP_{adjusted}$ in Step 9 for each control point on the mock-up wall, which reflects the difference between the true position of control points relative to the new setup of the multi-station and the calculated position based on the resection algorithm.

2.3.2 Experiment on rotation

Experiments were conducted to test the accuracy of the algorithm for rigid-body transformation (Algorithm 2) in computing the rotation angle around a single axis. These simplified experiments isolated the rotation around one axis from translation and rotation around the other two axes, which serve as a good starting point for evaluating the algorithm.

The experimental setup is shown in Figure 4. A 0.76 m high by 0.76 m wide prefabricated wall panel with a mass of 14 kg was installed on a heavy-duty desk monitor arm using a VESA mount. The monitor arm allows for rotating the panel around multiple joints within a range of at least 75° . Three Leica 360° mini prisms (Model GRZ101, 1.5 mm pointing accuracy) were installed on the top right, top left, and bottom left corners of the panel (Figure 4a). One reflective tape was placed at the center of the panel’s front face. The table that the monitor arm was attached to was level and provided a flat surface to place one leg of a digital protractor (Bosch GAM 220 MF, $\pm 0.1^\circ$ accuracy) to measure angles (Figure 4).

In each test, the table remained stationary, and the panel was only rotated around the joint on the VESA mount at the back face of the panel, which formed a rotational movement around the axis (hereafter referred to as *rotation axis_{arm}*) that was perpendicular to the panel face and went through the center of the panel’s front face as marked by the reflective tape. The north or y -axis of the multi-station needs to be parallel to *rotation axis_{arm}* to make sure that the angle around the y -axis obtained from Algorithm 2 is around the same axis as the actual rotation

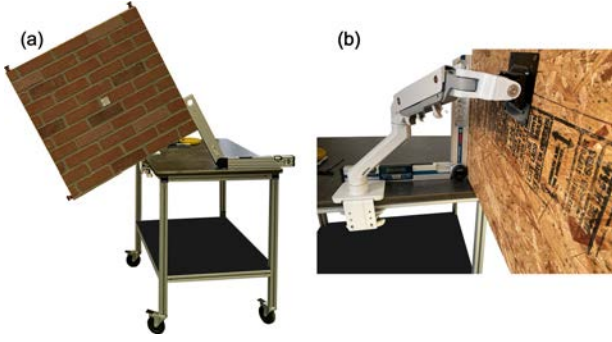


Figure 4. Experimental setup for testing rotation around a single axis with Algorithm 2.

to compare angles. To achieve this, the panel was set to plumb before all tests (a 90° between the table and the panel's back face shown in Figure 4b), and the north of the multi-station was set to perpendicular to the plane formed by the three prisms on the panel at the beginning of each test. The rotation angle around $rotation\ axis_{arm}$ was determined by measuring the angle between the edge on the right side of the panel and the table before and after the rotation with the digital protractor and then calculating the difference in the two angles (Figure 4a).

The center of rotation is essential in Algorithm 2 to determine translation accurately. While the true center of rotation was the monitor arm's joint at the center of the panel's back face, the reflective tape was used as the center of rotation when running Algorithm 2 because the center of rotation affects translation but not rotation angles. The tape and prisms were measured before all tests to determine the coefficients \mathbf{g} of the linear combination that can express the coordinates of the panel center using the coordinates of the three prisms. These coefficients were used to calculate the panel center. A total of seven tests were performed following the following procedure:

1. Set the panel at a position as the goal pose.
2. Measure the angle between the right edge of the panel and the table with the protractor (α_{goal}).
3. Measure the coordinates of the three prisms with the multi-station.
4. Set the north of the multi-station perpendicular to the plane formed by three prisms.
5. Measure the coordinates of the three prisms (P_{goal}^n) and the tape ($\mathbf{G}_{goal,measured}$) with the multi-station.
6. Rotate the panel around the $rotation\ axis_{arm}$ to the current pose.
7. Measure the angle between the right edge of the panel and the table with the protractor (α_{curr}).
8. Measure the coordinates of the three prisms ($P_{curr}^n(t)$) and the tape ($\mathbf{G}_{curr,measured}$) with the multi-station.
9. Run Algorithm 2 to obtain the angles between the

current panel pose and the goal pose.

The accuracy of the algorithm was evaluated by the difference between the rotation angles that were measured with the protractor and calculated with the algorithm. The measured rotation angle was obtained by subtracting α_{goal} in Step 2 from α_{curr} in Step 7. Additionally, the translation from the current panel position to the goal position was compared between the algorithm output and measurement. The measured translation was obtained by subtracting $\mathbf{G}_{curr,measured}$ in Step 8 from $\mathbf{G}_{goal,measured}$ in Step 5.

3 Results

This section presents the performance of the algorithm on resection (Algorithm 1) first, followed by the performance of the algorithm on rigid-body transformation (Algorithm 2).

3.1 Experiment on resection

Four tests where the multi-station measured four control points fixed on a mock-up wall were conducted, and the translation and the rotation angle of the north ("azimuth") of the multi-station are shown in Table 1.

In Test No. 1 where the multi-station stayed stationary, the maximum absolute error in translation was 1.0 mm and that the rotation angle in the north was 0.004° . These non-zero values were caused by the measurement error of the multi-station.

In Tests No. 2—4, the multi-station was translated in all three axes and rotated around the vertical axis (z -axis) by an angle in a wide range. The results of Tests No. 2—4 show that the translation calculated by Algorithm 1 was close to the true translation as determined by the full-size prism on a pole with a maximum absolute error of 6.5 mm on one axis. It should be noted that the true error in translation was likely smaller than 6.5 mm because the multi-station may have not been set up at the exact location of the pole due to human error. Therefore, the accuracy of Algorithm 1 is better reflected by the Euclidean difference in the measured and adjusted coordinates of the four control points on the mock-up wall as shown in Table 2.

Consistent with Table 1, Table 2 shows a small baseline error (an average of 0.6 mm point difference) in the coordinates of control points in Test No.1 due to instrument error. After the multi-station was translated and rotated (Tests No. 2—4), Algorithm 1 was able to predict the coordinates of control points in the current setup with an average error between 2.3 and 3.1 mm, which is lower than a threshold of 3.2 mm as recommended by industry partners for measuring panel positions [2]. The error varied among different control points and tests, but the variation was small. There was no noticeable difference between

Table 1. Performance of Algorithm 1: translation and azimuth.

Test No.			1	2	3	4
Measurement mode			Auto	Manual	Auto	Auto
Translation	True (m)	<i>x</i>	0	4.4909	-3.1040	-2.9615
		<i>y</i>	0	-1.8064	3.2602	3.8264
		<i>z</i>	0	0.0037	-0.0076	-0.3000
	Calculated (m)	<i>x</i>	0.0010	4.4898	-3.1105	-2.9658
		<i>y</i>	0.0001	-1.8122	3.2657	3.8299
		<i>z</i>	-0.0005	0.0046	-0.0066	-0.3017
	Absolute error (mm)	<i>x</i>	1.0	1.1	6.5	4.3
		<i>y</i>	0.1	5.8	5.5	3.5
		<i>z</i>	0.5	0.9	1.0	1.7
Azimuth	Calculated (°)		0.004	58.3	83.7	267.9

Table 2. Performance of Algorithm 1: Error in the coordinates of four control points (*CP*, mm).

Test No.	1	2	3	4
<i>CP</i> ₁	0.7	2.5	1.9	2.8
<i>CP</i> ₂	0.6	3.1	1.5	3.1
<i>CP</i> ₃	0.6	3.9	4.1	3.8
<i>CP</i> ₄	0.6	3.1	1.5	2.6
Average	0.6	3.1	2.3	3.1
Maximum	0.7	3.9	4.1	3.8

Test No.2 and the other two tests with a different measurement mode, which indicates that the auto mode can achieve an accuracy as good as the manual aiming mode.

While on average, the error in the predicted coordinates was acceptable for prefabricated construction, the maximum error was about 1 mm above the threshold of 3.2 mm (Table 2). Therefore, this error needs to be reduced further. One approach is to investigate the reason for the maximum error appearing at *CP*₃ which was the control point at the lower left corner of the mock-up wall in all three tests. It was likely that the tape of this control point was not measured as accurately as the other three tapes. Another approach is to study if the error can be reduced by using more control points as the input to Algorithm 1.

3.2 Experiment on rotation

The algorithm of rigid-body transformation (Algorithm 2) was tested by rotating a panel around a single *y*-axis, and the results are displayed in Table 3.

The test panel was rotated around the *y*-axis as set by the multi-station by an angle from $\sim 5^\circ$ to $\sim 26^\circ$ at an

interval of $\sim 5^\circ$ both in the clockwise and counterclockwise direction. An angular error smaller than 0.5° was achieved when comparing the rotation angle measured by the digital protractor and the angle around the *y*-axis calculated by Algorithm 2. The error in the angle tends to be larger when the angle is larger, which was likely caused by a larger human error in rotating the panel. In addition to the angle around the *y*-axis, angles less than 0.57° in absolute value were obtained for the *x* and *z*- axes. The non-zero angles around these two axes were caused by the measurement error of the multi-station and the possibility that the panel was slightly rotated around these two axes when it was rotated around the monitor arm's joint on the VESA mount due to human error. The experiment setup needs to be improved further to eliminate the human error.

The true translation of the panel is zero because the panel was only rotated, and the translation measured by the tape at the panel center and that calculated by using the coefficients of the linear combination of three prisms are both nearly zero. A maximum Euclidean distance error of 5.5 mm was observed in the measured and calculated translation. This error can be reduced further by increasing the accuracy of the calculated coefficients of the linear combination by taking the average of multiple measurements.

4 Conclusions and next steps

Two algorithms were developed to aid the real-time tracking and positioning of prefabricated components in construction. The resection algorithm enables the alignment of the digital twin with physical twin regardless of the position and orientation of the total station. Lab-scale tests show that an average error of less than 3.1 mm was achieved in point coordinates, which satisfies the industry recommendation of less than 3.2 mm. The second algo-

Table 3. Performance of Algorithm 2

Test No.	Measured angle (°)			Rotation angle from Algorithm 2 (°)			Error	
	Current	Goal	Rotation angle	y-axis	x-axis	z-axis	Rotation angle (°)	Translation (mm)
1	95.4	90.2	5.2	5.13	0.03	0.29	0.07	1.0
2	100.9	91	9.9	9.91	0.02	0.12	0.01	5.5
3	105.3	90.8	14.5	14.42	-0.01	-0.29	0.08	4.0
4	111.7	91	20.7	20.45	0.01	-0.33	0.25	4.8
5	115.7	89.4	26.3	26.13	-0.24	-0.57	0.17	1.9
6	90.8	111.6	-20.8	-20.45	-0.23	0.43	0.35	4.4
7	91.1	117.9	-26.8	-26.37	-0.22	0.32	0.43	3.0

rithm transforms the component from the current pose to the intended goal pose. The tests of rotating a component around a single axis show that the error in the rotation angle was less than 0.5° , which is acceptable given the large rotation angles of the component.

Several limitations in this paper require future research. First, the errors observed in the performance of the two algorithms need to be reduced. Potential approaches include reducing instrument error (e.g., selecting different measurement settings, using reflectors with higher accuracy), revising the algorithms to decrease their sensitivity to instrument error, and eliminating human error by designing better experiments. Second, experiments that can test the transformation algorithm for rotations around all three axes need to be conducted. Third, the Euler angles around the x , y , and z -axes in a Cartesian coordinate system set by the total station are currently used in the transformation algorithm. The choice of using the prefabricated component as the local coordinate system should be included in the algorithm. Ultimately, these two algorithms will be implemented on the Real-Time Evaluator (RTE) to optimize and automate the process of installing prefabricated components with high accuracy.

5 Acknowledgements

This research was supported by the DOE Office of Energy Efficiency and Renewable Energy (EERE), Building Technologies Office, under the guidance of Sven Mumme, and used resources at the Building Technologies Research and Integration Center (BTRIC), a DOE-EERE User Facility at Oak Ridge National Laboratory.

References

- [1] Tarek Omar and Moncef L Nehdi. Data acquisition technologies for construction progress tracking. *Automation in Construction*, 70:143–155, 2016.
- [2] Nolan W. Hayes, Bryan P. Maldonado, Diana Hun, and Peter Wang. Automated tracking of prefabricated components for a real-time evaluator to optimize and automate installation. In *Proceedings of the 40th International Symposium on Automation and Robotics in Construction*, pages 192–199, Chennai, India, 2023.
- [3] Bryan P. Maldonado, Nolan W. Hayes, and Diana Hun. Automatic point Cloud Building Envelope Segmentation (Auto-CuBES) using Machine Learning. In *Proceedings of the 40th International Symposium on Automation and Robotics in Construction*, pages 48–55, Chennai, India, 2023.
- [4] Guobin Chang. On least-squares solution to 3d similarity transformation problem under gauss–helmert model. *Journal of Geodesy*, 89(6):573–576, 2015.
- [5] Leon C Price, Jingdao Chen, Jisoo Park, and Yong K Cho. Multisensor-driven real-time crane monitoring system for blind lift operations: Lessons learned from a case study. *Automation in Construction*, 124:103552, 2021.
- [6] Yan Xu, Yi Luo, and Jian Zhang. Laser-scan based pose monitoring for guiding erection of precast concrete bridge piers. *Automation in Construction*, 140:104347, 2022.
- [7] Jim Lawrence, Javier Bernal, and Christoph Witzgall. A purely algebraic justification of the kabsch-umeyama algorithm. *Journal of research of the National Institute of Standards and Technology*, 124:1–6, 2019.
- [8] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.
- [9] Jian S Dai. Euler–rodrigues formula variations, quaternion conjugation and intrinsic connections. *Mechanism and Machine Theory*, 92:144–152, 2015.
- [10] David Eberly. Euler angle formulas. Online: <https://www.geometrictools.com/Documentation/Documentation.html>. Accessed: 15/12/2023.

Assessment of Traditional and Robotic Approaches to Interior Construction Layout: A Framework and Comparative Study

Catherine Caputo¹, Ashtarout Ammar¹, and Ashley Johnson¹

¹Myers-Lawson School of Construction, Virginia Polytechnic Institute and State University, United States
catherine10@vt.edu, aammar@vt.edu, alj@vt.edu

Abstract –

As the construction industry witnesses a growing integration of robots and automated systems on complex construction sites, project teams exhibit varying definitions of successful robot employment. Notably, there is an absence of standard criteria for stakeholders to assess the impacts of these technologies on productivity, cost, safety, and pertinent human factors. Existing studies suggested key framework elements, yet none provided a comprehensive, quantitative means to assess on-site construction robots. In response, this study introduces a holistic framework of Key Performance Indicators (KPIs) as a reference for researchers to evaluate single-task robots. A case study was conducted using a set of extracted KPIs, comparing the traditional construction interior layout method with the performance of a single-task mobile layout robot. The study demonstrates a comparative approach that project teams can adopt to maximize robot benefits and meet project-specific goals.

The case study accounts for the unpredictability of robot implementation that project teams may need to adapt to. The results highlight drawbacks of recent automation, such as technological inefficiencies. Depending on the application, these challenges can increase project completion time and affect space utilization. This research presents a comprehensive productivity analysis of a recently introduced mobile layout robot. Additionally, we highlighted robot advancements in comparison to previous layout robots assessed in past studies. These advancements provide positive cost and safety implications. The conclusions offer insights into the feasibility of adopting these technologies and considerations for stakeholders seeking to implement them.

Keywords –

Robotics, Automation, Single-task Robots, On-site, Construction, Interior Layout, Key Performance Indicators (KPIs), Efficiency, Productivity

1 Introduction

The construction industry is a leading driving force in any nation's economy. In the United States (US), the construction industry contributes 4% to the Gross Domestic Product (GDP) [1]. In 2022, the value added to the US GDP by the construction industry amounted to one trillion US dollars [2]. A look ahead to 2024, and as the global market is witnessing inflation rates, it is evident that the construction industry will face challenges with wage cost increases, supply chain disruptions, skilled labor shortages, and rising construction costs [3]. In addition to the prolonged safety and productivity issues the industry has suffered, these forecasted challenges add another layer of complexity for decision-makers to integrate and implement emerging technologies and automation into the construction workflows and to upskill the workforce. Thus, construction automation can be viewed simultaneously as an opportunity and a challenge.

There are three major opportunities for construction automation: 1) automating traditional redundant physical tasks on sites, for instance, robots laying bricks; 2) off-site modular construction such as the use of 3D-printed construction components; and 3) digitization of the design, planning, and management procedures [4].

Adopting robots in construction requires an upfront investment, making it prohibitive for small construction organizations that lack the necessary resources. Also, there are concerns about the safety risks of integrating robots into a volatile and hazardous environment such as construction sites. The construction industry is a dynamic environment that requires immediate interventions and workflow changes; however, including robots might restrain the needed flexibility. As construction companies adopt automation and try to use robots on and off construction sites, barriers among stakeholders limit their usage, mainly due to the resistance to change and lack of skilled workforce [5]. Thus, the degree to which robotics are adopted in the construction industry relies on the awareness and perceptions of their advantages and

disadvantages concerning these barriers [6]. Conversely, understanding robot types and their applications among stakeholders can expand their adoption.

The types of construction robotics can be generally classified into four categories: 1) off-site prefabrication, 2) on-site automation, 3) drones and autonomous vehicles, and 4) exoskeletons [7]. This broad categorization can cause an overlap of robot identification terms and definitions [8]. Robots can be further categorized according to their applications. Robot technology applications in the construction industry vary, including, but not limited to, painting and spraying, demolition, brick and concrete laying and plastering, construction welding, drilling, bolting, drywall and façade installation, steel-truss assembly, transportation, inlaying, surveying, inspection and monitoring, roadwork, excavating and earth moving, and interior decoration [7, 9–11].

The opportunities for robotic employment across the entirety of the construction industry are extensive because of their growing versatility, which can cover a wide range of on-site and off-site activities [12]. However, implementing automation and robotics on construction sites is more challenging since they are implemented in highly unstructured, uncontrolled, and congested environments [13]. For instance, the need for error mitigation in exterior and interior layout amidst site congestion is critical, considering these tasks ensure that the project is accurately built according to design plans and budget [14]. Despite inconsistent site conditions, the layout precision and efficiency are contingent upon the ability of surveyors to provide accurate positioning according to acceptable tolerances with the use of a total station and Global Positioning System (GPS), among other geomatic instruments [15]. This process is fundamental to a project's short and long-term quality and schedule, as it establishes the structural integrity and upstanding of the designed structure [15]. As a result, the safety of on-site workers, building occupants, and other relevant stakeholders is an essential consideration of the layout planning phase [14]. Due to the emphasis on precise execution, surveyors are to be heedful; as a result, the time associated with rework and error mitigation affects the project's efficiency, cost, and schedule overruns [15]. Additionally, on sites that demand complex design plans and coordination between various trades, optimization of space and schedule is critical, as it combats the growing issue of low productivity and high costs [16].

The application of factory-based automation in construction can be viewed as a technology transfer from the manufacturing sector, albeit with some exceptions. In this context, automated tooling is adapted to manufacture building elements instead of traditional products [17]. Unlike factory automation, implementing automation on

construction sites involves unique hurdles and prospects. It necessitates developing and deploying specialized equipment and processes, marking a departure from direct technology transfer [9]. For instance, mobile layout robots have recently been introduced to construction sites to mitigate the risks and limitations of traditional layout methods on complex project sites [18]. This implementation creates a fertile ground for research, new business ventures, and the emergence of start-ups in this innovative field.

Moreover, with the unceasing advancement of automation and robots, developing continuous assessment criteria and implementation frameworks is critical to evaluate the advantages and associated risks of implementing robots on construction sites [19]. As such, this study aims to investigate the metrics used by researchers and practitioners to evaluate the implementation of automation and robotics on construction sites and provide standard criteria to help the construction industry evaluate this implementation. Furthermore, a case study will be presented to test the applicable metrics by comparing the use of a construction robot versus the traditional method in conducting an interior construction layout.

2 Background

Considering the desire to integrate construction automation, the shortage of skills, and the emphasis on enhancing sustainability in the construction industry, it is probable that widespread acceptance of automation and robotics will be commonplace in construction soon [20]. This is largely because construction tasks can be repetitive and tiresome, and the collaboration of robots and humans can alleviate workload and exhaust [9, 21]. However, the industry is considered dangerous, complex, and unpredictable [22]. Therefore, anticipation and preparation are the keys to supporting the construction sector as it seeks new, tangible uses for automation [4].

Implementing construction automation equipment for on-site tasks necessitates the design of this equipment to be portable for transportation to various job sites, where it can be set up, utilized, and dismantled for relocation to the next assignment [20]. For instance, conventional equipment like heavy earthmoving machinery has undergone retrofitting, and there is a growing trend of manufacturing new equipment with a focus on an automated or semi-automated future [9]. Initially, on-site automation led to the creation of building systems tailored specifically to integrate with those automated construction systems. However, in many instances, this approach reduced the distinctiveness of each building [20]. In today's context, there is a renewed effort in automated construction that accommodates variations across units while incorporating standardized elements [20]. For instance, automated equipment designed for

constructing concrete reinforcement reduces the labor demand of repetitive tasks on the construction site. It enables performance-driven variability in rebar tying. Moreover, precisely fastening material where needed helps minimize schedule without incurring additional costs [23].

The extent to which companies and project teams implement on-site construction automation is contingent on known success indicators in similar application areas. However, construction companies cannot rely on historical data to evaluate and assess on-site construction automation due to the limited history of construction industry engagement with robots, and managers cannot determine the optimal robotic solution for a specific project [24]. Conversely, a substantial number of Key Performance Indicators (KPIs) can be applied to evaluate robot performance; however, it's important to identify indicators that align with the set goals of the company or involved teams [25]. The indicators chosen to measure automated technology are specific to the objectives of each study and depend on the focus of the task and the robot used. Yet, there is an overarching framework of KPIs that are repeatedly addressed. In direct response to the main issues that the industry faces when integrating robots, improvement levels are often measured in terms of productivity, cost, safety, and human factors [26].

Robotics and automation in the manufacturing and production settings have become increasingly common [20]. Notably, the authors of [25, 31] provided a general structure of productivity metrics for manufacturing robots that could be manipulated and adopted to assess on-site construction robots. Given the unique challenges of construction sites, the authors of [10, 17, 27, 32] assessed the performance of on-site construction robots. Compared to the manual method, the robots provided higher quality with less working time and labor intensity, thereby providing insight into how robots can alleviate the issue of low efficiency within the industry, as it relieves skilled workers of tedious tasks and exposure to safety hazards [32]. The specific metrics used by researchers and practitioners to evaluate productivity improvement for off and on-site robot implementation are summarized in Table 1.

Table 1. Productivity KPIs for assessing automation in construction.

Productivity	
KPI	Reference
Cycle time	[10, 17, 25–27, 29, 31]
Cycles/Jobs completed	[25, 27, 31]
Efficiency	[17, 25, 26, 28, 31, 32]
Manpower	[27, 28, 32]
Quality/Accuracy (rework, repeatability, material waste)	[28, 32]

Table 1. (Continued)

Productivity	
KPI	Reference
Set up time	[26]
Training hours	[26]
Utilization	[25, 26, 31]
Wait/Disconnected time	[25, 26, 31]
Working time/Speed	[17, 27, 28, 32]
Yield	[25, 28, 31]

The large economic output of the construction industry contributes to its significance in the global industrial sector [5]. Therefore, perceptions of high costs can negatively influence robot adoption [33]. According to the schedule impacts of robot deployment, productivity assessment data can be monetarily quantified to compare manual and robotic work. The monetary transformation of performance indicators into financial indicators using engineering principles, such as return on investment, cash flow analyses, and benefit-cost ratio, enables companies to determine whether robot investment is economically beneficial [25, 27]. As such, the authors of [10, 27, 29], using a cost-benefit analysis of single-task construction robots, were able to monetize the impacts of productivity and speed. The direct costs associated with robot acquisition, operation, and maintenance, such as unit costs, savings in estimated schedule, and rework reduction, have been compared with traditional work's labor and material costs [26, 28]. Additionally, indirect costs, including costs of workforce training, waste mitigation, health damage, and other demands of robot deployment, are to be included in a comprehensive assessment of automation in construction [8, 30, 34]. The quantified measures of direct and indirect cost in existing studies, as Table 2 shows, are strategically chosen to support the project's financial goals.

Table 2. Cost KPIs for assessing automation in construction.

Cost	
KPI	Reference
Benefit-cost ratio	[27, 28]
Health damage cost	[34]
Innovation (training, technology acquisition)	[8, 30]
Labor cost (number of employees, function, salary)	[8, 27–30]
Operational cost (maintenance, license, energy, resource costs)	[8, 27, 30]

Table 2. (Continued)

Cost	
KPI	Reference
Payback period	[8, 27]
Productivity (unit cost, construction cost, construction time, project dimensions)	[26, 30]
Profitability (revenue, market share)	[30]
Quality (cost of rework, delay)	[8, 30]
Return on investment	[8, 25, 27, 28]
Technology cost (hardware and software costs)	[8, 29, 30]

Besides cost and productivity, automated construction can further improve workers' health and working conditions – a highly prioritized objective of the construction industry. As such, authors of [10, 34] evaluated the health damage caused by robotic and manual methods in interior construction. The studies found that robot adoption yielded healthier working conditions [10, 34]. Despite the growing utilization of construction automation and robotics, there have yet to be specific Occupational Safety and Health Administration (OSHA) standards for the robotics industry [35]. Even though safety impacts may not be identifiable on the project level, their long-term effects and cost impacts are to be considered [36]. To a greater degree, robotic performance should be evaluated against the large-scale, global pushes toward sustainability [8]. Metrics of safety KPIs are presented in Table 3 as measuring progress towards achieving the prevention of existing or possible hazards.

Table 3. Safety KPIs for assessing automation in construction.

Safety	
KPI	Reference
Harmful byproduct production (air pollution, dust or chemical concentration, noise levels, etc.)	[8, 10, 28, 32, 34]
Identifiable safety concerns	[32]
Muscle strain	[28]
Number of incidents, injuries, fatalities or hazards	[8, 26, 28, 30]
Safety inspection time	[28]

In conjunction with the quantitative performance indicators, the low level of automation in construction compared to other industries is also due to human and social factors, particularly perceptions and attitudes toward robots. Typical social barriers include lack of knowledge and familiarity, lack of training, fear of job loss, situation awareness, and distrust [5, 9]. Experts in related fields ranked “current work culture/aversion to change” as the 4th most significant factor limiting the adoption of robotics in the construction industry, behind cost and productivity factors [7]. Relevant human factors, as seen in Table 4, can lead to hesitancy and negatively influence the robot's improvement of construction efficiency.

Table 4. Human Factor KPIs for assessing automation in construction.

Human Factors	
KPI	Reference
Acceptability	[8]
Adaptability	[8]
Comprehensibility	[37]
Fatigue	[21]
Operator's average stress	[25]
Reliability	[8, 37]
Safety Perception	[21]
Situation awareness	[26]
Stakeholder satisfaction	[6, 8]
Trust	[17, 26]

3 Methodology

Given the lack of recent literature and case studies assessing robot task performance on construction sites, particularly in interior construction layout, this study aims to 1) develop a multifaceted framework of Key Performance Indicators according to metrics of productivity, cost, quality, and safety; 2) assess the performance of a single-task layout robot; and 3) conduct a comparative analysis of manual and robotic interior layout methods. In accordance with the sequence of steps shown in Figure 1, this section identifies the applicable KPIs extracted from existing literature to assess the robotic interior layout method compared to the traditional method.

The tasks preceding point data collection or layout are classified as set-up tasks. For interior layout, set-up tasks aim to mitigate the need for rework through proper instrumental and methodological steps to ensure precise locations of partitions and systems according to specified

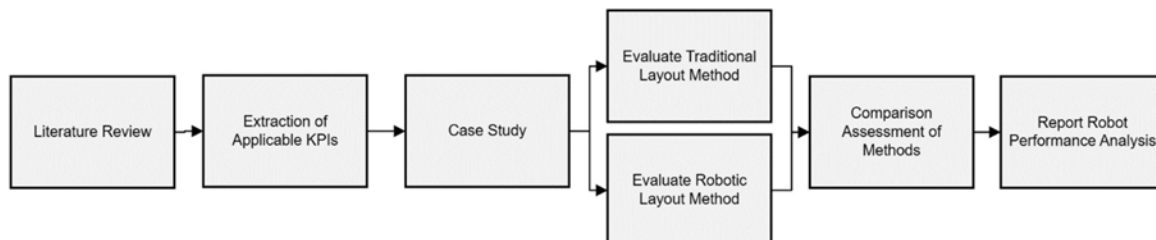


Figure 1. Research Methodology

design plans. The time associated with set up is often minimal, as it's completed once before the task begins. Therefore, set-up time is often not considered in large-scale productivity analyses, but it is a determinant for schedule efficiency depending on the frequency of rework.

The consistency of robot implementation, including higher levels of repeatability and speed, with proper rework mitigation, can reduce the time it takes to complete a given task. As a result, the time saved during a robot's working time indicates increased productivity and a basis for future robot process improvement.

Other studies did not address the time the robot was static or did not perform productive work. In response to this limitation, this study separately assesses wait time as an individual KPI to provide an accurate account of a limitation of layout robots. This cannot be mitigated like other set-up preventative measures and can contribute to increased working time. Thereby, these unforeseeable disadvantages of implementation can pose hindrances to productivity. Congested sites, often those with small areas or complex design plans, result in similar impacts when using the traditional methodology.

Similarly, reducing wait time increases efficiency. On a larger project level, efficiency studies can quickly assess if the project's schedule and cost are well optimized. Measures of robot efficiency assess productivity from workflow initiation to completion and improvement points expected to expedite a project's schedule if implemented.

Of the extracted KPIs, this study adopted 4 KPIs that assess productivity: 1) set-up time, 2) working time, 3) wait time, and 4) efficiency. For additional clarification on how this study categorized the adopted productivity KPIs, the formulaic definitions are provided in Table 5.

Table 5. List of KPIs and their corresponding formulas used to assess performance quantitatively.

KPI	Equation	Reference
Set Up Time	$\text{Set up Time} = \sum \text{Task set up time}$	[26]
Working Time	$\text{Working Time} = \text{Time task ends} - \text{Time task started}$	[26]
Wait Time	$\text{Wait Time} = \sum \text{Static Times}$	[31]
Efficiency	$\text{Efficiency} = \frac{\text{Working Time}}{\text{Total Time}}$	[31]

4 Case Study Comparison Results

We conducted a case study to analyze HP's single-task layout robot compared to the manual layout method. The case study intends to provide a practical means of highlighting the impact and challenges that can further contribute to adopting robots in interior layouts on construction sites.

The *HPSitePrint*, an autonomous printing robot size $50.5 \times 31.7 \times 26.1$ cm, is a portable device that a single operator can transport easily. The performance process was evaluated to determine if the robot was feasible and fit for task employment. Figure 2 details how this study categorized the robot's workflow process according to the adopted KPIs. The implemented workflow provides an organizational framework of productivity metrics for direct comparison to the traditional method.

Unlike the traditional method, the cloud-based management system controlled and organized the robot

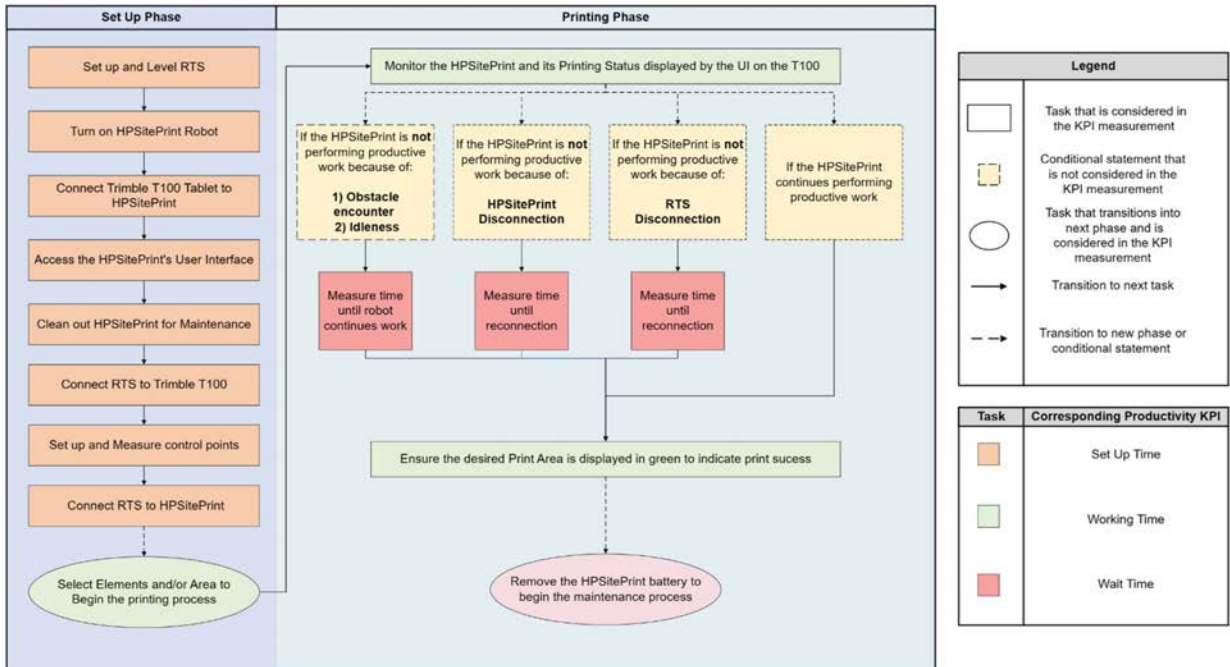


Figure 2. *HPSitePrint* workflow and breakdown of tasks into KPIs to assess productivity.

layout process. This system uses a CAD file with three layers specifying the objects to be printed, objects to be considered obstacles, and objects representing the total station reference points and lines. To initiate the layout, the Trimble Robotic Total Station's (RTS) 573 laser was manually aimed at control points to ensure correct orientation. The operator of the T100, a data collection and processing tablet, could then select the area or components to be printed. The robot autonomously printed the lines and texts based on its self-determined path, tracked by the RTS. However, the autonomous printing path made the implementation of safety controls difficult.

Robots implemented on-site can face interruptions from human and material traffic. Known obstacles from the CAD file can be addressed, but an advancement of the *HPSitePrint* is the ability to add unanticipated obstacles after arrival on-site. The operator could manipulate the robot's movement using the remote-controlled joysticks to avoid oncoming or unanticipated obstacles.

The robot was tasked with printing wall and window lines, door arcs, and text over a 9.17 m² area within an enclosed, controlled room. The robot is compatible with ink fluids and can overcome irregular surfaces up to 2 cm thick. Therefore, it can be employed on a wide selection of porous and nonporous surfaces, including builders' paper, the printing surface used in this study. The robot's performance was compared to that of the traditional method using the same CAD file. Once the workflows of both processes were established, we could quantify productivity, cost, and quality measures by completing three trials for each method. There were no measurable safety issues encountered. However, the robot's safety attributes will be briefly discussed in the subsequent safety analysis.

Productivity: Unlike existing studies assessing mobile layout robots, this study analyzed productivity in terms of four separate KPIs instead of dividing total time over a given area. When the robot arrived at the layout area, the time taken to activate the robot was recorded. Once printing began, we recorded the time it took for the robot to finish the desired layout. If the robot were to disconnect from the RTS or T100, encounter an obstacle, or remain idle, the time until reprinting began was measured. At the beginning of each trial, the robot underwent a self-calibration process. This wait time minimally affected the robot's productivity, as the idleness lasted an average of 19 seconds. The total time (i.e., the set-up, working, and waiting times) was calculated to determine the robot's efficiency, defined as the ratio of working time to total time. Compared to the average efficiency of the manual method, the robot method was 17% less efficient. The robot was tasked with a small print area of architectural elements. If tasked

with a larger layout containing plans of multiple trades, the robot is anticipated to increase efficiency. The averages of the three trials for the robot and manual methods in terms of the four applied productivity KPIs are shown in Table 6.

Table 6. Productivity comparison results of the manual and robotic layout methods.

KPI	Manual	Robot
Set Up Time	8:08	15:28
Working Time	26:59	23:29
Wait Time	---	0.19
Total Time	35:07	39:16
Efficiency	76.8%	59.8%

Cost: Secondly, we considered the costs required to print the layout. Depending on the need, the robot can be rented instead of bought. To provide a way for project teams to predict and allocate costs, *HPSitePrint* includes all costs in a single fee per square foot of executed layout. Once the layout is completely printed or a month has passed, the cloud software marks the CAD file as completed, and users are billed according to the layout area(s). If implemented, trades can combine their layouts into a single CAD file to be printed within one month to avoid duplicate charges and accelerate the schedule. HP charged 20 cents per square foot or 0.093 per square meter. The printing cost for the 9.17 m² area was \$19.74. To compare, the mean hourly wage of a construction laborer, including those who operate surveying and measuring equipment, in 2022 was \$22.29 in Virginia, according to occupational employment and wage estimates [38].

Safety: Comparable to previous case studies of ergonomic measures, robot implementation alleviated bodily demand, as the conventional approach involved frequently bending down to establish the chalk line. Due to the unpredictability of construction sites, robots need the capability to detect real-time changes in the environment, respond promptly, and navigate present obstacles [11]. An advantage of the *HPSitePrint* is its obstacle avoidance and cliff safety sensors, as shown in Figure 3. These reduce risks by stopping robot movement when encountering elements not included in the CAD file. Ensure the sensors correctly identify obstacles; a daily check of sensor functionality is mandatory.

Quality: Lastly, we evaluated the quality of the robot's print. We overlaid two layout prints using the



Figure 3. *HPSitePrint* Safety Sensors

same control line and RTS orientation. The robot's repeatability in printing sequential layouts ranged from 0 to 6.35 mm. This variance may pose an issue on-site, depending on the allowable tolerance. Traditionally, the process required two workers to mark point locations within a 6.35 mm tolerance, which the software audibly confirmed. The workers were prompted to relocate the point if staked out of tolerance. Similarly, the workers re-joined the points if chalk lines failed to connect points or were not visible.

5 Conclusion, Limitations, & Future Work

The fast evolution of robots poses challenges to adoption, emphasizing the need for current studies on practical ways to measure their impacts. The absence of a definitive evaluation method underscores the need for a systematic guide when comparing robotic and traditional approaches.

To address this gap, the study's first aim was to address the lack of a recent performance assessment method for on-site construction robots. We achieved this by consolidating KPIs from existing literature into a concise framework to quantify the impacts of robot deployment. Using applicable KPIs, the second aim was to analyse the application of an interior layout robot compared to the traditional method. The analysis showed the robot offers preventative safety measures, expanded printing abilities, and cost-saving opportunities. Also, it identified limitations of the robot, such as inconsistent repeatability.

Given that robots are recent additions to the industry, challenges in revolutionizing assessment methods are anticipated. In this study, our assessment of the robot's performance was limited due to the off-site location. For instance, data collection pertaining to safety and human factors was not feasible without interruptions and interactions with on-site activities. Subsequent studies will allow us to further assess these factors on a larger-scale construction site. Also, the use of trade-collaborated plans can be evaluated in a longer case study period. The results indicated that the traditional method is more efficient, but these changes are anticipated to alter the outcome in favor of the robot.

Based on the insight of the case study, we demonstrated the significance of using a multi-metric framework. Future research should validate the proposed framework's reliability, inclusiveness, and efficacy.

References

- [1] Johnson A. Using Construction As An Economic Indicator, <https://www.forbes.com/sites/forbesbusinesscouncil/2023/08/16/using-construction-as-an-economic-indicator/?sh=5c07266a7bfa> (2023, accessed 12 December 2023).
- [2] Value added to gross domestic product by the construction industry in the United States from 2000 to 2022, <https://www.statista.com/statistics/785445> (2023, accessed 12 December 2023).
- [3] Reynolds A. Headwinds and headaches for global construction industry, <https://www.rlb.com/americas/insight/headwinds-and-headaches-for-global-construction-industry/> (2023, accessed 12 December 2023).
- [4] Chui M, Mischke J. The impact and opportunities of automation in construction, <https://www.mckinsey.com/capabilities/operations/our-insights/the-impact-and-opportunities-of-automation-in-construction> (2019).
- [5] Jäkel J-I, Rahnama S, Klemm-Albert K. Construction Robotics Excellence Model: A framework to overcome existing barriers for the implementation of robotics in the construction industry. Epub ahead of print 15 July 2022. DOI: 10.22260/ISARC2022/0085.
- [6] Charlesraj VP, Nijalingamurthy R. Stakeholder Perspectives on the Adoption of Drones in Construction Projects. 2020. Epub ahead of print 26 October 2020. DOI: 10.22260/ISARC2020/0168.
- [7] Davila Delgado JM, Oyedele L, Ajayi A, et al. Robotics and automated systems in construction: Understanding industry-specific challenges for adoption. *J Build Eng* 2019; 26: 100868.
- [8] Pan M, Linner T, Pan W, et al. A framework of indicators for assessing construction automation and robotics in the sustainability context. *J Clean Prod* 2018; 182: 82–95.
- [9] Adepoju O. Robotic Construction Technology. In: Adepoju O, Aigbavboa C, Nwulu N, et al. (eds) *Reskilling Human Resources for Construction 4.0: Implications for Industry, Academia and Government*. Cham: Springer International Publishing, pp. 141–169.
- [10] Brosque C, Skeie G, Fischer M. Comparative Analysis of Manual and Robotic Concrete Drilling for Installation Hangers. *J Constr Eng Manag* 2021; 147: 05021001.
- [11] Brosque C, Galbally E, Khatib O, et al. Human-Robot Collaboration in Construction: Opportunities and Challenges. 2020, pp. 1–8.
- [12] Bruun EPG, Pastrana R, Paris V, et al. Three Cooperative Robotic Fabrication Methods for the Scaffold-Free Construction of a Masonry Arch, <http://arxiv.org/abs/2104.04856> (2021, accessed 1 December 2023).
- [13] Saidi K, Bock T, Georgoulas C. Springer Handbook of Robotics. In: *Robotics in Construction*. Springer, Cham, 2016, pp. 1493–1519.
- [14] The Role of Layout in Delivering Efficient, Successful Building Projects - Civil + Structural

- Engineer magazine. <https://csengineermag.com/>, <https://csengineermag.com/the-role-of-layout-in-delivering-efficient-successful-building-projects/> (accessed 1 December 2023).
- [15] Sestras P. Methodological and On-Site Applied Construction Layout Plan with Batter Boards Stake-Out Methods Comparison: A Case Study of Romania. *Appl Sci* 2021; 11: 4331.
- [16] Hawarneh AA, Bendak S, Ghanim F. Construction site layout planning problem: Past, present and future. *Expert Syst Appl* 2021; 168: 114247.
- [17] Wang L, Fukuda H, Shi X. A Preliminary Comparison Between Manual and Robotic Construction of Wooden Structure Architecture. Kitakyushu, Japan. Epub ahead of print 14 October 2020. DOI: 10.22260/ISARC2020/0218.
- [18] HP Revolutionizes Construction Layout Process With New SitePrint Robotic Solution, <https://press.hp.com/us/en/press-releases/2022/hp-new-siteprint-robotic-solution.html> (accessed 28 November 2023).
- [19] Brosque C, Skeie G, Örn J, et al. Comparison of Construction Robots and Traditional Methods for Drilling, Drywall, and Layout Tasks. In: *2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, pp. 1–14.
- [20] Davis M. What Is Construction Automation, and How Will It Drive the Future of Building?, <https://www.cmaanet.org/sites/default/files/resource/Construction%20Automation.pdf> (accessed 15 December 2023).
- [21] Hopko S, Wang J, Mehta R. Human Factors Considerations and Metrics in Shared Space Human-Robot Collaboration: A Systematic Review. *Front Robot AI*; 9, <https://www.frontiersin.org/articles/10.3389/frobt.2022.799522> (2022, accessed 30 October 2023).
- [22] Pan Y, Zhang L. Roles of artificial intelligence in construction engineering and management: A critical review and future trends. *Autom Constr* 2021; 122: 103517.
- [23] TyBOT Case Study: Koppel Bridge, <https://www.constructionrobots.com/post/tybot-case-study-koppel-bridge> (2022, accessed 13 December 2023).
- [24] Skibniewski MJ, Nof SY. A framework for programmable and flexible construction systems, [https://doi.org/10.1016/0921-8890\(89\)90006-7](https://doi.org/10.1016/0921-8890(89)90006-7) (1989, accessed 15 December 2023).
- [25] Caiazza C, Nestić S, Savković M. A Systematic Classification of Key Performance Indicators in Human-Robot Collaboration. In: Mihić M, Jednak S, Savić G (eds) *Sustainable Business Management and Digital Transformation: Challenges and Opportunities in the Post-COVID Era*. Cham: Springer International Publishing, 2023, pp. 479–489.
- [26] D1.6 Key Performance Indicators (KPIs) for assessment I - FULL PAPER. *Sharework-project*, <https://sharework-project.eu/download/d1-6-key-performance-indicators-kpis-for-assessment-i-full-paper/> (2022, accessed 31 October 2023).
- [27] Hu R, Iturralde K, Linner T, et al. A Simple Framework for the Cost-Benefit Analysis of Single-Task Construction Robots Based on a Case Study of a Cable-Driven Facade Installation Robot. *Buildings* 2020; 11: 8.
- [28] Attalla A, Attalla O, Moussa A, et al. Construction robotics: review of intelligent features. *Int J Intell Robot Appl* 2023; 7: 535–555.
- [29] Epping K, Zhang H. A Sustainable Decision-Making Framework for Transitioning to Robotic Welding for Small and Medium Manufacturers. *Sustainability* 2018; 10: 3651.
- [30] Agustí-Juan I, Glass J, Pawar V. A Balanced Scorecard for Assessing Automation in Construction. In: *Proceedings of the Creative Construction Conference 2019*. Budapest University of Technology and Economics, pp. 155–163.
- [31] *The Top 5 Cobot KPIs - How to Measure and Improve the Performance of Collaborative Robots*. Robotiq, <https://www.hteautomation.com/data/siteshare/vendor/byid/1268/files/Top%205%20KPIs%20-%20How%20to%20measure.pdf> (2020, accessed 31 October 2023).
- [32] Asadi E, Li B, Chen I-M. Pictobot: A Cooperative Painting Robot for Interior Finishing of Industrial Developments. *IEEE Robot Autom Mag* 2018; 25: 82–94.
- [33] Pan M, Pan W. Understanding the Determinants of Construction Robot Adoption: Perspective of Building Contractors. *J Constr Eng Manag* 2020; 146: 04020040.
- [34] Chen C, Li X, Yao W, et al. Analysis of the impact of construction robots on workers' health. *Build Environ* 2022; 225: 109595.
- [35] Robotics - Overview | Occupational Safety and Health Administration, <https://www.osha.gov/robotics> (accessed 6 November 2023).
- [36] Brosque C, Fischer M. A robot evaluation framework comparing on-site robots with traditional construction methods. *Constr Robot* 2022; 6: 187–206.
- [37] Halder S, Afsari K, Chiou E, et al. Construction inspection & monitoring with quadruped robots in future human-robot teaming: A preliminary study. *J Build Eng* 2023; 65: 105814.
- [38] Virginia - May 2022 OEWS State Occupational Employment and Wage Estimates, https://www.bls.gov/oes/current/oes_va.htm (accessed 19 December 2023).

Development of automated transport system

Shunsuke Igarashi¹, Yuji Kinoshita², Taku Tani², Takayoshi Hachijo¹ and Masahiro Indo³

¹Institute of technology, Shimizu Corporation, Japan

²Construction technology Division, Shimizu Corporation, Japan

³NOVARE, Shimizu Corporation, Japan

igarashi@shimz.co.jp

Abstract

We have developed an automated transport system for construction sites, manufacturing plants, and distribution warehouses, such as an automated transport trolley, an automated transport forklift, and a transport management system. It can be instructed by a tablet to make a delivery and it can move autonomously along a given route by self-localization using LiDAR. In addition, a camera sensor can recognize the pallet or roll box pallet for loading and unloading. We have operated the automated transport trolley in construction sites and tested the automated transport forklift in distribution warehouses. In this paper, we report on the transport system and the result of operation and testing.

Keywords –

Automated transport trolley; Automated transport forklift; AMR; LiDAR; Camera sensor; Path planning

1 Introduction

Materials are transported in manufacturing plants, construction sites, and distribution warehouses using equipment such as trolleys and forklifts. Material handling operations are becoming increasingly mechanized and automated in manufacturing plants and distribution warehouses, where specialized equipment is often introduced. In facilities requiring flexibility in equipment to handle small lots of many different products, however, dedicated equipment is often not applicable. Reliance on manual labor is expected in construction sites due to the difficulty of installing equipment. Manual trolley transportation can be

hazardous and physically demanding, especially when handling heavy items, posing caught-in or load collapse risks. Even when forklifts or other vehicles are used, the work also involves dangers such as collisions and load collapse, requiring the operator's attention and concentration. In order to improve the safety and productivity of these transport operations and to save the workforce, we have developed a system to transport materials automatically.

It is said that the first transport robot [1], appeared in the 1950s. Since then, various transport trolleys have been developed [2]-[5], but many are optimized for specific facilities, such as receiving loads from production lines or towing dedicated carts or basket trolleys. Few automated transport trolleys can load and unload pallets placed on the ground or handle harsher environments than factories or warehouses, and most have limited travel speeds. Forklifts have also been studied [6][7], but they mostly move on predetermined routes and require guides like fixed frames due to their low tolerance to material misalignment, or the predetermined stacking height, leaving little room for flexibility.

Given this background, automating transport operations requires capital investments, such as changing the work method by installing specialized equipment rather than automating the current work processes without changing them.

In the construction, there were examples of automation by installing facilities such as line tape and material storage before 2000, but they were not suitable for construction sites where work progressed day by day. Since 2010, LiDAR and camera sensors have been developed, and a method of movement without the facilities has been proposed, but it was not suitable for

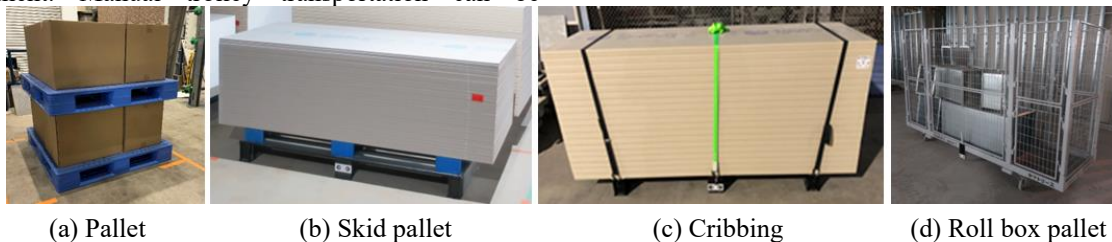


Figure 1 Packing mode

construction sites where the environment was changing and difficult to mapping and detect the self-position, and it was not an environment in which anyone could give work instructions.

To solve this problem, we have developed a more general-purpose transport system that requires minimal modification or expansion of facilities and equipment. The general transport work often uses transport equipment such as (a) pallets, (b) skid pallets, (c) cribbing, and (d) roll box pallet, as shown in Figure 1. Storage spaces, racks, and elevators are often designed according to the equipment standards. We developed two transport robots: automated transport trolley (Robo-Carrier) and automated transport forklift (Robo-Carrier Fork). These transport robots use LiDAR (Light Detection and Range) sensors for self-position detection and use camera sensors to detect objects such as materials, eliminating the need for building the environment, such as transport guides and equipment for specific materials. In addition, automating the elevator system at construction sites enables automatic material transportation to designated floors. We have operated automated transport trolleys at multiple construction sites and tested automated transport forklifts at logistics facilities. This paper describes the transport system and robots and reports on the results of operation and testing.

2 Transport System

2.1 System Overview

As shown in Figure 2, the transport system consists of a server called an integrated management system, a user interface tablet terminal, an automated transport trolley and forklift as transport robots, and a construction elevator specifically for construction sites.

The integrated management system receives the operation instructions from the user via the tablet terminal, generates the operation commands, and sends them to the transport robots. In addition, it receives the operation status of the transport robots and sends the status information to the tablet terminal. The integrated management system is connected to transport robots and tablet terminals via public networks or Wi-Fi.

2.2 Route Map

The worker uses the tablet terminal to register the transport work to be carried out as a task. The tablet displays a site plan and a route map, as shown in Figure 3, which includes the paths of the transport robots, material storage locations, elevators, and vehicle waiting areas. The route map shows a straight line connecting the pathways along which the transfer robot can travel and the site's material storage areas

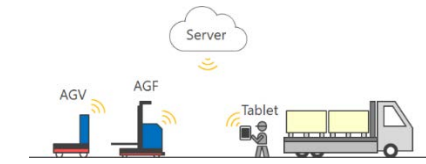


Figure 2 Transport system overview

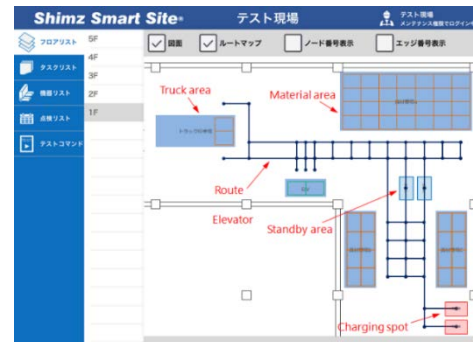


Figure 3 Route map

Task	Status	Progress	1st destination mode
60095 1F → 5F 開始日: 2022/ /	実行中	0 / 4	[Play] [Stop] [Cancel] [Refresh]
60096 1F → 5F 開始日: 2022/ /	実行中	0 / 4	[Play] [Stop] [Cancel] [Refresh]
60097 1F → 5F 開始日: 2022/ /	実行中	0 / 4	[Play] [Stop] [Cancel] [Refresh]
60098 1F → 5F 開始日: 2022/ /	実行中	0 / 4	[Play] [Stop] [Cancel] [Refresh]

Figure 4 Task list

and vehicle waiting areas. A vertex or intersection of the pathways is called a node, and the straight line connecting the nodes is called an edge. The robot's orientation for loading and unloading materials can be specified for each storage location. For stacked storage, the number of stacks can be specified at the same location, and for racks, the number of materials and their respective heights can be specified. In addition, multiple storage areas can be grouped as an area so that the destination of materials can be indicated collectively as an area instead of specifying each storage area. This route map can be drawn by CAD add-in software, and program code can be generated. Codes saved on the server can be displayed on the tablet.

2.3 Task Registration and Operation Instructions

For material transport instructions, select the type of material, destination floor, destination area, etc., and register them as a task. At the time of execution, pressing the play button in the task list, as shown in Figure 4, sends operation instructions from the integrated management system to the transport robots and construction elevators.

The motion instructions for the transfer robots include moving, loading, unloading, etc., and instructions sent to the construction elevators are grouped into tasks, such as opening/closing doors, moving floors, etc. The integrated management system sends the order and timing of sending motion instructions according to the status of the transport robots and construction elevators and sends commands in a timely manner.

2.4 Path Finding

Once a command is executed, the system searches for a travel route to the destination on the route map. Transport robots can move forwards, backwards, and even sideways and perform actions like rotating and turning. Since it is also essential to align the orientation of the vehicle body for the material storage area, it is necessary to find not only the path it travels but also which direction and method to travel along the edge, such as forward/backward or sideways, and which actions to perform at the nodes, such as spinning, turning, and maintaining the direction. To ensure the appropriate orientation of the transport robot, every possible route, including actions, is calculated to select the shortest travel-time route. When multiple robots are active in the same route map, based on the theory of multi-agent path planning [8], the expected passing time of each node is calculated, and one node is reserved for one vehicle to avoid collision. The other vehicle searches for a route to avoid collision by calculating a route that does not pass the reserved node at that time, as shown in Figure 5. The integrated management system monitors the positions of the robots during movement and, if interference between vehicles is predicted, re-executes the path finding and changes the route as necessary.

3 Automated Transport Robot

3.1 Automated Transport Trolley

As shown in Figure 6, the automated transport trolley is equipped with a mechanism to lift the material to its bed by the fork. It can transport flat-laid cargo and is primarily intended for use in construction sites. Its specifications are: load capacity 1,000 kg, body dimensions W1,400mm x L2,000mm x H1,930mm, body weight 1,160kg, and lithium-ion rechargeable battery-powered. The wheels consist of two drive steering wheels and two caster wheels, with the drive steering wheels being independently controllable. This allows the robot to move forwards, backwards, turn, and move sideways and rotate. The lower part of the fork has a scissor-type load-bearing mechanism, allowing it to support the weight of materials on the ground. This can eliminate the need for a counterweight, making the body lighter than a conventional forklift with a capacity of 1 ton. As a result,

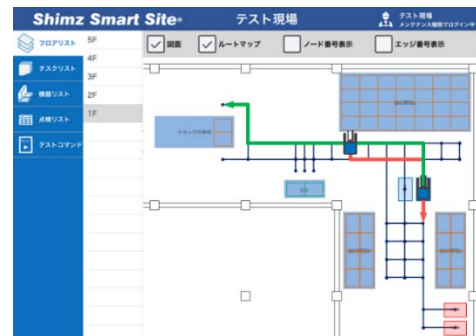


Figure 5 Multi-unit routing

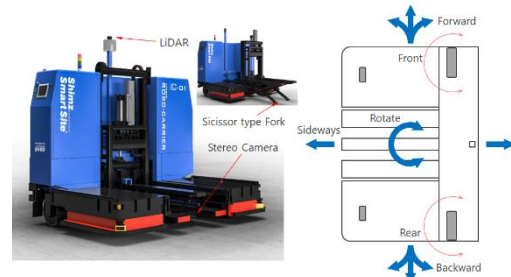


Figure 6 Automated transport trolley

it can be used on floors with general structures, such as construction sites. However, since the wheels need to roll under the fork, using skid pallets without bottom boards is necessary instead of standard pallets. During loading operations, a contact sensor located at the base of the fork can detect when the fork is fully inserted into the material.

3.2 Automated Transport Forklift

As shown in Figure 7, the automated transport forklift is based on a reach-type forklift as the base machine. It can unload trucks, stack materials, and load them onto racks at logistics facilities, manufacturing plants, and construction sites. Its specifications are: load capacity 1,000 kg, lifting height 3,000 mm, body dimensions W1,290mm x L2,450mm x H2,530mm, body weight 2,550kg, and lead-acid battery-powered. The wheels consist of one rear drive steering wheel and two front steering wheels, each independently controllable. It can move forwards, backwards, and sideways, and turn and spin. In addition, translation and rotation can be

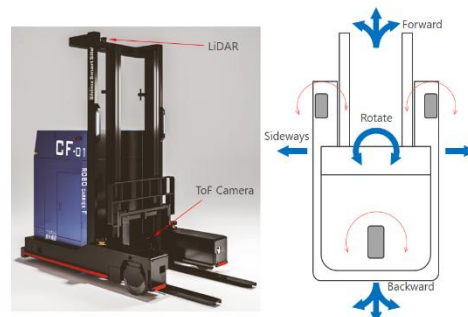


Figure 7 Automated transport forklift

controlled simultaneously when driving or positioning, eliminating the need to turn the vehicle body around. Due to space constraints, there may not be enough space for turning around at construction sites, and a small turning radius for changing direction is a significant advantage.

3.3 Equipment Configuration

As shown in Figure 8, the control equipment configuration of the automated transport trolley and forklift consists of an integrated control PLC (Programmable Logic Controller) that performs control calculations and a vehicle control PLC that performs steering, driving, and lifting operations according to commands from the former. The system consists of a communication device to receive commands from a server, SLAM (Simultaneous Localization and Mapping) controller and LiDAR sensor to obtain the self-position, and a camera controller and camera sensor to acquire material information. The automated transport trolley uses a stereo camera, and the forklift uses a TOF (Time of Flight) camera. The integrated control PLC and vehicle control PLC share status information. Based on the operation instructions from the server, the integrated control PLC performs calculations of instruction values required for wheel steering angle, travel speed, lift movements, etc., using information such as self-position from the SLAM controller and material coordinates from the camera controller. The calculated results are instructed to the vehicle control PLC, which sends operation commands to the motor driver and lift to perform driving control and loading/unloading operations. The vehicle control PLC manages safety sensors, and in case those safety devices are activated, it executes emergency stops or temporary halts based on its determination. The integrated control PLC shares the current operating status with the vehicle control PLC and sends the status information, including its current position, to the server.

3.4 Self-Position Detection

A LiDAR sensor for self-position detection is mounted on top of the vehicle for autonomous driving. This sensor measures the shape of surrounding objects in a building in a 270-degree horizontal plane. Generally, a map for self-position estimation is created from the measurement data, but in this system, map data is generated based on CAD data of architectural drawings. Among the architectural drawings, the system uses two superimposed drawings, one containing structural elements such as columns and beams and finishing materials such as interior walls and doors, and the other containing temporary equipment such as construction elevators. This drawing data is converted to map data using dedicated software, eliminating the need for prior

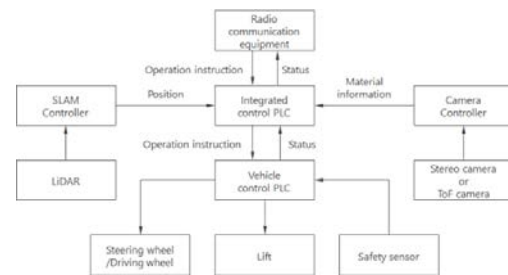


Figure 8 In-vehicle equipment configuration

map creation and origin alignment tasks. In addition, objects that are not included in the drawing data may be measured at the site. If these objects occupy a large proportion of the area, errors may occur in self-position estimation, so the system has a function that overwrites the map data with data obtained from measurements to generate a map according to the site environment. This enables travel control by obtaining information on the planar position and direction of the vehicle body.

3.5 Material Position Detection

Travel control based on self-position estimation enables the robot to travel autonomously within the site, but when loading and unloading materials, the materials are not always in the proper position, so alignment with the object is necessary. To insert the fork appropriately without causing collapses or collisions, the transport robots have camera sensors that detect the position of materials, allowing them to face the material in the correct position. The transport trolley uses a stereo camera, while the forklift uses a ToF camera, so the basic system differs. The trolley's camera, which detects dedicated markers on the pallets, is installed at the fork storage position underneath the body. The marker is attached to the center front of the pallet, as shown in Figure 9. Although there are other methods like affixing markers at both ends of the pallet [9][10], a small marker is placed in the center of the pallet to ensure they fit accurately within the camera's field of view even when the vehicle is close to the pallet. By detecting markers with a stereo camera, the robot can detect the planar position of the pallet.

The camera sensor of the automated transport forklift is mounted at the base of the fork. The ToF camera, as shown in Figure 10, can acquire 3D data of the target material. From the data obtained, the fork pockets of the pallet can be detected, as shown in Figure 11. It detects the space below the material for basket trolleys and



Figure 9 Skid pallet marker

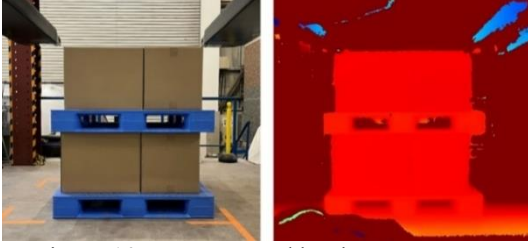


Figure 10 Data captured by the ToF camera



Figure 11 Pallet and material detection

similar items. Furthermore, information on rotation around the vertical axis can be obtained from the depth distance in the flat area of the material or pallet. The camera sensor allows the vehicle to face the material's center. It can also recognize stacked materials and detect the top part of the load, enabling operations like stacking multiple materials or taking materials from the top of a stack.

3.6 Travel Control

When the transport robot moves, the integrated management system sends the node coordinates of the travel path, the orientation of the robot, and the passing method at the node. When traveling a straight line connected by two points P_i, P_{i+1} , the system controls the robot to follow this line. As shown in Figure 12, the robot's self-position (X_C, Y_C, θ_C) and the distance d_n from the target line can be expressed using the line equation as:

$$d_n = \frac{aX_C + bY_C + c}{\sqrt{a^2 + b^2}} \quad (1)$$

$$\begin{cases} a = -Y_{i+1} + Y_i \\ b = X_{i+1} - X_i \\ c = X_i Y_{i+1} - X_{i+1} Y_i \end{cases} \quad (2)$$

and the angular difference between the vehicle and the target line can be expressed as:

$$(\gamma - \theta_C) \quad (3)$$

Using the translational control input gain K_{dp} , and rotational control input gain $K_{\theta p}$, and the translational control input φ_d and rotational control input φ_θ are expressed as:

$$\varphi_d = -K_{dp} \tan^{-1} \left(\frac{d_n}{1000} \right) + (\gamma - \theta_C) \quad (4)$$

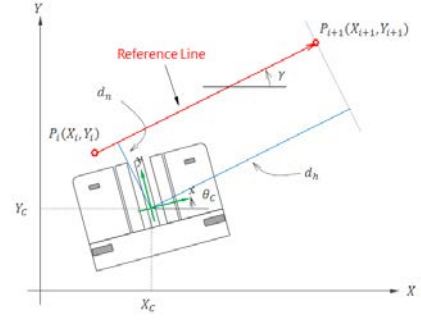


Figure 12 Line tracking control model

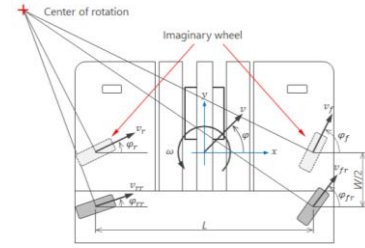


Figure 13 Steering angle calculation model for trolley

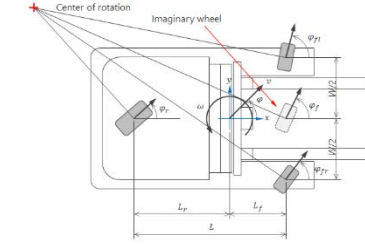


Figure 14 Steering angle calculation model for forklift

$$\varphi_\theta = K_{\theta p}(\gamma - \theta_C) \quad (5)$$

Here, \tan^{-1} is used in Equation (4) to prevent large control inputs proportional to position deviation by asymptotically approaching $\pi/2$ because the control input is handled in the steering angle dimension. As a result, when the distance is greater than a certain level, an almost maximum rudder angle is input.

Here, when the distance between the target coordinate point P_{i+1} and the line perpendicular to the target straight line is d_h , the following equation is obtained:

$$d_h = \sqrt{(X_{i+1} - X_C)^2 + (Y_{i+1} - Y_C)^2 - d_n^2} \quad (6)$$

The speed is decelerated in steps according to the distance to the destination by calculating the speed sequentially during driving, and the stopping and turning operations are performed.

Since the driving wheels of the automated transport trolley are located off the center of the vehicle, the two virtual wheels placed in the center of the vehicle are considered, as shown in Figure 13, to simplify the calculation and provide symmetry. For the calculated control inputs, when the rudder angle of the front and rear virtual wheels is φ_f, φ_r , the rudder angle of the virtual wheels when moving forward is as follows:

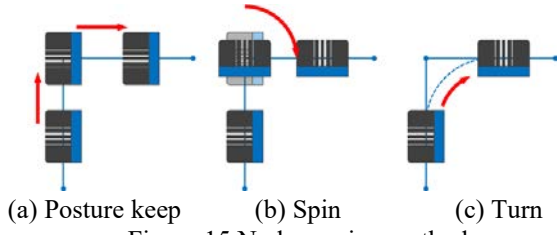


Figure 15 Node passing method

$$\begin{cases} \varphi_f = \varphi_d + \varphi_\theta \\ \varphi_r = \varphi_d - \varphi_\theta \end{cases} \quad (7)$$

the rudder angle of the virtual wheel when moving backward is as follows:

$$\begin{cases} \varphi_f = \varphi_d - \varphi_\theta \\ \varphi_r = \varphi_d + \varphi_\theta \end{cases} \quad (8)$$

The actual rudder angles of the vehicle's driving wheels φ_{fr} , φ_{rr} are calculated as:

$$\varphi_{fr} = \tan^{-1} \left(\frac{2L \sin \varphi_f \cos \varphi_r}{2L \cos \varphi_f \cos \varphi_r + W \sin(\varphi_f - \varphi_r)} \right) \quad (9)$$

$$\varphi_{rr} = \tan^{-1} \left(\frac{2L \cos \varphi_f \sin \varphi_r}{2L \cos \varphi_f \cos \varphi_r + W \sin(\varphi_f - \varphi_r)} \right) \quad (10)$$

In this way, the control system can track the target in a straight line. In the case of sideways movement, two wheels are driven separately, but the control input is calculated in the same way as above, and the straight-line tracking control is performed based on the speed difference of each driving wheel based on the sum of the translation control input and the turning control input. However, since translation and turning are dependent, it takes time to converge compared to traveling.

Similarly, the automated transfer forklift is modeled using virtual wheels, and how to calculate control inputs is the same as for the trolley. As shown in Figure 14, the steering angle of the virtual wheels during forward movement is

$$\begin{cases} \varphi_f = \varphi_d + \tan^{-1} \left(\frac{2L_f}{L} \tan \varphi_\theta \right) \\ \varphi_r = \varphi_d - \tan^{-1} \left(\frac{2L_r}{L} \tan \varphi_\theta \right) \end{cases} \quad (11)$$

and the steering angle of the virtual wheels when moving backward is

$$\begin{cases} \varphi_f = \varphi_d - \tan^{-1} \left(\frac{2L_f}{L} \tan \varphi_\theta \right) \\ \varphi_r = \varphi_d + \tan^{-1} \left(\frac{2L_r}{L} \tan \varphi_\theta \right) \end{cases} \quad (12)$$

Where, the \tan^{-1} term is corrected for the misalignment of the center of rotation due to the geometric relationship between the front and rear wheels. Similarly, if the steering angle of the actual wheels is φ_{fr} , φ_{fl} ,

$$\varphi_{fl} = \tan^{-1} \left(\frac{2L \sin \varphi_f \cos \varphi_r}{2L \cos \varphi_f \cos \varphi_r - W \sin(\varphi_f - \varphi_r)} \right) \quad (13)$$

$$\varphi_{fr} = \tan^{-1} \left(\frac{2L \sin \varphi_f \cos \varphi_r}{2L \cos \varphi_f \cos \varphi_r + W \sin(\varphi_f - \varphi_r)} \right) \quad (14)$$

The rear wheel steering angle is obtained using the virtual wheels. Since all three wheels of an automated transport forklift can be steered, translation and turning can be controlled independently, even when traveling sideways.

As shown in Figure 15, there are three types of node passing methods: (a) maintaining posture, (b) spinning, and (c) turning. In maintaining posture, the robot stops at a node and steers the wheels to change the direction of travel from traveling forward to traveling sideways. In spinning, the robot stops at the node and rotates in place to a predetermined angle. In turn, the robot follows an arc connecting two straight lines to pass the node. Turning is the quickest way to pass through the three passing methods, while spinning takes the most time.

3.7 Material Positioning Control

The self-position estimation using LiDAR has an error of approximately $\pm 50\text{mm}$ and $\pm 1\text{deg}$. It is necessary to allow some errors when traveling. For tasks like loading and unloading, where positioning relative to the material is critical, a camera sensor is used to ensure accuracy to prevent misalignment of materials. As shown in Figure 16, the camera sensor provides a target coordinate $P_V(X_V, Y_V, \theta_V)$ for a pallet or other target. This provides a perpendicular line from the center surface of the target object, and by controlling the machine to follow this line, it can face the object's center.

For the automated transport trolley, the trolley moves sideways, when the sensor located at the base of the fork detects contact with a material, the robot stops moving forward and places it onto the vehicle by lifting, moving back, and lifting down the fork. This action allows the material to be taken in, and the reverse action allows it to be unloaded.

For the automated transport forklift, as shown in Figure 17 left side, after moving to the front of the material storage area, the lift is raised, assuming that the material is stacked, and the top edge of the material is searched with a camera sensor. When the top edge is detected, it aligns the lift to the height of the fork pocket, aligns the body in front of the material while detecting the position and rotation of the pallet, and picks up the

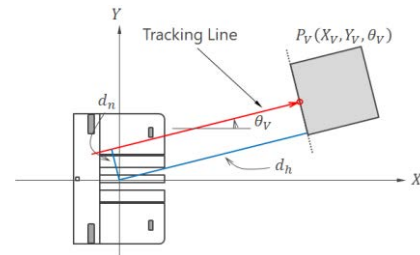


Figure 16 Positioning control with the camera sensor

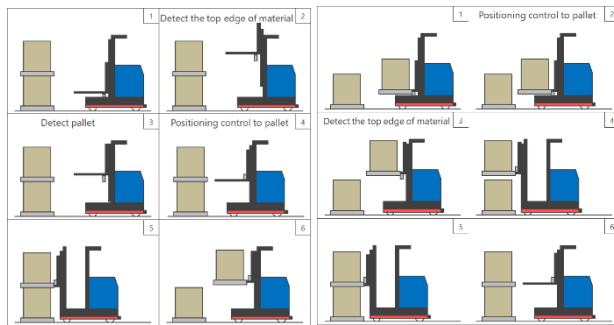


Figure 17 Loading and unloading control

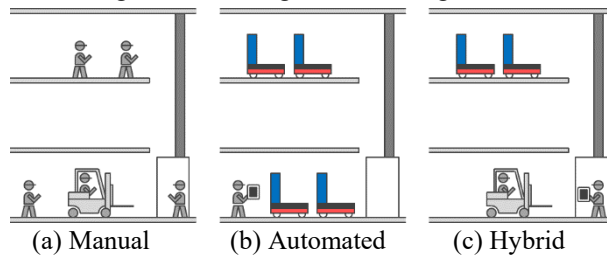


Fig.18 Transportation structure

material. The pick-up task is completed as follows: the robot moves forward while extending out the fork, stops when the sensor at the fork base detects contact with the material, lifts the fork until the sensor on the top surface of the fork detects contact with the material, stops the lift, then moves the fork back, and moves backward.

For unloading, as shown in Figure 17 right side, it detects the lowest pallet in the storage area, detects the position and rotation of the pallet, and aligns the body in front of the material. After that, it searches the top edge of the material and lifts the load higher than it to some extent before placing the load on it. If no material is in the storage area, the robot places the load directly on the space. The unloading task is completed as follows: the robot moves forward to a predetermined distance while moving the fork forward, stops, lifts the fork down, and after the sensor on the fork top detects the separation from the material, stops the lift, moves the fork back, and moves backward. These operations enable placing flat or stacking materials.

In the case of a material rack, the robot receives information on the height of the material storage area from the integrated management system, and the lift is raised to a predetermined height in advance to perform loading and unloading tasks.

4 Testing and Operation

4.1 Operation of Automated Transport Trolleys

An automated transport trolleys was applied to a construction site, and many materials were transported. We are transported at night when no workers are present, rather than during the daytime when there is heavy

worker traffic, because safety laws and rules have not yet been established. While this ensures safety, it is also more efficient because the transport area is larger, elevators can be occupied in an environment where no workers are present.

Conventional transport operations, materials are brought in by truck on the first floor, they are loaded onto the elevator using a forklift. The elevator is operated to move to the upper floors, and an operator is required because the elevator is for construction work. On the upper floors, due to floor weight restrictions, forklifts are not generally used, and hand lifts or dollies are used to remove materials from the elevator and place them in the storage area. As shown in Figure 18, (a) manual work often requires a total of five workers: one forklift operator, one loading operator on the first floor, one elevator operator, and two workers on the upper floor. In contrast, (b) when an automated transport system is used, one system operation manager and two to four automated transport trolley are used: one to two on the loading floor and one to two on the upper floors.

The time required for manual transfer was 24 pallets/hour, or 2.5 minutes per pallet, while that for the automated transport system was 12 pallets/hour, or 5 minutes per pallet. Therefore, although automatic transfer takes longer than manual transfer, it requires only one person, thus saving labor. The number of pallets transported per person is approximately 5 pallets for manual transport and 12 pallets for automated transport, resulting in a higher workload per person. On the other hand, the automated transport system requires the cost of automated transport trolleys and automated elevator equipment. The cost of an automated transport trolley is about the same as the cost of a worker, but the full cost of the elevator automation system is additional. Therefore, the cost can be reduced by using manual operation of the elevator. In addition, since forklifts are generally located on the first floor, the work on the loading floor can be done manually. As shown in Figure 16(c), if only unloading on the upper floor is automated as a hybrid system of automated and manual work, work can be performed by two people: one forklift operator and one elevator operator who also manages the system operation. Under this system, work can be performed at a rate of 20 pallets/hour, 3 min./pallet. This is a cost-effective balance between the speed of manual work and the manpower savings of automated transport.

Manual loading and unloading requires many people to load and unload the elevator. On the other hand, automated transport reduces the number of workers, but it is slower than manual transport. Having only one automated transport trolley on a floor causes significant delays, so having two trolleys on a floor is more efficient. This reduces the time delay for unloading on the upper floor compared to manual operation. However, loading

on the first floor is more efficient without automation because of the time difference compared to manually operated forklifts.

4.2 Testing of Automated Transport Forklifts

The automated transport forklifts were tested in a test site and logistics facility, moving stacked materials, materials stored in racks, and unloading materials from trucks. In the truck unloading test, pieces of material loaded on a truck were taken out from the left and right sides and stacked in two or three stages at a storage area. The automated transport forklift took 45 minutes for this task, while a manually operated forklift took about 20 minutes, depending on the operator's skill level. The work time is roughly twice as long as that of a forklift operated by a person. At the logistics facility, the prescribed transportation was performed while moving over distances of several tens of meters. Although the loading and unloading operations took longer with the robot, there was little difference in the travel time. Therefore, when the traveling distance is long, the performance difference between humans and robots is further reduced.

5 Conclusion

We built a transport system to automate transport work on site, where the management system, operation terminals, and transport robots all work together to automate the transfer of materials. The system has been proven applicable even in environments without adequate facilities, such as construction sites. Although the work of a single transfer robot is slower than that of a human operator, production efficiency can be increased by operating multiple robots with fewer personnel. Transport robots follow predetermined routes and perform searches and positioning in front of the material during loading and unloading, while the human operator flexibly selects a route according to the surrounding conditions, checks material conditions in advance, adjusts the height and position and makes decisions. This makes a difference in work speed. In some cases, due to safety concerns, the traveling speed limit is applied to automated transfer robots. While there is ample room for further improvement, achieving the same operation time as humans with current technology is challenging, and operating multiple units with a small workforce is the current solution.

Automation enables efficient use of production time, improves safety and productivity, and helps solve problems such as labor shortages. However, changing jobs from performing material transportation to operating transport robots will change the human resources required, so personnel development, training, and system usability will also become essential. In terms of operation,

it will be necessary to build work processes and facility layouts to incorporate the automated transport system, indicating the necessity for solutions and improvements in both technology and operation.

References

- [1] Worker Shortages in Manufacturing Make AGVs More Beneficial Than Ever, <https://www.fredagv.com/news/worker-shortages-in-manufacturing-make-agvs-more-beneficial-than-ever/>, 2022.
- [2] Jianqi Zhang, Xu Yang, Wei Wang, Jinchao Guan, Ling Ding, Vincent C.S. Lee, "Automated guided vehicles and autonomous mobile robots for recognition and tracking in civil engineering", *Automation in Construction* Vol.146, pages 1–24, 2022.
- [3] Guang Yang, Shuoyu Wang, Hajime Okamura, Bo Shen, Yasuhiro Ueda, Toshiaki Yasui, Tetsuya Yamada, Yuki Miyazawa, Satomi Yoshida, Yuta Inada, Shingo Ino, Kazuo Okuhata, Yoshinobu Mizobuchi, "Hallway exploration-inspired guidance: applications in autonomous material transportation in construction sites", *Automation in Construction* Vol.128, pages 1-14, 2021.
- [4] Murata Machinery, LTD. <https://logistics.muratec.net/products/xio/>
- [5] DAIHUKU, <https://www.daifuku.com/solution/intralogistics/products/vehicle/>
- [6] Akhilesh Bhat, Natsuki Kai, Takayuki Suzuki, Takahiro Shiroshima, Hiroshi Yoshida, "An advanced autonomous forklift based on a networked control system", *IFAC-PapersOnLine*, Vol.56, Issue 2, pages 11444-11449, 2023.
- [7] Vicent Girbés, Leopoldo Armesto, Josep Tornero, "Path following hybrid control for vehicle stability applied to industrial forklifts", *Robotics and Autonomous Systems*, Vol.62, Issue 6, Pages910-922,2014.
- [8] Yuki Miyashita, Tomoki Yamauchi, Toshiharu Sugawara, "Distributed Planning with Asynchronous Execution with Local Navigation for Multi-agent Pickup and Delivery Problem", *AAMAS 2023*, pages914-922, 2023.
- [9] Michael Seelinger, John-David Yoder, "Automatic visual guidance of a forklift engaging a pallet", *Robotics and Autonomous Systems* Vol.54, pages1026-1038, 2006.
- [10] N. Bellomo, E. Marcuzzi, L. Baglivo, M. Pertile, E. Bertolazzi, M. De Cecco, "Pallet Pose Estimation with LIDAR and Vision for Autonomous Forklifts", *IFAC Proceedings Volumes*, Vol.42, Issue 4, pages612-617, 2009.

Automated As-built 3D Reconstruction Using Quadraped Robot and 3D LiDAR Sensor

Vincent J.L. Gan^{1,2}, Difeng Hu¹, Ruoming Zhai^{1,3}, Yushuo Wang¹

¹ Department of the Built Environment, National University of Singapore, Singapore

² Centre of 5G Digital Building Technology, National University of Singapore, Singapore

³ School of Geodesy and Geomatics, Wuhan University, China

vincent.gan@nus.edu.sg (correspondence), e0788081@u.nus.edu,
ruomingzhai@whu.edu.cn, wangyushuo@u.nus.edu

Abstract –

Advancements in automated 3D scene reconstruction are crucial for accurately documenting and sharing knowledge about the current state of the building and infrastructure facilities. Traditional methods of 3D reconstruction involve laser scanning to capture the building's as-built conditions, which is often a tedious and time-consuming process. This study introduces an automated as-built 3D reconstruction method, with the aid of quadraped robot and 3D LiDAR sensor. The core of this approach is the integration of LiDAR Simultaneous Localisation and Mapping (SLAM) algorithm with robotic control, aiming to achieve detailed 3D/2D mapping and complete point cloud data acquisition. A practical field experiment was conducted to test this approach in reconstructing the building interiors. The results demonstrate the effectiveness of this novel 3D scanning method in achieving high completeness and density in point cloud data.

Keywords –

Quadraped Robot; LiDAR; 3D Reconstruction; Mobile 3D Scanning; Dense Measurement

1 Introduction

Condition monitoring, inspection and quality assessment of architecture, engineering, and construction (AEC) projects requires accurate data with respect to the as-built conditions. This process requires laser scanning to automatically capture 3D as-built point cloud data, followed by manual processing of the captured point clouds to create 3D scene, which is laborious, time-consuming, and error-prone. While automatic 3D BIM reconstruction is an emerging vision, the development still faces exceptional difficulties. Several studies [1-5]

attempted to use Light Detection And Ranging (LiDAR) integrated with quadraped robots or unmanned ground vehicles for 3D scene reconstruction of the built environment.

UGVs, equipped with LiDAR, were able to devise a route, navigate towards specific targets, and gather point cloud data from the built environment. Unmanned Aerial Vehicles (UAVs) were tasked with the efficient collection of radiation data by surveying outdoor areas, a method proving more efficient than that of Unmanned Ground Vehicles (UGVs). In a notable development, Kandath et al. [6] introduced a UAV-UGV collaborative system designed to enhance UGV navigation within indoor settings through UAV-provided obstacle location data. Similarly, Peterson et al. [7] crafted a multi-robot system integrating a UGV with a multirotor drone, aimed at improving exploration and navigation efficiencies in outdoor environments. Kim et al. [8] furthered this innovation by creating a UAV-aided automated framework for task execution within cluttered spaces, deploying UAVs initially to generate a 3D terrain map that includes obstacle details through the processing of photographic imagery, thereby optimising ground robot navigation paths.

Despite its success in outdoor point cloud acquisition, LiDAR-equipped mobile laser scanning faces significant challenges in indoor environments due to GPS limitations, restricting its application in these settings. To overcome this, researchers are exploring the use of Simultaneous Localisation and Mapping (SLAM) algorithms as an alternative to global navigation for indoor 3D reconstruction [9, 10]. SLAM functions by determining the scanner's location within an environment while concurrently mapping the 3D scene using LiDAR sensors. This approach is advantageous as it does not rely on GPS signals, allowing the scanning system to be robot-mounted for real-time tracking and 3D scene creation of interior spaces. Specifically, incorporating a

robotic dog into this system could significantly enhance indoor navigation capabilities, thanks to its advanced maneuverability and stability in varied terrains. This integration promises to refine scan planning algorithms substantially, leading to the generation of higher quality point cloud data. However, current research in this domain is still evolving and has yet to achieve optimal scanning quality [11]. Existing robot-assisted LiDAR sensing method can hardly capture the completed dense 3D measurement of the building's as-built conditions because the SLAM algorithms have yet fully incorporated the scanning requirement for dense measurement. Integration of the point cloud data quality and completeness over 3D scanning into robot navigation and control for optimising the scanning process is a central task.

This study introduces a new automated as-built 3D reconstruction method, with the aid of quadruped robot and 3D LiDAR sensor to obtain completed and accurate scanned data in regard to the as-built conditions. This requires the use of autonomous robot navigation, localisation, and mapping algorithms in a GPS-denied environment while LiDAR sensor captures dense 3D measurement of buildings. The application provides potential added values to automated smart city reconstruction, constant monitoring, surveillance of built environment, and smart facilities maintenance management.

In line with academia's research agenda, this study aims to develop a new 3D BIM reconstruction approach for automatic point cloud data acquisition. This contributes to the creation of new knowledge, method, and algorithm in the field of automated 3D scanning and as-built BIM reconstruction. The conventional approach separates the 3D scanning and mapping. The study introduces a quadruped robot-assisted automatic scanning and mapping approach utilising SLAM to capture updated as-built information. This research enriches our understanding by integrating robot navigation and optimal pathfinding with SLAM algorithms, enabling comprehensive 3D measurements of as-built conditions.

2 Methodology

As discussed previously, integration of the robotic navigation with LiDAR SLAM is needed to reconstruct the completed and accurate point clouds with respect to the as-built conditions. The requirement for LiDAR scanning to capture the as-built conditions is integrated with robot navigation. The proposed new method controls robot navigation by taking account of 3D scanning, which is different from previous studies.

In this study, a quadruped robot, enhanced with the high-resolution Ouster OS1 LiDAR sensor, is developed

for detailed 3D data capture in complex settings. This research proposes two 3D scanning approaches: the first integrates 3D reconstruction with 2D mapping using only the 3D LiDAR sensor, while the second employs an additional 2D LiDAR sensor for navigation, decoupling 3D reconstruction from 2D mapping. Then, an experiment is undertaken in NUS campus to compare and validate the proposed quadruped robot-based LiDAR scanning as a viable approach. Detailed methodology is presented in the following sections.

2.1 Quadruped robot-based LiDAR scanning

Figure 1 presents the configuration of the quadruped robot equipped in this study, which is augmented with the Ouster OS1 LiDAR sensor. Each of the robot's legs is powered by three high torque density electric motors—referred to as the knee motor, thigh motor, and hip motor—which govern the movement of the corresponding joints. This control system allows the quadruped robot to perform a variety of maneuvers such as walking, leaping, and executing pitch, roll, and yaw actions to facilitate 3D scanning in confined and cluttered environments such as indoor. The Ouster OS1 LiDAR sensor is a state-of-the-art device designed for high-resolution 3D data acquisition, providing unparalleled precision and range for various applications. The LiDAR sensor is able to capture detailed spatial information, making it an ideal component for integrating with robotic systems in complex environments.



Figure 1. Quadruped robot equipped with 3D LiDAR sensor for automated 3D scanning

In this study, two 3D scanning solutions are proposed. The first solution couple 3D reconstruction and 2D mapping (for robot navigation) together using solely the 3D LiDAR sensor. Alternatively, the second solution decouples the 3D reconstruction from the 2D mapping by employing an additional 2D LiDAR sensor dedicated

exclusively to robot navigation. The objective of contrasting these two solutions is to pinpoint any hardware or software challenges that may arise during the scanning process. Detailed methods are discussed in the followings.

2.1.1 Coupled 3D Reconstruction and 2D Mapping

Figure 2 depicts the coupled 3D reconstruction and 2D mapping for data acquisition and navigation. In this dual-processing technique, a 3D LiDAR scanner collects point cloud data to create a detailed three-dimensional reconstruction of the environment. Concurrently, this 3D point cloud data is processed and translated into a two-dimensional map. This 2D map, while less detailed in terms of vertical information, is crucial for navigation purposes as it provides a simplified and accessible overview of the environment, highlighting pathways and obstacles relevant for guiding quadruped robot.

The key advantage of this coupled method is that it allows for simultaneous localisation and mapping in both three and two dimensions. The 3D reconstruction offers a high-fidelity representation of the environment, which is essential for tasks that require depth information. Meanwhile, the 2D map facilitates efficient path planning and real-time navigation, providing a clear layout of the environment that can be easily interpreted by robotic navigation algorithms.

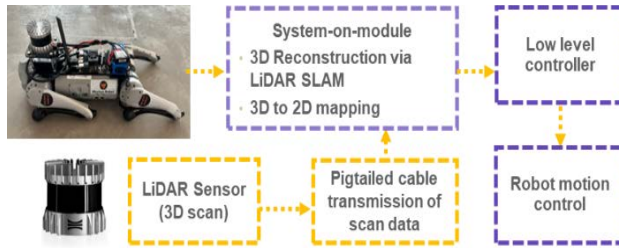


Figure 2. Proposed solution based on coupled 3D reconstruction and 2D mapping

2.1.2 Decoupled 3D Reconstruction and 2D Mapping

Figure 3 displays decoupled 3D reconstruction and 2D mapping where the generation of a three-dimensional model and the facilitation of navigation are separated. Here, a 3D LiDAR sensor gathers point cloud data to construct a three-dimensional representation of the environment. In parallel, a separate 2D LiDAR sensor (e.g., RPLIDAR) operates independently, scanning the environment at a single elevation to collect planar point cloud data. This data is then processed through mapping algorithms to create a two-dimensional map. Unlike the

complex 3D model, the 2D map offers a streamlined, birds-eye view of the terrain, instrumental for navigation purposes. The map simplifies the environment to its basic contours and obstacles, providing the necessary information for pathfinding and movement within a space.

This decoupled approach ensures that the 3D reconstruction can be rich in detail and structure without being encumbered by the requirements of real-time navigation. Meanwhile, the dedicated 2D mapping allows for quick and responsive navigation solutions, optimized for real-time operational needs. Together, they form a dual system that caters to the high-resolution modelling and efficient navigational demands of the quadruped robot.

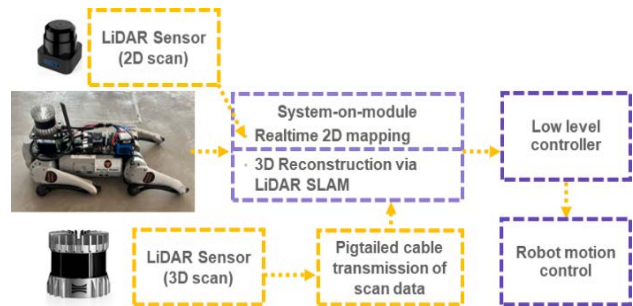


Figure 3. Proposed solution based on decoupled 3D reconstruction and 2D mapping

2.2 3D Reconstruction Using LiDAR SLAM

In this study, LiDAR SLAM is leveraged to reconstruct the point clouds for representing the as-built conditions. Specifically, Lidarslam_ros2 [12] is used to reconstruct the 3D scene of point clouds. It features offline reconstruction that allows for processing all current and past observation data to estimate LiDAR sensor poses and the map, which is advantageous to online reconstruction that relies on current data.

The present study uses Lidarslam_ros2 because comprehensive data analysis is required to produce a precise 3D model. By considering the full spectrum of LiDAR data, the algorithm allows for a more thorough and potentially more accurate pose estimation of the LiDAR sensor at different points in time, leading to a refined and cohesive map. This is particularly beneficial in environments where the LiDAR sensor might experience occasional obstructions or in scenarios where every detail is critical for the creation of an accurate 3D representation. It follows a general LiDAR slam to reconstruct the 3D scene:

- 3D LiDAR sensor collects 3D point clouds of

the environment (initial scan).

- Odometry (abbreviated as odom) information is calculated to estimate change in position over time. It is used to estimate the 3D LiDAR sensor or quadruped robot position relative to the starting location.
- When travelling to the next positions, new 3D point clouds are collected and matched with previously collected data using the odometry information to form partially the 3D scene.
- Error accumulates when mapping the point clouds. Thus, when the sensor revisits a previously mapped area, loop closure is detected to optimise the 3D scene.

2.3 2D Mapping for Robot Navigation

2D mapping aims to localise the robot's positions and guide its movement. In this study, the movement of the quadruped robot is controlled remotely through either by entering specific coordinates or by interactively selecting the destination on a 2D map, which then prompts the robot to move accordingly.

Since two contrastive solutions are proposed in this study, the source of information for 2D mapping might vary. For **coupled 3D reconstruction and 2D mapping**, the scan data from the 3D LiDAR sensor are converted to 2D for the first instance, and then analysed with the odometry obtained from the quadruped robot to generate the 2D map. This is because Lidarslam_ros2 in this study is an offline 3D reconstruction algorithm, which couldn't produce the odometry in real-time to support the 2D mapping.

For **decoupled 3D reconstruction and 2D mapping**, a 2D LiDAR sensor (e.g., RPLIDAR) is used to localise the robot's positions and guide its movement. It follows the same regime as 3D reconstruction, which first generates point clouds for one scan, and then automatically registers point clouds from subsequent scans in the shared global frame of coordinates by estimating the scanner's real-time position and orientation. Figure 4 shows the transformation of the information from 2D LiDAR sensor into the 2D map. Typically, rplidar_node is meant to interface with the RPLIDAR sensor and provide scan data. However, it doesn't inherently provide odometry data, as odometry usually comes from motion sensors, encoders, or other motion-related sources in robots. That said, the RPLIDAR sensor derive "pseudo" odometry from the 2D scan data. This is usually achieved through scan matching techniques that estimate the robot's motion by aligning consecutive LIDAR scans. In this study, a ROS package called GMapping is used to undertake the 2D mapping, which then guides the quadruped robot in real-time navigation in the built environment.

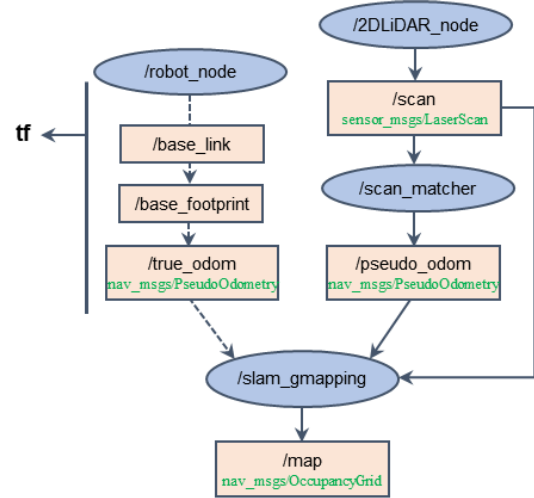


Figure 4. tf tree for 2D mapping

2.4 Field Testing

The Quadruped robot-based LiDAR scanning in this study is validated through field tests conducted on the National University of Singapore (NUS) campus, particularly focusing on the scanning of one educational building. For these tests, a 3D LiDAR sensor is installed and configured on the robot, which then navigate autonomously and perform 3D measurements of the built environment (see Figure 5). This setup aims to rigorously assess the robot's capabilities and its effectiveness in capturing detailed spatial data.



Figure 5. 3D view of the experimental site

Following the completion of the experiments, the gathered point clouds undergo a thorough evaluation, focusing on two key parameters: the density of the point clouds and their completeness. The success of these tests demonstrates the potential of this approach in providing accurate as-built 3D measurements of the built environment, using robotics and LiDAR sensor.

3 Results and Discussion

3.1 Scanned 3D Scenes

Figure 6 shows the 2D map and the quadruped robot movement trajectory. As mentioned, the movement of the quadruped robot is controlled remotely through entering specific coordinates or by interactively selecting the destination on a 2D map, which then prompts the robot to move. The experimental site is the corridor outside a series of research offices and laboratories. Once inputting the destination, the quadruped robot moves along the corridor while scanning the built environment. Figure 7 depicts the point clouds in the reconstructed 3D scene offline using Lidarslam_ros2. The results demonstrate that mobile 3D scanning with the quadruped robot and 3D LiDAR sensor yields decent point cloud quality that surpasses that of previous studies, notably without any point cloud mismatching.

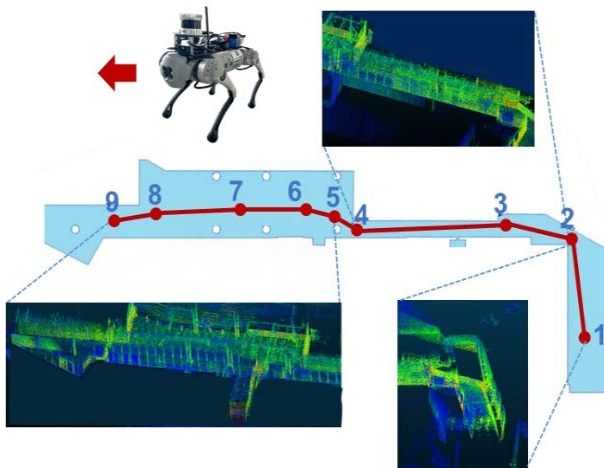


Figure 6. 2D map and the quadruped robot movement trajectory

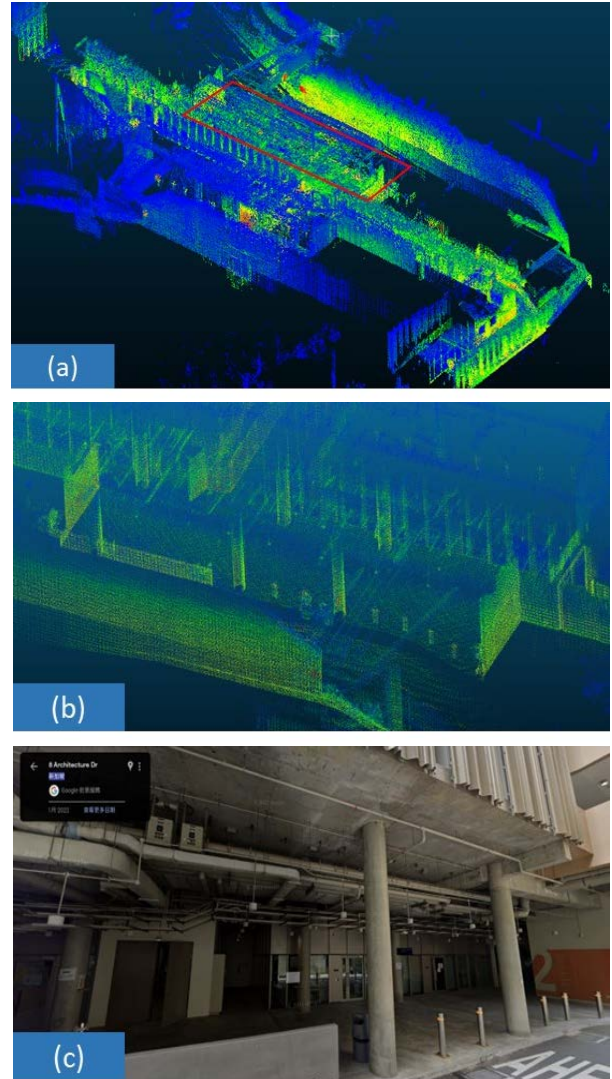


Figure 7. (a) Reconstructed 3D scene offline using Lidarslam_ros2; (b) As-built conditions of the building components; (c) Overview of the 3D scene

3.2 Comparison of Coupled and Decoupled 3D Reconstruction & Navigation

While promising results are obtained using quadruped robot and 3D LiDAR sensor. The system is not without issues. For example, integrating 3D scan to 2D mapping has encountered problems because the objective of 3D scan, which emphasises accuracy, is far different from 2D mapping, which focuses on real time processing and perception. The issue arises in terms of hardware and software perspectives.

When using a wireless connection between the 3D LiDAR sensor (through an additional NX Xavier) and the robot's system-on-module, data transfer is restricted due to wireless bandwidth limitations. The NX Xavier might

also face constraints in caching the scanned data due to its limited computational capacity.

Transitioning to a direct cable connection for scan data to the robot's system module improves data transmission. Yet, challenges arise when converting 3D scan data to 2D and performing real-time mapping. The resultant 2D map lacks proper organisation due to insufficient odometry details from the 3D LiDAR, making it unsuitable for navigation. This is because the LiDAR SLAM algorithm in this study is offline 3D reconstruction, and the “pseudo” odometry information is calculated after scanning is completed which cannot be used for the real-time 2D mapping. Consequently, odometry data from an alternative source, such as the quadruped robot, is utilised for creating the 2D map as shown in Figure 8. However, this approach may compromise the map's quality due to potential discrepancies in data integration. A preferable solution is to use separate 3D and 2D LiDAR sensors for independent 3D reconstruction and 2D mapping.

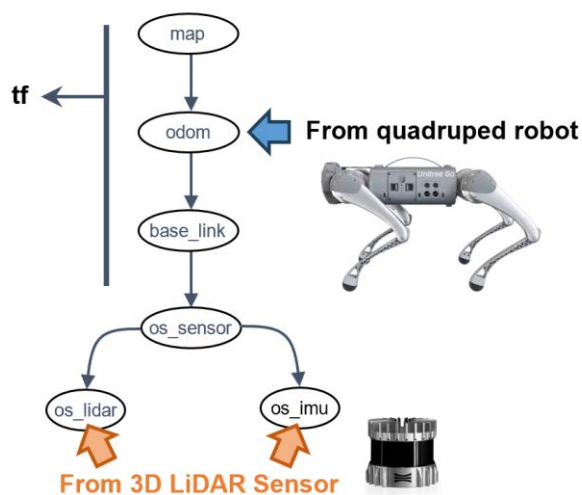


Figure 8. Use of quadruped robot odometry with 3D LiDAR information for 2D mapping

4 Conclusion

This research presents a novel automated as-built 3D reconstruction technique utilising a quadruped robot and a 3D LiDAR sensor. At the forefront of this method is the integration of the SLAM algorithm with enhanced robot motion control for automatic 3D scanning, which facilitate the generation of as-built point cloud data for buildings or infrastructure facilities. This synergy aims to facilitate comprehensive 3D mapping and ensure the thorough collection of point cloud data. A field experiment focusing on the reconstruction of building

interiors was conducted. The outcomes of this experiment highlight the mobile scanning method's capability in delivering point cloud data with remarkable completeness and density, showcasing its potential in advanced 3D reconstruction applications.

However, this study has certain limitations, which the research team will further address in future studies. Firstly, the optimal scan positions and robot control strategy need meticulously crafted to optimise the point clouds for both density and completeness. A holistic scan planning optimisation approach, with the consideration of different sensor perception (such as 3D LiDAR sensor, terrestrial laser scanner), will be needed to model the scanning operation and optimise the scan positions. Additionally, the quadruped robot's low-angle viewpoint affects the accurate mapping of features on vertical surfaces and elevated horizontal areas, including MEP utilities and ceiling. The ambient lighting and the limitations in detecting transparent materials such as glass are essential for thorough mapping. The deployment and testing of 3D LiDAR sensors are crucial to verify the proposed scanning method as a robust approach for point cloud data acquisition. This process necessitates a quantitative evaluation, specifically aimed at gauging the completeness and density of the scanned point clouds against benchmarks set by advanced terrestrial laser scans. This evaluation process will focus on assessing the completeness and density of the scanned point clouds. A key method in this evaluation will be the orthographic projection of the point clouds, derived from building elements in 3D space, onto a plane. The projected data will then be converted into a regular grid. This grid-based representation will align with the predetermined minimum requirements for point cloud density in this study, allowing for a more structured and measurable analysis.

The present study evaluates `Lidarslam_ros2`, an offline reconstruction algorithm. Recognising the challenge associated with real-time point cloud processing, this research aims to advance towards potential solutions including online 3D scene reconstruction algorithms and/or ensuring robust bandwidth of connections to facilitate uninterrupted data transmission to a remote server for point cloud processing. Addressing these technical considerations is essential, as it enables future iterations of this research. The outcome of the present study will contribute to the automation and precision of point cloud data acquisition in the built environment. This advancement will enable us to extend our research into the realms of scan-to-BIM and as-built BIM generation.

Acknowledgements

This research is supported by the Ministry of Education Singapore under the Academic Research Fund Tier 1 (A-8001207-00-00) and (A-8002138-00-00). Any opinions and findings are those of the authors and do not necessarily reflect the views of the grantor. The authors are grateful to the NUS Centre of 5G Digital Building Technology for its support of the experiments.

References

- [1] El-Halawany, S.I. and Lichti, D.D., "Detecting road poles from mobile terrestrial laser scanning data." *GIScience & Remote Sensing*, 2013. 50(6): p. 704-722.
- [2] Yen, K.S., Lasky, T.A., and Ravani, B., "Cost-benefit analysis of mobile terrestrial laser scanning applications for highway infrastructure." *Journal of Infrastructure Systems*, 2014. 20(4): p. 04014022.
- [3] Gené-Mola, J., et al., "Fruit detection in an apple orchard using a mobile terrestrial laser scanner." *Biosystems Engineering*, 2019. 187: p. 171-184.
- [4] Rosell-Polo, J.R., et al., "Kinect v2 sensor-based mobile terrestrial laser scanner for agricultural outdoor applications." *IEEE/ASME Transactions on Mechatronics*, 2017. 22(6): p. 2420-2427.
- [5] Hu, D., Gan, V.J.L., and Yin, C., "Robot-assisted mobile scanning for automated 3D reconstruction and point cloud semantic segmentation of building interiors." *Automation in Construction*, 2023. 152: p. 104949.
- [6] Kandath, H., Bera, T., Bardhan, R., and Sundaram, S., "Autonomous navigation and sensorless obstacle avoidance for UGV with environment information from UAV." In *2018 Second IEEE International Conference on Robotic Computing (IRC)*, 2018. IEEE.
- [7] Peterson, J., et al., "Online aerial terrain mapping for ground robot navigation." *Sensors*, 2018. 18(2): p. 630.
- [8] Kim, P., Park, J., Cho, Y.K., and Kang, J., "UAV-assisted autonomous mobile robot navigation for as-is 3D data collection and registration in cluttered environments." *Automation in Construction*, 2019. 106: p. 102918.
- [9] Chong, Z.J., et al., "Synthetic 2D LIDAR for precise vehicle localization in 3D urban environment." In *IEEE International Conference on Robotics and Automation*, 2013, IEEE: Karlsruhe, Germany. p. 1554–1559.
- [10] Hu, D., Gan, V.J., Wang, T., and Ma, L., "Multi-agent robotic system (MARS) for UAV-UGV path planning and automatic sensory data collection in cluttered environments." *Building and Environment*, 2022. 221: p. 109349.
- [11] Aryan, A., Bosché, F., and Tang, P., "Planning for terrestrial laser scanning in construction: A review." *Automation in Construction*, 2021. 125: p. 103551.
- [12] Sasaki, R., "lidar slam ros2." Available at https://github.com/rsasaki0109/lidar slam_ros2, 2022.

Enhancing UAV Efficiency in Bridge Inspection through BIM-based Flight Planning and Image Quality Assurance

Feng Wang¹, Yang Zou^{1*}, Enrique del Rey Castillo¹, James B.P. Lim^{1,2}

¹Department of Civil and Environmental Engineering, University of Auckland, New Zealand

²School of Engineering, University of Waikato, New Zealand

fwan133@aucklanduni.ac.nz, yang.zou@auckland.ac.nz, e.delrey@auckland.ac.nz, james.lim@waikato.ac.nz

Abstract –

Unmanned Aerial Vehicles (UAVs) have been widely applied for bridge inspection, and significant progress has been made to automate the UAV-enabled inspection process. However, current solutions face two primary challenges: (1) they typically rely on pilots' experience to create a pre-defined flight path for automated image collection, and (2) the potential for localisation errors and depth-varying bridge environments may lead to image quality issues such as blurriness, improper exposure, insufficient resolution, and coverage. To overcome these challenges, this paper introduces a novel framework that combines an objective Building Information Modelling (BIM)-based flight mission planning tool with an in-flight image-quality assessment module to ensure the comprehensive capture of high-quality images. A field experiment was carried out on a standard girder bridge to validate the effectiveness of this framework. The results show that the proposed framework: (1) can enable the collection of bridge images via a UAV with minor human intervention and (2) can ensure that the captured UAV images meet the inspection requirements, allowing for the reconstruction of a high-quality digital bridge model and effective detection and quantification of potential defects. The framework marks a pioneering approach for highly automated and efficient bridge inspection using UAVs and is versatile enough for application in diverse real-world bridge scenarios.

Keywords –

Bridge Inspection; Unmanned Aerial Vehicles (UAVs); High-quality Image Collection; Automation; BIM

recommendations [1,2]. In recent years, camera-mounted UAVs have become increasingly popular for bridge visual inspections, potentially replacing the traditional inspection equipment (e.g., rigging, scaffoldings, or heavy bucket trucks) and offering a safer, more cost-effective, and more efficient solution for bridge condition assessment [3].

After years of development, substantial progress has been achieved in UAV-enabled bridge inspection (UBI). Despite efforts made in this field, some aspects remain unaddressed. While current studies predominantly emphasise the automation of image collection, the planning process still relies on human experience, lacking a principled or automated tool for objective image capture plans [4,5]. Furthermore, the quality of images is challenging to predict, as it is influenced by unforeseen factors and can be easily affected by the bridge environment, such as uneven light conditions and varying depth [6,7].

This paper presents a systematic framework designed to automate the image collection process for UBI, which is expected to capture high-quality images sufficient for identifying potential defects and generating a photogrammetric 3D model at a metric scale. The proposed framework employs a BIM model of the target bridge and consists of two stages for achieving high-quality and efficient image acquisition.

The remaining part of this paper is structured as follows. Section 2 provides a comprehensive review of the development of UBI with a particular focus on the image collection process. In Section 3, the details of the proposed method are introduced. Section 4 describes a field experiment to validate the proposed framework. Section 5 discusses the applicability of this framework and suggests future research directions. Finally, we conclude this paper in Section 6.

1 Introduction

Bridges are a vital component of infrastructure systems, which require regular inspections to identify potential defects and provide maintenance

2 Literature Review

This section begins with an overview of UAV-enabled bridge inspection. It then delves into a detailed analysis of the image collection process, with a particular

emphasis on the level of automation (LoA).

2.1 UAV-enabled Bridge Inspection

Over the past few years, UAVs have attracted increasing interest in bridge inspection and made substantial advancements in this field. One of the representative studies was conducted by Seo, et al. [8], who evaluated the efficacy of camera-mounted UAVs as a novel tool for bridge inspection. This study successfully validated the effectiveness of UAV imagery for damage identification and demonstrated that UAVs can perform bridge inspection at a significantly reduced cost compared to conventional methods. Furthermore, Morgenthal, et al. [9] introduced a complete and coherent framework for UAV-based inspections of large bridges to support bridge condition assessment and data management. This framework stands as the pioneering guide for the UBI, encompassing the entire process from image collection and processing to management.

Many research interests have also been paid to investigate the performance of other types of UAV-mounted remote sensing systems in bridge inspection, including thermal cameras [10] and Light Detection and Ranging (LiDAR) sensors [11,12]. For example, Omar and Nehdi [10] utilised UAV infrared thermography to detect subsurface anomalies in concrete bridge decks. Bolourian and Hammad [11] employed a LiDAR-equipped UAV and proposed a 3D path planning method for automatically collecting precise point cloud data of bridges. Among the above-mentioned sensors, visual cameras are recognised as the most cost-effective and widely-used tool for UBI, particularly due to their outstanding performance in detecting various defects and considering the constraints of UAV onboard weight, power, and computing capabilities [7].

Existing studies on bridge inspection using a camera-mounted UAV can be classified into two approaches in terms of how the images are processed. The first approach is 2D image-based damage identification, which is to directly identify the defects from 2D images captured by the UAV using various image processing techniques such as edge detection methods [13], object-based analysis [14], and/or deep learning methods [15]. Through these methods, defects can be detected automatically, and their specifications can also be quantified via these image processing algorithms. Recent studies [16,17] have focused on investigating lightweight algorithms for the real-time detection and quantification of multiple-type defects at the same time.

The other approach is 3D model-based damage identification, which involves first reconstructing a photogrammetric 3D bridge model from the collected overlapping 2D images and then identifying and mapping the identified defects onto the 3D model, as demonstrated in various studies [5,18,19]. The 3D model-based

damage identification is not only beneficial for providing superior 3D models with the accuracy to resolve defects but also can improve the interpretive capabilities and support the needs of infrastructure inspection [18].

2.2 UAV Image Collection

Image collection is one of the most critical processes in UBI, as the quality of the images directly affects the accuracy and effectiveness of bridge condition assessment. In this section, a specific review was conducted to evaluate the image collection in current UBI practice, with a particular emphasis on LoA. Referring to the criteria established in the field of autonomous cars [20], we classified the LoA of UAV-enabled bridge image collection into five levels: L1-Manual, L2-Partial Automation, L3-Conditional Automation, L4-High Automation, and L5-Full Automation.

Most of the studies [5,8,18] were identified to be at the lowest level of automation, i.e., L1. At this level, UAV pilots should be present and manually control the UAV all the time to capture the desired images at pre-determined areas of interest. The UAV is featured with a multi-directional obstacle awareness and emergency braking system, enabling it to automatically stop immediately in case of an emergency. A significant drawback of manual image collection is that the image collection process relies on pilots' experience, making the process time-consuming and susceptible to pilots' subjective errors.

To address the above issues, a feasible solution is to automate the image collection at a series of pre-determined viewpoints. Several studies [4,21] utilised the waypoint flight mission of the UAV to automatically capture images at a sequence of pre-defined waypoints for image collection. However, this automation depends heavily on precise UAV localisation throughout the flight. Most commercial UAVs (e.g., DJI, Skydio) support Global Positioning System (GPS)-based waypoint flight missions for automatic image collection for bridges. For partial critical inspection areas such as underneath bridge girders, the UAV must be manually operated to capture images due to the loss of GPS signals. Consequently, studies [4,21] employing GPS-based automated navigation are classified as L2 in terms of automation.

To facilitate automated image collection across the entire bridge environment, recent studies have proposed novel UAV localisation methods for estimating UAV's global pose in GPS-denied bridge areas, such as ultra-wideband (UWB) beacon-based [22] or fiducial marker-corrected stereo visual-inertial [23] localisation methods. With accurate global pose, these UAVs can autonomously navigate a pre-defined flight path covering the entire bridge environment. As such studies rely on a pre-existing environmental model to plan a pre-defined flight path, they fall into conditional automation, i.e., L3.

This level represents the most advanced automation achieved to date in UAV-enabled bridge inspection.

While recent technological advancements have significantly enhanced the utilisation of UAVs for automated image collection in bridge environments, existing methodologies, as highlighted in various studies [15,22,24,25], still grapple with some limitations. Primarily, these studies necessitate a pre-defined flight path, which is determined based on pilots' or inspectors'

experience. Additionally, persistent challenges in image quality, encompassing issues like out-of-focus blurriness, improper exposure, and insufficient resolution and coverage, persist due to inherent errors in UAV localisation and the intricate nature of bridge environments, including uneven lighting conditions and UAV vibrations. These kinds of quality issues can significantly affect the performance of post-damage detection and quantification from images.

Table 1. LoA assessment of image collection process in UAV-enabled bridge inspection

References	LoA	Specifications	Limitations
[5,8,18]	L1	<ul style="list-style-type: none"> Pilots manually operate the UAV to capture desired images all the time. UAVs are featured an emergency braking system for added safety. 	<ul style="list-style-type: none"> Subjective error by UAV pilots. Time consuming.
[4,21]	L2	<ul style="list-style-type: none"> Requires a pre-defined obstacle-free flight path. Automatic image collection is only achieved in GPS-available bridge areas. 	<ul style="list-style-type: none"> Significant human intervention is required for image collection in GPS-denied bridge areas.
[15,22,24,25]	L3	<ul style="list-style-type: none"> Requires a pre-defined obstacle-free flight path. The UAV should be capable of locating itself in the entire bridge environment. 	<ul style="list-style-type: none"> Human intervention is required in preparing the flight path for complex bridge structure. Image quality issues exist.

3 Methodology

To address the identified gaps, this paper aims to propose a comprehensive framework for automated high-quality image collection, as presented in Figure 1. This framework employs a BIM model as input and comprises two primary stages: Preparation and Image Collection. Different from existing semi-automated [21] or automated image collection methods [25], this proposed framework incorporates several key features: (1) An

objective and automated flight path planning algorithm to generate a flight mission based on inspection requirements for guiding the image capture of all target bridge areas; (2) An in-flight image quality check (IIQC) algorithm to rapidly assess the quality of UAV-captured images and to display the results with a multi-level colour-rendering system; (3) A supplemental image collection (SIC) step to ensure full coverage of the entire bridge with high-resolution UAV images. Detailed explanations of all components are described below.

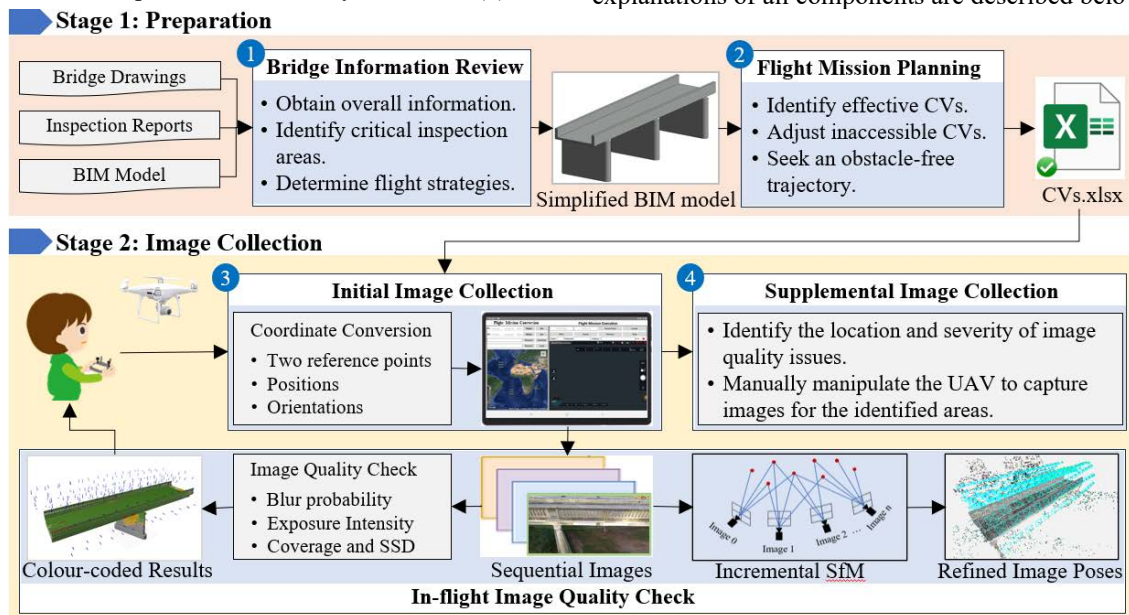


Figure 1. Overall framework for automated high-quality image collection in UAV-enabled bridge inspection

3.1 Preparation

The Preparation stage is essential for successful image collection using UAVs. It involves two key steps: Bridge Information Review and Waypoint Flight Mission Planning.

3.1.1 Bridge Information Review

Following the recommendation in [8], our framework starts with the Bridge Information Review step. This includes examining as-built plans, previous inspection reports, and other relevant documents to guarantee a comprehensive understanding of the target bridge. For example, the review of previous inspection reports can help inspectors identify critical inspection areas before planning the image capture plan, such as the bottom surfaces of girders and pier caps. The information gained during this step allows the pilot to develop reasonable flight strategies in bridge areas with limited accessibility, identify previous defects, and monitor or update critical damage, such as concrete cracks on the target bridge.

3.1.2 Flight Mission Planning

The Flight Mission Planning step is designed to identify a sequence of effective camera viewpoints (CVs) containing positions and orientations and seek an obstacle-free flight path through all the identified viewpoints. In our framework, an automated 3D flight mission planning tool (i.e., a Revit Plug-in), developed in our previous study [4], is employed to generate a waypoint flight mission based on a simplified BIM model of the bridge. Specific steps include (1) planning effective CVs for scanning each bridge part with sufficient resolution under photogrammetry constraints; (2) adjusting those inaccessible CVs due to UAV safety considerations and seeking alternative CVs that can achieve as much effectiveness as the original ones; (3) finding an obstacle-free trajectory through all the identified viewpoints. Figure 2 showcases the developed flight mission planning tool, which has been successfully validated on a range of bridge structures with varying shapes and sizes, as evidenced in the study [4].

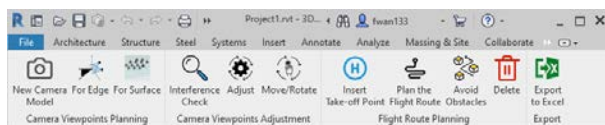


Figure 2. Developed Revit Plug-in for automated flight mission planning

3.2 Image Collection

Once the flight mission is created, in-field image collection can commence. This stage initially involves the UAV automatically executing the waypoint flight mission to collect desired images at those pre-defined

viewpoints. This is followed by supplemental image collection (SIC) for areas not adequately captured in the initial step, conducted manually by the pilot. To offer effective guidance to the pilot, an in-flight image quality check module runs throughout this stage to rapidly assess the quality of UAV-captured images and display the results.

3.2.1 Initial Image Collection

Many UAV systems, such as DJI, Skydio, or those using the Pixhawk flight controller, feature an intelligent flight mode, enabling them to execute waypoint flight missions to fly to designated waypoints automatically and capture images. However, these systems rely heavily on GPS localisation data, and current drone mapping software (e.g., DJI Pilot, DroneDeploy, Pix4Dcapture) does not allow for the execution of custom waypoint flight missions. Therefore, we developed an Android app utilising the DJI Software Development Kit (SDK), as shown in Figure 3, to facilitate automated image collection in bridge areas where GPS signals are reliable. For autonomous flight in GPS-denied areas beneath bridge girders, existing commercial UAVs prove inadequate. Our previous study [23] introduced a custom-built UAV capable of accurately estimating its position in both GPS-available and GPS-denied bridge areas. Such a UAV is essential for automatically capturing all desired images according to a pre-generated waypoint flight mission.

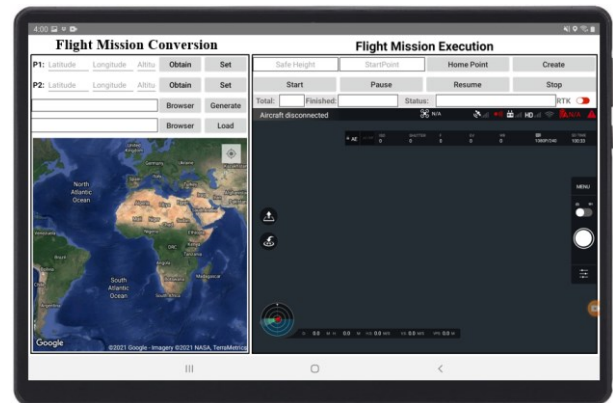


Figure 3. Developed Android App for automated image collection using DJI UAVs

3.2.2 In-flight Image Quality Check

Although the Initial Image Collection step achieves autonomous image capture, navigating a UAV accurately to pre-determined viewpoints remains challenging due to localisation and control errors. These difficulties can lead to images with insufficient resolution and coverage. Furthermore, bridge images captured by UAVs often suffer from other quality issues like blurriness and improper exposure, exacerbated by uneven lighting

conditions and UAV vibrations. These issues can significantly hinder damage detection from images and bridge condition assessment. To tackle these problems and enhance image quality, we incorporate an image check module, named IIQC, into our framework, aiming at rapidly assessing the quality of UAV-captured images against the inspection requirements and providing immediate feedback to the UAV pilot, thereby guiding the collection of supplementary images.

The IIQC module utilises the BIM model as a reference to support image quality assessment and encompasses three key modules: (1) a coarse-to-fine image pose estimation module for obtaining precise image poses relative to the reference bridge model, (2) an image quality assessment module for comprehensively evaluating image quality in terms of blurriness, exposure, coverage, and surface sampling distance (SSD) at the pixel level. Figure 4 presents the pipeline of the image quality check and multi-level colour rendering system.

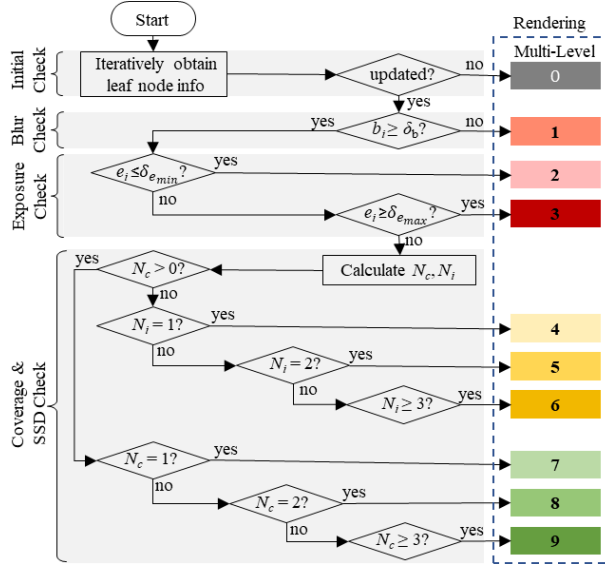


Figure 4. Pipeline of the in-flight image quality check and rendering system

3.2.3 Supplemental Image Collection

The colour-coded results from the IIQC allow the UAV pilot to effortlessly identify the location and severity of issues. Table 2 illustrates the recommended actions for different levels of evaluation results. With these guidelines, the UAV pilot can adeptly operate the UAV to specific areas for the re-collection of images. These newly captured images are then further evaluated and correlated with previous ones. This step stops until the target bridge is well captured. Finally, all images will be processed using commercial photogrammetry software (i.e., iTwin Captrue Modeler) to create a complete 3D digital model of the bridge and further processed to identify and quantify potential defects.

Table 2. Recommended actions for supplemental image collection

Level & Status	Recommended Actions
0 Missing areas	<ul style="list-style-type: none"> Check the image capture plan and analyse the reasons. Capture necessary images.
1 Blurry areas	<ul style="list-style-type: none"> Recapture this area and keep an eye on the vibrations of the UAV.
2 Underexposure	<ul style="list-style-type: none"> Check illumination condition. Recapture this area.
3 Overexposure	<ul style="list-style-type: none"> Adjust capture orientation to avoid areas that are too bright. Recapture this area.
4 One insufficient capture	<ul style="list-style-type: none"> Fly closer to the target surface. Capture more images of this area.
5 Two insufficient captures	<ul style="list-style-type: none"> Fly closer to the target surface. Capture more images of this area.
6 Over three insufficient captures	<ul style="list-style-type: none"> Fly closer to the target surface. Recapture images of this area.
7 One valid capture	<ul style="list-style-type: none"> Capture two more images of this area.
8 Two valid captures	<ul style="list-style-type: none"> Capture one more image of this area.
9 Over three valid captures	<ul style="list-style-type: none"> No further actions.

4 Experiment and Results

In this part, a field experiment was conducted to validate the feasibility of the proposed method.

4.1 Experimental Setup

The target bridge for testing is a typical girder bridge located in New Zealand, as shown in Figure 5. The UAV used in this experiment is DJI Phantom 4 Pro v2.0, equipped with a 20-megapixel sensor with the image resolution of 5472×3648 pixels. Since the bridge includes several similar spans, we only selected the middle area of the superstructure to validate the feasibility of the proposed framework.



Figure 5. Target bridge environment

4.2 Experimental Results

This experiment strictly follows the workflow in the proposed framework. Figure 6 illustrates the results of each step. Firstly, critical information regarding the target bridge, such as the 2D drawings and historical inspection reports, was reviewed. According to the drawings, a 3D BIM model of the bridge was created. Secondly, a waypoint flight mission was created to scan the tested areas using the developed Revit Plug-in. The expected spatial resolution was set to 1.5 mm/pixel with an overlap of 66.7%, as suggested by [26]. Subsequently, we moved to on-site image collection. The developed Android app was utilised to facilitate the UAV to automatically capture the expected images within the generated flight mission, with 365 images being automatically captured with a time cost of 55 minutes. Meanwhile, the IIQC module ran on a standard laptop and provided image quality metrics results in near real-time. Finally, according to the colour-coded results, the UAV pilot controlled the UAV to capture more images to ensure all areas of interest were well captured by high-quality images. Finally, a total of 510 images were captured. The results show that the proposed framework is feasible for practical bridge image collection.

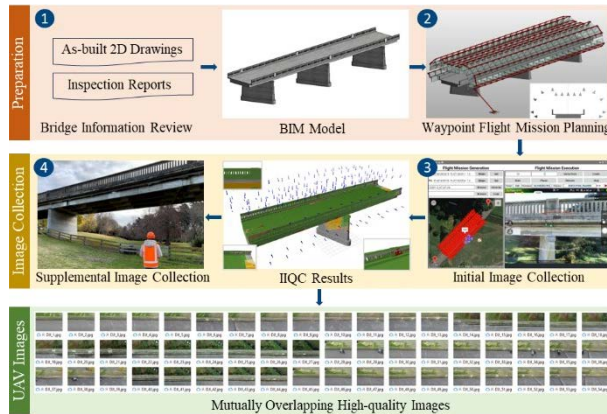


Figure 6. Results of the collected images using the proposed framework

To further evaluate the effectiveness of the newly involved IIQC module, we further processed the captured images using commercial photogrammetry software, i.e., iTwin Capture Modeler, to generate a 3D digital model. We compared the quality of the reconstructed models using only the images captured during the initial image collection and all images, including the supplemental images. The example outcomes of the comparative experiment are depicted in Figure 7. The findings from this experiment validate the value of the SIC step, which involves capturing supplementary images in areas that may not have been adequately covered during the initial image collection and ensuring the reconstructed model with high quality to identify and map the potential defects.

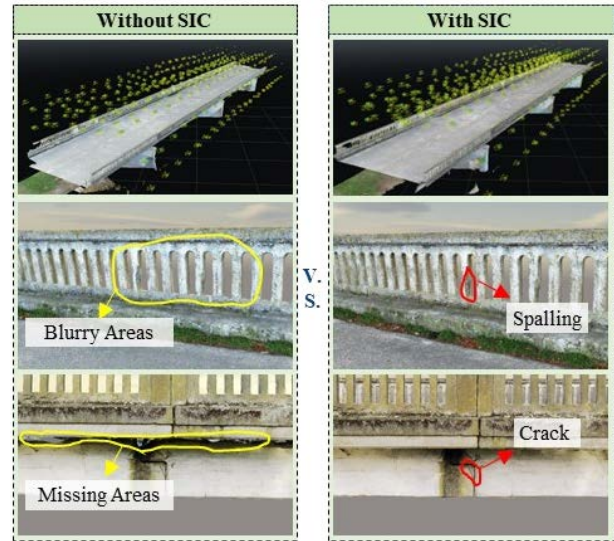


Figure 7. Comparison between the pre-supplemental collection and post-supplemental collection

5 Discussion

This paper introduces a BIM-supported framework for automated, high-quality image collection using a UAV. By integrating automated flight mission planning, automated image collection, an in-flight image quality check module, and supplemental image collection, the framework enables UAVs to capture high-quality images for bridge inspections with minimal human intervention. The proposed framework is versatile and can be readily applied to most bridge structures that have clear surroundings.

While the framework shows promising results, it still has limitations. It operates under the assumption of a clear bridge environment or that surrounding obstacles are pre-modelled for flight mission planning. In complex bridge environments with trees or power lines, significant human input is still necessary. Additionally, the framework depends on a pre-existing bridge model for flight path planning and in-flight image quality check, which poses a challenge for ancient bridges. In the future, UAV-enabled image acquisition for bridge inspection is anticipated to progress towards full autonomy, eliminating the need for a pre-determined flight path and reliance on a bridge model as prior knowledge, as shown in Figure 8.

To achieve the level of high automation, the UAV will be enabled with real-time environment sensing and adaptive path replanning capabilities by integrating advanced SLAM algorithms like RTAB-MAP [27] and deep reinforcement learning-based path planning algorithms [28]. Regarding the ultimate goal of full automation, research can focus on enhancing the UAV's

capability for semantic environmental modelling by integrating advanced image-based [29] or point cloud-based [30] semantic segmentation algorithms. In this stage, UAVs will be able to autonomously execute

inspection tasks without any human intervention during flight, making them suitable for rapid post-disaster bridge condition assessments.

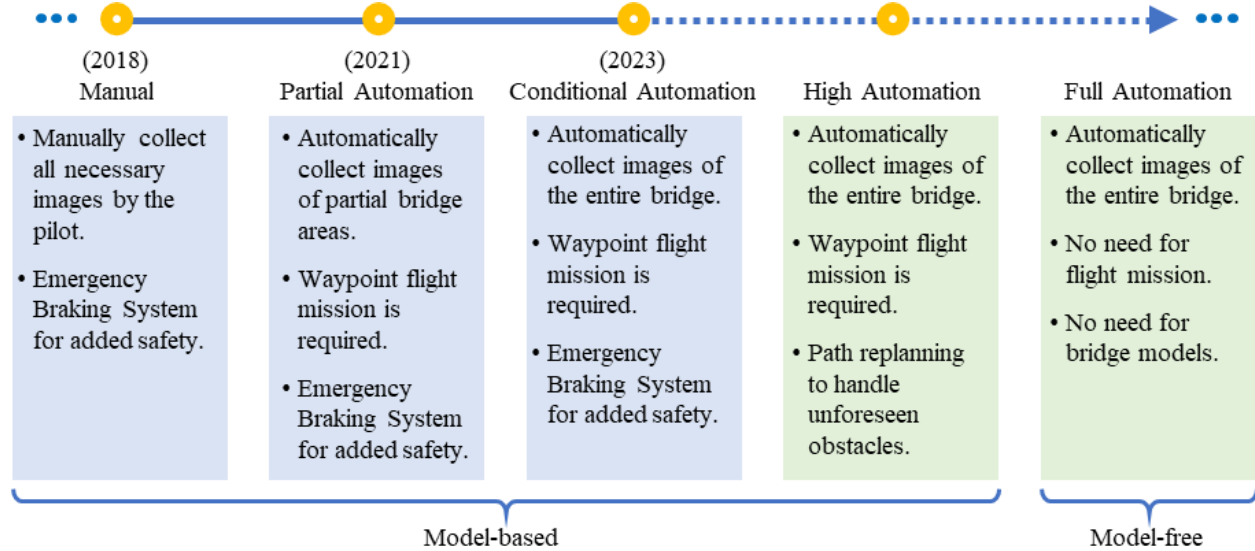


Figure 8. A vision of automated UAV image collection for bridge inspection

6 Conclusions

This paper presents a holistic framework for automated UAV image collection for bridge inspection by incorporating a set of tools to enable high automation and ensure the collection of high-quality images. The feasibility of the framework has been validated through a field experiment, leading to the following conclusions:

1. The tools incorporated within this framework can significantly diminish the necessity for human intervention and streamline all sequences of UAV-enabled bridge image collection, from the initial preparation to on-site image collection.
2. The IIQC module, coupled with the supplemental image collection within this framework, can guarantee that the UAV-captured images are of high quality for reconstructing a 3D digital model of the bridge and for the effective identification and quantification of potential defects.

The main contributions of this paper are listed below:

1. A systematic framework has been proposed for fully automated high-quality image collection for bridge inspection using UAVs and BIM (Figure 1).
2. All algorithms and tools developed to enhance UAV efficiency in bridge inspection have been made available as open-source resources ([Github](#)).
3. A vision towards fully autonomous UAV image collection for bridge inspection has been presented, charting a direction for future research (Figure 8).

Code Availability

The Revit Plug-in for automated path planning is available at <https://github.com/fwan133/MissionPlanner>.

The Android App for automated image collection is available at https://github.com/fwan133/DJI_App.

The IIQC project for rapid image quality check is available at <https://github.com/fwan133/IIQC>.

References

- [1] Dorafshan, S. and M. Maguire, Bridge inspection: human performance, unmanned aerial systems and automation. *Journal of Civil Structural Health Monitoring*, 8(3):443-476, 2018.
- [2] Gattulli, V. and L. Chiaramonte, Condition Assessment by Visual Inspection for a Bridge Management System. *Computer-Aided Civil and Infrastructure Engineering*, 20(2):95-107, 2005.
- [3] Jeong, E., J. Seo, and J. Wacker, Literature review and technical survey on bridge inspection using unmanned aerial vehicles. *Journal of Performance of Constructed Facilities*, 34(6):04020113, 2020.
- [4] Wang, F., et al., Automated UAV path-planning for high-quality photogrammetric 3D bridge reconstruction. *Structure and Infrastructure Engineering*:1-20, 2022.

- [5] Chen, S., et al., UAV bridge inspection through evaluated 3d reconstructions. *Journal of Bridge Engineering*, 24(4):05019001, 2019.
- [6] Lee, J.H., et al., A new image-quality evaluating and enhancing methodology for bridge inspection using an unmanned aerial vehicle. *Smart Structures and Systems*, 27(2):209-226, 2021.
- [7] Zhang, C., et al., Towards fully automated unmanned aerial vehicle-enabled bridge inspection: Where are we at? *Construction and Building Materials*, 347:128543, 2022.
- [8] Seo, J., L. Duque, and J. Wacker, Drone-enabled bridge inspection methodology and application. *Automation in Construction*, 94:112-126, 2018.
- [9] Morgenthal, G., et al., Framework for automated UAS-based structural condition assessment of bridges. *Automation in Construction*, 97:77-95, 2019.
- [10] Omar, T. and M.L. Nehdi, Remote sensing of concrete bridge decks using unmanned aerial vehicle infrared thermography. *Automation in Construction*, 83:360-371, 2017.
- [11] Bolourian, N. and A. Hammad, LiDAR-equipped UAV path planning considering potential locations of defects for bridge inspection. *Automation in Construction*, 117:103250, 2020.
- [12] Gaspari, F., et al., Integration of UAV-LiDAR and UAV-photogrammetry for infrastructure monitoring and bridge assessment. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 43B2:995-1002, 2022.
- [13] Jena, K.K., et al. Unmanned aerial vehicle assisted bridge crack severity inspection using edge detection methods. In *2019 Third International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*. pages 284-289, Palladam, India, 2019.
- [14] Zollini, S., et al., UAV photogrammetry for concrete bridge inspection using Object-Based Image Analysis (OBIA). *Remote Sensing*, 12(19):3180, 2020.
- [15] Kang, D. and Y.-J. Cha, Autonomous UAVs for structural health monitoring using deep learning and an ultrasonic beacon system with geo-tagging. *Computer-Aided Civil and Infrastructure Engineering*, 33(10):885-902, 2018.
- [16] Ding, W., et al., Crack detection and quantification for concrete structures using UAV and transformer. *Automation in Construction*, 152:104929, 2023.
- [17] Jiang, S., Y. Cheng, and J. Zhang, Vision-guided unmanned aerial system for rapid multiple-type damage detection and localization. *Structural Health Monitoring*, 22(1):319-337, 2023.
- [18] Khaloo, A., et al., Unmanned aerial vehicle inspection of the Placer River Trail Bridge through image-based 3D modelling. *Structure and Infrastructure Engineering*, 14(1):124-136, 2018.
- [19] Wang, X., et al., Rapid seismic risk assessment of bridges using UAV aerial photogrammetry. *Engineering Structures*, 279:115589, 2023.
- [20] Khan, M.A., et al., Level-5 autonomous driving—are we there yet? A review of research literature. *ACM Comput. Surv.*, 55(2):Article 27, 2022.
- [21] Lin, J.J., et al., Bridge inspection with aerial robots: Automating the entire pipeline of visual data capture, 3D mapping, defect detection, analysis, and reporting. 35(2):04020064, 2021.
- [22] Jiang, S., Y. Wu, and J. Zhang, Bridge coating inspection based on two-stage automatic method and collision-tolerant unmanned aerial system. *Automation in Construction*, 146:104685, 2023.
- [23] Wang, F., et al., UAV navigation in large-scale GPS-denied bridge environments using fiducial marker-corrected stereo visual-inertial localisation. *Automation in Construction*, 156:105139, 2023.
- [24] Jung, S., et al. Toward autonomous bridge Inspection: A framework and experimental results. In *16th International Conference on Ubiquitous Robots (UR)*. pages 208-211, Jeju, Korea (South), 2019.
- [25] Ali, R., et al., Real-time multiple damage mapping using autonomous UAV and deep faster region-based neural networks for GPS-denied structures. *Automation in Construction*, 130:103831, 2021.
- [26] Wang, F., et al. Optimal UAV image overlap for photogrammetric 3d reconstruction of bridges. In *IOP Conference Series: Earth and Environmental Science*. pages 022052, Melbourne, Australia, 2022.
- [27] Labbé, M. and F. Michaud, RTAB-Map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation. *Journal of Field Robotics*, 36(2):416-446, 2019.
- [28] Yan, C., X. Xiang, and C. Wang, Towards real-time path planning through deep reinforcement learning for a uav in dynamic environments. *Journal of Intelligent & Robotic Systems*, 98(2):297-309, 2020.
- [29] Kim, H., Y. Narazaki, and B.F. Spencer Jr, Automated bridge component recognition using close-range images from unmanned aerial vehicles. *Engineering Structures*, 274:115184, 2023.
- [30] Yang, X., et al., Semantic segmentation of bridge point clouds with a synthetic data augmentation strategy and graph-structured deep metric learning. *Automation in Construction*, 150:104838, 2023.

MPC-Based Proactive Swing Attenuation for Double-Pendulum Overhead Cranes

Juhao Su^{1,2}, Shichen Sun^{1,2}, Ching-Wei Chang², and Siwei Chang²

¹Academy of Interdisciplinary Studies, The Hong Kong University of Science and Technology, Hong Kong S.A.R.

²Hong Kong Center for Construction Robotics, Hong Kong S.A.R.

jsuan@connect.ust.hk, ssunbb@connect.ust.hk, ccw@ust.hk, siweichang@ust.hk

Abstract -

Overhead cranes, which are traditionally classified as underactuated systems, face a persistent challenge in balancing the assertiveness of trolley movement and the amplitude of the payload oscillation. This difficulty is exacerbated by the crane's inherent double-pendulum characteristics. To improve the operational efficiency and fault tolerance of cranes when lifting large and heavy construction components, this research proposed a novel and proactive swing attenuation crane controlling approach. The method was designed to integrate two auxiliary control inputs extracted from the complex and coupled dynamics using a customized Model Predictive Controller (MPC). The performance of the proposed approach was validated by comparing it to common controllers such as the linear quadratic regulator (LQR) and sliding mode controller (SMC). It is proven that the proposed approach makes cranes more resilient to specific types of uncertainties and adverse conditions, resulting in safe, efficient, and intelligent overhead crane transportation.

Keywords -

Double-Pendulum; Overhead Crane; Modular integrated Construction (MiC); Swing Attenuation; Proactive Hook-load Stabilization; Model Predictive Control (MPC); Sliding Mode Control (SMC)

1 Introduction

In the Architecture, Engineering, and Construction (AEC) industry, efficient transportation of substantial payloads is a routine necessity, demanding machinery such as crane systems with exceptional maneuverability and adaptability. Recent research highlighted that these crane systems often exhibit the characteristics of underactuated systems, providing increased degrees of freedom with a limited number of control inputs [1], [2], [3]. Consequently, careful treatment of trolley translation in the horizontal plane is crucial to avoid undesirable payload swings, which not only reduce efficiency but also lead to potential on-site damages [4]. In practical applications, the hook is commonly not positioned at the exact end of the hoisting rope; instead, additional ropes connect the load to the



(a) Demonstration of connection between hook and MiC (b) The self-built 2-D overhead crane model

Figure 1. Onsite photos of MiC and 2-D crane model

hook as shown in Figure 1a. However, this results in a coupled and complex double pendulum swing, making the dynamic model of crane-load challenging to control [5].

Traditionally, skilled operators effectively mitigate load oscillations based on their environmental perception and empirical judgment. However, given the substantial costs associated with workforce training and potential accidents during manual crane operations, there is a compelling need for advanced automated control mechanisms tailored for crane systems or a novice-friendly auxiliary operating system, especially for transporting giant and heavy payloads like MiC modules [6].

The pendulum model poses a classical oscillatory control problem, prevalent in all cranes. Previous researchers have applied various open-loop control methods to eliminate hook swings. Two common strategies include input shaping [7], [8] and trajectory planning-based control [9], [10]. These methods design multiple inputs based on the system's dynamic model to counteract oscillations. However, vulnerability may arise if the system identification during modeling is incomplete, or if unknown external disturbances occur during execution [11]. In such cases, aligning the system's response with the envisioned motion trajectory becomes challenging. Concurrently, numerous

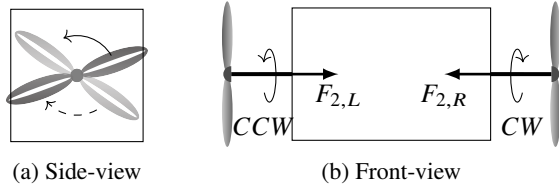


Figure 2. Illustration of spinning propellers mounted on the MiC block (CW for clock-wise, CCW for counter-clock-wise), total force acted on the block would be $F_2 = F_{2,L} - F_{2,R}$, where the propellers mounted on the hook were identical

studies have explored closed-loop control, employing advanced controllers like Fuzzy Control and its variant [12], [13], [14], as well as Model Predictive Control (MPC) [15], [16]. These controllers effectively mitigate swing angles in both transient and steady-state phases. However, when the model is intricately delineated as a double pendulum, the pronounced nonlinear dynamics pose challenges, especially considering external factors like air resistance and wind, which could compromise the reliability of the controller's model. Despite adopting a closed-loop methodology, the scarcity of controllable variables in underactuated systems remains a regrettable limitation.

Therefore, to enable rapid and significant oscillation suppression during various trolley movement tasks, a novel approach adding two sets of auxiliary control inputs is introduced to this underactuated system. Given the strong coupling behavior among inputs and dynamics, the application of single-input single-output (SISO) controllers, such as ordinary PID controllers, may encounter challenges [17]. This study emphasizes the application of multi-input multi-output (MIMO) controllers instead, for achieving desired control outcomes.

Maintaining the integrity of the system, the aforementioned control inputs rely on the thrust differentials generated by propellers' rotations, as both the hook and MiC are suspended in the air and connected using flexible cables and joints, without moment transmission. This configuration, as illustrated in Figure 1b and 2, allows for the desired execution of intended objectives meanwhile minimizing response instability and lag during swing angle zero-crossing transitions, where the need for abrupt stop-reverse rotations of propellers arises. Since an overhead crane involves movements in two perpendicular planes [18], the study formulates a 2-D spatial model for preliminary validation and presents the following contributions:

1. Analyze the system dynamics of a 2-D overhead crane with two additional control inputs.
2. Utilizing states of the hook and measurable trolley movement, design a model predictive-based controller to tune the inputs under constraints.

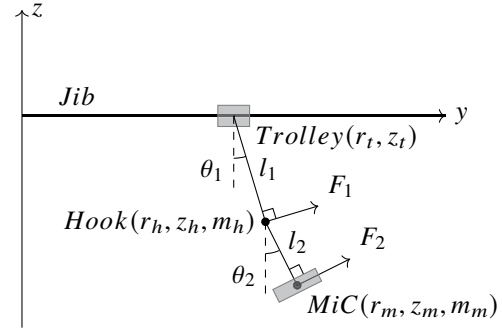


Figure 3. Schematic of the double-pendulum system

3. Achieve noticeable efficacy and show strong resilience against assigned adverse impacts and uncertainty through simulations.

The rest of the paper is organized as follows. Section 2 analyzes full-state dynamics for controller design under the hypothesis of single-axis trolley movement along the jib and forces acting at the center of mass. Simulations in Section 3 verify performance and robustness. The conclusion and future work are presented in the last section.

2 Methodology

2.1 System dynamics

For the 2-D overhead crane scenario interested in this study, its schematic is shown in Figure 3, all directions and angular polarity are specified and applied consistently throughout all simulations and controller designs. If the first set of steel ropes connecting the trolley and the hook is l_1 , the second set connecting the hook and MiC block is l_2 , then all coordinates can be properly expressed using $r_t, \theta_1, \theta_2, l_1, l_2$, and trigonometry as:

$$\begin{aligned} r_h &= r_t + l_1 \sin \theta_1, & r_m &= r_t + l_1 \sin \theta_1 + l_2 \sin \theta_2, \\ z_h &= 0 - l_1 \cos \theta_1, & z_m &= 0 - l_1 \cos \theta_1 - l_2 \cos \theta_2. \end{aligned}$$

Then the dynamics of this system can be represented using Lagrangian Mechanics. The kinetic energy (\mathcal{T}), potential energy (\mathcal{V}) and Lagrangian (\mathcal{L}) of the system can be expressed as below:

$$\begin{aligned} \mathcal{T} &= \frac{1}{2} m_m ((\dot{r}_t + l_1 \cos \theta_1 \dot{\theta}_1 + l_2 \cos \theta_2 \dot{\theta}_2)^2 \\ &\quad + (l_1 \sin \theta_1 \dot{\theta}_1 + l_2 \sin \theta_2 \dot{\theta}_2)^2) \end{aligned} \quad (1)$$

$$+ \frac{1}{2} m_h ((\dot{r}_t + l_1 \cos \theta_1 \dot{\theta}_1)^2 + l_1^2 \sin^2 \theta_1 \dot{\theta}_1^2),$$

$$\mathcal{V} = -m_m g (l_1 \cos \theta_1 + l_2 \cos \theta_2) - m_h g l_1 \cos \theta_1, \quad (2)$$

$$\mathcal{L} = \mathcal{T} - \mathcal{V}. \quad (3)$$

For the forces acting on the hook and MiC block, they are marked as F_1 and F_2 , respectively. Since their directions are always perpendicular to the rope sets (l_1, l_2) and within the plane of y-z axes, their interference against θ_1 and θ_2 can be derived as below:

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\theta}_1} - \frac{\partial \mathcal{L}}{\partial \theta_1} = F_1 l_1, \quad (4)$$

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\theta}_2} - \frac{\partial \mathcal{L}}{\partial \theta_2} = F_2 l_2, \quad (5)$$

substituting (3) to (4) and (5), the shortened form can be obtained:

$$\begin{aligned} F_1 = & l_1(m_h + m_m)\ddot{\theta}_1 + (m_h + m_m)\cos\theta_1\ddot{r}_t \\ & + (m_h + m_m)g\sin\theta_1 + l_2m_m\sin(\theta_1 - \theta_2)\ddot{\theta}_2^2 \\ & + l_2m_m\cos(\theta_1 - \theta_2)\ddot{\theta}_2, \end{aligned} \quad (6)$$

$$\begin{aligned} F_2 = & l_2m_m\ddot{\theta}_2 + m_m\cos\theta_2\ddot{r}_t + m_mg\sin\theta_2 \\ & - l_1m_m\sin(\theta_1 - \theta_2)\ddot{\theta}_1^2 + l_1m_m\cos(\theta_1 - \theta_2)\ddot{\theta}_1. \end{aligned} \quad (7)$$

2.2 Linearization and state-space function

For crane dynamics, independent of the categories they belong to, a widely adopted and well-established practice is to linearize the dynamic equations through the small angle assumption [19], [20]. For Equation (6) and (7), they can be linearized and rearranged as below:

$$\begin{aligned} \ddot{\theta}_1 = & -\frac{g(m_h + m_m)}{l_1m_h}\theta_1 + \frac{gm_m}{l_1m_h}\theta_2 - \frac{F_2 - m_m\ddot{r}_t}{l_1m_h} \\ & + \frac{F_1 - (m_h + m_m)\ddot{r}_t}{l_1m_h}, \end{aligned} \quad (8)$$

$$\begin{aligned} \ddot{\theta}_2 = & \frac{g(m_h + m_m)}{l_2m_h}\theta_1 + \frac{(m_h + m_m)(F_2 - m_m\ddot{r}_t)}{l_2m_hm_m} \\ & - \frac{F_1 - (m_h + m_m)\ddot{r}_t}{l_2m_h} - \frac{g(m_h + m_m)}{l_2m_h}\theta_2. \end{aligned} \quad (9)$$

The primary objective of the controllers is to mitigate the oscillations of both the hook and MiC module near the normal projection of the trolley position. Thus, the full states vector in this study will be designated as $x(t) = [\theta_1, \dot{\theta}_1, \theta_2, \dot{\theta}_2]^T$ and the desired state should be $x_{des} = [0, 0, 0, 0]^T$. Inspired from Equation (8) and (9), let

$$v_1 = F_1 - (m_h + m_m)\ddot{r}_t, \quad (10)$$

$$v_2 = F_2 - m_m\ddot{r}_t, \quad (11)$$

thus a tailored control inputs vector $u(t) = [v_1, v_2]^T$ can be obtained to form the following state-space system:

$$\dot{x} = Ax + Bu, \quad (12)$$

$$y = x, \quad (13)$$

where $y(t) \in \mathbb{R}^{4 \times 1}$ denotes the output vector, $A \in \mathbb{R}^{4 \times 4}$ and $B \in \mathbb{R}^{4 \times 2}$ are the system parameter matrices, where can be derived from Equation (8) - (13) as below:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{g(m_h+m_m)}{l_1m_h} & 0 & \frac{gm_m}{l_1m_h} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{g(m_h+m_m)}{l_2m_h} & 0 & -\frac{g(m_h+m_m)}{l_2m_h} & 0 \end{bmatrix}, \quad (14)$$

$$B = \begin{bmatrix} 0 & 0 \\ \frac{1}{l_1m_h} & -\frac{1}{l_1m_h} \\ 0 & 0 \\ -\frac{1}{l_2m_h} & \frac{m_h+m_m}{l_2m_hm_m} \end{bmatrix}. \quad (15)$$

It is worth noting that \ddot{r}_t is neither an independent control input nor a component of the system states. Given that the trolley movement is considered arbitrary, acceleration serves as a parameter readily obtainable through measurement. They could be treated as compensation in the controller for the dynamic impact of external motion on the system output.

Take a time interval T_s as sampling time, the system can be properly discretized from Equation (12) and (13):

$$x(k+1) = A_dx(k) + B_d u(k), \quad (16)$$

$$y(k) = x(k), \quad (17)$$

where $A_d = (1 + AT_s)|_{x=x(k)}$, $B_d = (BT_s)|_{x=x(k)}$.

2.3 MPC design

Model Predictive Controller (MPC), employing the receding horizon approach, is a control strategy that predicts the dynamics within h steps (horizon length) and optimizes the control input array to minimize the cost between predicted and reference values. Eventually, only the first element of this array is applied. Therefore, a reference generator for prediction at time instant k within the horizons is set as:

$$r_f(k)_i = x(k)_1 + (i-1)\frac{x_{des} - x(k)_1}{h-1}, \quad i = 1, 2, 3, \dots, h, \quad (18)$$

where $r_f(k)_i$ is the i^{th} generated reference state at time instant k , $x(k)_1$ denotes state at time instant k , and the discretized system follows:

$$x(k)_i = x(k)_{i-1} + \left(\frac{T_d}{h} + T_s\right)\dot{x}(k)_{i-1}, \quad i = 2, 3, 4, \dots, h, \quad (19)$$

where T_d is the inherent time delay of the system, T_s is the time step for prediction, and the cost function is defined

as:

$$J_M = (r_f(k)_h - x(k)_h)^T F_M (r_f(k)_h - x(k)_h) + \sum_{i=1}^{h-1} \{(r_f(k)_i - x(k)_i)^T Q_M (r_f(k)_i - x(k)_i) + u_i^T R_M u_i\}, \quad (20)$$

where $Q_M \in \mathbb{R}^{4 \times 4}$, $F_M \in \mathbb{R}^{4 \times 4}$ and $R_M \in \mathbb{R}^{2 \times 2}$ denotes weight matrices for tuning control preference on process error, final error, and control effort, respectively.

One leading advantage of adopting MPC lies in its ability to impose constraints not only on the generation of the system's output but also on the state variables in upcoming time steps. The constraints on input selection typically relate to the mechanical capabilities of the device, such as acceleration limit and jerk limit, while the state constraint prevents the system from exceeding undesirable states regarding safety. In this study, the quadratic programming (QP) problem with constraints was defined as:

$$\begin{aligned} \min \quad & J_M \\ \text{s.t.} \quad & \dot{x} = f(x(t), u(t)) \\ & |\theta_i| \leq \theta_{max}, \quad i = 1, 2 \\ & u(t) \in Y(t), \end{aligned} \quad (21)$$

where $f(\cdot)$ denotes the predict function and can be discretized into Equation (19) and further derived by substituting Equation (8) and (9), Y represents the input constraints including jerk and acceleration according to motor limitation.

Conservatively setting these constraints can mitigate actuator saturation in hardware experiments [18], which enables MPC to more effectively anticipate future state changes in a well-defined dynamic system.

3 Simulations

3.1 Comparative controller design

3.1.1 LQR

According to the state-space derived in Section 2.2, a Linear Quadratic Regulator (LQR) can be established using Equation (16) and (17). A standard LQR cost function is introduced as below:

$$J_L = \sum_{i=0}^{\infty} \{(x_{curr} - x_{des})^T Q_L (x_{curr} - x_{des}) + u^T R_L u\}, \quad (22)$$

where $(x_{curr} - x_{des})$ denotes the error between the current state and desired state, $u = -K_{lqr}(x_{curr} - x_{des})$, $Q_L \in \mathbb{R}^{4 \times 4}$ and $R_L \in \mathbb{R}^{2 \times 2}$ denote two weight matrices for precision and control effort trade-off. An optimum

coefficient K_{lqr} can be then obtained from solving the optimizing problem of $\min J_L$ and implemented to compute the optimal control vector as a constant matrix.

3.1.2 SMC

Sliding Mode Controller (SMC) employs a sliding surface, which is a hyperplane in the space constituted of state error and its derivative, and utilizes discontinuous control to make the system slide towards this plane and converge towards the origin of the space, which implies no state error exists and so as the trend it induces. Let $x_s(t) = [\theta_1, \theta_2]^T$, control objective is $x_{s,des} = [0, 0]^T$ and let

$$\ddot{x}_s = \delta + \Lambda u, \quad (23)$$

where δ and Λ can be derived from Equation (8) and (9) as:

$$\delta = \begin{bmatrix} -\frac{g(m_h+m_m)}{l_1 m_h} \theta_1 + \frac{g m_m}{l_1 m_h} \theta_2 \\ \frac{g(m_h+m_m)}{l_2 m_h} \theta_1 - \frac{g(m_h+m_m)}{l_2 m_h} \theta_2 \end{bmatrix}, \quad (24)$$

$$\Lambda = \begin{bmatrix} \frac{1}{l_1 m_h} & -\frac{1}{l_1 m_h} \\ -\frac{1}{l_2 m_h} & \frac{1}{l_2 m_h} \end{bmatrix}. \quad (25)$$

Set the sliding mode function as

$$s(t) = \Omega(x_s - x_{s,des}) + (\dot{x}_s - \dot{x}_{s,des}), \quad (26)$$

where $\Omega = \text{diag}[\varpi_1, \varpi_2]$ must satisfy Hurwitz condition of $\varpi_1 > 0, \varpi_2 > 0$,

$$\dot{s} = \Omega(\dot{x}_s - \dot{x}_{s,des}) + (\ddot{x}_s - \ddot{x}_{s,des}), \quad (27)$$

$$\dot{s} = -\epsilon \tanh(s) - \rho s, \quad (28)$$

where Equation (28) is the defined expression to achieve a nearly exponential reaching law toward the sliding plane with $\epsilon > 0, \rho > 0$ so that the Lyapunov function meets its criterion:

$$\begin{aligned} V &= \frac{1}{2} s^2 \geq 0, \\ \dot{V} &= s \dot{s} \leq 0. \end{aligned}$$

From Equation (23) - (28), the control input u of SMC could be derived as:

$$u = (\Lambda^T \Lambda)^{-1} (\Lambda^T (\delta + \Omega(\dot{x}_s - \dot{x}_{s,des}) - \epsilon \tanh(s) - \rho s)). \quad (29)$$

3.2 Comparative simulation results

To validate the effectiveness of proposed proactive control method and the designed controller, a comparative software in the loop (SITL) simulation is conducted using MATLAB & Simulink by creating non-linear dynamic model plant based on Equation (6) and (7), the constant

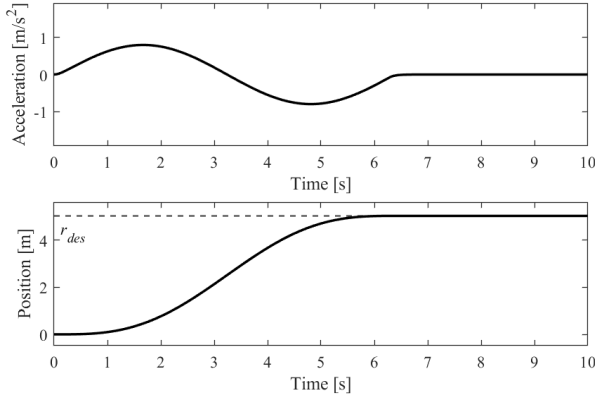
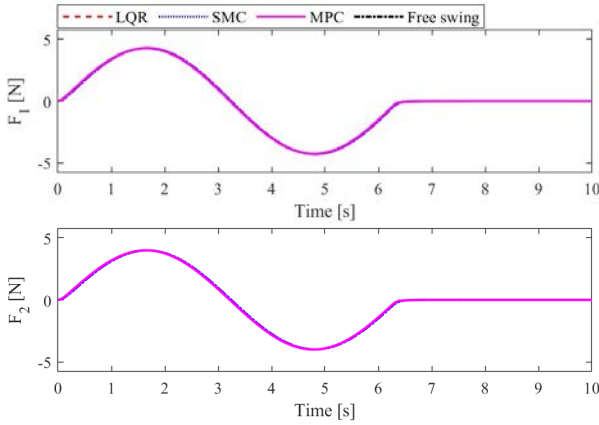


Figure 4. Trolley movement pattern

Figure 5. Simulation result of applied force (F_1, F_1)

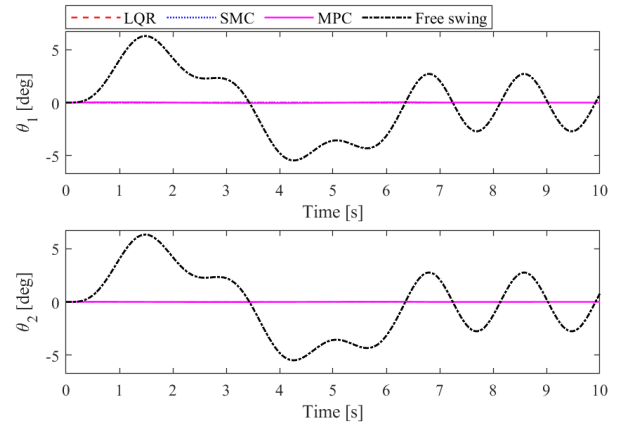
parameters introduced in the preceding formula derivation will be assigned according to the actual measurements from the model in Figure 1b to aids experiments in the future:

$$\begin{aligned} m_h &= 0.38 \text{ [kg]}, & l_1 &= 0.6 \text{ [m]}, & g &= 9.81 \text{ [N/kg]}, \\ m_m &= 4 \text{ [kg]}, & l_2 &= 0.2 \text{ [m]}. \end{aligned}$$

3.2.1 Ideal condition

The simulated scenario began at time instant $t = 0[s]$ with the initial condition of staying at equilibrium. Meanwhile, a smooth initiation was provided for the trolley with starting jerk $\ddot{r}_t = 0$ by applying a first-order low-pass filter to a sinusoidal acceleration as shown in Figure 4. The trolley continued its motion until $t = 2\pi[s]$, reaching its desired position r_{des} . Throughout the trolley's motion and after the motion terminated, each controller was aimed to minimize the oscillations in the angles θ_1 and θ_2 .

The simulated configuration for all controllers is identical, all the sampling frequencies were set to be $100[Hz]$,

Figure 6. Simulation result of swing angle (θ_1, θ_2)

and the T_s in MPC was set to be $0.05[s]$ with horizon $h = 10$.

The LQR weight matrices and coefficients in SMC were fixed to

$$Q_L = \text{diag}[5, 0.2, 5, 0.2], \quad \epsilon = 15, \quad \rho = 150,$$

$$R_L = \text{diag}[5 \times 10^{-3}, 5 \times 10^{-3}], \quad \Omega = \text{diag}[2, 4],$$

as both LQR and SMC rely on state-space where u is not directly equivalent to the applied force, the compensation of \ddot{r}_t term according to Equation (10) and (11) is needed:

$$F_1 = v_1 + (m_h + m_m)\ddot{r}_t, \quad (30)$$

$$F_2 = v_2 + (m_m)\ddot{r}_t. \quad (31)$$

Weight matrices in MPC are defined as:

$$Q_M = \text{diag}[5, 0.18, 5, 0.18],$$

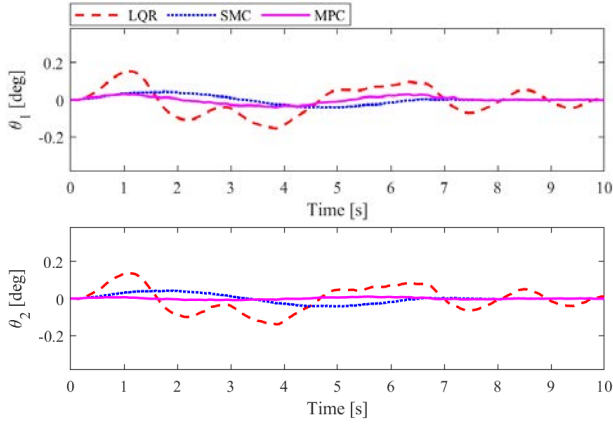
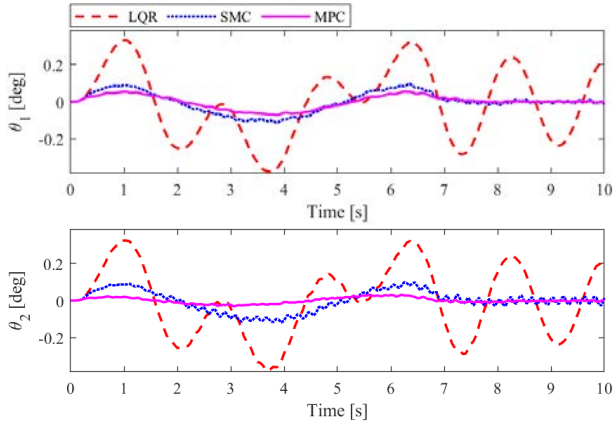
$$F_M = \text{diag}[10, 1, 10, 1],$$

$$R_M = \text{diag}[1 \times 10^{-7}, 1 \times 10^{-7}],$$

and the OCP in (21) is solved by a multi-shooting method with each time interval solved by an interior point optimizer (IPOPT) within CasADi [21].

To avoid transient and excessively high outputs, all forces are constrained by a saturation module within the threshold of $|F_1|, |F_2| \leq 5[N]$. The performance of different controllers was evaluated by observing swing angle θ_1 and θ_2 as shown in Figure 6.

The results indicate that the proposed proactive approach with different controllers can make the system keep a near equilibrium during the acceleration and deceleration of the trolley. This reduction brings the oscillations from potentially hazardous angles ($6.5[deg]$) to virtually $0[deg]$. According to Figure 5, after the trolley moved, the forces are dominated by the compensation term of \ddot{r}_t instead of states, therefore exhibiting subtle distinctions among different controller behaviors.

Figure 7. Swing angle (θ_1, θ_2) with 0.02[s] delayFigure 8. Swing angle (θ_1, θ_2) with 0.04[s] delay

3.2.2 Condition with delay and noise

However, trolley acceleration (\ddot{r}_t) obtained from sensors may have delay and noise, so as the states obtained. Thus in this group of tests, the signal transport delay is considered and the noise in the sensors is simulated by adding Gaussian noise to the original states vector $x(t)$ and \ddot{r}_t :

$$N(t) = [\theta_1, \dot{\theta}_1, \theta_2, \dot{\theta}_2, \ddot{r}_t]^T + n, \quad (32)$$

where $n = [n_1, n_2, n_3, n_4, n_5]^T \sim \mathcal{N}(0, \sigma_i)$ is a vector that each element is independent and identically distributed Gaussian random variables with a mean of zero and follow a normal distribution with standard deviation σ_i , given the magnitude of n_{1-4} is on the order of 10^{-6} and n_5 is on the order of 10^{-4} .

Since the latency is not a constant and instead, fluctuates around a baseline. To illustrate this, the signal delay is modeled as the sum of a constant part and a random variable with a variance of 0.2. The T_d of MPC in Equation (19) can only access the constant part of the delay and is unaware of the current magnitude of the random component.

After some preliminary numerical simulations, LQR and SMC exhibited substantial oscillations and signs of non-convergence in the presence of delay, leading to the implementation of fractional gains to u (K_1 for LQR, K_2 for SMC), the applied forces are in the form below:

$$F_{1,fi} = v_1 K_i + (m_h + m_m) \ddot{r}_t, \quad (33)$$

$$F_{2,fi} = v_2 K_i + (m_m) \ddot{r}_t, \quad i = 1, 2. \quad (34)$$

Table 1. Quantified results with increased m_m

Method	$\theta_{1,max}$	$\theta_{2,max}$	$\theta_{1,RMSE}$	$\theta_{2,RMSE}$
LQR	-2.38	-2.27	0.86	0.78
SMC	-2.40	-2.23	0.85	0.81
MPC	-1.31	-1.42	0.78	0.84

Through iterative adjustments, these gains were eventually set at critical values where the controllers were prone to collapse, as outlined below:

$$\text{delay} = 0.02[s], \quad K_1 = 0.2, \quad K_2 = 0.25,$$

$$\text{delay} = 0.04[s], \quad K_1 = 0.04, \quad K_2 = 0.05.$$

The data depicted in Figures 7 and 8 illustrates the impact of introducing delay and noise on controller performance. However, with the aid of fractional gain, it does not lead to a severe failure. Notably, LQR exhibits high sensitivity, with decreasing gain causing the persistent squeezing of state effects. It fails to achieve state convergence as the delay continues to increase. In comparison, SMC outperforms LQR, showcasing its resilience to uncertainty and achieving a considerable extent of convergence. MPC emerges as the most trustful option, requiring no adjustment like fractional gains. It exhibits only minor fluctuations in the vicinity of the zero axis.

However, the capacity of LQR and SMC to control pure state changes is significantly diminished by the fractional gain, particularly in facing unknown disturbances directly affecting the load and hook.

3.2.3 Condition with actuator saturation

According to the findings presented in Figure 6, the supplied forces F_1 and F_2 already approached the assumed actuator limit, and further increment may lead to actuator saturation. Therefore, the MiC mass (m_m) was modified to 8[kg], while maintaining the output saturation operative within the range of $[-5, 5][N]$ to test the saturated performances. Before initiating the simulation, the input constraints (Y) of the MPC are defined as:

$$|F_i| \leq 5, \quad i = 1, 2. \quad (35)$$

However, Equation (30) and (31) imply that adjustments applied to the LQR and SMC weight matrices affect the

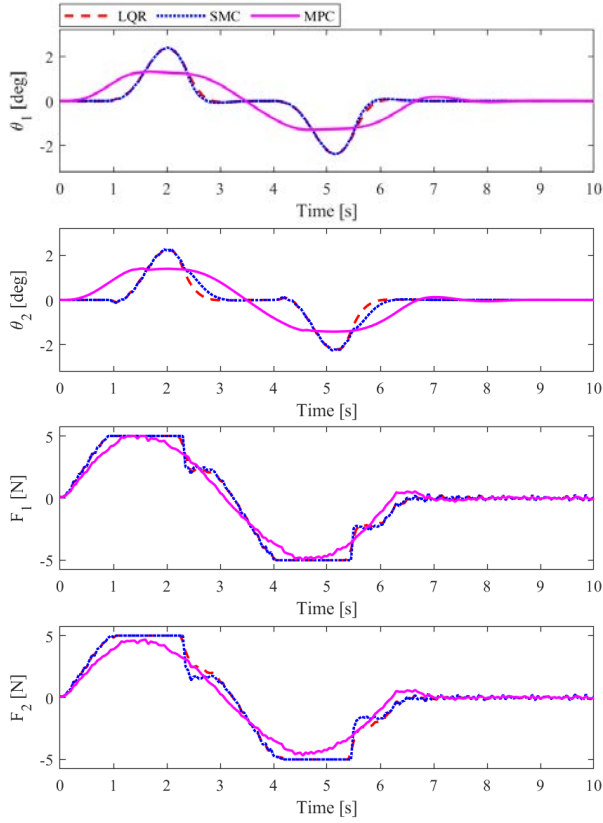


Figure 9. Simulation result of swing angle (θ_1 , θ_2) and applied force (F_1 , F_2) with increased m_m

magnitude of u , which only has a relatively minor impact on F_1 , F_2 compared to the influence of acceleration compensation when mass is considerable.

The performance of these controllers, as illustrated in Figure 9 and Table 1 demonstrated how MPC accurately anticipates future saturation conditions and adjusts input magnitudes in advance to minimize the maximum swing angle. The forces applied are recorded in Figure 9, clearly highlighting that after the initiation of the simulation, both the LQR and SMC quickly reached their upper limits, consistently driving the MiC and hook with maximum output. As the trolley begins to decelerate, there is a sudden drop in force. In contrast, the output curve of the MPC remains smooth and continuous, demonstrating a more amicable interaction with the actuator. The results indicate that, while maintaining a similar overall RMSE, MPC exhibits a more conservative swing angle and is more actuator-friendly. This attribute makes MPC the most suitable proactive controller for MiC, which inherently carries a substantial load.

4 Conclusion

This study introduced a proactive hook-load stabilization system to address the inherent limitations of under-actuated overhead cranes by adding two auxiliary inputs. With the focus on single-axis trolley motion along the jib, three Multiple-Input Multiple-Output (MIMO) controllers were compared in simulations, and the main outcomes are:

- Efficacy is verified through three comparative simulations, demonstrating superior performance against a control-less scenario.
- Resilience to uncertainty and undesired conditions is confirmed by considering factors like control signal delays, noisy sensor readings, and actuator saturation.

Results indicated the proposed MPC-based method enables more aggressive trolley motions, enhancing fault tolerance for novice operators and eventually maximizing operational efficiency.

Future works of this study involve conducting scaled experiments and observer design to enhance robustness under complex integrated disturbances. Meanwhile, the dimensionality of the system will be expanded to 3-D, considering the movement of the trolley in the complete horizontal plane.

References

- [1] G. H. Kim and K. S. Hong. Adaptive sliding-mode control of an offshore container crane with unknown disturbances. *IEEE/ASME Transactions On Mechatronics*, 24(6):2850–2861, 2019. doi:10.1109/TMECH.2019.2946083.
- [2] T. Yang, N. Sun, H. Chen, and Y. Fang. Adaptive optimal motion control of uncertain underactuated mechatronic systems with actuator constraints. *IEEE/ASME Transactions on Mechatronics*, 28(1):210–222, 2022. doi:10.1109/TMECH.2022.3192002.
- [3] N. Sun, Y. Wu, Y. Fang, and H. Chen. Nonlinear antiswing control for crane systems with double-pendulum swing effects and uncertain parameters: Design and experiments. *IEEE Transactions on Automation Science and Engineering*, 15(3):1413–1422, 2017. doi:10.1109/TASE.2017.2723539.
- [4] L. Ramli, Z. Mohamed, A. M. Abdullahi, H. I. Jaafar, and I. M. Lazim. Control strategies for crane systems: A comprehensive review. *Mechanical Systems and Signal Processing*, 95:1–23, 2017. doi:10.1016/j.ymssp.2017.03.015.

- [5] B. Zhao, H. Ouyang, and M. Iwasaki. Motion trajectory tracking and sway reduction for double-pendulum overhead cranes using improved adaptive control without velocity feedback. *IEEE/ASME Transactions on Mechatronics*, 27(5):3648–3659, 2021. doi:10.1109/TMECH.2021.3126665.
- [6] G. Li, X. Ma, Z. Li, and Y. Li. Time-polynomial-based optimal trajectory planning for double-pendulum tower crane with full-state constraints and obstacle avoidance. *IEEE/ASME Transactions on Mechatronics*, 28(2):919–932, 2022. doi:10.1109/TMECH.2022.3210536.
- [7] T. Yang, N. Sun, H. Chen, and Y. Fang. Neural network-based adaptive antiswing control of an underactuated ship-mounted crane with roll motions and input dead zones. *IEEE Transactions on Neural Networks and Learning Systems*, 31(3):901–914, 2019. doi:10.1109/TNNLS.2019.2910580.
- [8] H. Ouyang, J. Hu, G. Zhang, L. Mei, and X. Deng. Decoupled linear model and s-shaped curve motion trajectory for load sway reduction control in overhead cranes with double-pendulum effect. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 233(10):3678–3689, 2019. doi:10.1177/0954406218819029.
- [9] H-H. Lee. Motion planning for three-dimensional overhead cranes with high-speed load hoisting. *International Journal of Control*, 78(12):875–886, 2005. doi:10.1080/00207170500197571.
- [10] N. Sun, Y. Fang, Y. Zhang, and B. Ma. A novel kinematic coupling-based trajectory planning method for overhead cranes. *IEEE/ASME Transactions on Mechatronics*, 17(1):166–173, 2011. doi:10.1109/TMECH.2010.2103085.
- [11] M. Zhang, Y. Zhang, B. Ji, C. Ma, and X. Cheng. Modeling and energy-based sway reduction control for tower crane systems with double-pendulum and spherical-pendulum effects. *Measurement and Control*, 53(1-2):141–150, 2020. doi:10.1177/0020294019877492.
- [12] C. Aguiar, D. Leite, D. Pereira, G. Andonovski, and I. Škrjanc. Nonlinear modeling and robust lmi fuzzy control of overhead crane systems. *Journal of the Franklin Institute*, 358(2):1376–1402, 2021. doi:10.1016/j.jfranklin.2020.12.003.
- [13] Y. Zhao and H. Gao. Fuzzy-model-based control of an overhead crane with input delay and actuator saturation. *IEEE Transactions on Fuzzy Systems*, 20(1):181–186, 2011. doi:10.1109/TFUZZ.2011.2164083.
- [14] R. Roman, R. Precup, and E. M. Petriu. Hybrid data-driven fuzzy active disturbance rejection control for tower crane systems. *European Journal of Control*, 58:373–387, 2021. doi:10.1016/j.ejcon.2020.08.001.
- [15] M. Vukov, W. Van Loock, B. Houska, H. J. Ferreau, J. Swevers, and M. Diehl. Experimental validation of nonlinear mpc on an overhead crane using automatic code generation. In *2012 American Control Conference (ACC)*, pages 6264–6269. IEEE, 2012. doi:10.1109/ACC.2012.6315390.
- [16] D. Schindele and H. Aschemann. Fast nonlinear mpc for an overhead travelling crane. *IFAC proceedings volumes*, 44(1):7963–7968, 2011. doi:10.3182/20110828-6-IT-1002.03510.
- [17] P. Bauer and J. Bokor. Performance comparison of siso and mimo low-level controllers in a special trajectory tracking application. In *22nd Mediterranean Conference on Control and Automation*, pages 1293–1298. IEEE, 2014. doi:10.1109/MED.2014.6961554.
- [18] H. Chen, Y. Fang, and N. Sun. A swing constraint guaranteed mpc algorithm for underactuated overhead cranes. *IEEE/ASME Transactions on Mechatronics*, 21(5):2543–2555, 2016. doi:10.1109/TMECH.2016.2558202.
- [19] H. Chen and N. Sun. Nonlinear control of underactuated systems subject to both actuated and unactuated state constraints with experimental verification. *IEEE Transactions on Industrial Electronics*, 67(9):7702–7714, 2019. doi:10.1109/TIE.2019.2946541.
- [20] J. Vaughan, D. Kim, and W. Singhose. Control of tower cranes with double-pendulum payload dynamics. *IEEE Transactions on Control Systems Technology*, 18(6):1345–1358, 2010. doi:10.1109/TCST.2010.2040178.
- [21] J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl. CasADi – A software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation*, 11(1):1–36, 2019. doi:10.1007/s12532-018-0139-4.

Reinvent reinforced concrete with robotics and 3D printing

Jean-Francois Caron¹, Nicolas Ducoulombier³, Léo Demont¹, Victor de Bono^{1,3} and Romain Mesnil^{1,2}

¹Navier Laboratory, Ecole des Ponts ParisTech-Univ. Gustave Eiffel-CNRS, Marne la Vallée, France

²Build'in, Co-Innovation Lab École des Ponts ParisTech, Marne la Vallée, France

³XtreeE, 18 Rue du Jura, Rungis, France

jean-francois.caron@enpc.fr,

Abstract -

3D concrete has undergone significant development over the last few decades. Yet unreinforced printed elements generally do not comply with existing building standards or regulations and are therefore not used as load-bearing elements. There is still a gap between research and use, and despite several proposals, standard commercial solutions for the reinforcement of 3D-printed structural elements are still some way off. The proposed concept is inspired by the composites industry and echoing Pier Luigi Nervi's last century Ferrocement, uses 3D Concrete Printing (3DCP) and a patented technology called FBP for *Flow-Based Pultrusion for additive manufacturing*. The reinforcement is provided by long, aligned fibers, and produces a transverse isotropic composite mortar. The first experiments demonstrate an increase in tensile strength and ductility, and an industrial prototype, in collaboration with XtreeE company, is currently under development to push away the Reinforced Concrete limits proposing a disruptive way to design and build with concrete.

Keywords -

Cementitious composite; 3D printing; Long fibers

1 Introduction

Possessing a tensile strength within the range of a few MPa and a fragile behavior, unreinforced cementitious mortars are unsuitable for tensile configurations. Microcracks, initiating at an early stage, swiftly propagate, leading to rapid rupture. The remedy lies in intimately "stitching" these cracks with a suitable reinforcement so that stresses can continue flowing across the crack, limiting the rupture process. For concretes and mortars, the most developed solutions are structural ones, using macro systems as rebars and cables, usually steel ones. This idea is attributed to the French engineer Joseph Lambot in 1848, who first thought of reinforcing the cement of his agricultural tanks and a small boat with an iron lattice. He patented this concept as "Ferciment" in 1855, which later became "Ferrocement" [1]. Numerous devel-

opments followed, particularly for building and civil engineering, with 262 patents filed between 1886 and 1906, leading to the first regulations in France that same year. Then, two approaches emerged. The first approach derived from Lambot's one is promoted by Cottencin, and notably adopted by the architect Anatole de Baudot for the Saint-Jean de Montmartre church (1894-1904) and the vaulted roof of Tulle Theater. The second proposed construction systems like beam and column structures with the progressive use of larger-section steel profiles, known as "rebars". Initially promoted by François Hennebique, this approach eventually prevailed and still constitutes a significant part of current constructions, more "readable" and much easier to build. Many references contributed quickly to its reputation, and Hennebique's system closed the way to the more optimal Cottencin's system. However the famous Italian architect-engineer Pier Luigi Nervi, in the mid-last century, continued to develop this concept of fine steel mesh. "Ferrocement" then became a technique enabling architectural challenges, such as in the Orbetello Aerodrome hangar (Italy) completed in 1940, with astonishing efficiency and aesthetics (see Figure 1). Nervi prefabricated small structural elements and joined them by keying rebars with concrete on site [18], restoring the structural continuity of the work. This solution embodies "resistance through form", a concept Nervi advocates for generalization, making the complexity and beauty of buildings possible. It's time to revisit these choices, with today's tools, robotics, given the visionary concepts and their relevance to today's resources and challenges.



Figure 1. Orbetello Aerodrome in Italy by Pier Luigi Nervi in 1940 (Credits : Hulton Getty)



Figure 2. Post-reinforcement: a) Pile with printed formworks [10] b) traditional reinforcement [2] c) reservations for prestressing cables [21]

2 Reinforcement and 3D printing

There exist three ways to reinforce a 3D printed element : before, during or after printing. The following sum-up these different approaches, starting with the most obvious and easy one, reinforcing the already printed element.

2.1 Reinforce after printing

The structure is reinforced after printing, meaning once it's already in place. Traditional techniques such as reinforced or prestressed concrete, as mentioned earlier, are applied. For instance, it's possible to create structures with printed cementitious materials that include openings for inserting reinforcement bars or prestressing cables, which are secured by injecting a grout, as shown in Figure 2-c. This method can be quickly deployed on construction sites since existing design methods, integrated into regulatory frameworks, are directly applicable. Another approach to achieving complex-shaped structures compliant with regulations using 3D printing is to print the formwork into which traditional concrete, reinforced by traditional methods, will be cast. This is demonstrated by the Chinese company Winsun (see Figure 2-b), positioning reinforcements in a printed mold. XtreeE, on the other hand, prefers the use of ultra-high-performance fiber-reinforced concrete, which is easier to work with, for the construction of the column illustrated in Figure 2-a. In [22, 23] another example concerns a methodology to optimize a Eurocode-compliant beam, using such printed formworks. The potential concrete savings are significant: through a parametric study, they were estimated at 48% on average for beams of housing buildings. A prototype is presented in Figure 3 where the formworks and the robotic setups are also shown.

Other proposals involve reinforcing 3D-printed structures with external reinforcements. In [1], a beam is suggested in which the compression elements are 3D printed in cementitious material using robotic extrusion, while the tensile forces are absorbed by an external metallic framework. In [33], a bridge is constructed with several



Figure 3. Prototype of an optimized reinforced concrete beam with printed formworks [23].

printed mortar voussoirs assembled using post-tensioned prestressing cables.

2.2 Reinforce before the printing

It is well known that another way to reinforce concrete is to introduce a modest proportion of fibers with high tensile strength enhances tensile behavior, allowing for the sewing of micro or macro-cracks and averting sudden, localized rupture. Typically, reinforcement is achieved through short fibers added directly to the mortar in the mixer with a limited volume fraction (a few percent by volume). Hence are produced high-performance fiber-reinforced concrete (HPFRCC) with metal fibers —the only ones specified in structural design regulations—while preserving favorable workability [32]. A natural idea for 3D printing is to integrate reinforcement in the same way, into the cementitious material from the initial mixing and extrude fiber-reinforced mortar. In order to obtain pumpable mortar, the fibers must be flexible and therefore have a limited diameter. The most commonly used fibers are synthetic fibers (carbon, polyvinyl alcohol: PVA, Polyethylene PE) or mineral fibers (glass or basalt). To maintain a mode of failure by decohesion of the interface and/or not to block the pumping and extrusion system, the length of the fibers is also limited (approximately 10 mm). Finally, to ensure a sufficiently low shear yield stress of the fresh mortar, the volumetric proportion is limited, generally less than or equal to 2% [30]. In [34], such a volumetric fiber rate of PE fibers allowed for a very significant deformation at rupture in the deposition direction of about 10% for a flexural strength of 20 MPa. Despite extrusion difficulties and relatively poor volume ratio of fibers, the effectiveness of this type of reinforcement for printed mortars will see significant improvement in the coming years.

2.3 Reinforce during the printing with robotics

The idea here is to make the mortar printing and the reinforcement placement concurrent. These approaches are more sophisticated and generally developed for non-standard concrete structures within an integrated digital process based on advanced robotic manufacturing strategies.

Several research efforts have focused on the direct adaptation of the reinforced concrete principle using steel bars or rebars, as seen in [8, 31]. This approach is interesting as it provides quick perspectives on acceptability and use by professionals. Most solutions involve robotically inserting small bars, somewhat like nails, into the printed stack, either perpendicular to the stacking direction or at an angle. This has the advantage of reinforcing the crossed interfaces, which can be weak points (cold joints) for the structure if there is poor adhesion between layers due to, for example, too rapid drying of the layer receiving the new deposit.

In [8], a process called Additive Manufacturing of Reinforced Concrete (AMoRC) is proposed. It is a hybrid process involving intermittent stud welding and continuous concrete extrusion. Segmented steel reinforcement bars are assembled to form a three-dimensional reinforcement mesh, simultaneously integrated into the extruded concrete (Figure 4).

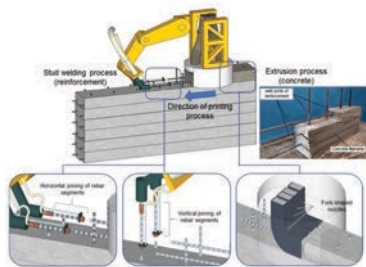


Figure 4. 3D Printing process (AMoRC), concrete extrusion and welding of reinforcements [8].

In [31], experiments with this type of nailing reinforcement on various mortars, already printed, demonstrated effective reinforcement if the nail orientation is correctly chosen, and if the nail surface is rough enough to ensure a good interface with the mortar.

In [16], U-shaped reinforcement elements, called "riders", staple vertically through the layers simultaneously during printing. The staples penetrate multiple layers but also interlock to form a reinforcement network. Finally, in [15], a device with a dispenser (Figure 5) deposits needles horizontally between the layers to achieve reinforcement along the axis of the strands, approaching traditional reinforced concrete. Mechanically integrating continuous and unrolled reinforcement at the nozzle level, such as micro-cables, chains, or metal grids [4, 20, 3, 24] (Figure 6), is also a possibility. Particular attention must be paid to the interface quality between the reinforcement and the cementitious material, which determines the composite behavior. Mineral fiber elements are widely investigated, for example, in [26] or by the authors of this chapter in [14]. The former implements a thin and narrow strip of Mineral-

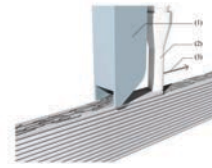


Figure 5. Deposition of needles between printed strands [15].

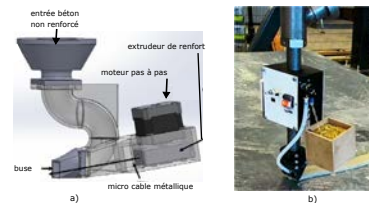


Figure 6. Co-extrusion a) with a metal cable [20] b) and with a metal chain [4].

Impregnated Carbon Fiber (MCF) unrolled downstream of the nozzle on an already printed layer, covered by the current deposit. In [28], to ensure good impregnation of the carbon fiber wick, a pre-impregnation line using very fine mortar is set up upstream to prepare the strip to be deposited (Figure 7). The threads are then mechanically placed between the layers upstream of the printed strand (Figure 8). A variant of the MCF process called ProfiCarb

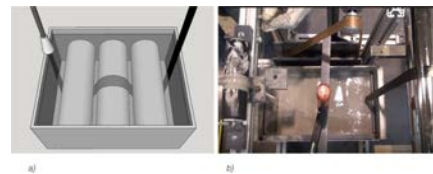


Figure 7. Pre-impregnation system of the MCF process [27].

[29] repackages pre-impregnated threads into coils that will be directly driven by the flow of a single-component mortar (Figure 9). Several reinforcements, here 6 twisted carbon fiber yarns, can thus be dispersed in the section of the strand (Figure 9). This latter process is very close to the Flow-Based-Pultrusion concept proposed by the authors of this paper [14], which will be detailed in the following paragraph. This approach allows more homogeneous, "in-line" reinforcement and avoids technologies and motorizations for guiding and driving reinforcement, which complicate the process and reduce formal possibilities, such as pronounced curves, for example.

All these continuous reinforcement solutions naturally align with traditional reinforced concrete techniques. They



Figure 8. MCF nozzle with motorized routing of reinforcement under the extruded strand [27].



Figure 9. ProfiCarb process and view of the nozzle with 6 yarns [17, 29].

reinforce the material in a preferred direction [2], induce anisotropy, and are therefore very effective for the structure's bending behavior, even for moderate volumetric fractions of reinforcement. To go further, they must fit as much as possible into the formal freedom offered by robotics.

The next paragraph focuses on the Flow-Based Pultrusion concept [14, 6], which seems to meet the imperative of efficiency, simplicity, and adaptability to new construction system proposals.

3 Details on a Specific In-Line Reinforcement: The Flow-Based Pultrusion Concept

What is proposed here is an in-line approach and material strategy, heavily inspired by long-fiber composite material technologies. On the contrary of localized rebars in reinforced concrete, the extruded material is reinforced in a more uniform manner, with continuous aligned fibers distributed inside the extruded lace. We will briefly describe the technology, experiments, and the performance of the reinforced material.

3.1 Technology

The technology described in the patent [5] is called FBP, which stands for Flow-Based Pultrusion. Control of the rheological behavior of the cement matrix ensures the routing and impregnation of continuous small-diameter strands (glass, basalt, etc.) without any motorization (10 left). Such technological simplicity has numerous advantages: it is cost-effective, more reliable since the absence of motors decreases potential points of failure, and re-

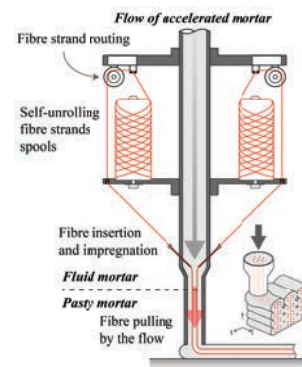


Figure 10. Left: Fiber printing device: 1/coils 2/pulleys 3/guides 4/drive head. Bottom right: Transverse isotropic arrangement in the printing direction via the patented Flow-Based Pultrusion process [5]

duces the complexity of the system control. The resulting material, called *Anisotropic Concrete*, is thus uniformly reinforced in a single direction, offering new possibilities (10 right). It can improve the strength and ductility of the cured material and contribute to better handling of fresh laces during deposition, allowing for more complex paths on slopes or cantilevers, as shown later (Figure 12 right).

3.2 Prototypes and Devices

A first prototype was developed at the Navier laboratory and adapted to a two-component printing head from XtreeE, mounted on a 6-axis ABB 6200 robot (Figure 11). This current prototype works with fine-grained mortar, and not with large aggregates. An important challenge in this technology is that the reinforcement should be sufficiently flexible to match the curvature of a printed lace, and should provide sufficient anchoring capacity in the printed material. Not all reinforcement strands are compatible due to these challenges: first experiments using “piano cord” metallic wires were not satisfactory due to poor adhesion with the mortar and bending radius limitations [13]. So-called roving yarns are more favorable. They are constituted of multiple microfilaments, are much more flexible, and easier to bond to the mortar due to their microfilamentary structure.

In 3D concrete printing by extrusion, there are two main technologies: single-component or two-component. The first deposits a lace directly pumped from a batch of mortar prepared and having the right consistency to be stacked. A powerful pump is necessary because the mixture is relatively thick, having a high yield stress (a few thousand KPa) that evolves quite slowly but irreversibly until a consistency too high to be printed is reached. The two-component technology prints a much more fluid material (a few hun-

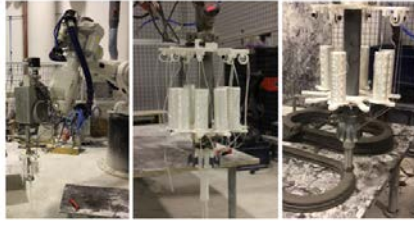


Figure 11. Prototype 1, Navier laboratory, with fiber-glass printing

dred KPa) accelerated strongly by an additive just before printing. The initial mixture is therefore easier to pump, more stable, and allows more freedom since the operator can continuously modify it based on what is sought or needs correction. Our FBP process is closely linked to this technology, which allows easy introduction of fibers into a fluid material that quickly solidifies. The impregnation is very effective, and the drive is assured because the mortar quickly structures around the fiber. Initial prints have been made with glass, basalt, and carbon fibers, up to 6% by volume. Details on the process technology can be found in [9]. To hope to go further, a second prototype for industrial purposes is under development in collaboration with XtreeE. In Figure 12 is shown from left to right : the

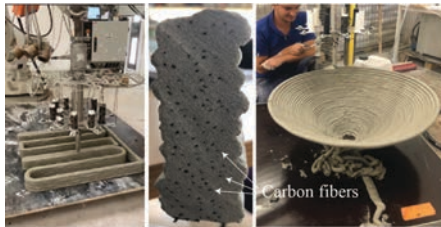


Figure 12. Prototype 2, XtreeE:carbon fibers, uniformly distributed in the cross-section (center). A fiber-reinforced corolla structure (right).

XtreeE FBP prototype, the cross-section of a cured cord reinforced with 1% carbon fibers in the center, and the print of a corolla with an overhang angle of more than 60 degrees, that would not be printable with a regular tool-path slicing and unreinforced printing mortar because the rings quickly reach a state of uniform axial tension that the mortar alone cannot sustain.

3.3 Characteristics of the Reinforced Cured Material

Tensile and flexural tests were conducted on reinforced laces with different rates of glass and basalt fibers. Cracking evolution and strains were monitored with Digital Image Correlation (DIC) using a custom MATLAB code called *NavDIC* [25]. Other tests are underway with other

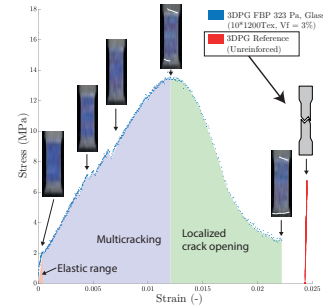


Figure 13. Tensile test of a 3% glass fiber-reinforced tensile specimen, and image correlation revealing the progressive appearance of cracks. On the left in red, a non-reinforced specimen.

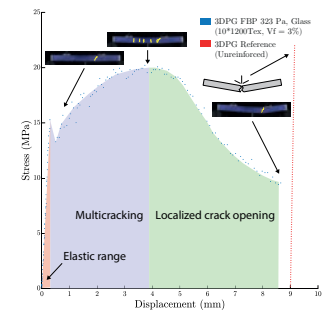


Figure 14. 4-point bending test of a 3% glass fiber-reinforced tensile specimen, and image correlation revealing the progressive appearance of cracks. On the left in red, a non-reinforced specimen.

fibers such as carbon or PVA. The specimens consist of one layer for tension and multiple stacked layers for flexion. Figure 13 shows the type of tensile behavior obtained with 3% fibers (here glass), and the high ductility obtained thanks to the diffuse microcracking made possible by the presence of fibers [19]. Flexural behavior is very similar, as seen on Figure 14, and more details can be found in [6]. Meter-scale tests on beams with a reinforced intrados are underway, the initial results are very promising [7] and suggest that the pseudo-ductile diffuse microcracking behaviour works at the structure scale.

4 Conclusion and Perspectives

Considering the low mechanical characteristics obtained in tension and the regulatory framework, reinforcing 3D printed mortar is imperative. It is an highly strategic research and industrial issue for the commercial development of the technology. While there is currently no off-the-shelf solution, several initiatives exist and already provide possible solutions that now need to be developed further. Often inspired by the long and rich history of ce-

ment and cast concrete reinforcement technologies, they also try to leverage new tools available today, such as robotics, parametric design, or materials science. A brief reminder of existing initiatives is proposed and a classification is made according to whether they post-reinforce the structure, reinforce the initial fresh batch or in-line. The first one has the advantage of being quickly implemented in real situations, through demonstrators, and thus promoting the technology. The second one is very promising, as this type of fiber mortar is already widely accepted and regulated in the community. Technical progress is still needed to improve achievable fiber rates and the workability of mortars, but the initial results in terms of strength and ductility are very encouraging. Finally, in-line reinforcement is certainly an interesting avenue, perhaps more in tune with the spirit of technology and construction 4.0. This should allow for stronger reinforcement, especially when a composite material approach is adopted with small and numerous fibres, in competition with steel systems for conventional concretes, more flexible in use, and adaptable to the needs and objects to be reinforced. However, these latter systems still pose some problems, anchoring and flexibility of reinforcements, technological simplicity, and compatibility with the formal performance expected from 3D printing, which should enable new structural systems to be proposed. A closer look is given at the initiative of the authors of this paper, an in-line reinforcement method called Flow-Based Pultrusion, which could address several challenges: good anchoring of fibers in the cement matrix, technological simplicity, and inspiration for new construction systems. As an example an original concrete truss concept, illustrated in Figure 15, and presented in [11] is realized through robotically free-form deposition [12]. In this type of structure, the elements work only in tension or compression. The anisotropy brought by a robotized in-line fiber extrusion process would therefore be particularly suitable and disruptive in terms of concrete construction systems : It leverages the opportunity to go beyond the specific performance of additive manufacturing (*putting the right material in the right place*) by putting *the right reinforcement in the right direction*.

In conclusion, the competition between two reinforcement systems proposed at the end of the 19th century and mentioned in the introduction, the *Hennebique* system (reinforced posts and beams) and the *Cottencin* system (thin ferrocement slabs), could well be replayed in favor of the latter thanks to new printing processes and in-line reinforcement as proposed by Flow-Based Pultrusion. This could reinvent more free and lighter structures such as those proposed by Candela or Pier Luigi Nervi (Figure 1). The latter wanted, in his own words, to “get out of a formwork architecture” by making the use of wooden formwork unnecessary and serving a vision that coupled

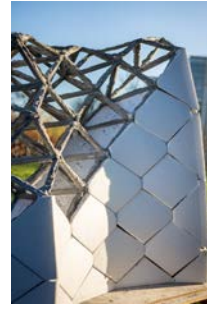


Figure 15. Curved concrete truss [11], Photo credit Stefano Borghi

material, form, and structure for more virtuous constructions.

References

- [1] D. Asprone, F. Auricchio, C. Menna, and V. Mercuri. 3d printing of reinforced concrete elements: Technology and design approach. *Construction and Building Materials*, 165:218–231, Mar. 2018. ISSN 09500618. doi:10.1016/j.conbuildmat.2018.01.018. URL <https://linkinghub.elsevier.com/retrieve/pii/S0950061818300187>.
- [2] D. Asprone, C. Menna, F. P. Bos, T. A. Salet, J. Mata-Falcón, and W. Kaufmann. Rethinking reinforcement for digital fabrication with concrete. *Cement and Concrete Research*, 112:111–121, Oct. 2018. ISSN 00088846. doi:10.1016/j.cemconres.2018.05.020. URL <https://linkinghub.elsevier.com/retrieve/pii/S0008884618300309>.
- [3] F. Bos, Z. Ahmed, E. Jutinov, and T. Salet. Experimental Exploration of Metal Cable as Reinforcement in 3d Printed Concrete. *Materials*, 10(11):1314, Nov. 2017. ISSN 1996-1944. doi:10.3390/ma10111314. URL <http://www.mdpi.com/1996-1944/10/11/1314>.
- [4] F. P. Bos, E. Bosco, and T. A. M. Salet. Ductility of 3d printed concrete reinforced with short straight steel fibers. *Virtual and Physical Prototyping*, 14(2):160–174, Apr. 2019. ISSN 1745-2759, 1745-2767. doi:10.1080/17452759.2018.1548069. URL <https://www.tandfonline.com/doi/full/10.1080/17452759.2018.1548069>.
- [5] J. Caron and N. Ducoulombier. Method and device for manufacturing anisotropic concrete. *Patent*, wo2020249913/fr3097152, 2020.
- [6] J.-F. Caron, L. Demont, N. Ducoulombier, and R. Mesnil. 3d printing of mortar with

- continuous fibres: Principle, properties and potential for application. *Automation in Construction*, 129:103806, 2021. ISSN 0926-5805. doi:<https://doi.org/10.1016/j.autcon.2021.103806>. URL <https://www.sciencedirect.com/science/article/pii/S0926580521002570>.
- [7] J.-F. Caron, N. Ducoulombier, L. Demont, V. de Bono, and R. Mesnil. 3D Printing of Continuous-Fibers Cementitious Composites: Anisotropic 3D Mortar. *Open Conference Proceedings*, 3, Dec. 2023. ISSN 2749-5841. doi:10.52825/ocp.v3i.193.
- [8] M. Classen, J. Ungermann, and R. Sharma. Additive manufacturing of reinforced concrete—development of a 3d printing technology for cementitious composites with metallic reinforcement. *Applied Sciences*, 10(11), 2020. ISSN 2076-3417. doi:10.3390/app10113791. URL <https://www.mdpi.com/2076-3417/10/11/3791>.
- [9] L. Demont, N. Ducoulombier, R. Mesnil, and J.-F. Caron. Flow-based pultrusion of continuous fibers for cement-based composite material and additive manufacturing: rheological and technological requirements. *Composite Structures*, 262:113564, 2021. ISSN 0263-8223. doi:<https://doi.org/10.1016/j.compstruct.2021.113564>. URL <https://www.sciencedirect.com/science/article/pii/S0263822321000258>.
- [10] R. Duballet. *Building systems in robotic extrusion of cementitious materials*. PhD thesis, Université Paris-est, Dec. 2018.
- [11] R. Duballet, O. Baverel, and J. Dirrenberger. Space Truss Masonry Walls With Robotic Mortar Extrusion. *Structures*, 18:41–47, Apr. 2019. ISSN 23520124. doi:10.1016/j.istruc.2018.11.003. URL <https://linkinghub.elsevier.com/retrieve/pii/S2352012418301309>.
- [12] R. Duballet, R. Mesnil, N. Ducoulombier, P. Carneau, L. Demont, M. Motamedi, O. Baverel, J.-F. Caron, and J. Dirrenberger. Free Deposition Printing for Space Truss Structures. In F. P. Bos, S. S. Lucas, R. J. Wolfs, and T. A. Salet, editors, *Second RILEM International Conference on Concrete and Digital Fabrication*, RILEM Bookseries, pages 873–882, Cham, 2020. Springer International Publishing. ISBN 978-3-030-49916-7. doi:10.1007/978-3-030-49916-7_85.
- [13] N. Ducoulombier, C. Chateau, M. Bornert, J.-F. Caron, T. Weitkamp, and J. Perrin. Characterisation and modelling of interfacial damage in fibre-reinforced concrete for 3D printing in construction. In *10th International Conference on Fracture Mechanics of Concrete and Concrete Structures*. IA-FraMCoS, June 2019. doi:10.21012/FC10.235562.
- [14] N. Ducoulombier, L. Demont, C. Chateau, M. Bornert, and J.-F. Caron. Additive manufacturing of anisotropic concrete: a flow-based pultrusion of continuous fibers in a cementitious matrix. *Procedia Manufacturing*, 47:1070–1077, 2020. ISSN 2351-9789. doi:<https://doi.org/10.1016/j.promfg.2020.04.117>. URL <https://www.sciencedirect.com/science/article/pii/S2351978920311732>. 23rd International Conference on Material Forming.
- [15] L. Gebhard and al. Aligned interlayer fibre reinforcement and post-tensioning as a reinforcement strategy for digital fabrication. In *Second RILEM International Conference on Concrete and Digital Fabrication*, pages 622–631, Cham, 2020. Springer International Publishing. ISBN 978-3-030-49916-7.
- [16] O. Geneidy, S. Kumarji, A. Dubor, and A. Sollazzo. Simultaneous reinforcement of concrete while 3d printing. In F. P. Bos, S. S. Lucas, R. J. Wolfs, and T. A. Salet, editors, *Second RILEM International Conference on Concrete and Digital Fabrication*, pages 895–905, Cham, 2020. Springer International Publishing. ISBN 978-3-030-49916-7.
- [17] E. Ivaniuk, S. Müller, T. Neef, and V. Mechtcherine. Strategies for Integrating Reinforcement Into 3D Concrete Printing at the TU Dresden. *Open Conference Proceedings*, 1:23–34, Feb. 2022. ISSN 2749-5841. doi:10.52825/ocp.v1i.73.
- [18] T. Leslie. ‘Laborious and difficult’: The evolution of Pier Luigi Nervi’s hangar roofs (1935–41). In I. Wouters, S. Van De Voorde, I. Bertels, B. Espion, K. De Jonge, and D. Zastavni, editors, *Building Knowledge, Constructing Histories*, pages 229–234. CRC Press, 1 edition, Sept. 2018. ISBN 978-0-429-50620-8. doi:10.1201/9780429506208-32.
- [19] V. C. Li and H.-C. Wu. Conditions for Pseudo Strain-Hardening in Fiber Reinforced Brittle Matrix Composites. *Applied Mechanics Reviews*, 45(8): 390–398, Aug. 1992. ISSN 0003-6900, 2379-0407. doi:10.1115/1.3119767.
- [20] J. H. Lim, B. Panda, and Q.-C. Pham. Improving flexural characteristics of 3d printed geopolymer composites with in-process steel

- cable reinforcement. *Construction and Building Materials*, 178:32–41, July 2018. ISSN 09500618. doi:10.1016/j.conbuildmat.2018.05.010. URL <https://linkinghub.elsevier.com/retrieve/pii/S0950061818310675>.
- [21] S. Lim, R. Buswell, T. Le, S. Austin, A. Gibb, and T. Thorpe. Developments in construction-scale additive manufacturing processes. *Automation in Construction*, 21:262–268, Jan. 2012. ISSN 09265805. doi:10.1016/j.autcon.2011.06.010. URL <https://linkinghub.elsevier.com/retrieve/pii/S0926580511001221>.
- [22] S. Maitenaz, M. Charrier, R. Mesnil, P. Onfroy, N. Metge, A. Feraille, and J.-F. Caron. Fabrication of a truss-like beam casted with 3d printed clay moulds. In *Proceedings of the 38th International Symposium on Automation and Robotics in Construction (ISARC)*, pages 712–716, Dubai, UAE, November 2021. International Association for Automation and Robotics in Construction (IAARC). ISBN 978-952-69524-1-3. doi:10.22260/ISARC2021/0096.
- [23] S. Maitenaz, R. Mesnil, P. Onfroy, N. Metge, A. Feraille, and J.-F. Caron. Materialising structural optimisation of reinforced concrete beams through digital fabrication. *Accepted in Structures*, Nov. 2023.
- [24] T. Marchment and J. Sanjayan. Mesh reinforcing method for 3D Concrete Printing. *Automation in Construction*, 109:102992, Jan. 2020. ISSN 09265805. doi:10.1016/j.autcon.2019.102992. URL <https://linkinghub.elsevier.com/retrieve/pii/S0926580519306132>.
- [25] P. Margerit. Navdic, an open source matlab dic app, July 2022. URL https://github.com/MargeritPierre/navDIC_v0.
- [26] V. Mechtcherine, A. Michel, M. Liebscher, and T. Schmeier. Extrusion-based additive manufacturing with carbon reinforced concrete: Concept and feasibility study. *Materials*, 13(11), 2020. ISSN 1996-1944. doi:10.3390/ma13112568. URL <https://www.mdpi.com/1996-1944/13/11/2568>.
- [27] V. Mechtcherine, A. Michel, M. Liebscher, and T. Schmeier. Extrusion-Based Additive Manufacturing with Carbon Reinforced Concrete: Concept and Feasibility Study. *Materials*, 13(11):2568, Jan. 2020. doi:10.3390/ma13112568.
- [28] V. Mechtcherine, A. Michel, M. Liebscher, K. Schneider, and C. Großmann. Mineral-impregnated carbon fiber composites as novel reinforcement for concrete construction: Material and automation perspectives. *Automation in Construction*, 110:103002, Feb. 2020. ISSN 0926-5805. doi:10.1016/j.autcon.2019.103002.
- [29] T. Neef and V. Mechtcherine. Simultaneous Integration of Continuous Mineral-Bonded Carbon Reinforcement Into Additive Manufacturing With Concrete. *Open Conference Proceedings*, 1:73–81, Feb. 2022. ISSN 2749-5841. doi:10.52825/ocp.v1i.80.
- [30] H. Ogura, V. N. Nerella, and V. Mechtcherine. Developing and Testing of Strain-Hardening Cement-Based Composites (SHCC) in the Context of 3D-Printing. *Materials*, 11(8):1375, Aug. 2018. ISSN 1996-1944. doi:10.3390/ma11081375.
- [31] A. Perrot, Y. Jacquet, D. Rangeard, E. Courteille, and M. Sonebi. Nailing of layers: A promising way to reinforce concrete 3d printing structures. *Materials*, 13(7), 2020. ISSN 1996-1944. doi:10.3390/ma13071518. URL <https://www.mdpi.com/1996-1944/13/7/1518>.
- [32] P. Rossi. Ultra-high-performance fiber-reinforced concretes. *Concrete international*, December 1982: 46–52, Dec. 1982.
- [33] T. A. M. Salet, Z. Y. Ahmed, F. P. Bos, and H. L. M. Laagland. Design of a 3d printed concrete bridge by testing. *Virtual and Physical Prototyping*, 13(3):222–236, 2018. doi:10.1080/17452759.2018.1476064. URL <https://doi.org/10.1080/17452759.2018.1476064>.
- [34] B. Zhu, J. Pan, B. Nematollahi, Z. Zhou, Y. Zhang, and J. Sanjayan. Development of 3D printable engineered cementitious composites with ultra-high tensile ductility for digital construction. *Materials & Design*, 181:108088, Nov. 2019. ISSN 02641275. doi:10.1016/j.matdes.2019.108088. URL <https://linkinghub.elsevier.com/retrieve/pii/S026412751930526X>.

Unified framework for mixed-reality assisted situational adaptive robotic path planning enabled by 5G networks for deconstruction tasks

Chu Han Wu¹, Marit Zöcklein² and Sigrid Brell-Cokcan³

¹Chair for Individualized Production in Architecture, RWTH Aachen University, Germany

²Center Construction Robotics, Aachen, Germany

wu@ip.rwth-aachen.de

Abstract –

The development of 5G networks is expected to accelerate the realization of Industry 4.0 with its extended capabilities such as network slicing, reduced latency, increased bandwidth and much more. Leveraging these improved network performances, real time machine to machine communication could be realized. Moreover, in situations where human-machine interactions take place, safety protocols and systems can also be revisited and enhanced. In this paper, a unified framework between ROS, a widely adopted robotic middleware, and Unity, a versatile game development engine is developed. The system utilizes ROS' adaptability to operate robots, while Unity offers a visually engaging user interface. Through this unified approach, industrial mobile communication between machines and human operators is elevated, allowing for efficient coordination and synchronization of robotic tasks during complex deconstruction processes. Thus, increasing safety on construction sites on the one hand and improving general working conditions on the other.

Keywords –

5G network; Unity; ROS; Human-machine interaction; deconstruction

1 Introduction

In construction projects, on-site workers face significant occupational hazards. Workers in this industry are more susceptible to injuries and accidents compared to those in other industries, as highlighted by German Social Accident Insurance on the yearly report of 2022 [1]. Even within the construction industry, deconstruction processes are among the more dangerous processes due to unforeseeable scenarios such as unnoticed loosened building elements, unpredictable

cracks and are therefore hazardous in more ways than one. This risk is especially pronounced during the deconstruction process, where deformation, cracks, and falling building elements are among the factors that may lead to unforeseeable accidents. Furthermore, there are pollutants such as silica dust [2] and building element particles [3] present during the construction and deconstruction processes. In which long-term exposure to these substances can have serious negative effects on the health of the workers. One improvement suggestion to this is to reduce the working time of construction workers on the site, particularly during critical processes.

In recent years, in an effort to reduce exposure of workers to highly hazardous processes, researches on teleoperation and remote controlling have been an increasingly popular theme especially in the field of deconstruction robotics [4]. In the context of research field, Robot Operating System (ROS2) [5] is one of the more popular frameworks in developing robotic applications for teleoperation as shown in following papers [6] [7].

However, this approach could be further enhanced by integrating it with Unity [8], a multifaceted game development platform. By adopting this consolidated framework, a more inclusive and customised user experience can be provided as Unity supports multiple protocols, including Message Queuing Telemetry Transport, (MQTT) [9], and enables the creation of mixed reality applications for an even more intuitive and immersive path planning and control [10]. With the implementation of a mixed reality functionality, the remote operator would have an enhanced situational awareness along with increased engagement and focus.

The use of ROS2 in conjunction with Unity is an established research direction explored in a number of comparative studies [11] [12]. These studies compare different simulation environments for vehicle control within different virtual simulation environments such as Unity, V-Rep [13], and Gazebo [14]. Additionally, such

a framework has been adopted into industrial robotic process monitoring [15], and unmanned underwater vehicle simulation for telepresence robots [16]. In this particular research paper, the connection between ROS2 nodes and TCP connection with Unity is established. The Unity-developed application functions as the user interface for visualization and path planning, with messages forwarded to ROS2 for device control and processing.

However, most related research was conducted in controlled environments or virtual world, which stands in contrast to the construction site's highly dynamic nature, making it susceptible to unexpected events. Therefore, the aim of this paper is to address the research question: "How can a unified framework of ROS and Unity be implemented for construction to improve communication and aid robotic planning for on-site deconstruction tasks through 5G network?"

1.1 Potential

The integration of Unity in ROS2 involves many intricate steps. Yet, the possibility for visualization and control is highly enhanced. Unity's interoperability with MQTT, a popular IoT communication standard [17], enables it to control devices using different communication protocols. This allows for complete control from a single application. Since this matches the nature of construction sites where stakeholders commonly change throughout the construction project and each one contributes their individual solutions for communication, the potential in streamlining communication between the stakeholders by a unified framework is remarkable.

Furthermore, there is an increasing potential and deployment of semi- and fully-autonomous machines such as Autonomous Guided Vehicles (AGVs) on the construction sites to assist with heavy duty processes [18]. These AGVs, can be equipped with GPS trackers that can communicate via MQTT to enable accurate real-time location display on a deconstruction site. Operators may now carefully arrange the actions of the deconstruction robot provided with this information.

In addition, Unity, being a prominent development engine for Mixed Reality applications, offering expansive potential for Virtual Reality (VR) and Augmented Reality (AR), this use case can be extended as a standalone AR application using Unity. One of the major potentials of mixed reality is the amalgamation of virtual environment with the physical one, providing visual cues from both side of the environment, allowing the user to more informed and better decisions. This is especially relevant for deconstruction, where in contrast to traditional demolition, where materials and building elements are to be recovered. Thus, any damage to the materials and building elements whilst deconstruction

operation is to be minimised. Besides that, during the machine operation, danger zones could also be defined to trigger responsive actions, prompting the motion of the machines to slow down or halt instantly upon the detection of any risks. The risks could include any humans or machines that enter the defined danger zones. This integrated framework showcases the versatility of Unity and ROS in orchestrating seamless communication, offering practical solutions for complex industrial scenarios such as construction sites.

Due to the transitory nature of the construction site, machines and workers are constantly moving around. A robust, stable and real time network is needed for the communication between the machines. The requirements for safety protocols are insufficient with the current WLAN performance [19]. An extensive research on the overview of Extended Reality (XR) in mobile networks, it is shown that 4G network is unable to provide the latency and uplink/downlink capacity for the necessary use cases [37]. This is where 5G networking comes into play, where the technical requirements for 5G networks are as described by 3GPP in this technical specification [20]. These requirements allow for a real-time safety and control communication protocol between human-machine and machine-machine.

2 Methods and implementations

The described and implemented use scenario originates from a use case for 5G enabled and mixed reality supported deconstruction planning developed in TARGET-X research project [21]. In the project 5G is being introduced as new wireless communication standard into key industries like construction and its performance is evaluated in large scale trials such as the 5G Industry Campus Europe (5G-ICE) at Campus Melaten Aachen [22].

2.1 Experimental Setup

In order to test the planned setup including equipment, processes and technology within the 5G network ecosystem, the Reference Construction Site in Aachen is utilized. The Reference Construction Site is a test bed that is officiated by the Center Construction Robotics. It serves as living lab for research in construction by simulating an actual construction site with real processes while still providing a controlled environment.

Moreover, for this test bed, a 5G network infrastructure has been developed in which the tower crane on the Reference Construction Site plays the role as a transmission tower to provide the site with a 5G network [23]. The tower crane was chosen because it is positioned centrally on the site and allows for elevated mounting of the antennas. Additionally, omni-antennas were installed such that any rotational maneuvers of the

crane during construction does not affect the network coverage. Hence, the developed 5G network infrastructure is able to reflect that construction sites are typically dynamic in terms of resource and spatial availability.



Figure 1. Construction Site operated by Center Construction Robotics in Aachen.

3 Implementation in Unity

For the setup of the unified framework, there are multiple software assets and packages that are used. These packages include assets from Unity:

- **ROS-TCP-Connector** [24] is responsible for the connection between ROS and Unity, this package is based on a previous from ROS#. A TCP endpoint that enables message transfer between Unity and ROS while operating as a ROS node is set up. It also generates C# classes, serializing and deserializing the messages from ROS.
- **Unity Robotics Visualization** [25] allows Unity to visualize incoming and outgoing messages, including navigation messages, sensor data, and more. Messages such as /tf [26] can be visualized as joints in Unity Editor.
- **URDF importer** [27] is used to import a URDF-formatted robot into a Unity scenario. The geometry, graphic meshes, kinematic and dynamic properties of a robot are defined by the URDF file. This package parses a URDF file and imports it into Unity with the help of PhyX 4.0 articulation bodies.
- **BioIK** [28] is an optimized inverse kinematic algorithm solver based on this research paper [29] that can be implemented in Unity directly onto a hierarchical object, including URDF files.
- **Bézier Path Creator** [30] is an asset in Unity that allows user to create paths in Unity dynamically and visualize them during runtime.
- **Vuforia Engine** [31] is an asset in Unity that allows developer to develop their application into

Augmented Reality application using anchoring methods such as plane detection, image recognition and object detection.

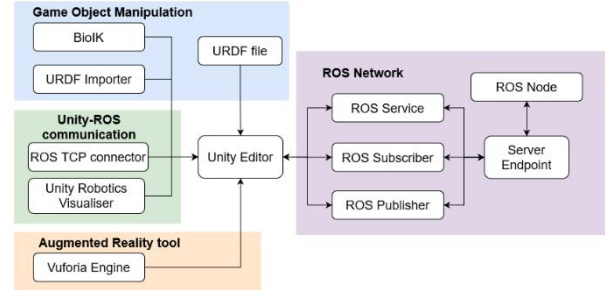


Figure 2. An overview of the categorized functionalities of the packages in Unity

3.1.1 Scene in Unity

The implementation of the packages into the Unity scene are comprised of the following steps:

1. The URDF importer is added into Unity and the URDF file of the robot model is adapted and imported into the scene. The robot is assumed to be stationary as the robot's base does not move during the operational task. In this scenario, a demolition robot, as shown in Figure 3, a BROKK 170 is used.

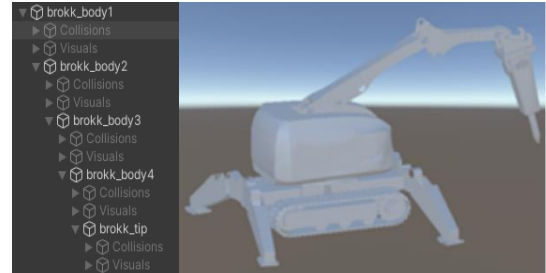


Figure 3. Screenshot of the imported URDF Brokk model.

2. After the model has been imported, inverse kinematics are implemented into the model. In this paper, the BioIK algorithm is deployed. The rotational limits of the joints are defined true to the physical robot's limitation and also to prevent any self-collision. The hierarchical order of the robotic arm is as follow: brokk_base, brokk_body1, brokk_body2, brokk_body3, brokk_boyd4, brokk_tip.
3. Then a target that defines the end position is added into the scene as well. This is necessary to realize the path planning. Using Bézier path creator, the target is programmed to follow the path that is created during runtime.
4. A predefined set of waypoints are included into the scene to define the path which the end effector

should follow. These waypoints can be manipulated by the user. The number of waypoints can be added dynamically during runtime, if the paths of the target should be more complicated.

5. At the parent of the robot of the URDF file, the script ROS Transform Tree Publisher from the ROS-TCP-Connector package is added. This script creates a /tf message and stamps them with the time and transform information from Unity and sends them to the ROS2 endpoint.
6. In ROS2, a TCP endpoint package that establishes the connection between Unity and ROS2 is created. The package is run in parallel with Unity, the communication is represented in Figure 2 in the green and purple areas.

3.1.2 Implementation of mixed reality

With the increasing computational power of handheld devices such as tablets and mobile devices, application that requires heavy computational tasks such as the presented one can be build and deployed to any mid-range devices. One of the advantages of a mixed reality is the amalgamation of virtual environment with the physical one, providing visual cues from both side of the environment, allowing the user to make more informed decisions. As Unity is one of the game engines that supports the development for mixed reality application, this use case can be extended as a standalone Augmented-Reality (AR) application. In addition to the aforementioned implementations, Vuforia Augmented Reality SDK is used to develop an AR application for an enhanced visualization and path planning. The implementation for the Augmented-Reality application in Unity is as follows:

1. Firstly, the Vuforia SDK is imported into Unity. An image target is trained and used as an anchor point. The image target is not limited to any markers such as Aruco markers [32]. Any image that is asymmetric and has sufficient feature points is valid as a target.
2. Alternatively, an object target is also trained to detect the particular shape of the object in real world.
3. The unity packages generated from by Vuforia is inserted into Unity and the hierarchical order of the game objects are rearranged with accordance to the SDK's documentation.
4. The virtual camera in the scene is also changed to the physical camera on the device to take in information from the physical world and superimposing 3D models onto the screen of the device.

3.1.3 5G network communication between different machines in different network

In order to exploit the potential of the 5G network that

was implemented on the construction site in this use case, the machine-machine communication is set up as shown in Figure 4.

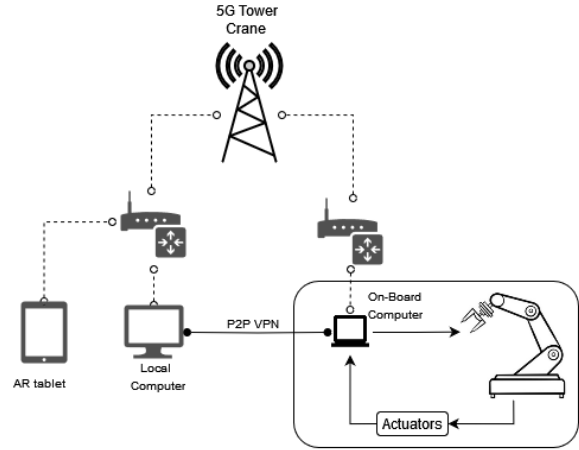


Figure 4. Established connection between multiple machines in the different networks.

The AR application is developed and built as described in section 2.2.1 and 2.2.2, the application is then installed in an Android tablet device. The tablet and a local computer are connected to the same network as shown on the left side of the Figure 4. The /tf messages are sent and received by the local computer.

In order to send ROS messages across different networks in a secure manner, VPN has to be setup. In this case, between the local computer and the on-board computer. A Peer-to-Peer (P2P) tunnel is set up using Husarnet VPN by Husarion. Husarnet uses the ChaCha20-Poly1305 to provide the encryption and authentication of data [33]. With both the machines' IP addresses being added to the peer list, the machines are able to communicate. Cyclone DDS (Data Distribution Service) are installed into both the computers as the recommended middleware in the official documentation from Husarnet. Lastly, the ports on which the ROS2 nodes run, are defined using the ROS_DOMAIN_ID command. In this case, the domain ID is set to be 0. It is necessary to define the domain ID used by DDS to restrict the range of the ports that will be used for discovery and communication. This also means that there can be multiple ROS systems running simultaneously, with an upper limit of 120 processes on one computer. Otherwise, it might spill over into other domain IDs or ephemeral ports [34].

4 Validation

In order to validate the implemented connections between the application and the machines and test them for their robustness, latency tests are carried out. For this

purpose, we examine the different components of the communication setup.

4.1 ROS-Unity connection and message flow

With the connection between Unity and ROS2 successfully set up, the `rqt_graph` looks as below in Figure 5 along when `rviz2` is executed. The messages are generated by the Unity AR application and sent to the local computer that is connected to the same network where the ROS server endpoint is running. This received messages are then shared across the VPN tunnel, making them accessible to the computer on board the machine.

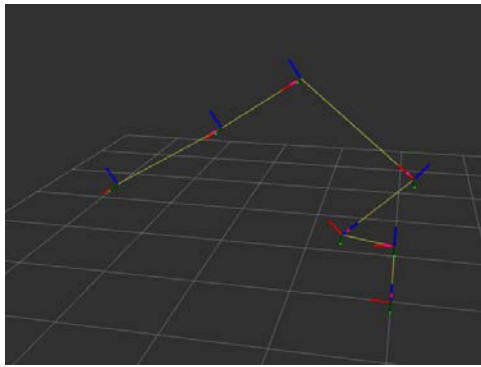


Figure 6. Visualisation of `/tf` messages of the on-board computer received from the AR application.

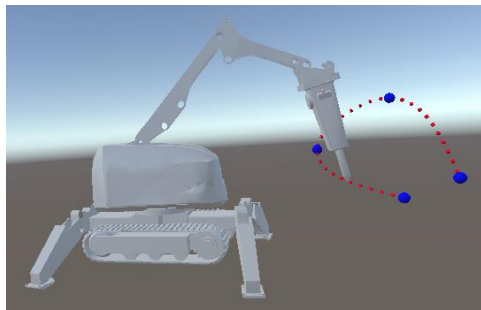


Figure 7. Screenshot of the path planning in Unity.

4.2 Deployed Augmented Reality application

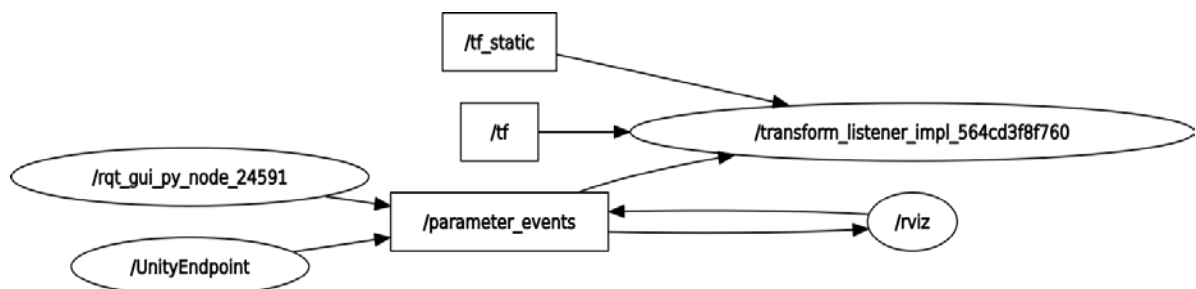


Figure 5. `rqt_graph` of the connections

As the connections between the application and the servers are set up, along with the successful built and deployed AR application. The application is run with an image as a target to allow the model to anchor into the digital environment. In Figure 8, the blue points are the waypoints that are used to define the pathway. The red dots show the generated pathway, for which the end effector will follow.

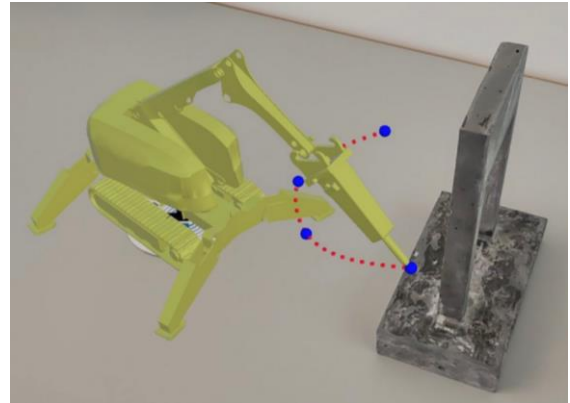


Figure 8. Screenshot of the application using image target as anchoring.

However, as the application is running, some jittering of the model occurred. This causes unwanted transformation, `/tf` messages of the model to be published. This phenomenon is observed to occur when the user moved the device too fast or when the lighting of the physical environment is insufficient. The jittering also occurred for the object target recognition, it is especially pronounced when the sunlight is too strong, where the reflective surface of the physical BROKK machine to cause glares to the camera.

4.3 Latency result

In order to evaluate the robustness of the 5G network in the construction site, a latency test is carried out with the set up as above. The latency is tested using My Traceroute (MTR) [35] using 100 packages of 64 bytes with a frequency of 10 Hz. The tests are done with varying publishing frequencies from the Unity

application using ICMP, TCP and UDP. The results are tabulated according to the respective protocols.

For all tested protocols, the results for the average latencies are compatible within their respected standard deviation. However, the distribution for the UDP and TCP protocol are noticeably broader when the ROS messages are sent at the frequency of 10 Hz. This is reflected by the comparably larger standard deviations.

Furthermore, it should be noted that the UDP protocol has a package loss of approximately 20%. The best average latency is observed at the publishing frequency of 50 Hz by the ICMP test. And the worst average latency is observed to be at 10 Hz by the UDP test

Table 1: ICMP test

ROS Freq. [Hz]	Loss %	Avg [ms]	Best [ms]	Worst [ms]	StDev
10	0	28.84	20.60	72.55	6.09
50	0	27.58	18.95	39.42	4.24
100	0	28.85	19.34	51.16	6.60

Table 2: UDP test

ROS Freq. [Hz]	Loss %	Avg [ms]	Best [ms]	Worst [ms]	StDev
10	21	34.48	21.61	120.34	19.92
50	19	29.90	21.11	37.36	3.62
100	21	31.37	19.78	49.94	6.06

Table 3: TCP test

ROS Freq. [Hz]	Loss %	Avg [ms]	Best [ms]	Worst [ms]	StDev
10	0	32.71	22.89	130.61	18.94
50	0	28.03	20.06	41.04	3.98
100	0	29.04	18.43	41.04	5.85

5 Discussion

This study shows the entire framework and workflow of developing an AR application with Unity and ROS2, connected via 5G campus network. Unity is chosen not only because of its ability to build a mixed reality application, but also because of its versatile ability to be used with other communication protocols such as MQTT, to develop an interactive user interface, to process and visualise data acquired from various sensors. Although the workflow and implementation of this application is rather complicated and it requires multiple components to realise this particular use case, the application's framework is versatile and can be adapted across a wide range of scenarios with minimum amendments.

In order to realise a multi-network communication between different machines, a P2P VPN is set up so that the machines can discover and communicate with each other. This can however also be achieved by using port forwarding but is advised against due to its potential security risk for it exposes the ports to the public. This is especially undesirable when considering the messages sent are used to control heavy duty robots and any mishaps or hijacking of the port could lead to unthinkable consequences. Using a VPN can highly reduce the risk, albeit causing an increase in latency. One of the reasons for the increased latency is due to the encryption of the packages during transmission. The packages are encrypted with a key and a nonce. This ensures that the packages are kept confidential and with an authentication tag, making sure that the messages are not tampered with.

In the latency test, there is some package loss in the UDP protocol. This is due to the nature of UDP where it does not check for errors nor receipt during transmission, causing some package to be lost during transmission. Looking at the average time of the tests, the messages have a latency of less than 50ms.

6 Conclusions and outlook

This paper shows the successful framework implementation of an augmented reality application enabled through a 5G network. However, the latency and the occasional unstable jitter might need to be improved with accordance to the recommended values as stated in [36]. With regards to the research questions, an AR application has been developed to aid in visualisation and planning of the robotics deconstruction planning. Additionally, the low latency requirement has been leveraged using the 5G network's capability. It is worth pointing out, that this framework of Unity and ROS2 is not limited to as such. The ability to define the domain ID opens the possibility to deploy multiple applications with only one machine to host and act as the control. This framework could also be extended to an AGV or other machines controlled via ROS. Furthermore, safety zones could be defined and visualised such that the machines could communicate with each other, slowing down or stopping immediately with accordance to the position, process and speed of the operating machines. Moreover, the framework allows sensors information e.g., from LiDAR scanners or RGBD cameras to be integrated with computer vision algorithms. This enables the machines to understand the environment and provide an action trigger to whenever something is detected within the vicinity of the danger zones. With the improving bandwidth and decreased latency of 5G network, the proposed extensions of the framework are promising. Subject to future research are cases where there is an ontology server set up within the construction project. Then these

actions or triggers detected by the sensors can be forwarded to the ontology. This process imparts semantic meaning to the identified actions and triggers, which in turn fosters effective information dissemination among relevant stakeholders and prompting subsequent actions as needed.

7 Acknowledgement

The content of this work is part of the research project TARGET-X, co-funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the other granting authorities. Neither the European Union nor the granting authority can be held re-sponsible for them. The TARGET-X project has received funding from the Smart Networks and Services Joint Undertaking (SNS JU) under the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101096614.



Co-funded by
the European Union



8 References

- [1] Deutsche Gesetzliche Unfallversicherung (DGUV). Statistik – Arbeitsunfallgeschehen 2022.
- [2] Normohammadi M, Kakooei H, Omid L, Yari S, Alimi R. Risk Assessment of Exposure to Silica Dust in Building Demolition Sites. *Safety and health at work* 2016;7(3):251–5.
- [3] Yu D, Duan H, Song Q, Li X, Zhang H, Zhang H et al. Characterizing the environmental impact of metals in construction and demolition waste. *Environmental science and pollution research international* 2018;25(14):13823–32.
- [4] Lee HJ, Brell-Cokcan S. Towards controlled semi-autonomous deconstruction. *Constr Robot* 2023.
- [5] Macenski S, Foote T, Gerkey B, Lalancette C, Woodall W. Robot Operating System 2: Design, architecture, and uses in the wild. *Science Robotics* 2022;7(66).
- [6] Baklouti S, Gallot G, Viaud J, Subrin K. On the Improvement of ROS-Based Control for Teleoperated Yaskawa Robots. *Applied Sciences* 2021;11(16):7190.
- [7] Mortimer M, Horan B, Joordens M. Kinect with ROS, interact with Oculus: Towards Dynamic User Interfaces for robotic teleoperation. In: 2016 11th System of Systems Engineering Conference (SoSE 2016): Kongsberg, Norway, 12-16 June 2016. Piscataway, NJ: IEEE; 2016, p. 1–6.
- [8] Messaoudi F, Simon G, Ksentini A. Dissecting games engines: The case of Unity3D. In: *Network and Systems Support for Games (NetGames)*, 2015 International Workshop on. [Place of publication not identified]: IEEE; 2015 - 2015, p. 1–6.
- [9] Message Queuing Telemetry Transport. [December 14, 2023]; Available from: <https://mqtt.org/>.
- [10] Delmerico J, Poranne R, Bogo F, Oleynikova H, Vollenweider E, Coros S et al. Spatial Computing and Intuitive Interaction: Bringing Mixed Reality and Robotics Together. *IEEE Robot. Automat. Mag.* 2022;29(1):45–57.
- [11] Platt J, Ricks K. Comparative Analysis of ROS-Unity3D and ROS-Gazebo for Mobile Ground Robot Simulation. *J Intell Robot Syst* 2022;106(4).
- [12] Santos Pessoa de Melo M, Da Gomes Silva Neto J, Da Jorge Lima Silva P, Natario Teixeira JMX, Teichrieb V. Analysis and Comparison of Robotics 3D Simulators. In: 2019 21st Symposium on Virtual and Augmented Reality (SVR): IEEE; 2019, p. 242–251.
- [13] Rohmer E, Singh SPN, Freese M. V-REP: A versatile and scalable robot simulation framework. In: 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems: IEEE; 2013.
- [14] Koenig N, Howard A. Design and use paradigms for gazebo, an open-source multi-robot simulator. In: p. 2149–2154.
- [15] Sita E, Horvath CM, Thomessen T, Korondi P, Pipe AG. ROS-Unity3D based system for monitoring of an industrial robotic process. In: IEEE/SICE International Symposium on System Integration Taipei T(editor. SII 2017 2017 IEEE: IEEE, p. 1047–1052.
- [16] Katara P, Khanna M, Nagar H, Panaiyappan A. Open Source Simulator for Unmanned Underwater Vehicles using ROS and Unity3D. In: *Pathway to get green in deep blue: UT'19 Kaohsiung April 16-19, 2019*, National Sun Yat-Sen University, Kaohsiung, Taiwan proceedings. [Piscataway, NJ]: IEEE; 2019, p. 1–7.
- [17] Mishra B, Kertesz A. The Use of MQTT in M2M and IoT Systems: A Survey. *IEEE Access* 2020;8:201071–86.
- [18] Yang Y, Pan W. Automated guided vehicles in modular integrated construction: potentials and future directions. *CI* 2021;21(1):85–104.
- [19] Lee HJ, Krishnan A, Brell-Cokcan S, Knußmann J, Brochhaus M, Schmitt RH et al. Importance of a 5G Network for Construction Sites: Limitation of WLAN in 3D Sensing Applications. In: Linner T, García de Soto B, Hu R, Brilakis I, Bock T, Pan W et al., editors. *Proceedings of the 39th International Symposium on Automation and Robotics in Construction: International Association for Automation and Robotics in Construction (IAARC)*; 2022.

- [20] 3GPP. Technical Specification Group Services and System Aspects; Service requirements for the 5G system;
- [21] reference to project
- [22] 5G Industry Campus Europe. [December 14, 2023]; Available from: <https://5g-industry-campus.com/infrastructure>.
- [23] Fraunhofer Institute for Production Technology IPT. 5G.NAMICO - Networked, Adaptive Mining and Construction. [December 14, 2023]; Available from: <https://www.ipt.fraunhofer.de/en/projects/5g-namico.html>.
- [24] Unity Technologies. ROS-TCP-Connector. [December 14, 2023]; Available from: <https://github.com/Unity-Technologies/ROS-TCP-Connector.git>.
- [25] Unity Technologies. Unity Robotics Visualizations Package. [December 14, 2023]; Available from: <https://github.com/Unity-Technologies/ROS-TCP-Connector/blob/main/com.unity.robotics.visualizations/>.
- [26] Foote T. tf: The transform library. In: 2013 IEEE Conference on Technologies for Practical Robot Applications (TePRA): IEEE; 2013.
- [27] Unity Technologies. URDF-Importer. [December 14, 2023]; Available from: <https://github.com/Unity-Technologies/URDF-Importer>.
- [28] Sebastian Starke. BioIK. [December 14, 2023]; Available from: <https://github.com/sebastianstarke/BioIK>.
- [29] Sebastian Starke, Norman Hendrich, Dennis Krupke, Jianwei Zhang. IROS Vancouver 2017: IEEE/RSJ International Conference on Intelligent Robots and Systems, Vancouver, BC, Canada September 24-28, 2017 conference digest. Piscataway, NJ: IEEE; 2017.
- [30] Sebastian Lague. Bézier Path Creator. [December 14, 2023]; Available from: <https://assetstore.unity.com/packages/tools/utilities/b-zier-path-creator-136082>.
- [31] PTC Inc. Vuforia Engine; 2023.
- [32] Kalaitzakis M, Cain B, Carroll S, Ambrosi A, Whitehead C, Vitzilaios N. Fiducial Markers for Pose Estimation. J Intell Robot Syst 2021;101(4):1–26.
- [33] Operate At The Edge Of Latency. [December 14, 2023]; Available from: <https://husarnet.com/>.
- [34] The ROS_DOMAIN_ID. [December 14, 2023]; Available from: <https://docs.ros.org/en/humble/Concepts/Intermediate/About-Domain-ID.html>.
- [35] My Traceroute. [December 14, 2023]; Available from: <https://www.bitwizard.nl/mtr/>.
- [36] 5G Americas. 5G Americas White Paper Communications for Automation in Vertical Domains November 2018.
- [37] Cardoso LFdS, Kimura BYL, Zorzal ER. Towards augmented and mixed reality on future mobile networks. Multimed Tools Appl 2024;83(3):9067–102.

Fast-Pixel-Matching Algorithm for Automated Shear Stud Welding Based on the Integration of 2D Drawings and Structured Light Cameras

Huiguang Wang¹, Lu Deng^{1,2,*}, Ran Cao^{1,2}, and Jingjing Guo^{1,2}

¹College of Civil Engineering, Hunan University, China

²Hunan Provincial Key Laboratory for Damage Diagnosis for Engineering Structures, Hunan University, China

*Corresponding Author

whg0917@hnu.edu.cn, [*denglul@hnu.edu.cn](mailto:denglul@hnu.edu.cn), rcaoh@hnu.edu.cn, guojingjing@hnu.edu.cn

Abstract –

Due to the lack of depth information, using 2D drawings to guide robotic manufacturing is difficult, such as stud welding and drilling. In this paper, a new algorithm, Fast-Pixel-Matching (FPM), is proposed to provide depth information for precise point positioning based on 2D drawings and structured light cameras. Based on our experiment, the efficiency of the FPM process is found to be at least 20 times higher than the brute-force search (BFS) method. Also, we demonstrated that the algorithm can be applied to automated stud welding, where the efficiency of positioning can be significantly improved and the positioning accuracy of the proposed method meets the code requirements.

Keywords –

Point positioning; 2D drawings; Supplementation of depth information; Stud welding; 3D vision

1 Introduction

Stud welding has been widely used in various scenarios, including nuclear power equipment [1], boiler equipment [2,3], bridge structures [4], building composite structures [5], and shipbuilding [6,7]. For example, the average consumption of studs per individual bridge amounts to 250,000 studs, and millions of studs are required for welding in the hull and compartments of ships [6]. Taking the studs on the upper flange surface of an I-beam as an example, before stud welding, workers are required to meticulously mark lines to locate studs on the surfaces of components based on Computer-Aided Design (CAD) drawings. This process is rather time-consuming, where positioning a stud typically takes around 10 seconds [8]. Moreover, workers will spend more time positioning studs on curved components, and the verticality of stud welding cannot be guaranteed.

During the process of positioning and welding studs, workers are required to maintain a prolonged and

extremely forward-bent posture [9]. According to [10], when the duration of this bending posture exceeds four hours, the risk of experiencing lower back pain is heightened. The daily working hours for welding workers would commonly exceed four hours.

In the current industry, the manufacturing process of steel beams has been highly automated in terms of steel plate cutting and welding. However, the welding of studs on the surfaces of components is still manually performed, which decreases the efficiency of the whole production line.

To increase the efficiency of stud welding, a robotic welding framework was proposed in this paper that involves a robot arm, a structured light camera, and a 2D drawing. In particular, a Fast-Pixel-Matching (FPM) algorithm was developed to convert the welding position from 2D drawings to 3D working space.

2 Related Work

With the evolution of CAD technology, abundant manufacturing information is embedded within drawings.

Taking computer numerical control (CNC) machine tools as an example, manufacturing 2D components often requires only a 2D Drawing Exchange Format (DXF) file, as seen in the case of CNC engraving machines [11]. However, if drilling is to be performed on curved surfaces, a 3D SolidWorks model becomes necessary [12]. This approach is viable in the field of mechanical engineering, where engineers commonly utilize 3D software for design. However, civil engineers tend to be accustomed to 2D design [13] and most of the structural designs were still represented using 2D drawings [14,15].

For example, existing automated stud welding equipment typically requires only 2D DXF drawings for 2D components [1]. During the shear stud welding process, the robotic arm descends automatically with its end-effector pointing vertically downward. The force sensor at the robotic arm's end-effector detects contact between the welding gun and the steel plate, triggering

the welding signal for the welding operation.

However, when dealing with curved components, due to the requirement of perpendicular installation of shear studs to the steel plate surface, the robotic arm's end-effector needs to maintain a perpendicular orientation at the welding point relative to the steel plate surface. Therefore, when dealing with 3D components, existing method requires the establishment of a 3D model based on 2D drawings to provide Cartesian coordinates and orientation information for shear stud welding[8], which proves to be quite time-consuming [16]. Hence, existing automated stud welding equipment does not align with the design conventions in civil engineering.

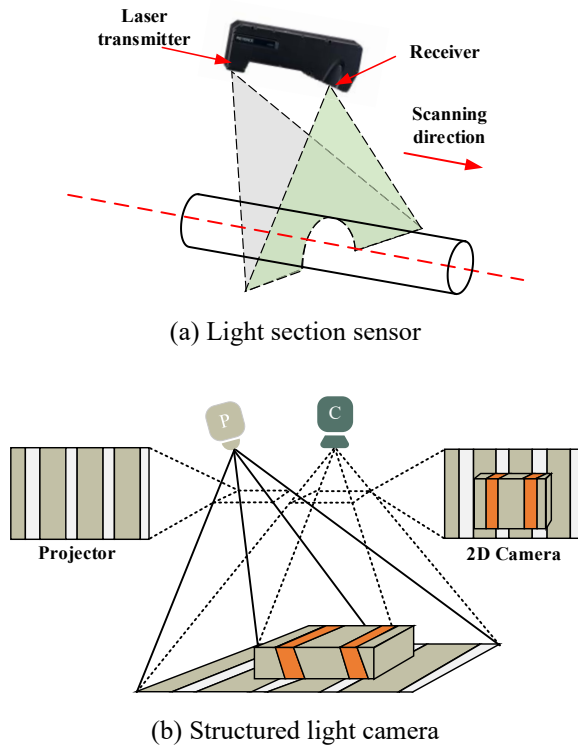


Figure 1. The method of point cloud data acquisition

In general, 2D drawings only contain planar coordinate information, thus necessitating the use of sensors to supplement missing Z-axis coordinates and orientation information of the target, such as studs. To acquire the missing information, in [2,3], the authors presented a measurement method that integrated 2D drawings with a light section sensor for the pose of the stud. The light section sensor captures point cloud data in the vicinity of the points to be welded, enabling the fitting of the orientation directional vector of the shear studs. However, the cost of the light section sensor is generally higher than that of a structured light camera. As illustrated in Figure 1(a), this method requires controlling the movement of the light section sensor in a specific

direction to obtain point cloud data. In contrast, the structured light camera shown in Figure 1(b) can capture point cloud data within an area with a simple overhead shot. Hence, using a structured light camera to supplement missing 3D coordinate information in 2D drawings would result in higher operational efficiency.

In the literature, 3D cameras have been used in several fields, such as fruit recognition [17], trajectory tracking [18], and pose tracking [19,20]. In most of these applications, the pixel coordinates $[u, v]^T$ and corresponding depth information z_{cam} can be directly obtained using the object detection algorithm and 3D camera, which enables the calculation of the object's world coordinates $[x_w, y_w, z_w]^T$ using Equation (1) (for robotic operation, such as picking a fruit).

$$\begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = z_{cam} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} f_x & 0 & p_x \\ 0 & f_y & p_y \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (1)$$

In the stud welding setting, however, the goal is to find the pixel coordinates, i.e., the welding point in the image, that corresponds to the given world coordinates $[x_w, y_w]^T$ from the 2D drawings. Due to the absence of world coordinates z_w and depth information z_{cam} , the pixel coordinates $[u, v]^T$ (guiding robotic arm) couldn't be directly determined based on $[x_w, y_w]^T$, using Equation (1).

To overcome this problem, this paper presents a new point-searching algorithm, i.e., FPM, which enables a rapid determination of corresponding pixel coordinates $[u, v]^T$ based on planar world coordinates $[x_w, y_w]^T$.

3 Method

The proposed framework comprises the following three steps.

1. The AutoCAD plug-in is developed to extract the planar world coordinates of studs and circular marks based on C# programming language and the .Net API [21].
2. Based on planar world coordinates $[x_w, y_w]^T$, FPM is proposed for a rapid search of pixel coordinates $[u, v]^T$ corresponding to planar world coordinates $[x_w, y_w]^T$, which allows for the supplementation of missing Z-axis world coordinates z_w using Equation (1).
3. Near the pixel coordinates $[u, v]^T$, select several

other pixel coordinates and their corresponding world coordinates for obtaining the orientation vector \vec{n} of the stud welding position. With the 3D world coordinates $[x_w, y_w, z_w]^T$ and the orientation vector \vec{n} , the pose of the robotic arm end effector can be calculated.

3.1 AutoCAD Plug-in

In 2D drawings, there are typically tens of thousands of studs, and manually marking lines for positioning is a labour-intensive and highly repetitive task. Therefore, we developed an AutoCAD plug-in to extract the coordinates of studs. The primary development approach is as follows:

1. Extracting the central coordinates of circles as the coordinates for studs. (shear studs are generally represented as circles in 2D drawings.)
2. As illustrated in Figure 2, it is necessary to draw circular marks O, A, and B on the 2D drawing and establish a local frame. Hence, the coordinates of studs $[x_l, y_l]^T$ in the local frame can be extracted.
3. Transforming the coordinates $[x_l, y_l]^T$ of studs in the local frame to the world frame, as detailed in Section 3.1.1.

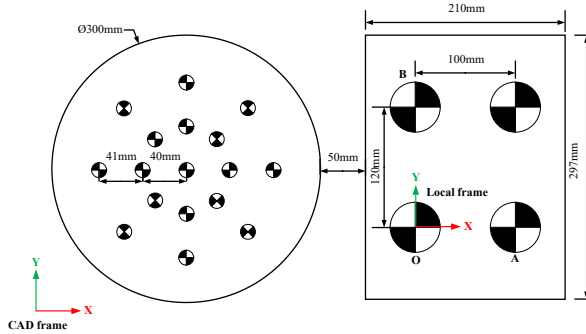


Figure 2. CAD drawing

A demo showing the extraction results of studs in 2D drawings can be found at: <https://youtu.be/nrdaXvO8dkE>.

3.1.1 Transformation of Planar Coordinates

The planar coordinates of studs require converting from the local frame to the world frame, and the method is as follows:

1. As shown in Figure 3, put the 3D printed component and marking board on the working plane according to their geometric relationships in the 2D drawing.
2. As shown in Figure 3, the structured light camera

mounted on the end of the robotic arm captures an overhead view of the marking board, obtaining RGB and depth images. Utilizing the Hough transform method [22], the pixel coordinates of the center points of circular marks O, A, and B are acquired. With Equation (1), the coordinates of the centers of circular marks O, A, and B in the world frame can be calculated.

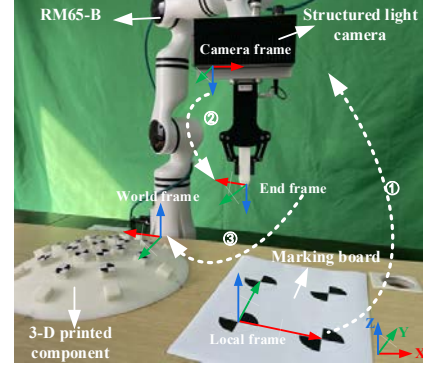


Figure 3. Experimental setup and the transformation between different frames

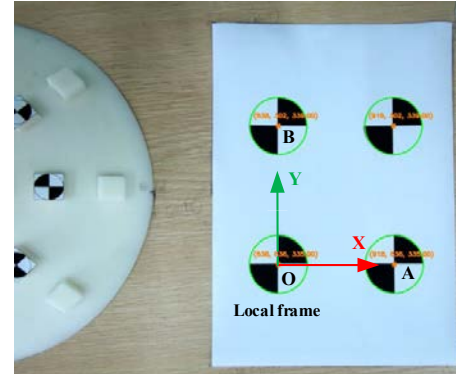


Figure 4. The center of circle recognition results using the Hough transform method

3. The center of circle recognition results $[u, v, z_{cam}]^T$ using the Hough transform method are shown in Figure 4, and substitute $[u, v, z_{cam}]^T$ into Equation (1) to obtain the world coordinates of marks O: $[x_w^O, y_w^O]^T$, A: $[x_w^A, y_w^A]^T$, B: $[x_w^B, y_w^B]^T$. Having obtained the coordinates of marks in the world frame, utilizing Equation (2) allows for the transformation of the coordinates of studs from the local frame to the world frame.

$$\begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix} = M_{local2world} \begin{bmatrix} x_l \\ y_l \\ 1 \end{bmatrix} \quad (2)$$

Where $M_{local2world}$ is the transformation matrix from local frame to world frame, and its expression is as follows:

$$M_{local2world} = \begin{bmatrix} \cos \langle \overrightarrow{OA}, \overrightarrow{X} \rangle & \cos \langle \overrightarrow{OB}, \overrightarrow{X} \rangle & x_w^O \\ \cos \langle \overrightarrow{OA}, \overrightarrow{Y} \rangle & \cos \langle \overrightarrow{OB}, \overrightarrow{Y} \rangle & y_w^O \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

Where $\overrightarrow{X} = [1, 0]^T$, $\overrightarrow{Y} = [0, 1]^T$.

3.2 Fast Pixel-Matching Process

After completing the coordinates transformation in Section 3.1.1, one can use the planar world coordinates of the studs to control the movement of the robot-mounted structured light camera, positioning above the 3D printed component, as shown in **Error! Reference source not found..** This enables the acquisition of RGB and depth images of the target area.

When transforming the coordinates of studs from the local frame to the world frame, with only knowledge of the planar world coordinates $[x_w, y_w]^T$, it is not possible to calculate the corresponding pixel coordinate $[u, v]^T$ using Equation (1).

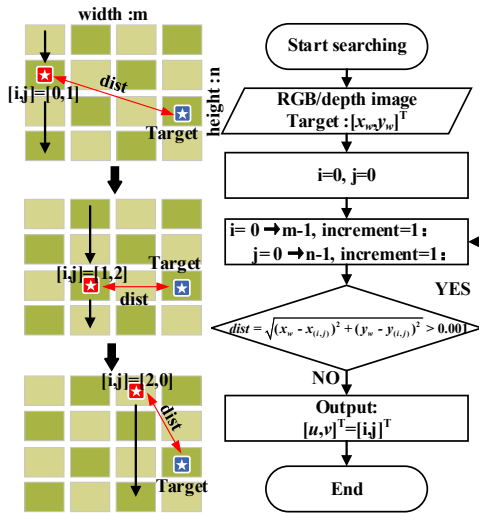


Figure 5. Iterative search process of BFS

However, it is feasible to adopt BFS for calculating each pixel coordinate corresponding to the world coordinate in the image. In Figure 5, the pixel coordinates currently being searched are $[i, j]^T$, and their corresponding planar world coordinates in the plane can be calculated by Equation (1) as $[x_{(i,j)}, y_{(i,j)}]^T$.

By comparing $[x_{(i,j)}, y_{(i,j)}]^T$ with $[x_w, y_w]^T$, when the

'dist' ($dist = \sqrt{(x_w - x_{(i,j)})^2 + (y_w - y_{(i,j)})^2}$) between the currently searched pixel (red star) and with target pixel (blue star) is less than 0.001m, the currently searched pixel coordinate $[i, j]^T$ can be considered as the pixel coordinate $[u, v]^T$ for the welding point of the stud.

However, it should be noted that BFS is quite time-consuming, and we proposed another fast point searching method, FPM, with details outlined below.

3.2.1 Algorithm Design

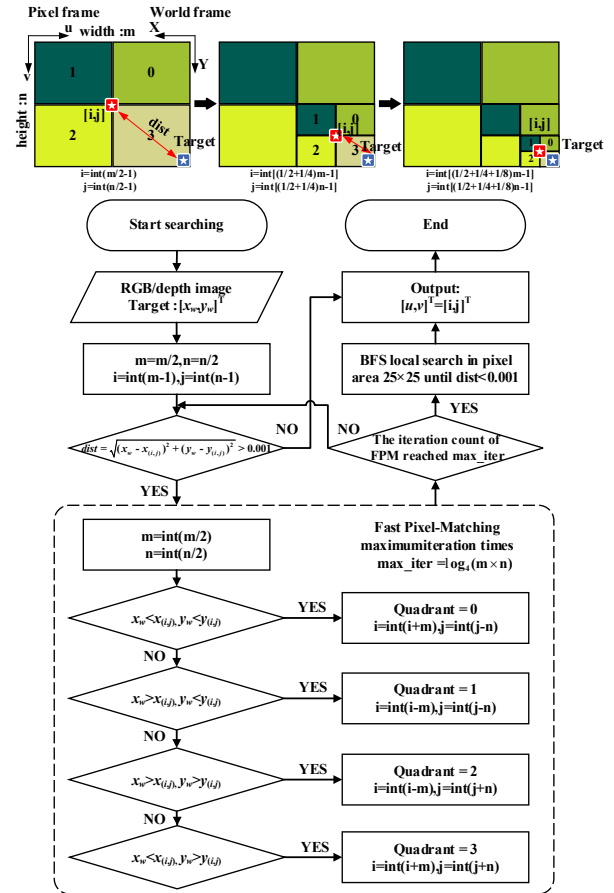


Figure 6. Iterative search process of the FPM process

It is important to note that, for the FPM application, as illustrated in Figure 6, the image pixel frame has its u-axis opposite to the X-axis of the world frame, while the v-axis aligns with the Y-axis of the world frame, dividing the image into four quadrants labeled 0, 1, 2, and 3.

Based on the above assumptions, Figure 6 illustrates the first three iterations of the algorithm's search process, and the computational iterative process of the algorithm is as follows.

As illustrated in Figure 6, when employing FPM for the search, during the first round of iteration, the world coordinate $[x_{(i,j)}, y_{(i,j)}]^T$ corresponding to the currently searched pixel coordinate $[i, j]^T$ at the center of the image is calculated. The coordinate $[x_{(i,j)}, y_{(i,j)}]^T$ is then compared with $[x_w, y_w]^T$, determining the next search quadrant based on the principle from Equation (3) to Equation (6). Through repeated cycles until 'dist' is less than 0.001m, the pixel coordinate $[u, v]^T$ corresponding to the planar world coordinate $[x_w, y_w]^T$ can be rapidly matched.

$$x_w < x_{(i,j)}, y_w < y_{(i,j)}, \text{Quadrant} = 0 \quad (3)$$

$$x_w > x_{(i,j)}, y_w < y_{(i,j)}, \text{Quadrant} = 1 \quad (4)$$

$$x_w > x_{(i,j)}, y_w > y_{(i,j)}, \text{Quadrant} = 2 \quad (5)$$

$$x_w < x_{(i,j)}, y_w > y_{(i,j)}, \text{Quadrant} = 3 \quad (6)$$

The space complexity of FPM is $O(\log_4 m \times n)$. For an image resolution of 1280×960 , the correspondence between $[x_w, y_w]^T$ and $[u, v]^T$ can be determined with a maximum of 11 ($\log_4 1280 \times 960 \approx 11$) iterations.

After 11 iterations of FPM, BFS is applied within a pixel range of 25×25 to obtain pixel coordinates, and 'dist' can converge to 0.001m. The camera's field of view and resolution are respectively $460\text{mm} \times 360\text{mm}$ and 1280×960 , so approximately 3 pixels correspond to each millimeter. Therefore, a pixel range of 25×25 (± 4 mm range) is considered appropriate.

For the sake of convenience in expression, the combined process of FPM and BFS will be referred to as the FPM process in the following content.

In theory, only $11 + 25 \times 25 = 636$ iterations are needed in the FPM process to accomplish the process. However, using only the BFS method would require iterations of $1280 \times 960 = 1,228,800$, which is almost 2000 times higher than that of the FPM method.

4 Experimental Setup and Results

4.1 Experimental Setup



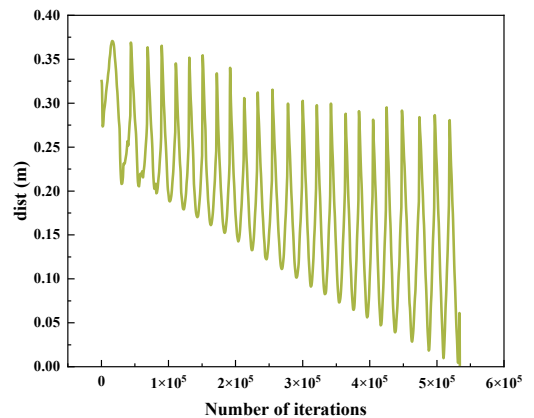
Figure 7. The dome of the LNG storage tank [23]

As shown in **Error! Reference source not found.**, the experimental setup is composed of a robotic arm (RM-65), structured light camera (PS800-E1, resolution 1280×960), 3D printed component, and a marking board. The prototype of the 3D printed component is shown in Figure 7, which is the steel structural dome at the top of the LNG storage tank. An experiment was conducted to assess the efficiency and accuracy of the FPM process.

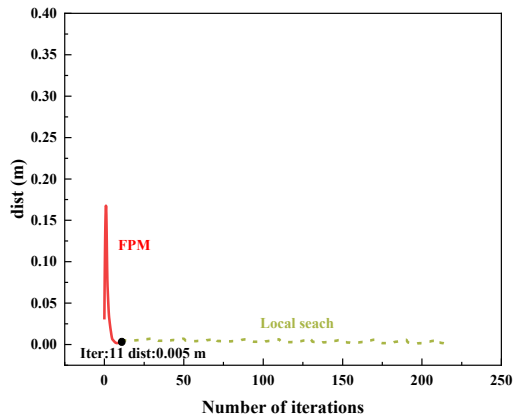
In the experiment, a total of 17 welding points shown in **Error! Reference source not found.**(a) were considered as the target for point searching. for precision comparison, in this paper, the ground truth of the depth information of these welding points was obtained by using the mouse to click on the pixel coordinates at the center of each mark in **Error! Reference source not found.**(a), and then transforming them into world coordinates using Equation (1).

4.2 Experimental Results

4.2.1 Efficiency



(a) BFS



(b) FPM process

Figure 8. The number of iterations for the pixel search in the BFS and FPM process

Based on the comparison shown in Figure 8, FPM achieved a much faster convergence than that of BFS. In particular, a total of 217 iterations were performed in the FPM process, while only BFS required 534,115 iterations, the comparative data can be found in Table 1.

The efficiency of the FPM process, BFS, and the manual method is detailed in Table 2 for comparison. FPM exhibits exponential convergence, with an iteration efficiency 47,604 times greater than BFS. In Table 2, the time required for manual positioning was calculated based on the assumption that locating a welding point would require 10 seconds by marking lines [8].

In terms of iteration time, the efficiency of the FPM process is 21 times that of BFS and 105 times that of manual methods, allowing for the rapid determination of pixel coordinates corresponding to planar world coordinates.

Table 1. Comparison of Iteration Steps between the FPM process and BFS

Method	FPM process	BFS
dist < 0.001m	217	534,115

Table 2. The time required for positioning using different methods

Method	Time (s)	Efficiency ratio
FPM process	1.62	105
BFS	34.4	5
Manual method	170.0	1

4.2.2 Accuracy

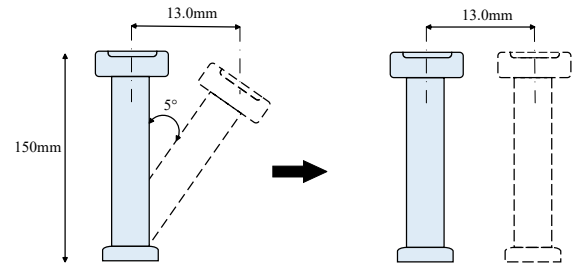


Figure 9. Conversion from verticality to offset

As shown in Figure 9, the 'Code for acceptance of construction quality of steel structures' only has a clear requirement for the verticality of studs. The inclination angle should not exceed 5 degrees [24]. Taking the commonly used size of stud M16 × 150 as an example, the maximum offset at the top of the stud can be calculated as 13.0mm. This calculation is used to determine the maximum allowable offset in the plane of the stud.

The positioning accuracy errors of the FPM process and BFS compared with the manual method, as shown in Figure 10, indicated that they fully meet the code requirement, and the maximum error was only 2.09 mm in the direction of Z axis.

Figure 11 presents the results of the 17-pixel coordinates mapped to the RGB image obtained through the FPM process. The green points represent the pixel coordinates corresponding to the planar world coordinates.

Additionally, comparing with related literature on x-y positioning accuracy, in [7], the proposed method has a maximum error of 8mm, while the FPM process has a maximum error of 7.1mm, representing a 9% improvement in accuracy. However, this approach relies on the manual use of an IMU device to designate the position of the welding points, while the method proposed in this paper can automatically search for the pixel coordinates of welding points without manual guidance.

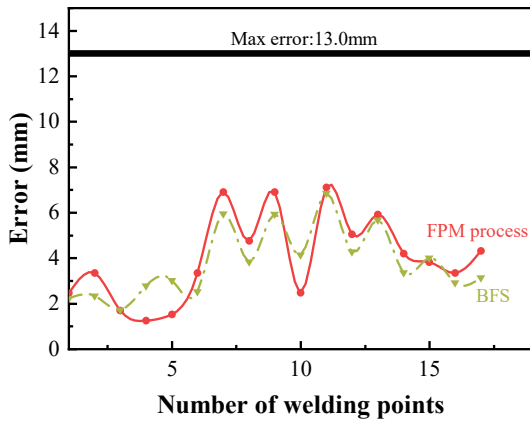


Figure 10. The x-y positioning accuracy errors of FPM process and BFS compared with the manual method

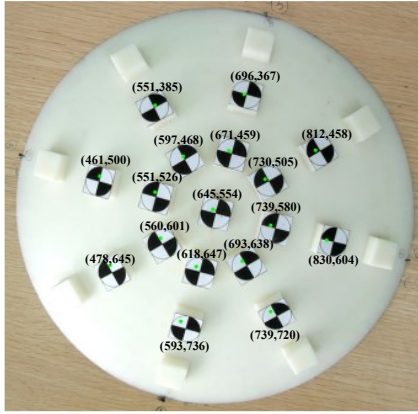


Figure 11. Pixel Coordinate Mapping by FPM process

4.2.3 Analysis of Experimental Results

This section conducts an analysis of the efficiency and errors in the iterative process of FPM process.

Efficiency:

As shown in Figure 8(b), the efficiency of FPM process is affected by the iterations of local search. Based on our experiments, the local search process can be significantly shortened or even avoided when the quality of the depth map is improved by using better cameras with higher accuracy in calculating the depth data.

The quality of the depth map is influenced by various factors, including but not limited to oversaturation caused by ambient light [25], interference from external active illumination source [26,27], and so forth. Therefore, missing depth (holes) are always observed in the depth map.

For the issue of missing depth, the following method is employed: the depth equivalent value for the currently searched pixel with missing depth is determined by calculating the average depth of non-zero points within a

7×7 pixel range centered on it.

The depth values for pixels with missing depth information are approximations, leading to errors when calculating world coordinates using Equation 1. And given that the maximum search iterations for FPM are set at $\log_4 1280 \times 960 \approx 11$ times, it becomes challenging to converge to values below 1mm, especially in cases of poor depth map quality. Hence, local search process is necessary for further refinement under such circumstances.

Therefore, the efficiency can be significantly improved by using better structured light camera with relatively fewer missing depth.

Positioning accuracy:

In our experiment, the errors in positioning mainly stem from three aspects that should be considered in further studies:

1. The quality of the depth map obtained from the structured light camera.
2. The accuracy of hand-eye calibration.
3. The relative positional relationship between the marking board and 3D printed components placed on the actual working table versus that plotted in the 2D drawings.

5 Conclusion and Future Research

To increase the efficiency of shear stud welding on the surface of 3D objects, a robotic manufacturing framework was proposed in this paper, which included a robotic arm, a structured light camera, and a 2D drawing. In particular, a fast point-searching algorithm, FPM, was developed and validated against a point-positioning experiment in our lab.

Based on our experiment, the efficiency of the FPM process is found to be 20 times higher than the traditional BFS method, and 100 times higher than the manual positioning through marking lines, and its positioning accuracy meets the code requirements.

Future research can be conducted to improve the proposed framework in the following aspects:

1. Based on the obtained world coordinates and the orientation vector at the welding point, the pose of the robotic arm's end-effector can be further calculated to accomplish the welding and evaluate its quality.
2. In the current setting, the robotic arm is fixed on the working plane. A more flexible welding platform can be developed by mounting the robotic arm on the linear axis or deploying it on an Unmanned Ground Vehicle (UGV) to accomplish the welding task on construction sites.

6 Acknowledgment

This research was supported by the National Natural Science Foundation of China (Grant No. 52278177, 52108136, and 52308312) and the National Key Research and Development Program of China (No. 2023YFC3806804).

References

- [1] Hongqi Jiang, Kang Du, Wen Sun, Guangchao Wang, Aining Jiang. The Design of a Large Automatic Welding Equipment for Shear Studs. *Machine Design and Research*, 36(05): 208-215, 2020.
- [2] Xudong Ji, Youdong Chen, Song Yang. Stud welding system using an industrial robot for membrane walls. *The International Journal of Advanced Manufacturing Technology*, 121(11-12): 8467-8477, 2022.
- [3] Youdong Chen, Qi Hu. Dual-robot stud welding system for membrane wall. *Industrial Robot: the international journal of robotics research*, 49(1): 132-140, 2022.
- [4] Alexander H Slocum, Andrew G Ziegler. An automated shear stud welding system. *Robotics autonomous systems*, 6(4): 367-382, 1990.
- [5] Andrew Glenn Ziegler. The design and fabrication of an automated shear stud welding system, Massachusetts Institute of Technology, 1988.
- [6] Rasmus S Andersen, Simon Bøgh, Thomas B Moeslund, Ole Madsen. Intuitive task programming of stud welding robots for ship construction. *2015 IEEE International Conference on Industrial Technology (ICIT)*, pages 3302-3307, 2015.
- [7] Rasmus S Andersen, Simon Bøgh, Thomas B Moeslund, Ole Madsen. Task space HRI for cooperative mobile robots in fit-out operations inside ship superstructures. *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 880-887, 2016.
- [8] Futao Diao. A new type of automatic shear studs welding workstation suitable for curved panels with dual welding guns. (*patent, in Chinese*)
- [9] Nathan B Fethke, Lauren C Gant, Fred Gerr. Comparison of biomechanical loading during use of conventional stud welding equipment and an alternate system. *Applied ergonomics*, 42(5): 725-734, 2011.
- [10] EB Holmström, J Lindell, U Moritz. Low back and neck/shoulder pain in construction workers: occupational workload and psychosocial risk factors. Part 2: Relationship to neck and shoulder pain. *Spine*, 17(6): 672-677, 1992.
- [11] Huibin Yang, Juan Yan, Automation. DXF file identification with C# for CNC engraving machine system. *Intelligent Control*, 6(01): 20-28, 2015.
- [12] Anastasios Tzotzis, Athanasios Manavis, Nikolaos Efklidis, Panagiotis Kyratsis. CAD-BASED AUTOMATED G-CODE GENERATION FOR DRILLING OPERATIONS. *International Journal of Modern Manufacturing Technologies*, 13 177-184, 2021.
- [13] Lucas Francisco Martins, Marcio Augusto Reolon Schmidt, André Luiz de Alencar Mendonça. Graphical representation analysis of complementary civil projects using "cad 2d", "bim" and "ra" and identification of interferences. *Boletim de Ciências Geodésicas*, 25, 2019.
- [14] Mohsen Foroughi Sabzevar, Masoud Gheisari, L James Lo. Improving access to design information of paper-based floor plans using augmented reality. *International Journal of Construction Education Research*, 17(2): 178-198, 2021.
- [15] Robert Eadie, Mike Browne, Henry Odeyinka, Clare Mckeown, Sean McNiff. BIM implementation throughout the UK construction project lifecycle: An analysis. *Automation in construction*, 36 145-151, 2013.
- [16] Bin Yang, Boda Liu, Dayu Zhu, Bingham Zhang, Zhichen Wang, Ke Lei. Semiautomatic structural BIM-model generation methodology using CAD construction drawings. *Journal of Computing in Civil Engineering*, 34(3): 04020006, 2020.
- [17] Hanwen Kang, Hongyu Zhou, Xing Wang, Chao Chen. Real-time fruit recognition and grasping estimation for robotic apple harvesting. *Sensors*, 20(19): 56-70, 2020.
- [18] Andrea Carron, Elena Arcari, Martin Wermelinger, Lukas Hewing, Marco Hutter, Melanie N Zeilinger, Automation Letters. Data-driven model predictive control for trajectory tracking with a robotic arm. *IEEE Robotics*, 4(4): 3758-3765, 2019.
- [19] Bowen Wen, Chaitanya Mitash, Baozhang Ren, Kostas E Bekris. se (3)-tracknet: Data-driven 6d pose tracking by calibrating image residuals in synthetic domains. *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 10367-10373, 2020.
- [20] Sulabh Kumra, Shirin Joshi, Ferat Sahin. Antipodal robotic grasping using generative residual convolutional neural network. *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 9626-9633, 2020.
- [21] Li Zhang, Peng Zhang. CAD secondary development technology based on. NET API. *IOP Conference Series: Materials Science and Engineering*, p. 072052, 2020.

- [22] HK Yuen, John Princen, John Illingworth, Josef Kittler. Comparative study of Hough transform methods for circle finding. *Image vision computing*, 8(1): 71-77, 1990.
- [23] SLS Bearings. The product catalog of shear studs. Online: https://www.slsbearings.com/slsbnn/files/SL%20BNN_Shear%20Stud%20catalogue%20%202020.pdf, Accessed: 22/11/2023.
- [24] GB 50205-2020, Code for acceptance of construction quality of steel structures, China Planning Press Beijing, China, 2020.
- [25] R. A. El-Laithy, J. Huang, M. Yeh. Study on the use of Microsoft Kinect for robotics applications. *Proceedings of the 2012 IEEE/ION Position, Location and Navigation Symposium*, pages 1280-1288, 2012.
- [26] Kai Berger, Kai Ruhl, Yannic Schroeder, Christian Bruemmer, Alexander Scholz, Marcus A Magnor. Markerless motion capture using multiple color-depth sensors. *VMV*, pages 317-324, 2011.
- [27] D Alex Butler, Shahram Izadi, Otmar Hilliges, David Molyneaux, Steve Hodges, David Kim. Shake'n'sense: reducing interference for overlapping structured light depth cameras. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1933-1936, 2012.

Implementation of a Robotic Manipulator End Effector for Construction Automation

Aarón Borgo^{1,2*}, Arturo Ruiz^{1,2*}, Jiansong Zhang, Ph.D.¹ and Luis C. Félix-Herrán^{2**}

¹ Polytechnic Institute, Purdue University, United States

² Tecnológico de Monterrey, School of Engineering and Sciences, Mexico.

* These authors contributed equally to this research work

** Corresponding author

A01251945@tec.mx, A01750408@tec.mx, zhan3062@purdue.edu, lcfelix@tec.mx

Abstract -

The construction industry presents very strong challenges such as being an industry where many fatalities occur, and delivery times have a great impact on the cost of the final product. In this context, an area of opportunity was found, which consist of automating the process of grabbing and nailing plywood sheets in construction site. To achieve this, our previous research proposed an end-effector for a robotic arm capable of performing these tasks. Built upon that end-effector design, a fundamental advancement was made in this paper. Specifically, our focus is placed on the selected robotic arm, and proposing new ways to integrate the electronic and pneumatic systems by making modifications to the initial design, as well as proposing a new manufacturing plan. The midterm goal is to manufacture a complete prototype that meets the requirements about grabbing and nailing the plywood sheets to potentially incorporate it into construction tasks. The use of robotic arms in grabbing and nailing plywood sheets could contribute to reducing construction times and accidents in this industrial sector.

Keywords -

End Effector, Robotic Arm, Plywood Grabbing and Nailing, Construction

1 Introduction

It is well known that the construction industry is one of the most dangerous economic sectors in the world. Based on the Occupational Safety and Health Administration (OSHA) statistics, of the 4,764 workers who died on the job in the United States of America during 2020, 976 were from the construction and extraction occupations, i.e., 28.48% of the total fatalities [1]. Moreover, according to reports from the Health and Safety Executive (HSE), the construction industry continues to hold the record for the highest number of fatalities among all economic sectors in the United Kingdom [2]. Actually, fatalities related to construction activities increased by 55% from April 2022 to March 2023.

Construction is one of the few industries that has not

been able to evolve with technology since most of the construction techniques used today have been used for many years, and most of the work still need to be performed manually. However, a study by Carra et al. [3] states that robots can vastly help in the installation, construction, maintenance, and inspection for the construction of a new building. This finding means that automation could potentially reduce the number of fatalities and could decrease the time and costs needed for a building to be finished; nevertheless this consideration relies on the capabilities of the end effector since it needs to be able to handle different types, sizes and weights of materials.

Nowadays, there is an increasing search for ways to implement robots in construction to support human activities and reduce accidents. This intervention includes from the design of the structure and support for the planning of the project to the visualization of the construction itself. For example, Son and Han reported image acquisition planning for image-based 3D reconstruction using a robotic arm [4]. In addition; Wang, Fukuda, and Shi did a preliminary comparison between manual and robotic construction of wooden structure architecture, verifying the much higher efficiency and accuracy of construction processes supported by robots compared with manual ones [5].

Examples of robots have been reported in the construction of wooden structures [6], but also in steel structures, where different technologies for the automation of steel beams assembly have been studied [7],[8],[9]. One inquiry includes a review of the state of the art with possible future opportunities for the application of robotic technologies in the construction industry, and highlights the applications of robots in support of construction [3]. Another approach has focused on the use of complex robotic systems [10] whereas others have been on mobile robots for construction applications, ranging from grasping small parts [11], to building reduced structures of heterogeneous brick patterns [12]. In all these cases, it has been observed that the end effector is a key element for the successful use of robots in the different construction tasks, so it is a topic of great interest for academia and industry.

There are many different designs of end effectors, but

they are still prototypes since it is really difficult to test designs in real life. One example is an anthropomorphic end effector developed by Firth et al. [13], which simulates a human hand that can hold heavy and dangerous items. Another instance is the design of Bae et al. [14], which is an end effector that can support an H-Beam alignment in high rise buildings, as well as the design of Liu et al. [15], which can handle glass substrate at high speeds. All of these innovative designs could potentially lead to safer and more efficient ways of construction.

As an evolution of using robots in the construction environment, where there is a significant amount of human interaction, collaborative robots emerge as a potential choice. The role of industrial robots includes: inspection, welding, spray painting, assembly, among others. In all these cases, human-robot interaction is vital and key challenges for human-robot collaboration (HRC) are task sharing, intent recognition and optimal trade of space and cost [16]. As it is important to understand these challenges, a critical review of HRC in on-site construction has been performed [17], and a new generation of collaborative robots for material handling for the automotive industry is shown in [18], and [19]. In addition, it is important to consider that collaborative mobile robots seek to combine the precision of a machine with the innate cognitive human skills to build structures that are not possible through conventional manual methods [20], [21], [22]. In other words, there are several benefits of HRC, which enables generalized robots to swiftly adjust to the complex and dynamic construction environments.

Furthermore there has been a lot of progress in evaluating the viability of having robotic arms perform simple tasks that occasionally cause accidents. These assignments, like nailing or bolting consume a lot of time, since the same movement need to be repeated many times. Chai et al. [23] created an end effector design which reinforces the beam network by nailing wood slabs. This realization was done with a mobile robotic arm equipped with a camera and a nail gun. First the camera detected where the wood slab needs to be placed, then the robotic arm placed the slab with a nail gun. Wang et al. [24] reported a similar design of end effector, but this design includes a glue dispenser and focuses more on a building method than on the capability of the end effector, so the results are much time saving achieved than if it was made by hand. Jung et al. [25], designed and created a similar end effector, but this time it's not made for wood and nails, but with bolts in steel beams. Their experimental results showed that the force emitted from the end effector was enough for the bolting task, but its time performance was not evaluated. There is also a different approach taken by Cheng et al. [26], who designed an end effector which could dispose nails on a wasted board. They focused more on

all the waste that is created when working in construction and how this waste can be reduced to avoid any unwanted accidents or injuries on the job. Each design in each robot focuses on some different objectives, and they have very promising results in their respective areas.

Even though plywood has gained widespread popularity as a construction material due to its capacity to endure moisture and provide strong support, it has little mention in research papers involving robots for construction. Because of its sturdiness and adaptability, plywood continues to be a cost-effective and long-lasting choice for small buildings [27]. Therefore, the herein research seeks to implement a new construction method involving plywood.

The contributions of this research effort are: the design adjustments, manufacture and assembly of the design reported by Zhang et al. [28], which consists of an end effector attached to a robotic arm, capable of grabbing and nailing plywood sheets in place, to take a first glance on how this process could be automated. This manuscript presents a Work-In-Progress regarding the methodology, robot setup, design and implementation of this innovative end effector. It should be noted that the patented design reported by Zhang et al. [28] focused on the design of the end-effector only. In contrast, this research effort focuses on how to adapt the design to a real robotic arm. This task thus requires further design (in more detail) and actual manufacturing of the components/parts to connect the end-effector and the robotic arm, and the final assembly of all components/parts to achieve a robotic system capable of grabbing and nailing plywood sheets.

2 Methodology

The main goal of the proposed method is to obtain a first working prototype for the autonomous construction robotic system newly invented [28]. A series of steps were proposed to reach a functional prototype, based on a need, a motivation and a reported patent. Figure 1 shows the stages of the applied methodology.

As shown in Figure 1, the red arrows represent the path followed for the implementation of the first prototype, leaving aside the vision system, to focus on the structure, electronic components and pneumatic system of the end effector. Besides, the first prototype of the robotic arm was performed, in order to make the necessary connections for its correct operation. As of now, design implementation and robotic arm debugging are being worked on (contoured in red in the flow diagram).

3 Robotic Arm

To know what characteristics the end effector needed to have, an inspection of the robot's characteristics was performed. For this purpose, as shown in Figure 2, the chosen

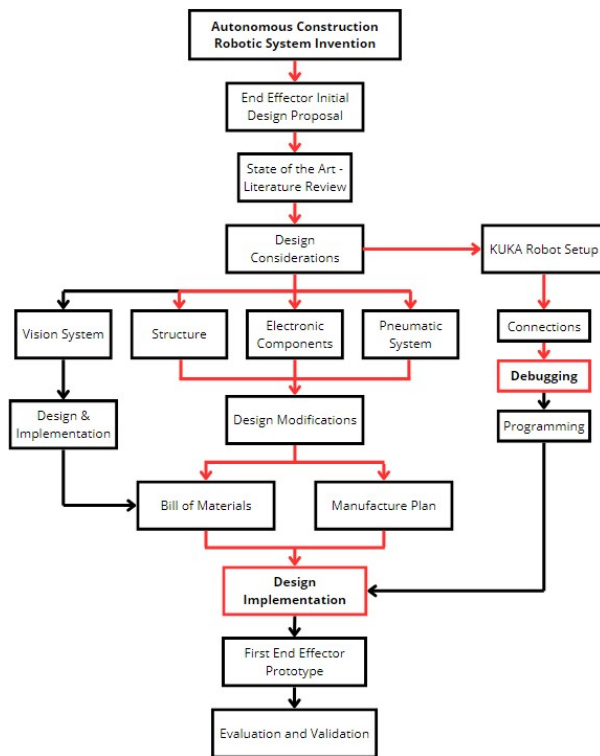


Figure 1. Applied Research Methodology

robotic arm was a KUKA KR 16 L6 with six degrees of freedom with a KR C2 controller. The mechanized system is capable of handling a maximum payload of sixteen kilograms when the center of mass of the payload is less than thirteen centimeters away from the robot's flange.



Figure 2. Robotic Arm KUKA KR 16 L6

Part of the work to be done was to fine-tune the robotic arm, since it has been used for other purposes, and subsequently, it was not used for a couple of years. There

were many loose cables and missing connections which indicate the robot had a lot of extra components, so an in-depth examination of this connections was made to ensure the robot could be used as intended, and could potentially add more extra components if needed. Some of the changes that have been made until now are: addition of an equipotential bonding cable, CMOS battery replacement and changing of controller batteries as well as removal of some unnecessary cables. The debugging phase is still in progress, as shown in Figure 3.

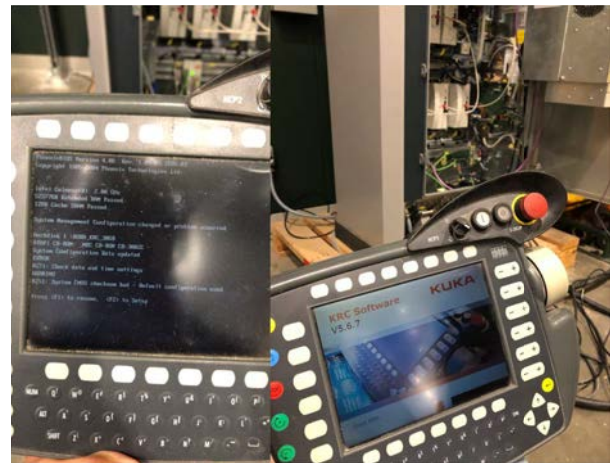


Figure 3. Debugging Process

4 The End effector

The design adjustments were performed on the patented design [28]. To achieve the job of grabbing and nailing plywood, the end effector needs to have two degrees-of-freedom (DOF), as shown in Figure 4. It is worth noting that one DOF rotates the plywood and the second DOF moves the nail gun. Moreover, the herein reported prototype focuses on the mechanical, electrical and pneumatic subsystems to ensure the correct movement of the whole system. Once this is done, the robot programming and the vision features of the end effector will be added.

4.1 Design

As mentioned before, the only strict conditions for the design of the end effector are the weight and the distance between the end effector's center of mass and the robot's flange. Under this constraint, the materials that were decided mainly concern toughness, weight and in the case of some special parts, the need to have very low friction. For this first prototype the properties that are expected to have are: the weight for the end effector is six kilograms without the plywood sheet, and with the plywood sheet is twelve kilograms, however this still doesn't take into account all of the electronic, and pneumatic components. These items should not add more than two kilograms to



Figure 4. End Effector Design in CAD

the total weight because many of them are very light and can be arranged in a way that does not add any weight to the payload. The center of mass of the end effector is located 6.6 centimeters away from the robot's flange. This design stage was done with a CAD modeling software.

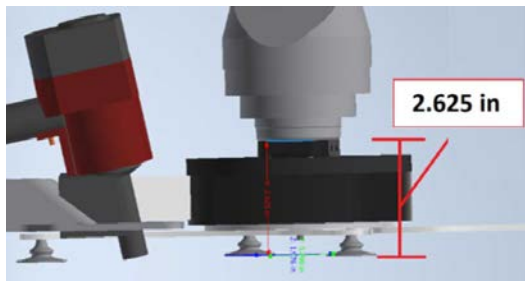


Figure 5. CAD Design of the End Effector. Enlarged View that Allows the Reader to Better Perceive the Elements of the Final Actuator in Figure 4.

The CAD shown in Figure 5 is flexible in terms of payload and the center of mass distance, which is helpful in case there are some major changes in the design. There were some key elements which the whole model was designed from, one of them are the DC motors. They are high RPM motors because these motors have a gearbox capable of delivering the necessary torque to move the components attached to it. This decision was made because these brushless motors are for long-endurance multirotor aircraft, but are designed to be lightweight with high efficiency and consistency, and are much cheaper than servo motors that can provide the desired torque. This way a cycloidal gear was designed for the first motor to be able to rotate the part where the plywood sheet will be attached

with a pneumatic system as shown in Figure 5. In addition, the cycloidal gear helps with the location for the center of mass to make sure it is in the required range of the robotic arm's payload diagram.

An aluminum arm is attached to the case that holds the cycloidal gear together. This arm connects to an elbow joint that is also powered with the same motor as the cycloidal gear, but in this case there is no cycloidal gear, instead a planetary gearbox is used. This is done to have less weight on this area, so the center of mass stays closer to the robot's flange. Also this planetary gearbox has a different gear ratio because it needs to have a different amount of torque. On the other end of the elbow joint, there is going to be a twenty three gauge pneumatic nail gun. This part of the design will let the nail gun move around the edge of the plywood sheet where the nails are supposed to be.

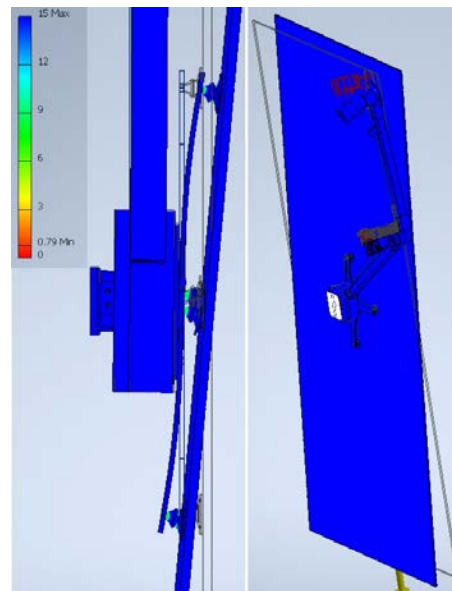


Figure 6. Stress Analysis. Safety Factor Results with Color Scale of End Effector with Plywood (Right) and a Closed-up Image of the Suction Cups (Left).

Furthermore, a stress analysis was conducted to ensure that the materials chosen for this design can handle the plywood's weight, therefore the only force interacting with the design is gravity. The focus in this analysis is the safety factor, and as shown in the scale in Figure 6, there are no elements of concern except for the suction cups. Nonetheless, this study does not incorporate the pressure exerted by the pneumatic system on the suction cups, as this additional pressure often contributes to their sturdiness. Moreover, the data sheet for the suction cups intended for use in the physical model confirms their capability to support the weight of the plywood. In Figure 6, it can also be seen that the design will slightly suffer defor-

mation caused by gravity, but it is greatly exaggerated in the figure since the deformity is less than 0.2 inches when the plywood is in vertical position.



Figure 7. CAD Pneumatic System Implementation

As shown in Figure 7, the pneumatic system involves four suction cups that are going to be connected to four vacuum generators connected to an air supply. The whole system will be connected with a series of adapters and tubes. These selected components are easy to obtain since those are commonly used in the industry. However, this system was designed to be as close as possible to the cycloidal gear because there must be enough space to fit all of the air tubes and adapters, so that the center of mass of the end effector will not be affected by this system, as well as not interfering with the cycloidal gear and the moving arms. This part of the end effector is designed to grab and let go of the plywood sheet.

For this first prototype, the robot, DC motors and pneumatic system, are controlled independently and cannot communicate with each other. Work is underway for making all of the systems controllable by a single laptop (or equivalent) to support the whole automated operation. Some additional adjustments to the design are not rejected either.

4.2 Manufacture

Once the design of the first prototype was finished, and the raw materials were acquired, a manufacture and assembly plan was developed. It was decided for this first prototype to make all of the pieces easy to assemble and dismount in case further modifications need to be done.

For this prototype most of the pieces were 3D printed because most of the parts are new designs, adjustments can be made in an easier and faster way. Polyoxymethylene (POM) or acetal resin was the first option for 3D filament to be used because of its key properties to offer better toughness and lower friction than Polylactic Acid (PLA) or Acrylonitrile Butadiene Styrene (ABS). However, having successful prints with this filament is not easy nor cheap because of the employed printer properties, and its difficulty for its layers to stick together. Work is underway to pursue successful printings with POM material.

In order to have a better understanding on how the end effector will be assembled, the 3D prints were done with PLA as shown in Figure 8. With a reliable 3D printer

filament, using Ender v3 3D printers with the help of Ultimate Cura software high quality prints were obtained. Parts 3D printed in this manner include the cycloidal gear components, robotic arm mount, nail gun mount, motors plus drives cases, elbow joint and rotation segment. All of the pins, bearings and electronics that go with them were left out from the 3D printing.



Figure 8. 3D Printed Component

Some pieces have very special material requirements. The elbow joint is planned to have a shaft that will be manufactured with stainless steel disks and rod, since it has to hold the weight of the pneumatic nail gun and an aluminum arm. For the end effector's arms, and suction cups mount, the chosen material was aluminum since there is no need for those pieces to be tougher than steel and they need to be as light as possible. First, the aluminum arms were cut to the desired size using a bandsaw machine. Then, for the manufacture of the elbow joint disks and suction cups mount, a laser cutter was used to make the complex figures needed, with high precision cutting on the stainless steel and aluminum sheets.

For the electronics and pneumatic systems, the selected components are easy to obtain since those are commonly used in the industry. For most of these components there were no modifications needed; however, the length of the cables and air tubes need to be modified to fit the design.

4.3 Assembly

To further discuss the missing details about unforeseen issues, the first prototype was built with the parts that were already manufactured and printed. In this case only the cycloidal gear, elbow joint, aluminum arms and part of the pneumatic system were assembled.



Figure 9. Cycloidal Gear Assembly

The cycloidal gear was the first section that was assembled together as shown in Figure 9. This is because it is the system's portion that may have more constraints once all of the parts were put together. The gear has to run smoothly, and with as less friction as possible to avoid wearing out once it is used with the DC motor. The initial testing with the assembling progress shows that everything is working as expected, but there will be some changes to the design to make the assembling process easier.



Figure 10. Elbow Joint Assembly. An In-depth Observation of the Elbow and Its Constituent Parts, for an Enhanced Understanding of Components that are Represented in Figure 4.

The elbow joint is the second section that was put together as it appears in Figure 10. This process consisted of joining every part with screws. As of this date the assembly process is almost done since the elbow shaft is the only missing piece to be manufactured.



Figure 11. Pneumatic Components Assembly. A Physical Representation of the Implementation of the Pneumatic System as Depicted in Figure 7.

The pneumatic part of the end effector has been the least complex part to assemble since it was outsourced. This component does not need any sort of manufacture process

except for the length of the air tubes. This section status is as shown in Figure 11 and there are no unforeseen issues.

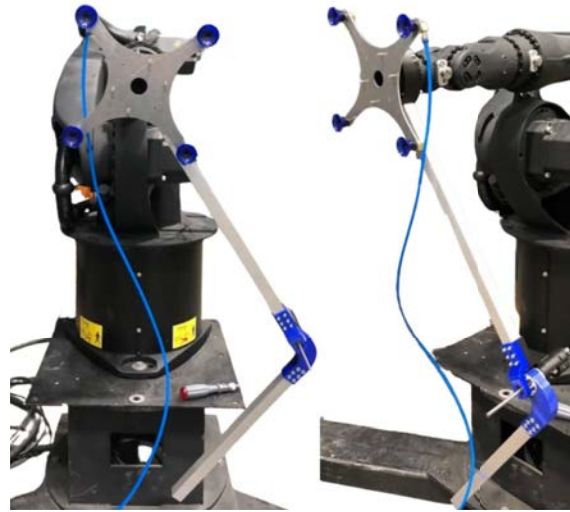


Figure 12. First End Effector Prototype. A Portrayal of the Implementation of the End Effector Illustrated in Figure 4.

Once these components were put together, a first glimpse of the physical end effector can be seen in Figure 12. As of now the majority of parts are within the expected characteristics and properties, and have not suffered major changes.

5 Conclusions and Future Work

Adaptations to a reported end-effector design are presented, as well as progress in the manufacturing and assembly of a robotic arm for grabbing and nailing of plywood sheets with a focus on construction tasks. The outcomes in this research allow to have a better physical visualization, integrated understanding and evaluation of all the components together to finish the implementation of the pneumatic and electronic system. Moreover, the expected specifications for the first prototype involve a weight of less than fourteen kilograms, accounting for the absence of cables and the pneumatic system tubes. Additionally, an evaluation will be conducted whether any material adjustments are deemed necessary once the assembly process is finished. This also allows to have the components ready to be elaborated with a more resistant material to withstand continuous tasks. Moreover, further considerations that need to be taken into account are the need of a cooling system for the motors in addition to having a micro-controller that can connect all of the electric, pneumatic and robotic arm components in order to automate this process once the testing phase is concluded; as well as validations with the complete prototype.

6 Acknowledgements

This work was supported by the School of Construction Management Technology at Purdue University. The authors would like to acknowledge: Dr. Jose Garcia Bravo for granting access to Banton Fluid Power Lab and supporting with the 3D printing process; Dave Kotterman, Built-It-At-Scale Lab Director, for having provided the required manufacturing equipment; As well as Rosita Lopez, Construction Lab Technician, for her help on the procurement process of materials.

References

- [1] U.S. Department of Labor. Occupational Safety and Health Administration (OSHA). Commonly used statistics. On-line: <https://www.osha.gov/data/commonstats>, Accessed: 10/25/2023.
- [2] Kanaris S. Construction industry fatalities have increased by 55% in the last year. new civil engineer. On-line: <https://goo.su/JtWY>, Accessed: 10/25/2023.
- [3] Carra G., Argiolas A., Bellissima A., Niccolini M., and Ragaglia M. Robotics in the construction industry: State of the art and future opportunities. In *Proceedings of the 35th International Symposium on Automation and Robotics in Construction (ISARC)*, pages 866–873, Taipei, Taiwan, 2018.
- [4] Hyo Son R. and K. Han. Image acquisition planning for image-based 3d reconstruction using a robotic arm. In *Proceedings of the 38th International Symposium on Automation and Robotics in Construction (ISARC)*, pages 769–775, Dubai, UAE, 2021.
- [5] Wang L., Fukuda H., and X. Shi. A preliminary comparison between manual and robotic construction of wooden structure architecture. In *Proceedings of the 37th International Symposium on Automation and Robotics in Construction (ISARC)*, pages 1568–1575, Kitakyushu, Japan, 2020.
- [6] Leng Y., Shi X., and Hiroatsu F. Application of robots to the construction of complex structures using standardized timbers. In *Proceedings of the 37th International Symposium on Automation and Robotics in Construction (ISARC)*, pages 1562–1567, Kitakyushu, Japan, 2020.
- [7] Chu B., Jung K., Chu Y., Hong D., Lim M.-T., Park S., Lee Y., Lee S.-U., Kim M. C., and Ko K. H. Robotic automation system for steel beam assembly in building construction. In *2009 4th International Conference on Autonomous Robots and Agents*, pages 38–43, Wellington, New Zealand, 2009.
- [8] Lee S.-K., Doh N. L., Park G.-T., Kang K.-I., Lim M.-T., Hong D.-H., Park S.-S., Lee U.-K., and Kang T.-K. Robotic technologies for the automatic assembly of massive beams in high-rise building. In *2007 International Conference on Control, Automation and Systems*, pages 1209–1212, Seoul, Korea (South), 2007.
- [9] Nam H., Choi W., Ryu D., Lee Y., Lee S.-H., and Ryu B. Design of a bolting robot for constructing steel structure. In *2007 International Conference on Control, Automation and Systems*, pages 1946–1949, Seoul, Korea (South), 2007.
- [10] Iturralde K., Pan W., Linner T., and Bock T. 19 - automation and robotic technologies in the construction context: research experiences in prefabricated façade modules. pages 475–493, 2022.
- [11] Asadi K., Haritsa V. R., Han K., and Ore J.-P. Automated object manipulation using vision-based mobile robotic system for construction applications. *Journal of Computing in Civil Engineering*, 35(1), 2021.
- [12] Basiri M., Gonçalves J., Rosa J., Vale A., and Lima P. An autonomous mobile manipulator to build outdoor structures consisting of heterogeneous brick patterns. *SN Applied Sciences*, 3:558, 2021.
- [13] Firth C., Dunn K., King M., and Haeusler M. H. Development of an anthropomorphic end-effector for collaborative use on construction sites. *CAADRIA proceedings*, 2(1):363–372, 2020.
- [14] Bae K., Chu B., Jung K., Lee Y., Hong D., Park S., and Lim M.-T. An end-effector design for h-beam alignment in high-rise building construction. In *2008 International Conference on Smart Manufacturing Application*, pages 465–469, Goyangi, Korea (South), 2008.
- [15] Liu Z., Chen Y., Song H., Xing Z., Tian H., and Shan X. High-speed handling robot with bionic end-effector for large glass substrate in clean environment. *Sensors*, 22(1), 2022.
- [16] Kumar A., Bahubalendruni M. V. A. R., Ashok D., and SankaranarayanaSamy K. Challenges and opportunities in human robot collaboration context of industry 4.0 -a state of the art review. *Industrial Robot*, 49(2):1–14, 2022.

- [17] Zhang M., Xu R., Wu H., Pan J., and Luo X. Human-robot collaboration for on-site construction. *Automation in Construction*, 150:104812, 2023.
- [18] Gambao E., Hernando M., and Surdilovic D. A new generation of collaborative robots for material handling. In *Proceedings of the 29th International Symposium on Automation and Robotics in Construction (ISARC)*, Eindhoven, The Netherlands, 2012.
- [19] Lee S. and Moon J. II. Introduction of human-robot cooperation technology at construction sites. In *Proceedings of the 31st International Symposium on Automation and Robotics in Construction and Mining (ISARC)*, pages 978–983, Sydney, Australia, 2014.
- [20] Sandy T., Giftthaler M., Dörfler K., Kohler M., and Buchli J. Autonomous repositioning and localization of an in situ fabricator. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2852–2858, Stockholm, Sweden, 2016.
- [21] Dörfler K., Sandy T., Giftthaler M., Gramazio F., Kohler M., and Buchli J. Mobile robotic brickwork. In *Robotic Fabrication in Architecture, Art and Design 2016*, pages 204–217. Springer, Cham., 2016.
- [22] Helm V., Ercan S., Gramazio F., and Kohler M. Mobile robotic fabrication on construction sites: Dim-rob. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4335–4341, Vilamoura-Algarve, Portugal, 2012.
- [23] Chai H., Wagner H. J., Guo Z., Qi Y., Menges A., and Yuan P. F. Computational design and on-site mobile robotic construction of an adaptive reinforcement beam network for cross-laminated timber slab panels. *Automation in Construction*, 142:104536, 2022.
- [24] Wang L., Naito T., Leng Y., Fukuda H., and Zhang T. Research on construction performance evaluation of robot in wooden structure building method. *Buildings*, 12(9), 2022.
- [25] Jung K., Chu Y., Chu B., Hong D., Park S., Lim M.-T., Lee Y., Lee S.-U., Ko K. H., and Kim M. C. Experimental evaluation of a robotic bolting device in steel beam assembly. In *Proceedings of the 2009 International Symposium on Automation and Robotics in Construction (ISARC 2009)*, pages 245–251, Austin, USA, 2009.
- [26] Cheng C., Wu M., Pan Y., and Shang H. A self-designed rotating end-effector based on robotic system for disposing of nails in wasted board. In *2021 International Conference on Networking Systems of AI (INSAI)*, pages 234–238, Shanghai, China, 2021.
- [27] MP MORAN. The use of plywood in construction. On-line: <https://www.mpmoran.co.uk/blog/post/the-use-of-plywood-in-construction>, Accessed: 10/25/2023.
- [28] Zhang J., M Lacny C., and Reardonet N. Autonomous robotic system for placing and fastening paneling material for building construction operations, September 05, 2023. U.S. Patent Number: 11745356.

Low center-of-gravity movement robotic arm with kinematic optimization algorithm for on-site construction

Ruiqi Jiang¹ and Xiao Li¹

¹Department of Civil Engineering, University of Hong Kong, HKSAR, China
richj233@connect.hku.hk, shell.x.li@hku.hk

Abstract –

Robotic arms have increasingly been applied to onsite building construction (e.g., bricklaying, welding, and 3D printing). However, engineers should make a great effort to customize different mobile platforms (e.g., quadruped, hexaploid, tracked robot) for robotic arms in the complex ground environment of the construction site. The cable-driven platform under the existing crane system has the potential to address this problem. Nevertheless, the robotic arm operation would cause the swing of the cable-driven platform. Therefore, this paper aims to propose a novel structure for the robotic arm, namely the center of gravity control margin (CoG-CM) robotic arm and its control algorithm to reduce the CoG movement in a cable-driven platform. Compared with the conventional robotic arm, the main contribution of the proposed one is that the arm consists of four parallel joints, which could provide the control capability of the system's CoG. This structure has infinite solutions for a determined target, and conventional control algorithms are unsuitable for this system. So, we formulate its CoG, and the robotic arm CoG pose control (ArmCoG-PC) is proposed to solve the kinematics of the CoG-CM robotic arm. Finally, the experiment on CoG-CM and typical 6-degree-of-freedom robotic arms validate our proposal.

Keywords –

Construction robots; Optimization algorithm; Robotic arm

1 Introduction

The contemporary construction site heavily depends on manual labor, exposing workers to substantial risks. In contrast, utilizing construction robots offers a solution for reducing human involvement in hazardous environments and presents the potential for productivity improvements in the construction industry [1]. Among the widely employed robotic technologies, the robotic arm is among the most prevalent [2]. Numerous researchers have made valuable outcomes of on-site construction applications, including additive manufacturing [3], automated installation [4-5], and

robotic bricklaying [6].

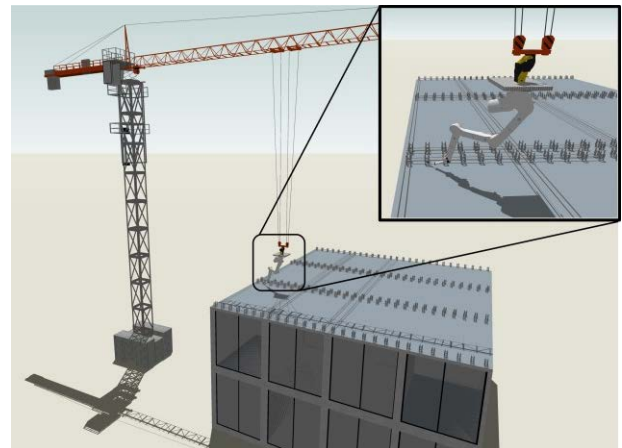


Figure 1. The robotic arm is mounted on a cable-driven platform under the existing tower crane.

While robotic arms show the capability in diverse construction tasks, their effective deployment relies on specially designed platforms for their application environments. Among the existing construction robots, wheeled platforms stand out as commonly utilized to support operational robotic arms [7]. Despite their adaptability to different environments, wheeled platforms face challenges when confronted with complex construction sites [8]. Therefore, these wheeled robots are mainly applied in the decoration stage with standard site environments, such as plastering [9] and fitting-out tasks [10]. To address these challenges, some researchers have innovatively designed rebar-tying robots sliding along tracks [11] and inspection robots moving on steel structures [12]. However, these designs require careful consideration of site conditions and constraints. As an alternative, researchers have explored the integration of aerial platform robotic arms into construction processes [13]. Nevertheless, drones' limited payload capacity restricts aerial robots' applicability in construction contexts.

Recently, the cable-driven platform (CDP) for robotic arms has been introduced into the construction domain [14]. The CDP is typically mounted at a tower crane above the working area, as shown in Figure 1. The CDP exhibits characteristics conducive to adapting to ground

conditions and supporting a promotive payload capacity [15]. Despite these advantages, the inherent elasticity of cables and the continual shifting of the center of gravity (CoG) during operations would cause a noticeable oscillation. This situation presents a challenge for robotic arms in achieving high-precision operations.

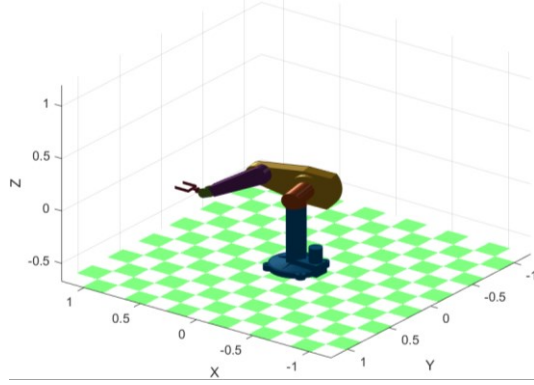


Figure 2. PUMA560 robotic arm.

To address the issue mentioned above, we propose a novel configuration of the robotic arm, namely the CoG control margin (CoG-CM) robotic arm, with a different joint sequence and amount from the conventional arm, as shown in Figure 2, along with its receding horizon control-based method, to minimize the movement of arm's CoG during operation. Specifically, the arm has four sequential parallel joints, providing a controllable margin for the CoG. However, the design causes higher computational complexity of joint pose. Traditional robotic arm pose control algorithms, e.g., [16] and [17], are not well-suited for solving this problem due to the highly coupled and nonlinear nature of the CoG formulation. In response, we propose the robotic arm CoG pose control (ArmCoG-PC) algorithm to tackle this issue. Within the ArmCoG-PC algorithm, the Particle Swarm Optimization (PSO) algorithm is introduced as a solver for optimal results, leveraging the principles of receding horizon control to achieve optimal outcomes within a sliding horizon. Additionally, we integrate constraints to effectively restrict the solution space to reduce the computational burden during the horizon solution stage.

The main contributions of this paper are as follows:

1. The CoG-CM robotic arm is proposed to reduce the CoG movement of the overall system, which is mounted on a cable-driven platform under the tower crane during the operation.
2. The control algorithm, namely robotic arm CoG pose control (ArmCoG-PC), is proposed. We also design constraints to limit the space of optimal results to reduce computational costs.
3. The comparison experiments are arranged, and the experimental results illustrate the effectiveness of the CoG-CM robotic arm and ArmCoG-PC algorithm.

2 Methodology

This section will first model the CoG Control Margin (CoG-CM) robotic arm using Denavit-Hartenberg (DH) representation parameters. Subsequently, the pose control algorithm, denoted as the robotic arm CoG pose control (ArmCoG-PC) algorithm, will be introduced. This algorithm is developed based on the CoG analysis and the state transition matrix.

2.1 Modeling CoG-CM robotic arm

To achieve a control margin for the CoG in the robotic arm, we propose a novel joint configuration called CoG-CM, which is different from wide-used robotic arms (e.g., PUMA560, as shown in Figure 2). In the CoG-CM robotic arm configuration, the first three joints are arranged sequentially perpendicular to each other, while the rotation directions of the third to sixth joints are parallel. Enhanced degrees of freedom (DoF) for the end effector are realized through the seventh and eighth joints.

The CoG-CM robotic arm could be denoted by DH representation [18]. The DH parameters are divided into standard and modified parameters and modeled based on links and joints as coordinate systems. The position of CoG is generally related to the positions of joints and the end of the links. Therefore, using the coordinate system at the end of the link, i.e., standard DH parameters, facilitates modeling the center of gravity for the CoG-CM robotic arm. Thus, the DH representation method in this paper represents the standard DH parameters. The CoG-CM robotic arm is shown in Figure 3, and its DH parameters are detailed in Table 1, where θ_i is the rotation angle of the i th joint, d_i is the distance along the z-axis from the i th joint to the $i + 1$ th joint, α_i is the twist about the x-axis between the i th and $(i + 1)$, and a_i is the link length of the i th joint. To simplify the calculation, d_i are all set as 0.

Table 1 DH parameters of the CoG-CM robotic arm

Joint No.	θ	d	α	a
1	θ_1	0	0.5π	0
2	θ_2	0	0.5π	a_2
3	θ_3	0	0	a_3
4	θ_4	0	0	a_4
5	θ_5	0	0	a_5
6	θ_6	0	0.5π	a_6
7	θ_7	0	0.5π	0
8	θ_8	0	0.5π	0

The general form of the transformation matrix could be obtained from DH parameters as follows:

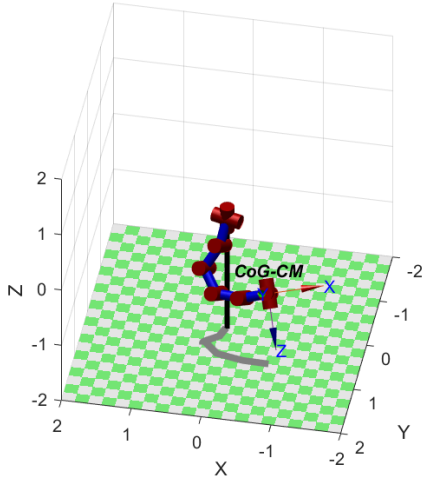


Figure 3. Center of gravity control margin (CoG-CM) robotic arm.

$$A_i^{i+1}(\theta_i) = \begin{bmatrix} R_i & O_i \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} c\theta_i & -s\theta_i c\alpha_i & s\theta_i s\alpha_i & a_i c\theta_i \\ s\theta_i & c\theta_i c\alpha_i & -c\theta_i s\alpha_i & a_i s\theta_i \\ 0 & s\alpha_i & c\alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (1)$$

where matrix $A_i^{i+1}(\theta_i)$ describes the coordinate transformation between the i th and $(i+1)$ th joints, R_i is the i th rotation matrix, O_i is the coordinate vector, $s(\cdot)$ and $c(\cdot)$ are sine and cosine functions, respectively.

Therefore, the transformation from the i th joint to i' th joint could be obtained by multiplying the transformation matrices as

$$A_i^{i'} = A_i^{i+1} \cdots A. \quad (2)$$

This paper mainly studies the rigidly connected robotic arms, so the CoG for the link between two adjacent joints remains relatively constant. Hence, the CoG for each link could be simplified. For ease of analysis, we consider each link as a uniformly distributed mass rod, and its CoG is located at the midpoint between two rotational joints. Specifically, for the link associated with the i th joint, with coordinates $O_i = [x_i \ y_i \ z_i]^T$, the CoG's coordinate O_{CoG_i} could be expressed as

$$O_{CoG_i} = \frac{1}{2}(O_i + O_{i+1}), \quad (3)$$

$$\begin{bmatrix} R_{i+1} & O_{i+1} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_i & O_i \\ 0 & 1 \end{bmatrix} A_i^{i+1}. \quad (4)$$

The position of the robotic arm's CoG could be

denoted as

$$O_{CoG} = \frac{m_1 O_1 + m_2 O_2 + \cdots + m_8 O_8}{m_1 + m_2 + \cdots + m_8} \quad (5)$$

where m_i is the mass of the i th link.

2.2 ArmCoG-PC algorithm

The four parallel joints in the CoG Control Margin (CoG-CM) robotic arm will lead to an infinite kinematic solution for reaching one fixed target. Moreover, the Eq. (4) indicates the relationship between the CoG (O_{CoG_i}) of the i th link and the preceding joints (from θ_1 to θ_{i-1}), making the direct calculation of algebraic solutions challenging. Moreover, analyzing the Center of Gravity's (CoG) movement during operation requires a global perspective, making it challenging to constrain the CoG movement in a single-step planning. Consequently, the proposed control algorithm must be able to solve complex optimization problems and demonstrate a certain level of predictability. Receding horizon control, also known as model predictive control, involves establishing a prediction horizon and utilizing an optimizer to solve for the sequence of outputs within the horizon based on optimization objectives. Hence, we adopt the concept of receding horizon control to compute the kinematics of the CoG-CM robotic arm.

Firstly, the discrete state-space function of the robotic arm could be expressed as

$$\Theta(t+1) = A\Theta(t) + BU(t)$$

$$\Theta(t) = [\theta_1(t) \ \theta_2(t) \ \cdots \ \theta_8(t)]^T, \quad (6)$$

$$u_i(t) = \begin{cases} 1, \text{Clockwise} \\ 0, \text{No Rotation} \\ -1, \text{Counterclockwise} \end{cases}$$

where $\theta_i(t)$ represents the angle of the i th joint at time t , A is the identity matrix, B is a diagonal matrix, and $U(t) = [u_1(t) \ \cdots \ u_8(t)]^T$ is the rotation direction of joints at time t .

To maneuver the robotic arm end effector $P_{ed}(t) = [x_e(t) \ y_e(t) \ z_e(t)]$ towards the target point $P_{tr} = [x_p \ y_p \ z_p]$ while simultaneously minimizing the movement of CoG, the objective function could be designed as follows:

$$\min k_1 \text{dis}(P_{ed}(t), P_{tr}) + k_2 \sum_{t=0} \text{dis}[O_{CoG}(t), O_{CoG}(t-1)], \quad (7)$$

where $\text{dis}(\cdot, \cdot)$ represents the Euclidean distance between two points, and k_1 and k_2 are adjustable coefficients for the distance to the target and the movement of CoG, respectively. It is worth noting that

constraining the movement of the CoG throughout operation should be analyzed from the whole movement process. Only considering the output sequence obtained within a single step cannot guarantee that the result is optimal within a longer horizon.

Based on the above analysis, the pose control algorithm necessitates the capacity to consider every output in a certain horizon. Consequently, receding horizon control is introduced to minimize the movement of the robotic arm during operation. The receding horizon control could be outlined in the following steps:

Step 1 (Prediction): The system's future states are predicted using the current state and state-space function in a certain span (also called prediction horizon).

Step 2 (Optimization): Based on the predicted states, the control problem could be transferred to an optimization problem with constraints, and the aim is to find the optimal sequence of control inputs.

Step 3 (Implementation): Only the control input for the first step is implemented. This step-by-step approach ensures that the calculated control inputs are responsive to the evolving dynamics of the system.

The receding horizon control is widely used in certain systems owing to its robustness and low sensitivity. However, the challenges arise in our system due to multiple optimization objectives and high coupling, making the formulation of the optimization function a complex task. Therefore, it is necessary to modify the computation process for our model.

We propose a sequential computation approach to establish a predictive horizon for pose control and alleviate the complexity involved in formulating from the steps in the predictive horizon. Firstly, the pose for the last step is computed. Subsequently, the computation progresses backward to the first step within the constraints of each step's reachable solution. The ArmCoG-PC algorithm could be summarized as follows:

Step 1: At time t and the horizon length l , obtain the optimal results $\hat{\Theta}$ of $(t + l)$ th step based on Eq.(6) and objective function.

Remark 1: It is noteworthy that, due to the constraints of window length l , the maximum available distribution range for the pose at the time $(t + l)$ is limited, i.e.,

$$\Theta(t) - lB \leq \hat{\Theta}(t + l) \leq \Theta(t) + lB. \quad (8)$$

The optimal solution distribution space would be constrained. The objective function could be denoted as

$$\arg \min_{\hat{\Theta}(t+l)} f(t, l), \quad (9)$$

$$f(t, l) = k_1 \text{dis}(P_{ed}(t + l), P_r) + k_2 \text{dis}[O_{\text{CoG}}(t), O_{\text{CoG}}(t + l)]. \quad (10)$$

Algorithm 1 The working procedure of PSO-RD

```

1: procedure PSO-RD(Input:  $\Theta(t)$ ,  $l_i$ , and parameters of PSO-RD Output:  $\hat{\Theta}(t + l_i)$ )
2:   Initialize parameters
3:   According to Eq. 8 or 11 to generate solution space  $X$ .
4:   while Not meet the set iteration do
5:     Update particle speed  $V$ , as equation .
6:     Constrain  $V$  with parameters.
7:     Update particle state  $S$ , as equation .
8:     Constrain  $S$  with solution space  $X$ .
9:     Update the best information of particle swarm.
10:  end while
11:  Output  $\hat{\Theta}(t + l_i)$  with the best information of particle swarm.
12: end procedure

```

Step 2: Based on the initial state $\Theta(t)$ and the obtained result $\hat{\Theta}(t + l)$, the result of time $(t + l - 1)$ is also limited and could be solved according to Eq. (9) and (10).

Remark 2: As the movement of the robotic arm from state $\Theta(t)$ to state $\hat{\Theta}(t + l)$ is a continuous process, the solution space for $\hat{\Theta}(t + l - 1)$ should be constrained by the two states, as illustrated in Figure 4, and it could be formulated as follows.

$$\begin{cases} \Theta(t) - (l - 1)B \leq \Theta(t + l - 1) \leq \Theta(t) + (l - 1)B \\ \hat{\Theta}(t + l) - B \leq \Theta(t + l - 1) \leq \hat{\Theta}(t + l) + B. \end{cases} \quad (11)$$

Step 3: Similar to the **Step 2**, compute the result of $\hat{\Theta}(t + l - 2)$ with $\Theta(t)$ and $\hat{\Theta}(t + l - 1)$ as constraints. Repeat this process iteratively until $\hat{\Theta}(t + 1)$ is obtained.

Step 4: Implement the output for time t with the Eq. (6) and $\hat{\Theta}(t + 1)$.

In practice, the primary controllers for robotic arms are digital and operate discretely. The control commands for the joints of the robotic arm are expressed as discrete sequences. Therefore, the distribution of each joint could be regarded as points in solution space. Heuristic or evolution algorithms could be adopted to solve the optimization problem of Eq.(9). This paper uses particle swarm optimization (PSO) as the optimizer. PSO algorithms are primarily designed for continuous solution spaces [19]. To adapt to the resolution of robotic arm's state, adjustments are necessary in terms of discretization and constraints. The PSO for robotic arm discrete problems (PSO-RD) is outlined in Algorithm 1, where l_i denotes the number of optimized steps. The speed and space update procedure could be expressed as Eq. (12) and (13), respectively, where v_i^n and s_i^n represent the velocity and state of particle i in the n th iteration, s_i^{best} and s^{gbest} signify the best states during the search process for the i th particle, and the swarm's best state, B' denotes the diagonal elements of B , and w_1 , w_2 , and w_3 are weight parameters.

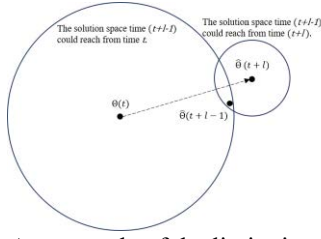


Figure 4. An example of the limitation of solution space.

$$v_i^n = \text{round}(w_1 v_i^{n-1} + w_2 \text{rand}(s_i^{\text{best}} - s_i^{n-1}) + w_3 \text{rand}(s_i^{\text{gbest}} - s_i^{n-1})), \quad (12)$$

$$s_i^n = s_i^{n-1} + B' v_i^n \quad (13)$$

The ArmCoG-PC algorithm could be expressed as Algorithm 2, where the symbol \oslash represents each element of two vectors divided element-wise.

Algorithm 2 The working procedure of ArmCoG-CP optimization algorithm

```

1: procedure PSO-RD(Input:  $\Theta(t)$ ,  $l$  Output:  $\hat{\Theta}(t+1)$ )
2:   Initialize parameters
3:   Compute  $\Theta(t+l)$  with  $\Theta(t)$  and Algorithm 1.
4:   for  $(l-i) \geq 1$  do
5:     Compute  $\Theta(t+l-i)$  with  $\Theta(t+l-i+1)$ ,  $\Theta(t)$ 
       and Algorithm 1.
6:      $i = i - 1$ .
7:   end for
8:   Set the output  $U(t) = B' \oslash (\Theta(t+l) - \Theta(t))$ .
9: end procedure

```

3 Experiment

This section conducts two sets of experimental comparisons: experiments involving the CoG-CM robotic arm model and a typical 6-DoF robotic arm model. For the CoG-CM robotic arm experiments, the comparison is made between the ArmCoG-PC algorithm and a direct joint rotation approach. Additionally, a comparison is made between the ArmCoG-PC algorithm and the commonly used inverse kinematics solution method on the typical 6-DoF robotic arm. Experiments only consider the mass of links to simplify the mass of robotic arms. The robotic toolbox in Matlab is used as the simulation environment [20].

3.1 Experiments on the CoG-CM robotic arm model

In this subsection, experiments based on the CoG-CM robotic arm model are analyzed, and the DH parameters of the CoG-CM robotic arm are shown in Table 2, where offset θ is the initial joint states expressed in radians.

The comparative process first employs the ArmCoG-PC algorithm to solve the kinematics of the CoG-CM robotic arm to reach the target. Subsequently, the joints are directly linearly rotated to achieve the final pose. The parameters of the PSO-RD comprise two types: particle

parameters and coefficient parameters. The particle parameters significantly impact computational efficiency. Hence, we utilize an enumeration method, incrementing each parameter. Specifically, at each moment $t = 0$, **Step 1** of the ArmCoG-PC algorithm is used to assess the setting of particle parameters. For each iteration, 10-time-repeat experiments are conducted until the cost is minimized across all 10 experiments. The number of particles is set at 10, with 40 iterations. For enhancing the rounding precision in such a discrete system, the coefficients in PSO-RD, i.e., w_1 , w_2 and w_3 are set as 0.8, 0.5, and 0.5, respectively. The coefficients of CoG movement and distance to the target k_1 and k_2 are all set as 1. The horizon length determines the algorithm's predictive capability and computational burden. We choose the value at which the motion trajectory appears as coherent as possible, i.e., 4. The initial states of the CoG-CM robotic arm $\Theta(0) = [-0.3, -0.3, -0.3, -0.3, -0.3, -0.3, -0.3, -0.3]$. The target coordinates are $[-0.7, 0.7, -0.7]$ and $[-0.5, 0.5, -0.5]$. The resolution of rotation is set as 0.02 rad.

The experimental results are presented in Figures 5, 6, and Table 3. Figures 5 and 6 depict the motion of the CoG-CM robotic arm optimized by the ArmCoG-PC algorithm and the direct linear rotation. The depicted blue lines represent the trajectory of the arm's CoG movement. The observations from Figures 5 and 6 illustrate that the poses devised by the ArmCoG-PC algorithm facilitate a more direct CoG movement than the linear direct rotation method, which follows a relatively longer curved path. Additionally, the CoG movement trajectory of the robotic arm, when planned with the ArmCoG-PC algorithm, exhibits a fluctuation as the endpoint of the robotic arm moves from one quadrant of the x-y plane to another, which means the proposed algorithm enables the arm to adjust its joints to maintain a low CoG movement distance. This observation also demonstrates the high CoG movement control margin of the CoG-CM robotic arm. Table 3 provides the total distance of the CoG movement for both methods. The results show the effectiveness of the ArmCoG-PC algorithm in significantly reducing the distance of CoG movement for the CoG-CM robotic arm.

Table 2 DH parameters of the CoG-CM robotic arm in experiments

Joint No.	Offset θ	d	α	a
1	-0.3	0	0.5π	0
2	-0.3	0	0.5π	0.4
3	-0.3	0	0	0.4
4	-0.3	0	0	0.4
5	-0.3	0	0	0.4
6	-0.3	0	0.5π	0.4
7	-0.3	0	0.5π	0
8	-0.3	0	0.5π	0

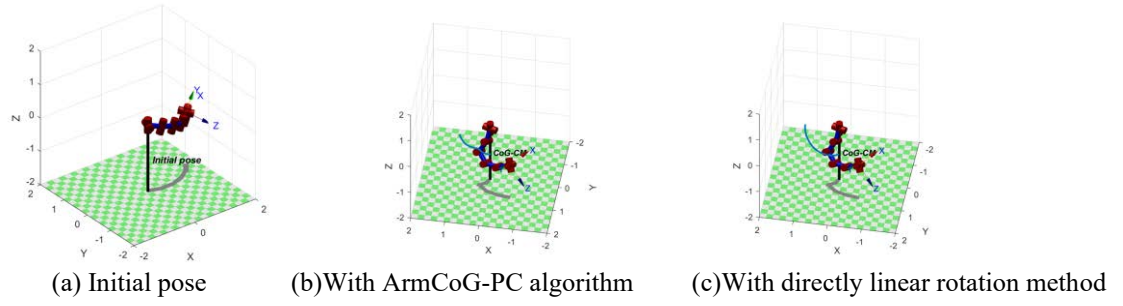
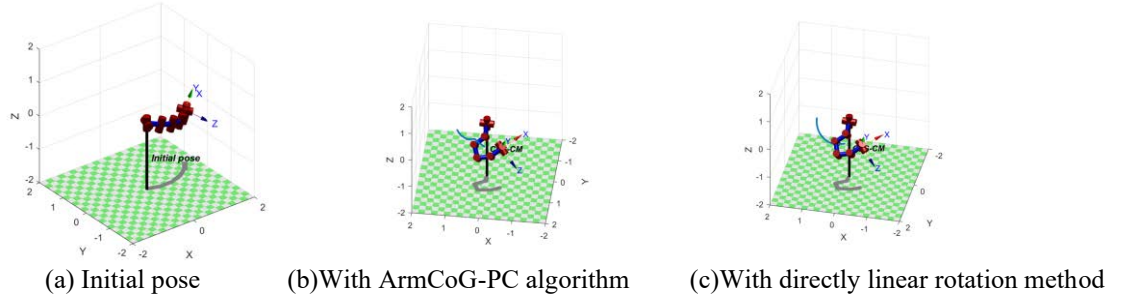
Figure 5. Operation results of CoG-CM robotic arm to $[-0.7, 0.7, -0.7]$.Figure 6. Operation results of CoG-CM robotic arm to $[-0.5, 0.5, -0.5]$.

Table 3 The movement distance of the CoG-CM robotic arm in experiments

Target	ArmCoG-PC	Linear Movement
$[-0.7, 0.7, -0.7]$	1.3250	1.5855
$[-0.5, 0.5, -0.5]$	1.2536	1.4843

3.2 Experiments on the typical 6-DoF robotic arm model

Table 4 The DH parameters of typical 6 DoF robotic arm in experiments

Joint No.	Offset θ	d	α	a
1	-0.3	0	0.5π	0.5
2	-0.3	0	0	0.5
3	-0.3	0	0.5π	0.5
4	-0.3	0	0.5π	0
5	-0.3	0	0.5π	0
6	-0.3	0	0	0

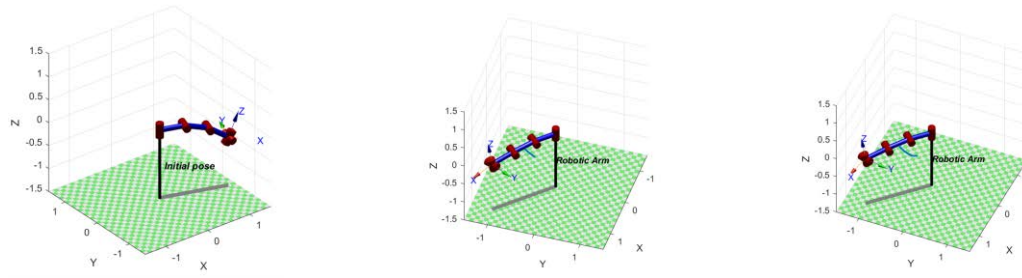
In this subsection, experiments based on the typical 6-DoF robotic arm are analyzed, and the DH parameters of the typical 6-DoF robotic arm are shown in Table 4. The setting of parameters is the same as in subsection 3.1. The initial states of the CoG-CM robotic arm $\theta(0) = [-0.3, -0.3, -0.3, -0.3, -0.3, -0.3]$. Due to the workspace limitation, the targets are set as $[1, -1, -0.2]$ and $[-0.2, 0.2, -0.4]$.

Figures 7 and 8 illustrate the movement of CoG to

two different targets in two methods. As shown in Figure 7, the ArmCoG-PC optimization algorithm has an insignificant reduction of CoG movement due to the DoF constraints of the robotic arm. However, according to the data in Table 5, the ArmCoG-PC optimization algorithm still works in this situation, reducing almost 42% movement distance of the arm's CoG, compared with directly moving the joints. In Figure 8, for situations that require a larger rotation angle, the effectiveness of the ArmCoG-PC algorithm becomes more pronounced. With the direct linear rotation method, the trajectory of the arm's CoG movement corresponding to the overall motion follows a long curve, while the pose controlled by the ArmCoG-PC algorithm exhibits a straighter line. The case controlled by the ArmCoG-PC algorithm reduces the distance by 0.3979, approximately 37%. Compared to the cases controlled by the ArmCoG-PC algorithm in subsection 3.1, the CoG movement fluctuation in this subsection is much larger, further demonstrating the effectiveness of the CoG-CM robotic arm. Overall, these results show the notable effectiveness of the ArmCoG-PC algorithm in reducing CoG movement for the robotic arm under various operational conditions.

Table 5 The movement distance of the CoG-CM robotic arm in experiments

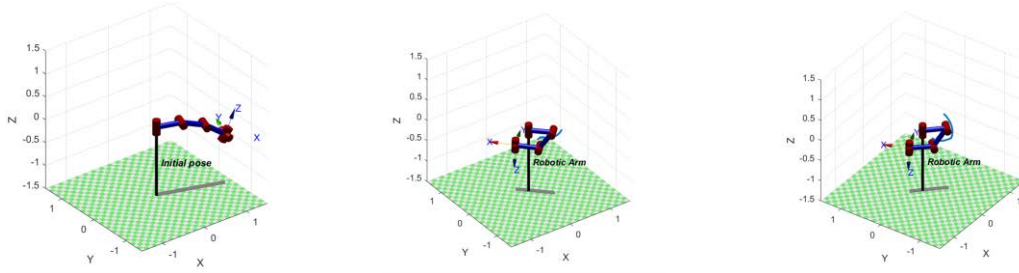
Target	ArmCoG-PC	Linear Movement
$[1, -1, -0.2]$	0.3400	0.5857
$[-0.2, 0.2, -0.4]$	0.6779	1.0758



(a) Initial pose

(b) With ArmCoG-PC algorithm

(c) With directly linear rotation method

Figure 7. Operation results of typical 6 DoF robotic arm to $[1, -1, -0.2]$.

(a) Initial pose

(b) With ArmCoG-PC algorithm

(c) With directly linear rotation method

Figure 8. Operation results of typical 6 DoF robotic arm to $[-0.2, 0.2, -0.4]$.

4 Conclusion

In this study, we addressed the challenge of oscillation induced by the Center of Gravity (CoG) movement in a robotic arm mounted on a cable-driven platform under the existing tower crane. Our study introduces the CoG Control Margin (CoG-CM) robotic arm configuration with a novel robotic arm CoG pose control (ArmCoG-PC) algorithm. The CoG-CM robotic arm, featuring four parallel joints, minimizes the CoG movement during operation. In the proposed configuration, the computational complexity and the requirement of predictive capability in motion control, as well as conventional algebraic kinematics methods, are unsuitable. Consequently, we propose the ArmCoG-PC algorithm to effectively solve the pose of the CoG-CM robotic arm. Experimental results, encompassing both the CoG-CM and a typical 6-DoF robotic arm, illustrate the robustness and efficiency of the ArmCoG-PC optimization algorithm.

While the effectiveness and advantages of the CoG-CM robotic arm and the ArmCoG-PC algorithm were demonstrated, certain aspects warrant further exploration in future studies. First, the experiments conducted thus far rely on simulations, and incorporating field testing would significantly bolster the applicability of our proposed approach to large-scale robotic arms in on-site construction. In addition, although longer link lengths in the CoG-CM robotic arm can improve the control margin,

they may simultaneously reduce the arm's load capacity. An in-depth analysis of the relationship between link length and actual load capacities is needed to optimize the trade-off between control margins and overall robotic arm efficiency. Furthermore, when dealing with complex environments, the predictive horizon of the ArmCoG-PC algorithm should be expanded to ensure comprehensively low CoG movement. This expansion may cause a high computational cost, so future research should focus on methods for enlarging the predictive horizon without a proportional increase in computational costs.

For the on-site construction, the CoG-CM robotic arm and ArmCoG-PC algorithm could provide opportunities for large-scale (heavy) robotic arms, especially empowering those robotic arms mounted on the cable-driven platform with high-precise operation in on-site construction.

Acknowledgment

This research is supported by grants from the University of Hong Kong (HKU Project No.109000053 and No.006010313).

References

- [1] Chen X, Chang-Richards A Y, Pelosi A, et al. Implementation of technologies in the construction industry: a systematic review[J]. Engineering,

- Construction and Architectural Management, 2022, 29(8): 3181-3209. DOI: 10.1108/ecam-02-2021-0172.
- [2] Gharbia M, Chang-Richards A, Lu Y, et al. Robotic technologies for on-site building construction: A systematic review[J]. Journal of Building Engineering, 2020, 32: 101584. DOI: 10.1016/j.jobbe.2020.101584.
- [3] Panda B, Lim J H, Mohamed N A N, et al. Automation of robotic concrete printing using feedback control system[C]. ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction. IAARC Publications, 2017, 34. DOI: 10.22260/isarc2017/0037.
- [4] Lee S, Gil M, Lee K, et al. Design of a ceiling glass installation robot[C]. Proceedings of the 24th International Symposium on Automation and Robotics in Construction. 2007: 247-252. DOI: 10.22260/isarc2007/0044.
- [5] Chu B, Jung K, Lim M T, et al. Robot-based construction automation: An application to steel beam assembly (Part I) [J]. Automation in construction, 2013, 32: 46-61. DOI: 10.1016/j.autcon.2012.12.016.
- [6] Yu S N, Ryu B G, Lim S J, et al. Feasibility verification of bricklaying robot using manipulation trajectory and the laying pattern optimization[J]. Automation in Construction, 2009, 18(5): 644-655. DOI: 10.1016/j.autcon.2008.12.008.
- [7] Ha Q P, Yen L, Balaguer C. Robotic autonomous systems for earthmoving in military applications[J]. Automation in Construction, 2019, 107: 102934. DOI: 10.1016/j.autcon.2019.102934.
- [8] Turner C J, Oyekan J, Stergioulas L, et al. Utilizing industry 4.0 on the construction site: Challenges and opportunities[J]. IEEE Transactions on Industrial Informatics, 2020, 17(2): 746-756. DOI: 10.1109/tii.2020.3002197.
- [9] Jenny S E, Pietrasik L L, Sounigo E, et al. Continuous Mobile Thin-Layer On-Site Printing[J]. Automation in Construction, 2023, 146: 104634. DOI: 10.1109/tii.2020.3002197.
- [10] Nlink. Drilly: Mobile drilling robot. On-line: <https://www.nlinkrobotics.com/projects/drilly-mobile-drilling-robot>, Accessed: 11/12/2023.
- [11] Momeni M, Relefors J, Khatry A, et al. Automated fabrication of reinforcement cages using a robotized production cell[J]. Automation in Construction, 2022, 133: 103990. DOI: 10.1016/j.autcon.2021.103990.
- [12] La H M, Dinh T H, Pham N H, et al. Automated robotic monitoring and inspection of steel structures and bridges[J]. Robotica, 2019, 37(5): 947-967. DOI: 10.1017/s0263574717000601.
- [13] Munoz-Morera J, Maza I, Fernandez-Aguera C J, et al. Assembly planning for the construction of structures with multiple UAS equipped with robotic arms[C]. 2015 International Conference on Unmanned Aircraft Systems (ICUAS). IEEE, 2015: 1049-1058. DOI: 10.1109/icuas.2015.7152396.
- [14] Hu R, Pan W, Iturralde K, et al. Construction Automation and Robotics for Concrete Construction: Case Studies on Research, Development, and Innovations[C]. ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction. IAARC Publications, 2023, 40: 683-690. DOI: 10.22260/isarc2023/0095.
- [15] Iturralde K, Feucht M, Illner D, et al. Cable-driven parallel robot for curtain wall module installation[J]. Automation in Construction, 2022, 138: 104235. DOI: 10.1016/j.autcon.2022.104235.
- [16] Fundamentals of robotic mechanical systems: theory, methods, and algorithms[M]. New York, NY: Springer New York, 2003. DOI: 10.1007/978-0-387-34580-2.
- [17] Gan J Q, Oyama E, Rosales E M, et al. A complete analytical solution to the inverse kinematics of the Pioneer 2 robotic arm[J]. Robotica, 2005, 23(1): 123-129. DOI: 10.1017/s0263574704000529.
- [18] Niku S B. Introduction to robotics: analysis, control, applications[M]. John Wiley & Sons, 2020. DOI: 10.1108/ir.2003.04930cae.002.
- [19] Kayhan A H, Ceylan H, Ayvaz M T, et al. PSOLVER: A new hybrid particle swarm optimization algorithm for solving continuous optimization problems[J]. Expert Systems with Applications, 2010, 37(10): 6798-6808. DOI: 10.1016/j.eswa.2010.03.046.
- [20] Corke P I. A robotics toolbox for MATLAB[J]. IEEE Robotics & Automation Magazine, 1996, 3(1): 24-32. DOI: 10.1109/100.486658.

System Integration of Construction Planning and Robots for a Joint Civil Engineering and Robotics Course

Ryosuke Yajima¹, Fumiya Matsushita¹, Keiji Nagatani¹ and Kazumasa Ozawa¹

¹School of Engineering, University of Tokyo

yajima@i-con.t.u-tokyo.ac.jp, matsushita@i-con.t.u-tokyo.ac.jp,
keiji@i-con.t.u-tokyo.ac.jp, ozawa@i-con.t.u-tokyo.ac.jp

Abstract -

This study describes the development of a system that links construction planning and construction robots for civil engineering and robotics-integrated education programs. In this program, it is possible to learn construction planning and automation using model construction robots. The major issue was that the information in the developed construction plans could not be used directly for the operation of the model construction robots. Therefore, to create innovations or ideas in an environment closer to future automated construction sites, where information, communication, and robot technologies are linked, a web application for construction planning was built and linked to model construction robots via API. Although some issues were revealed in the demonstration, the system worked as intended.

Keywords -

Education; Construction planning; Construction robots; Automation; Web applications; API; System integration

1 Introduction

Low productivity and labor shortages are major issues in Japan's construction industry. To solve these issues, the Ministry of Land, Infrastructure, Transport, and Tourism has set a policy called "i-Construction" and aims to improve productivity in the construction industry by 20% by 2025 [1]. One of the key points of "i-Construction" is to fully automate the construction site and transform it into a state-of-the-art factory. Furthermore, open innovation is emphasized to introduce new information, communication, and robot technologies into construction sites [2]. This trend has become important not only in Japan but also in all over the world [3]. Therefore, it is necessary to educate civil, ICT (Information and Communication Technology), and robotics engineers to collaborate across disciplines.

Our laboratory on Construction System Management for Innovation [4] has been advancing research and development and education to contribute to "i-Construction." One educational activity is the development of a civil engineering and robotics integrated education program [5].

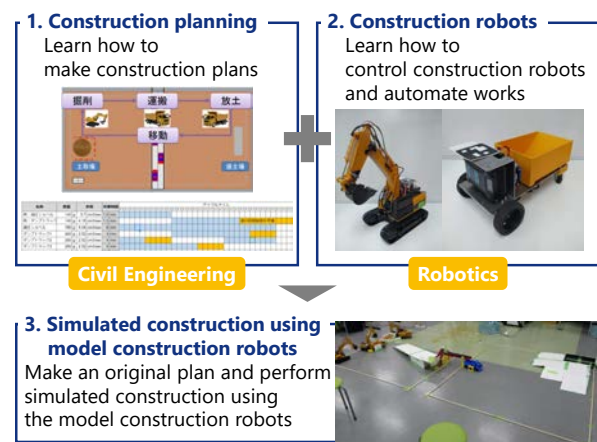


Figure 1. Basic configuration of the education program

The purpose of this educational program is to create innovations to improve productivity in earthworks and to develop human resources. In this program, it is possible to learn construction planning and automation using model construction robots. Participants can experience the process of creating innovation through a simulated automated construction that combines construction planning and model construction robots. In the field of civil engineering, there are programs to study BIM (Building Information Modeling) and ICT Construction as educational programs [6, 7]. There are also educational programs that introduce unmanned aerial vehicles [8] and develop IoT applications [9]. On the other hand, in the field of robotics, educational programs and competitions to create ideas and develop robots are widely conducted. The topics vary from game-like challenges to realistic problems [10, 11, 12]. The novelty of our educational program is that it focuses on the automation of earthworks, and integrates the knowledge of civil engineering, such as construction planning, and robotics, such as autonomous control.

This educational program has been conducted as a four-day intensive course for graduate students every year since 2021, with continuous improvements. In 2021, a basic

configuration was built, and it was confirmed that the participants created various ideas through practice. However, the only model construction robot that could be used in this educational program was an excavator. In 2022, although an environment that could handle full automation at construction sites was prepared by robotizing all excavators and dump trucks, some issues were revealed. The major issue was that the information in the developed construction plans could not be used directly for the operation of the model construction robots. When the construction plan was changed, it was necessary to modify the programs following the plan directly and certainly by hand. In other words, hard coding was required. This issue may occur not only in this educational program, but also in actual automated construction. In the research and development of the overall system of automated construction machines, higher-level systems have been built to connect the construction plan or tasks with low-level controllers of construction machines [13, 14, 15, 16]. Therefore, such kind of integration will be essential for future automated construction.

To address this issue, a system linking construction planning and construction robots was developed and introduced into our educational program. This changes the educational program to one that allows participants to experience the process of creating innovations in an environment that is more similar to a future construction site where information, communication, and robot technologies are linked. This paper describes the revealed issue, the developed linking system between construction planning and model construction robots, and a demonstration of the system through the practice of an educational program in 2023.

2 Basic Configuration of the Education Program

The basic configuration and features of our education program are explained (Figure 1). As mentioned above, this education program aims to create innovation to improve productivity in earthworks and develop human resources. This process was divided into the following three parts:

1. Construction planning

The method to build construction plans under site conditions and constraints is taught as a content of civil engineering course. Specifically, the layout of the construction equipment is planned, productivity is measured, the cycle time is calculated, and is examined to ensure the plan satisfies the required construction period for a simulated earthwork. Moreover, the direct construction cost for the built plan is calculated. This plan is built on an Excel sheet and a PowerPoint

file prepared as teaching tools.

2. Construction robots

The method to control construction robots and to automate construction works is taught as a content of robotics. Specifically, programs to control the angle of joints and the position or velocity of the body of the small model construction robots prepared as teaching tools are developed. The system of the robots is based on ROS (Robot Operating System) [17] and the programming language is Python.

3. Construction planning and simulated construction using model construction robots

In the simulated construction, it is required to transport iron beads instead of soil in a specified weight within a required construction period in the field of 2 m in length and 4 m in width. Two excavators, three dump trucks, conveyor belts, and temporary materials can be used for the simulated construction. Since 2022, all excavators and dump trucks have been robotized and the participants aim to realize fully automated construction. The time, quality, cost, and automation level are evaluated as scores, and the group competition is held. Each group builds an original construction plan that can receive a higher score and develops motion programs for the model construction robots to realize the plan.

The most significant feature of this educational program is the integration of civil engineering and robotics, considering future construction sites. Participants can experience the process of creating new ideas and innovations through practice and competition, not only on the table, but also using real equipment, although they are models and simulated environments.

3 Development of Linking System Between Construction Planning and Model Construction Robots

The major issues revealed by the implementation of the educational program in 2022 and their improvements are summarized. Other improvements are not discussed in this study.

3.1 Issue and Improvement Method

Although some issues were revealed from the practices of the previous education program in 2022, the major issue was that the information in the developed construction plans could not be used directly for the operation of model construction robots. As mentioned previously, the layout of the construction equipment was planned by drawing maps on a PowerPoint file. For example, to make the

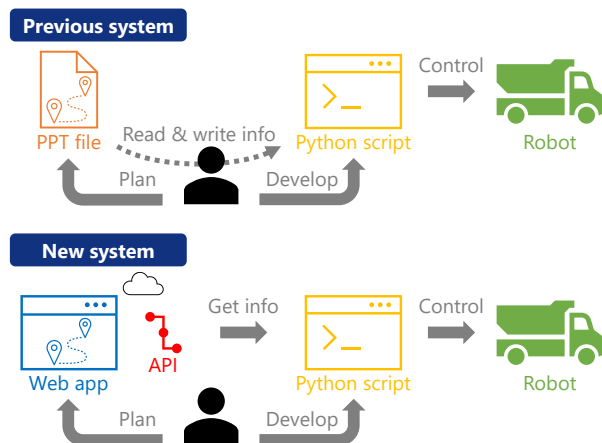


Figure 2. System for the education program

dump trucks follow a planned path using the implemented path-following control, it was necessary to develop control programs with hard-coded positions or paths based on the plan (upper side of Figure 2).

Therefore, a web application and an API (Application Programming Interface) are prepared to solve this issue. The web application is used for the layout planning of construction equipment. The API allows access to planned position or path information on the web application. This makes it possible to link the construction plan and model construction robots directly by developing programs that obtain information from the web application via the API and operate the model construction robots (lower side of Figure 2).

3.2 R-CDE and ServiceHUB

The R-CDE was used as a web application for construction planning, and ServiceHUB was used to link R-CDE and the model construction robots.

The R-CDE is a prototype of a common data platform for data and system collaboration in the construction phase, developed by the University of Tokyo and the Japan Federation of Construction Contractors [18]. This system visualizes the 3D models and 3D point clouds stored in the system.

ServiceHUB is a prototype API collaborative platform that links the R-CDE to various devices or applications. The API is implemented using the RESTful API. This allowed access to the above information of 3D models in the R-CDE.

These are web systems built on AWS. By connecting various applications and devices, such as total stations and GNSS on construction machines, to the R-CDE via the Internet and the ServiceHUB, information required for construction and inspection can be managed and shared. R-CDE also can be accessed through a browser (Figure 3).

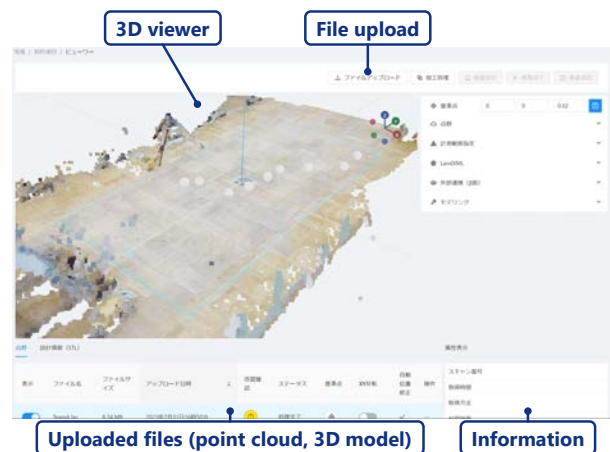


Figure 3. Operation screen of R-CDE

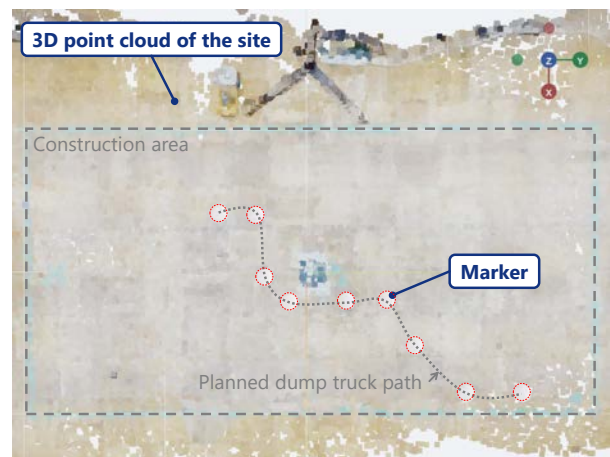


Figure 4. Details of the 3D viewer of R-CDE (top view)

From the browser, data can be uploaded and the information contained can be confirmed. In addition, a 3D viewer is provided to visualize uploaded 3D models and 3D point clouds. Users can place them in arbitrary positions in the 3D viewer.

In this educational program, the R-CDE is used to plan the layout of construction equipment. In addition, the information from the plan is directly used to control the model construction robot by connecting the R-CDE and the robot via the ServiceHUB. Specifically, first, a 3D point cloud of the simulated construction site is displayed on a 3D viewer of the R-CDE. Second, 3D models of a marker representing a robot's position and path, such as a sphere or cone, are placed at arbitrary positions on the 3D point cloud to plan the layout and path of the construction equipment (Figure 4). Third, the programs to control the robots send API requests to the ServiceHUB as needed and receive the marker position information in the R-CDE

as a response. Finally, the received information is used to move the robots.

4 Demonstration of the Developed System

The improved education program was conducted as a four-day intensive course at the University of Tokyo from July 31 to August 3, 2023. A system that links construction planning and model construction robots was demonstrated in this educational program. The prepared system and results of the demonstration are explained.

4.1 Prepared System

The system was prepared as follows.

- R-CDE and ServiceHUB
- Model construction robots
- PC
- Wireless LAN router

The R-CDE and ServiceHUB were built on AWS (Amazon Web Services). The model construction robot and PC were built using Ubuntu 20.04 LTS and ROS 2 Foxy, and connected to the same wireless LAN router. Programs to control the model construction robots were developed and executed on a PC. Moreover, the wireless LAN router was connected to the Internet via a mobile network. While creating a construction plan, the participants accessed the R-CDE from a web browser on a PC. When linking the construction plan and model construction robot, information on the R-CDE was obtained from the control program of the model construction robots via ServiceHUB. Because the educational program was conducted in three groups, three sets of this system were prepared.

4.2 Results

The methods used to obtain information on R-CDE via ServiceHUB in this educational program can be roughly divided into the following three methods.

1. From Postman
2. From Python scripts on Google Colaboratory
3. From Python scripts to control the model construction robots on the PC

Method 1 is a test to check the access and learn the mechanics of the API. The Postman is a platform for using or developing API. Method 2 is performed to learn how to access the API from a Python script. Method 3 connects the system to the model construction robots.

In Methods 1 and 2, information on the R-CDE can be obtained through ServiceHUB by sending requests. We confirmed that the R-CDE and ServiceHUB functioned normally.

Method 3 can be realized by integrating Method 2 into Python scripts to control the model construction robots. Note that the implementation of Method 3 and the use of the obtained information were not directly instructed and were left to the participants in the education program. One group developed a program that first performed API access when the program ran, obtained the position information of a group of markers placed on the R-CDE, and used it as a path for the dump truck. Although there were differences in usage, the developed system worked as intended, and the construction planning and model construction robots were integrated.

However, it is necessary to improve the usability of R-CDE. In particular, it seems difficult to arrange markers indicating the paths of dump trucks.

This system makes it possible to link the layout or route plan in the construction plan, i.e., the position information, to construction robots. However, there is other important information in the construction plan and control of construction robots, such as the timing or triggering of motion. Planning it in a system and linking it to construction robots is a topic for future work.

When focusing on the construction plans of each group, all groups built plans using only model construction robots in 2023, although there were differences in the details. Some groups used belt conveyors or bridges made of temporary materials until 2022. Thus, the variety of construction plans was reduced. The reduced variety may be because the participants concentrated on the software development of the construction robot and there was not enough time to create ideas for the construction plan. This was related to the fact that, unlike in previous years, the construction robot programming was taught first, followed by the construction planning. The additional development topics, such as API access, also affected it. Thus, the structure and duration of the educational program should be reconsidered.

5 Conclusion

This study describes the development of a system that links construction planning and construction robots for civil engineering and robotics-integrated education programs. To create innovations or ideas in an environment closer to future automated construction sites, where information, communication, and robot technologies are linked, a web application for construction planning was built and linked to model construction robots through API. Although some issues were identified, the developed system worked as intended.

In future, by solving the aforementioned issues, an environment in which participants can concentrate more on creating ideas for construction must be established. Furthermore, a similar setup to that of full-scale construction robots is provided, and the content is expanded to connect to real construction sites. The final goal is to improve construction sites by applying the valuable ideas obtained through this educational program to real construction sites.

Acknowledgements

The authors thank Professor Atsushi Yamashita of the University of Tokyo for his cooperation in conducting the course. We also thank Mr. Shogo Inatomi, Mr. Wakana Endo, and Mr. Chen Tong for their assistance with TAs.

References

- [1] Transport Ministry of Land, Infrastructure and Tourism. Committee report on i-construction (in Japanese). On-line: <https://www.mlit.go.jp/common/001127288.pdf>, Accessed: 12/12/2023.
- [2] Transport Ministry of Land, Infrastructure and Tourism. White paper on land, infrastructure, transport and tourism in Japan, 2016. 2017.
- [3] T. Bock. The future of construction automation: Technological disruption and the upcoming ubiquity of robotics. *Automation in Construction*, 59:113–121, 2015. doi:10.1016/j.autcon.2015.07.022.
- [4] Laboratory on construction system management for innovation. On-line: <http://www.i-con.t.u-tokyo.ac.jp>, Accessed: 12/12/2023.
- [5] F. Matsushita, R. Yajima, K. Ozawa, and K. Nagatani. Development of civil engineering / robotics integrated education program aimed at creating innovation in ict earthwork. *Journal of Japan Society of Civil Engineers, Ser. H (Engineering Education and Practice)*, 78(1):38–52, 2022. doi:10.2208/jscej.78.1.38.
- [6] F. Peterson, T. Hartmann, R. Fruchter, and M. Fischer. Teaching construction project management with bim support: Experience and lessons learned. *Automation in Construction*, 20(2):115–125, 2011. doi:10.1016/j.autcon.2010.09.009.
- [7] T. Kolli, R. Heikkilä, J. Röning, T. Sipilä, J. Erho, M. Hyyryläinen, and P. Lamassaari. Development of the education of open infra bim based construction automation. In *Proceedings of the 35th International Symposium on Automation and Robotics in Construction*, pages 791–797, Berlin, Germany, 2018. doi:10.22260/ISARC2018/0110.
- [8] I. M. P. Antonenko. Unmanned aerial vehicles as educational technology systems in construction engineering education. *Journal of information technology in construction*, 27, 2022. doi:10.36680/j.itcon.2022.014.
- [9] R. Chacón, H. Posada, Á. Toledo, and M. Gouveia. Development of iot applications in civil engineering classrooms using mobile devices. *Computer Applications in Engineering Education*, 26(5):1769–1781, 2018. doi:10.1002/cae.21985.
- [10] S. Evripidou, K. Georgiou, L. Doitsidis, A. A. Amanatiadis, Z. Zinonos, and S. A. Chatzichristofis. Educational robotics: Platforms, competitions and expected learning outcomes. *IEEE Access*, 8:219534–219562, 2020. doi:10.1109/ACCESS.2020.3042555.
- [11] K. Ishii, Y. Takemura, T. Matsuo, and T. Sonoda. Tomato harvesting robot competition. In *Proceedings of 2016 Joint 8th International Conference on Soft Computing and Intelligent Systems (SCIS) and 17th International Symposium on Advanced Intelligent Systems (ISIS)*, pages 537–542, 2016. doi:10.1109/SCIS-ISIS.2016.0118.
- [12] T. Doi, M. Shimaoka, and S. Suzuki. Creative robot contests for decommissioning as conceived by college of technology or kosen educators. *Journal of Robotics and Mechatronics*, 34(3):498–508, 2022. doi:10.20965/jrm.2022.p0498.
- [13] Q. Ha, M. Santos, Q. Nguyen, D. Rye, and H. Durrant-Whyte. Robotic excavation in construction automation. *IEEE Robotics Automation Magazine*, 9(1):20–28, 2002. doi:10.1109/100.993151.
- [14] Q. P. Ha and D. C. Rye. A control architecture for robotic excavation in construction. *Computer-Aided Civil and Infrastructure Engineering*, 19(1):28–41, 2004. doi:10.1111/j.1467-8667.2004.00335.x.
- [15] J. Seo, S. Lee, J. Kim, and S. Kim. Task planner design for an automated excavation system. *Automation in Construction*, 20(7):954–966, 2011. doi:10.1016/j.autcon.2011.03.013.
- [16] L. Zhang, J. Zhao, P. Long, L. Wang, L. Qian, F. Lu, X. Song, and D. Manocha. An autonomous excavator system for material loading tasks. *Science Robotics*, 6(55):eabc3164, 2021. doi:10.1126/scirobotics.abc3164.
- [17] Ros - robot operating system. On-line: <https://www.ros.org>, Accessed: 12/12/2023.

- [18] F. Matsushita, K. Miyaoka, F. Miyazaki, H. Kojima, J. Nobuto, and K. Ozawa. Development of a prototype of a common platform at construction stage and investigation of cooperative areas (in japanese). In *Proceedings of Civil Engineering and Construction Technology Presentations 2023*, 2023.

Development of Autonomous Robots for Construction

Takayoshi Hachijo¹, Shunsuke Igarashi¹, Taku Tani², and Masahiro Indo³

¹ Institute of Technology, Shimizu Corporation, Japan

² Construction Technology Division, Shimizu Corporation, Japan

³ NOVARE, Shimizu Corporation, Japan

hachijo_t@shimz.co.jp

Abstract –

We have developed two types of autonomous robots for construction works. One is an autonomous mobile robot composed of two industrial robotic arms for ceiling board construction work. It can attach plasterboard weighing approximately 14 kg to ceiling foundations with tapping screws. The other is also an autonomous mobile robot composed of two industrial robotic arms for raised floor construction work. It can install pedestals and panels to floor slabs. Moreover, we have developed assist robots which are autonomous mobile robots to supply materials for each construction work robot. In this paper, we describe overviews of these robots and case studies.

Keywords –

Autonomous mobile robot; Industrial robot; Ceiling board construction; Raised floor construction

the environment also changes daily, by using GNSS and 3D-LiDAR to determine self-position and the surrounding environment. Lauer et al. [8] developed a force control system for a construction machine equipped with a large manipulator to automatically position wood components when they were assembled on site. The system enabled positioning of large workpieces while allowing for divergence between the real construction site and the plan.

We have developed two types of robots to automate the ceiling board construction and the raised floor construction. In this paper, we report overviews and case studies of these autonomous robots. This paper is organized into 4 parts. We explain the overviews of the autonomous robots for ceiling board construction in Sec. 2, the overviews of the autonomous robots for raised floor construction in Sec. 3, and we give our conclusions in Sec. 4.

1 Introduction

In recent years, the number of skilled workers in Japan has declined by approximately 1.4 million over the past 25 years [1]. About 35% of workers in the construction industry are 55 years old or older, and only about 12% are 29 years old or younger.

Currently, in the construction industry, many examples of efforts to improve productivity by using robotics and digital technology have been reported [2-5]. The introduction of robots to construction sites is expected to improve productivity, alleviate labor shortage, and improve the working environment. However, several issues exist to install robots at construction sites. At the construction sites, the robots need mobility. Self-localization and alignment may be issues because the surrounding environment changes daily. Le et al. [6] developed a mobile robot that automatically fixed cross laminated timber (CLT) panels with screws, and determined the construction position by matching BIM data with the surrounding environment acquired by LiDAR sensors. Satoh et al. [7] developed a mobile robot equipped with a manipulator to patrol agricultural fields. The robot worked in farmland, where

2 Autonomous robots for ceiling board construction

We have developed autonomous robots for ceiling board construction. Ceiling board construction is one of interior construction works and the process of fixing boards made of gypsum or other materials to steel or wooden foundations using tapping screws or staples. The



Figure 1. Ceiling construction robot.

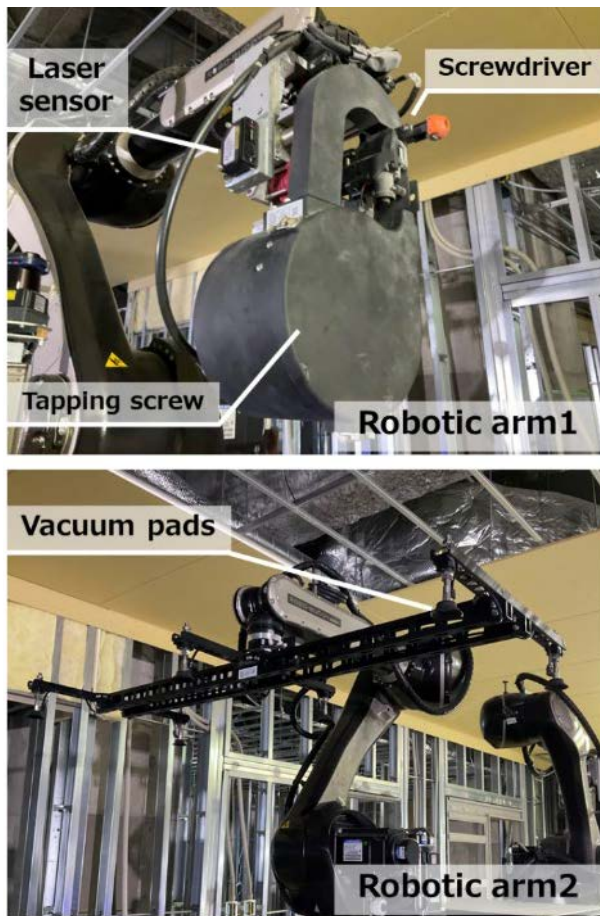


Figure 2. Robotic arms of ceiling construction robot.



Figure 3. Assist robot for ceiling construction robot.

ceiling board construction can conceal equipment materials (e.g., lightings, air conditionings, plumbing, codes) and improve thermal insulation, sound absorption, moisture proofing, and appearance. This is physically demanding work because it requires lifting and supporting a plasterboard and fixing it to the ceiling

foundation in an upward posture. In addition, this work requires some temporary scaffolds and a lot of materials. The installation and dismantling of temporary scaffolds are time-consuming, and it involves high-place work.

2.1 Specification

Figures 1 and 2 show a ceiling construction robot and its robotic arms, and figure 3 shows an assist robot for ceiling board construction. The specifications of these robots are shown in Table 1. The robotic arm1 has the end-effector equipped with a screwdriver, a tapping screw case, and a laser sensor to measure the plasterboards and the ceiling foundations. The tapping screw case can load with up to 400 tapping screws. The robotic arm2 has the end-effector equipped six suction pads for grasping the plasterboard.

Table 1 Specification of ceiling construction robot and assist robot for ceiling construction.

	Ceiling construction robot	Assist robot for ceiling construction
Dimensions [mm] L × W × H	2,400 × 800 × 2,600	2,100 × 800 × 2,100
Weight [kg]	2,160	1,115
Power supply	Battery-powered (Lithium-ion recharge battery)	
Control interface	Tablet	
Communication	Mobile or Wi-Fi	
Materials	Plasterboard	

The ceiling construction robot has three PLCs (Programmable Logic Controller). The integration PLC has control functions to communicate with the cloud server, send and receive robot status and commands, process the data obtained from the laser distance sensor and LiDAR sensor, and calculate the speed command value for wheels. The mobility PLC controls the wheels, the lifter, and the outriggers. The robotic arm PLC shares with both robotic arm controllers. These PLCs are connected to the same network and allow the ceiling construction robot to execute the commands autonomously. The assist robot for ceiling board construction can load up to 60 plasterboards and pass to the ceiling construction robot. Both robots have a lifter so they can work at height of up to 4 m. They can recognize self-location with 2D-LiDAR (Light Detection And Ranging) [9] and move autonomously to destinations; as a result, they can move without markers or magnetic tapes.

At construction sites, materials are moved frequently and the surrounding environment changes as construction progresses. Therefore, it is necessary to assume the

environment with dynamic obstacles such as workers and materials. To avoid these obstacles, the LiDAR sensor is mounted at the top of a mast which has a telescopic slider mechanism. It is retracted except when the robots move, so it does not interfere with the movement of the robotic arms. The controller can recognize self-location by comparing the map data created from the CAD drawing of the construction site and the point cloud data obtained by 2D-LiDAR. The wheels are mecanum wheels [10], which allow the robots to move in any direction. These robots are equipped with safety sensors (laser distance sensors, 2D-LiDARs, and bumper sensors) to ensure safety and allows them to slow down or stop if there is an obstacle, and they can stop all operations immediately in case of an imminent collision.

All our robots are managed by our cloud server. Our cloud server has a platform that integrates information such as robot status, site drawings, and robot travel routes. The cloud server can command the robots, monitor their status, and navigate the robots through their shortest paths. Since the robots are designed to operate in buildings under construction, we can control the robots without LAN infrastructure. By using a public mobile line, we enable the robots to communicate with the cloud server without any special preparation. And the robots can communicate with the cloud server directly to the

VPN through the mobile line and ensure security.

2.2 Construction procedure

Operators input tasks with a tablet. These tasks are sent to the cloud server via the mobile line connection and are converted into detailed commands on the basis of the maps and machine data. The detailed commands are sent to the integration PLC from the cloud server. The cloud server sequentially sends commands and the associated detailed information depending on the task. Then, the integration PLC sends move commands to the mobility PLC by referring to the self-location obtained by the 2D-LiDAR and the status. After arriving at the destination, the integration PLC sends work commands to the robotic arm PLC.

Figure 4 shows the flowchart of the sequence of robot construction and figure 5 shows the trajectories of robotic arm1. First, robotic arm1 measures the position of the plasterboard on the assist robot. Figure 5 (A) shows the sequence of measuring the position of the board supplied by the assist robot. Robotic arm1 measures the distance from its hand to the top of the stacked boards with a laser sensor at points Z_{A1} and Z_{A2} , then averages the obtained values to determine the distance of point C . Next, robotic arm1 detects the long side by moving its hand in the $-y_r$ direction on the robot's coordinate system O_r to obtain the edges X_{A1} and X_{A2} of the board. Similarly, the robotic arm1 detects the short side by moving its hand in the $-x_r$ direction to obtain the edges Y_{A1} and Y_{A2} of the board. The board coordinates system O can be obtained from the straight lines formed by connecting the edges.

Figure 5 (B) shows the sequence of measuring the position of the boards that have already been installed on the ceiling. Robotic arm1 detects the long side by moving its hand in the $-x_r$ direction on the O_r , to obtain the edges Y_{B1} and Y_{B2} of the board. Similarly, robotic arm1 detects the short side by moving its hand in the $-y_r$ direction to obtain two coordinates of the edges X_{B1} and X_{B2} of the board. Robotic arm1 measures the distance from its hand to the board at Z_{B1} to Z_{B4} in the sequence of detecting edges and averages the obtained values to determine the height of the ceiling. The long side line is calculated with X_{B1} and X_{B2} . The short side line is calculated with Y_{B1} and Y_{B2} . Then, the corner of the board and the relative coordinate system O' are calculated as the intersection of these two straight lines.

Figure 5 (C) shows the sequence of the measuring positions of the ceiling foundations. Robotic arm1 detects the edges of the metal foundations by moving in $-y'$ and $+y'$ directions on O' . The robot calculates straight lines for each foundation using the obtained P'_{n1} and P'_{n2} , and coordinates to fasten the tapping screws. These obtained coordinates are shared with both robotic arm controllers through the robotic arm PLC.

After the measurement of the board, robotic arm2

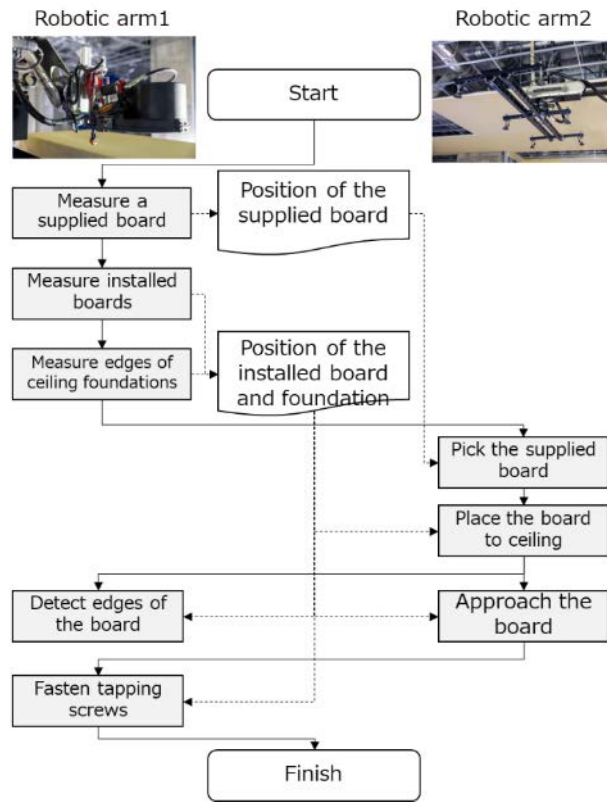


Figure 4. Flowchart of ceiling construction robot.

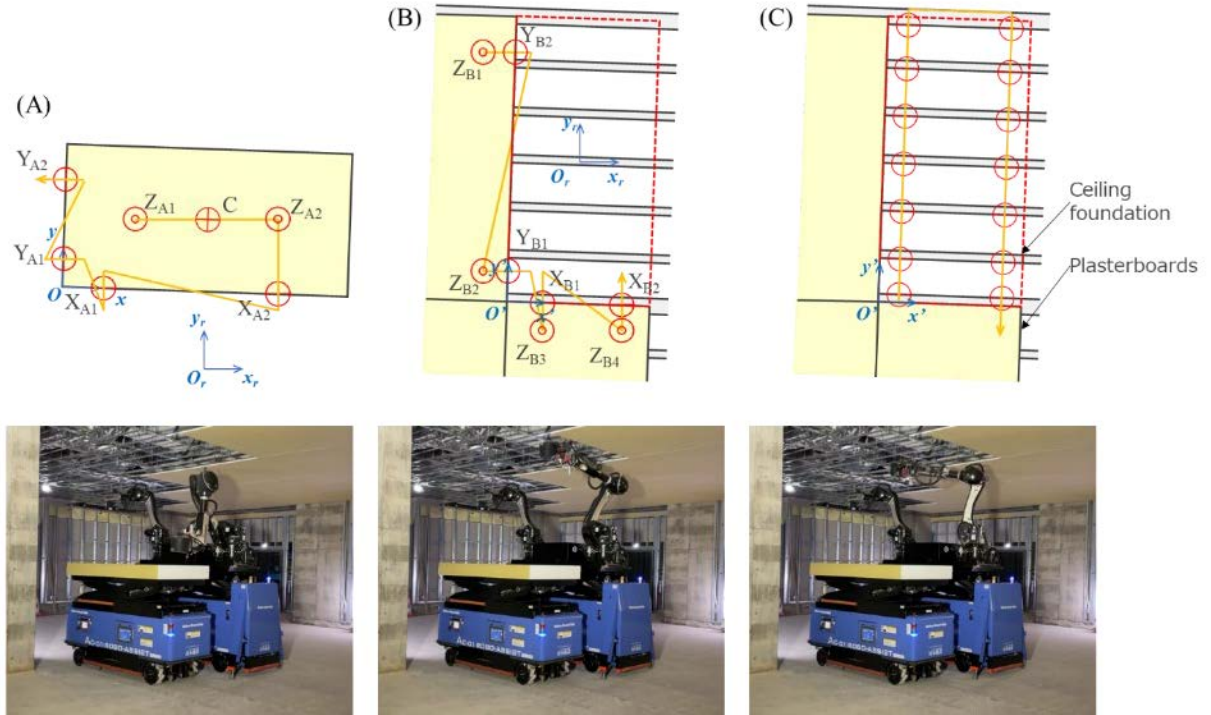


Figure 5. Trajectories of ceiling construction robot. (A) Process of measuring a board supplied by assist robot. Red circles indicate measurement points, and the yellow line represents the trajectory of the robot hand. (B) Process of measuring boards that have been installed on the ceiling and (C) Process of measuring the positions of the ceiling foundations.

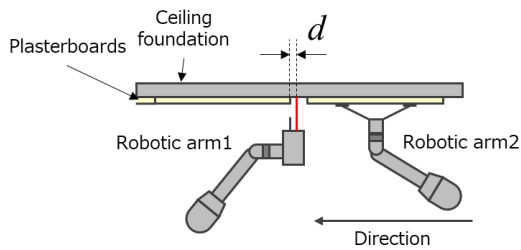


Figure 6. Board positioning process.

picks up the board on the basis of the measured position, turns it over and places it on the ceiling foundations. At this time, robotic arm2 presses the board against the adjacent one to prevent gaps in the $-x'$ and $-y'$ directions. Figure 6 shows a board positioning process. First, the robot sets observation coordinates at d away from the edge of the installed ceiling boards. Robotic arm1 moves to the observation point and starts observation with its laser sensor. Robotic arm2 continues to move in the pushing direction. If robotic arm1 detects the board, robotic arm2 moves in the pushing direction by d to adjust the board without gaps. The robots execute these processes for the long and short sides. This is an effective method for accurately positioning large objects using

only simple measurements.

After adjusting the board, robotic arm1 fixes the board with tapping screws. After it completely fixes the board, the robotic arm PLC sends a completion signal to the cloud server through the integration PLC. By repeating this flow and move commands, the robots can continuously install the ceiling boards.

2.3 Case study

We applied the ceiling board construction robot and the assist robot to a construction site in Japan. We constructed 74 boards (123 m²). Generally, a pair of workers install ceiling boards. The workers install one plasterboard in about 6 minutes, including setting up and taking down the scaffold. If the workers install 100 m² (60 plasterboards), it takes 6 hours a day. On the other hand, the robots install one plasterboard in about 8 minutes. If the robots install 100 m², it takes 8 hours a day. As long as the operator can charge the battery and supply materials, the robots can work day and night, and replace more work. The robots can be operated by a single operator and does not require special skills, thus reducing costs. In addition, the robot work does not require temporary scaffolding, eliminating the need to set up and remove scaffolding and reducing the need for

work at heights. This allows the robots to replace simple, repetitive, and hard labor, and workers can focus on tasks such as plasterboard processing and precision tasks that require accuracy. In this way, cooperative work between workers and robots can free the workers from hard labor and contribute to improved work efficiency.

3 Autonomous robots for raised floor construction

We have developed a floor construction robot to automate the raised floor construction work. It is a work to store wiring and piping by providing space on the floor slab. Raised floors offer easy access to wiring, and concealing cables prevent damage to cabling and tripping and falling hazards. Supports, generally called pedestals, are installed to adjust the height, and panels are laid on top of them. This work requires a considerable physical burden, for example, on the waist, because workers must constantly lean forward to perform the task.

3.1 Specification

Figures 7 and 8 show the floor construction robot and its robotic arms, and figure 9 shows the assist robot for the floor construction robot. The specifications of these robots are shown in Table 2. Similar to the ceiling construction robot, the floor construction robot is an autonomous mobile robot composed two robotic arms. The robotic arm1 has an end effector equipped with a camera, a laser distance sensor, a rotating laser receiver, and a gripper, which can locate floor line markings, pick up pedestals, dip into adhesive to the pedestals, and adjust the pedestal height by turning the screw. The robotic arm2 also has an end effector with a camera, a laser distance sensor, and vacuum pads, which can pick up panels. The assist robot for floor construction robot

has feeders for loading pedestals and panels. It can carry up to 36 panels and up to 88 pedestals. The upper part of the pedestal feeder has a feed mechanism to place the pedestals where the robot can pick them up so that they

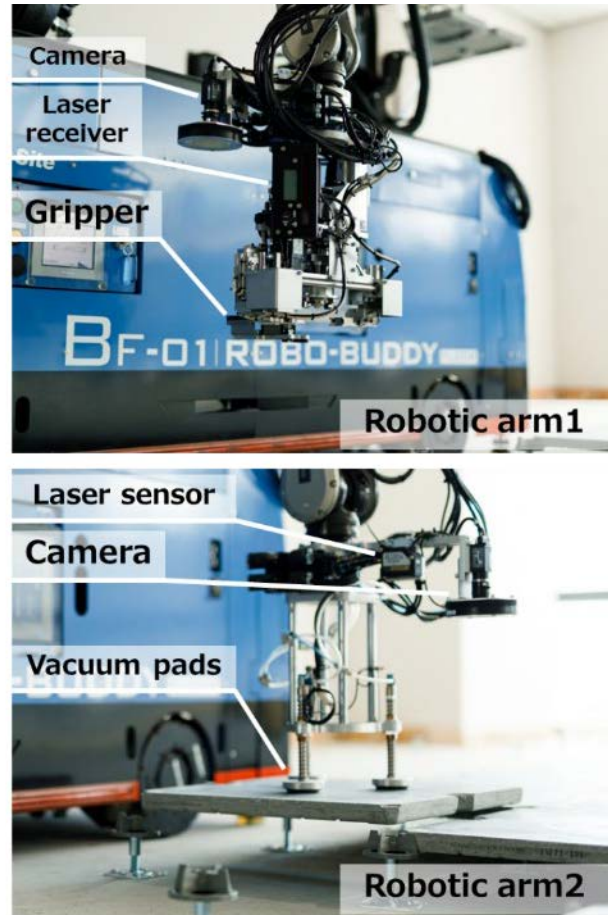


Figure 8. Robotic arms of floor construction robot.

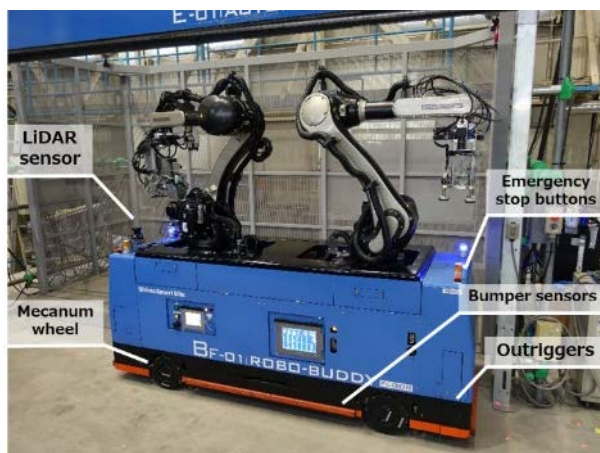


Figure 7. Floor construction robot.



Figure 9. Assist robot for floor construction robot.

are automatically supplied to the floor construction robot. In addition, a container is installed at the top to contain the adhesive to be applied to the bottom of the pedestals picked up by the robot.

Table 2 Specification of floor construction robot and assist robot for floor construction.

	Floor construction robot	Assist robot for floor construction
Dimensions [mm] L × W × H	2,500 × 800 × 1,750	2,500 × 800 × 1,700
Weight [kg]	1,460	1,125
Power supply	Battery-powered (Lithium-ion recharge battery)	
Control interface	Tablet	
Communication	Mobile or Wi-Fi	
Materials	Panels, pedestals, adhesive	

The components of the PLCs are the same as those of the ceiling construction robot and assist robot for ceiling construction robot. We have also been commonly developing other units, such as those for controlling traveling and sending and receiving commands from tablet applicable to these robots. The floor construction robot has a separate computer for processing the data obtained from the camera and laser distance sensor, and the processing result is sent to its PLC. In addition, the feeder of the assist robot for floor construction robot has

a PLC to control the device to feed the pedestals.

3.2 Construction procedure

In preparation for construction, first, perform floor marking on the floor slab for the raised floor installation. Since this floor marking is also necessary for the workers to perform their tasks, it can be done in the same work environment as the workers. Next, set up the laser level and adjust the height to the desired floor level.

Like ceiling construction robots, workers can generate work commands using a tablet and operate the robots through the cloud server. When the robot reaches the target position, it starts installing the raised floor.

Figure 10 shows a flowchart of the raised floor construction by the robot. When no pedestals are in place, only the robotic arm1 repeats the tasks of installing the pedestals. When four or more pedestals are placed, and the panels are ready to be laid, the robot performs the pedestal and panel installation simultaneously. First, the robotic arm1 takes pictures of the floor line markings and identifies the target position. Once the robotic arm1 completes taking pictures, the robotic arm 2 starts to take pictures of the floor markings. Since the robotic arms share their status and posture, they can move without interference from each other. The robotic arm1 picks up the pedestals supplied from the pedestal feeder of the assist robot. The position of the pedestals can be identified by recognizing the relative position from the marker on the feeder.

When the robotic arm1 picks up the pedestal, the robotic arm1 dips into the adhesive to bottom surface of the pedestal. A box containing adhesive is prepared on the assist robot. The adhesive is applied to the bottom surface based on information on the adhesive surface height measured by the laser distance sensor. In addition, a comb is prepared at the top of the box to remove excess adhesive.

Next, the pedestals are installed according to the floor line markings, and the pedestal height is adjusted according to the height of the laser level.

Once the pedestals are installed, the robotic arm2

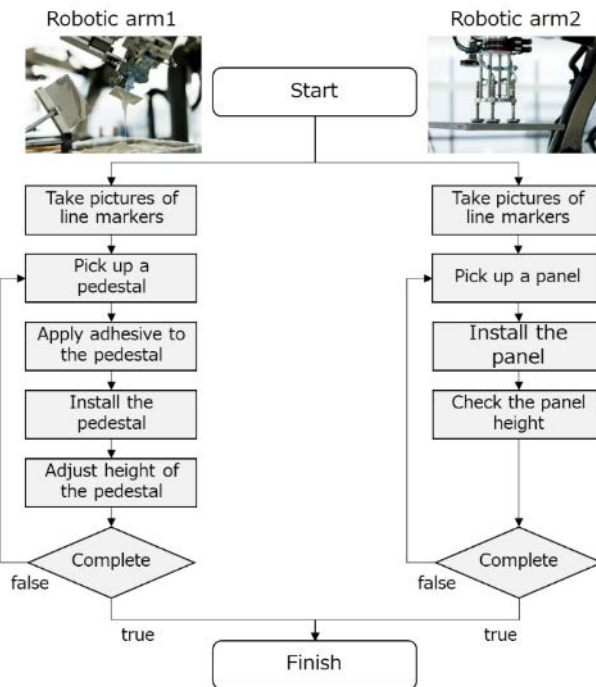


Figure 10. Flowchart of floor construction robot.

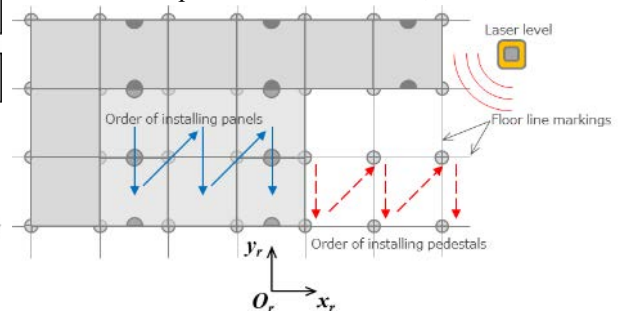


Figure 11. The order of the pedestal and panel installation

picks up the panels on the assist robot and installs the panel on the pedestals. Finally, measure the height of the panel edge installed to ensure correct installation. If not, skip the installation of the surrounding panels to avoid failure of the laying of adjacent panels. This process prevents the error propagation.

As the pedestals are installed, the target area to install the panels increases, so the robotic arms can communicate signals with each other using a PLC and work simultaneously, avoiding interference. Figure 11 shows the order of installing the pedestals and the panels when the robot works most efficiently. Each robotic arm places the panels and pedestals in the direction of each arrow. At this time, the center of the robot body is located on the center line of the panel to be installed last. Once the six panels and pedestals have been installed, the robot moves parallel in the x_r direction and starts the installation again.

3.3 Case study

The floor construction robot and the assist robot were applied to construction sites in Japan. They were applied at another site, and the area of 425 m² was covered in 10 days, excluding the perimeter of walls and pillars. The workers generally perform floor marking on the floor slab, install pedestals, and install panels step by step. Unlike the robotic construction procedure, the panels are usually installed over the entire work area after installing the pedestals instead of simultaneously laying the pedestals and panels. Therefore, it is not easy to compare work efficiency, but based on the production rates of all processes, the daily output per robot corresponds to the daily output per worker. The workers generally lay materials such as particleboard where the panel size does not fit, such as the area alongside the wall. Since these materials need to be sized according to the site dimensions, it is a labor-intensive work that requires measuring, cutting, and installing. By replacing workers with robots in simple tasks, workers can focus on these time-consuming and highly skilled tasks.

4 Conclusion

We have developed two types of autonomous robots. Ceiling construction robots can install ceiling boards automatically and floor construction robots can install the pedestals and panels automatically. These robots can move at construction sites using the self-localization system with 2D-LiDAR and safety devices. Each of the robots performed the installation at the construction sites and showed the usefulness of installation of the robots. The ceiling construction robots can perform the work of two operators in eight hours. The number of workers required is reduced to one operator. Moreover, the unit labor cost can be reduced as no special skills or physical

strength are required. The floor construction robots can perform the work of one worker in one day. As with the ceiling construction robots, this has the effect of reducing labor costs.

Construction sites involve a variety of tasks. In ceiling work construction, the works include a lot of tasks such as installing suspension bolts in the upper floor slab, constructing the ceiling foundation, fixing the plasterboards with screws, installing the finishing board, and installing the equipment. In floor construction, the works also include many tasks such as marking lines on the floor slab, installing the raised floor, laying the floor mat, and installing the cable. However, the amount of work that can be automated is small because our robots target only a single task. Therefore, it is still difficult to obtain cost advantages. Our future work is to expand the types of work that can be performed by the robots to realize more productive construction and more cost benefits.

References

- [1] Statistics Bureau of Japan Labor Force Survey, <https://www.stat.go.jp/english/data/roudou/index.html>
- [2] T. Bock, Construction Robotics, *Journal of Robotics and Mechatronics*, 28:116, 2016.
- [3] T. Bock, The future of construction automation: Technological disruption and the upcoming ubiquity of robotics, *Automation in Construction*, 59:113, 2015.
- [4] M. Daas, Toward a taxonomy of architectural robotics, *SIGRAFI 2014*, 1:623, 2018.
- [5] M. Daas and A. J. Wit, *Towards a robotic architecture*, Applied Research and Design Publishing, 2018.
- [6] D. D. Khoa Le, G. Hu, D. Liu, R. Zhao, S. Huang, P. Shrestha, and R. Belperio, The QUENDA-BOT: Autonomous Robot for Screw-Fixing Installation in Timber Building Construction, *2023 IEEE 19th CASE*, Auckland, New Zealand, 2023.
- [7] M. Satoh, A. Bhat, T. Usami, and K. Tatsumi, A Multi-Purpose Autonomous Mobile Robot as a Part of Agricultural Decision Support Systems, *2023 IEEE 19th CASE*, Auckland, New Zealand, 2023.
- [8] A. P. R. Lauer, T. Schürmann, A. Gienger and O. Sawodny, Force-Controlled On-Site Assembly Using Pose-Dependent Stiffness of Large-Scale Manipulators, *2023 IEEE 19th CASE*, Auckland, New Zealand, 2023.
- [9] Hokuyo UTM-30LX, <https://www.hokuyo-aut.jp/search/single.php?serial=169>
- [10] Nabtesco AGV Drive Unit, <https://precision.nabtesco.com/en/products/detail/RVW>

Autonomous Data Acquisition of Ground Penetrating Radar (GPR) Using LiDAR-based Mobile Robot

Amr Eldemiry^{1,2}, Muhammad Muddassir¹, and Tarek Zayed¹

¹Department of Building and Real Estate, The Hong Kong Polytechnic University, Hong Kong

²Public Works Department, Faculty of Engineering, Cairo University, Giza12613, Egypt

amr-m.eldemiry@polyu.edu.hk, muhammad.muddassir@polyu.edu.hk, tarek.zayed@polyu.edu.hk

Abstract – In this paper, we propose a ground mobile robot that can perform both surface mapping and subsurface mapping using a three-dimensional LiDAR Simultaneous Localization and Mapping System (3D LiDAR SLAM system) and a ground penetrating radar (GPR). The robot consists of a mobile platform equipped with a 3D LiDAR sensor and a GPR antenna mounted on a fixed chassis. The robot can autonomously navigate the environment and collect data from both the surface and the subsurface. The surface mapping is done by using the 3D LiDAR sensor of ± 3 cm range accuracy to observe the point cloud of the terrain, which is then processed to generate the 3D surface map. The subsurface mapping is done by using the GPR antenna to emit electromagnetic pulses into the soil and receive their reflections, which are then processed to generate a 3D subsurface map. Then, we can fuse the surface and subsurface maps to obtain a comprehensive representation of the terrain. We demonstrate the performance of our robot in real-world scenarios, such as bridges. We show that our robot can achieve high accuracy and efficiency in surface mapping tasks and GPR data acquisition.

Keywords –

Autonomous mobile robots; GPR; Surface mapping; LiDAR SLAM; subsurface mapping; 3D LiDAR

1 Introduction

Surface mapping and subsurface mapping are essential tasks for various applications, such as geology, engineering, archaeology, and environmental monitoring[1]. However, conventional surface and subsurface mapping (Hybrid Mapping) methods are often limited by their accuracy, efficiency, and accessibility. For example, surface mapping using aerial or satellite imagery may not capture the fine details of the terrain, while subsurface mapping using ground penetrating radar (GPR) may require the corresponding

surface mapping to access the required subsurface features accordingly. Therefore, there is a need for a novel method that can build hybrid mapping using a single mobile robot[2].

This paper presents a ground-mobile robot that can perform hybrid mapping using a three-dimensional LiDAR Simultaneous Localization and Mapping System (3D LiDAR SLAM system) and a GPR. The robot consists of a mobile platform equipped with a 3D LiDAR sensor, a camera, and a GPR antenna, which are mounted on a wheeled chassis. The robot can autonomously navigate the environment and collect data from both the surface and the subsurface. Surface mapping is done using the 3D LiDAR sensor to capture the point cloud of the terrain, which is then processed to generate a 2D orthophoto. The subsurface mapping is done by using the GPR antenna to emit electromagnetic pulses into the soil and receive their reflections, which are then processed to generate a 3D subsurface map. The robot can also fuse the hybrid maps to obtain a comprehensive representation of the terrain.

We demonstrate the performance of our robot in a real-world bridge scenario. We show that our robot can achieve high accuracy and efficiency in both surface and subsurface mapping tasks. We also discuss the advantages and limitations of our method compared to existing methods. Our paper contributes to the advancement of mobile robotics research by providing a novel solution for combining surface mapping and subsurface mapping using two complementary sensors.

This paper is organized as follows: the related works are introduced in section 2; section 3 explains the problem statement, the proposed robot design, the processing framework, and the experimental setup; section 4 presents the results; and finally, the conclusions are drawn in section 5.

2 Literature Review

Ground penetrating radar (GPR) is a geophysical technique that can produce high-resolution subsurface

images by measuring the backscattered electromagnetic waves from the objects in the ground. GPR has been widely used for various applications in civil engineering, such as detecting buried utilities, mapping soil layers, monitoring geothermal resources, and exploring archaeological sites. However, conventional GPR methods require manual data acquisition and processing, which are time-consuming, labour-intensive, and prone to human errors. Therefore, there is a need to develop automated and intelligent systems that can perform GPR data collection and analysis in real-time and provide accurate and reliable results[3].

One of the promising solutions for this challenge is using autonomous mobile robots (AMRs) equipped with GPR sensors. AMRs can operate autonomously or remotely in complex and hazardous environments, such as minefields, disaster zones, or inaccessible areas. To enhance their capabilities, AMRs can also carry additional sensors or devices, such as cameras and LiDARs. By integrating GPR with other sensors or devices, AMRs can perform multi-sensor data fusion and analysis to obtain more information about the subsurface features and objects.

This paper will introduce recent works[4-14] in GPR robots, demonstrating their potential and challenges in different domains. In [4], a novel robotic inspection system was proposed to combine impact-echo (IE) and ground penetrating radar (GPR) sensors to detect and map subsurface defects in concrete structures. The system leverages vision-based positioning and pose information to guide the robot to the target area and integrates learning-based and classical methods to process the IE data and reconstruct the underground objects. The system also uses GPR data processing techniques to create a 3D map of the subsurface features for better visualization. In addition, their work demonstrated the effectiveness and efficiency of the system through field testing on concrete slabs, showing that visual inspection alone can reveal shallow defects that are otherwise invisible.

Another recent work [5] was the proposed robotic system that automatically collects and analyses ground penetrating radar (GPR) data to detect and visualize underground utilities for construction surveys. The system consists of a mobile robot equipped with a GPR sensor, a localization module, and a data processing module. The localization module uses a visual-inertial odometry (VIO) algorithm to estimate the robot's pose and trajectory. The data processing module uses a convolutional neural network (CNN) to segment the GPR images and identify the underground objects. The system also employs a graph-based method to reconstruct the 3D shape and position of the underground utilities from the segmented GPR images. The authors evaluate their system on both synthetic and

real-world GPR data and demonstrate its advantages over existing methods in terms of accuracy, robustness, and usability.

In [8, 10], The main contribution of these two studies is to propose and demonstrate novel robotic systems that use non-destructive evaluation (NDE) methods, such as ground penetrating radar (GPR) and acoustic emission (AE), to inspect and evaluate the condition of bridge decks. [8] presents an autonomous robotic system that employs GPR, electrical resistivity (ER), and a camera for data collection. The system is capable of performing real-time, cost-effective bridge deck inspection and is comprised of a mechanical robot design, machine learning, and pattern recognition methods for automated steel rebar picking to provide real-time condition maps of the corrosive deck environments.

The second study [10] proposes a novel algorithm for automated rebar detection and analysis using GPR data. The algorithm integrates machine learning classification using image-based gradient features and robust curve fitting of the rebar hyperbolic signature. The approach avoids edge detection, thresholding, and template matching that require manual tuning and are known to perform poorly in the presence of noise and outliers. The detected hyperbolic signatures of rebars within the bridge deck are used to generate deterioration maps of the bridge deck. Both studies show that their robotic systems can achieve high accuracy, efficiency, and reliability in bridge deck inspection using NDE methods.

3 Problem statement and experiment

The problem this study addresses is the lack of efficient and reliable methods for combining surface and subsurface mapping (Hybrid Mapping) in complex and dynamic environments. Surface mapping is essential for understanding the topography, morphology, and features of the terrain, while subsurface mapping is essential for detecting anomalies, hazards, and resources in the soil. However, hybrid mapping methods are often limited by factors such as high cost, low accuracy, slow speed, and human intervention.

To overcome these limitations, this study proposes using an autonomous mobile robot mounted by ground penetrating radar (GPR) to perform hybrid mapping. GPR is a non-invasive technique that can penetrate various types of soil and rock layers and generate high-resolution images of the subsurface structure. The robot can also be used to map the surface features using a camera or a LiDAR sensor attached to it. By combining GPR with other sensors, such as LiDAR (light detection and ranging), this study aims to achieve high-precision

and high-efficiency hybrid mapping in challenging terrains.

The main challenges that this study faces are:

- Designing a robot that can drag the GPR antenna along predefined paths without damaging it or losing signal quality.
- Designing a LiDAR SLAM (simultaneous localization and mapping) system that can accurately locate and map the robot in three-dimensional space while avoiding obstacles and maintaining stability.
- Merging the surface and subsurface maps generated by GPR and LiDAR into a unified 3D model that can be used for further analysis or visualization.

The significance of this study lies in its potential applications for various fields such as health monitoring of structures, mining, construction, agriculture, defense, environmental monitoring, disaster management, archaeology, geology, hydrology, etc. Using an autonomous mobile robot mounted by GPR to perform hybrid mapping simultaneously; this study can provide valuable information that can enhance decision-making processes, improve operational efficiency, reduce risks, optimize resources utilization, etc.

3.1 Robot Design

The design of a ground-mobile robot for hybrid mapping is presented in this paper. As shown in Figure 1, the robot consists of a chassis, four wheels, a 3D LiDAR sensor (LSLiDAR C32), which can provide up to 150 m range with $\pm 3\text{cm}$ range accuracy and 31° vertical field of view, and a GPR antenna (GSSI 1600 MHz) which provide a depth range of 50 cm. The 3D LiDAR sensor provides high-resolution point cloud data of the surrounding environment, while the GPR antenna enables the detection and localization of buried features. The robot is powered by two batteries that can last for 4 to 6 hours, depending on the workload. The robot is equipped with a LiDAR SLAM system that processes the point clouds from the LiDAR sensor to model the surrounding surface environment. The robot can be controlled remotely or autonomously using wireless communication protocols. The robot's performance is evaluated on a real outdoor site.

The main challenge of designing such a robot is to balance the trade-off between accuracy, speed, reliability, and energy consumption. To achieve high accuracy, the robot needs to have a robust and efficient algorithm for point cloud registration and modeling. To achieve high speed, the robot needs to have a fast and lightweight hardware platform that can handle large-scale data processing. To achieve high reliability, the

robot needs to have a robust and adaptive navigation system that can cope with dynamic and uncertain environments.

The proposed solution addresses these challenges by using state-of-the-art robotics techniques. The robotics system uses LeGO-LOAM [15], a 3D-LiDAR SLAM system for surface mapping and robot localization. In addition to the A* path planner [16] as a global path planner and time-elastic-bands (teb)-local planner [17-19] for local path planning and obstacle avoidance.

The proposed solution has several advantages over existing methods for surface mapping using mobile robots. First, it can cover large areas with high resolution in real-time. Second, it can accurately detect buried features and utilities using GPR signals that penetrate different soil layers depending on their electrical conductivity.

In conclusion, this paper presents an innovative design of a ground mobile robot that a 3D LiDAR sensor and GPR antenna for hybrid mapping applications can be mounted and work for at least four hours.

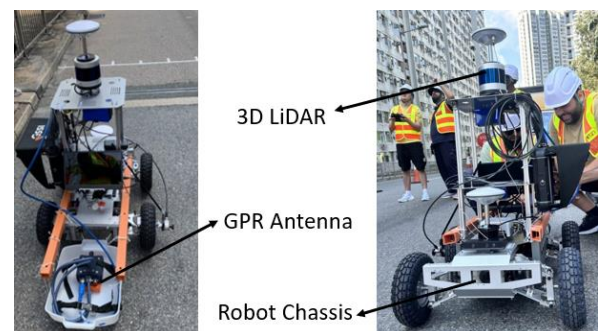


Figure 1. The main components of the robot are the chassis, 3D-LiDAR sensor, and GPR antenna.

The robot's battery system consists of four batteries: two batteries are for the robot and the other two are for the 3D-LiDAR sensor. The robot uses one battery at a time, while the other serves as a backup in case of power failure. Similarly, the 3D-LiDAR sensor switches between two batteries to ensure continuous operation. The battery system is designed to provide sufficient power for the robot and the sensor to perform at least 4 hours of continuous working.

3.2 Processing Framework

The process diagram in Figure 2 illustrates the steps involved in mapping the surface and subsurface of an environment using a robot equipped with a 3D-LiDAR sensor, a ground-penetrating radar (GPR), and a distance measurement indicator (DMI). The first step is

data acquisition, where the robot collects 3D-LiDAR and GPR data along with DMI data that indicates the distance travelled by the robot. The second step is data processing, where the 3D-LiDAR data is used to perform simultaneous localization and mapping (SLAM) to estimate the robot's pose and generate a point cloud representation of the surface. The GPR data is interpreted to produce a subsurface map that shows the location and depth of buried objects.

The DMI data is used to align the GPR data with the robot's trajectory. The third step is processing output, where the point cloud is converted into a surface map showing the surface's shape and texture. The robot's trajectory is also displayed on the surface map. The subsurface map is overlaid on the surface map using the DMI data as a reference. The final step is integration, where the surface and subsurface maps are combined to provide a comprehensive view of the environment.

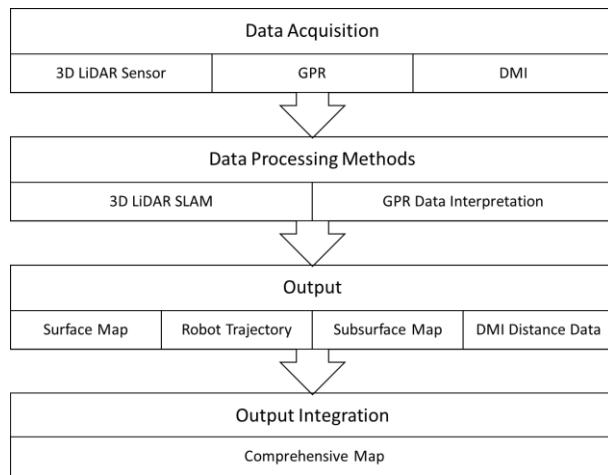


Figure 2. Process diagram of The GPR robot.

3.3 Experiment

The case study was conducted on a round 80-meter section of a bridge in Hong Kong, as shown in Figure 3, that was suspected of severe corrosion damage on its steel bars. The bridge was selected based on its age, traffic volume, environmental conditions, and structural design. The bridge was closed for inspection and preparation before the start of the case study. The GPR robot was deployed on both sides of the bridge at different locations to collect data from different depths and directions. The data were then transmitted back to a central server, where they were processed and analyzed.



Figure 3. Case study: corrosion inspection of a bridge in Hong Kong.

4 Results

The surface results of the GPR robot using the LiDAR SLAM method are presented in this section. The surface results show that the GPR robot successfully mapped the bridge deck features. The 3D point cloud of the bridge, as shown in Figures 4 to 6, was created using the point cloud data obtained from the 3D LiDAR sensor mounted on the GPR robot. The point cloud data were processed using the LiDAR SLAM method, which is a technique that simultaneously localizes the robot and maps the environment using the LiDAR measurements. The LiDAR SLAM method consists of three main steps: scan matching, loop closure detection, and pose graph optimization. Scan matching is the process of aligning consecutive scans to estimate the relative motion of the robot. Loop closure detection is the process of identifying previously visited places and correcting the accumulated drift. Pose graph optimization is the process of refining the robot poses and the map by minimizing the error between the measurements and the estimated state. The sharpness of the final point cloud of the bridge is shown in Figure 7.

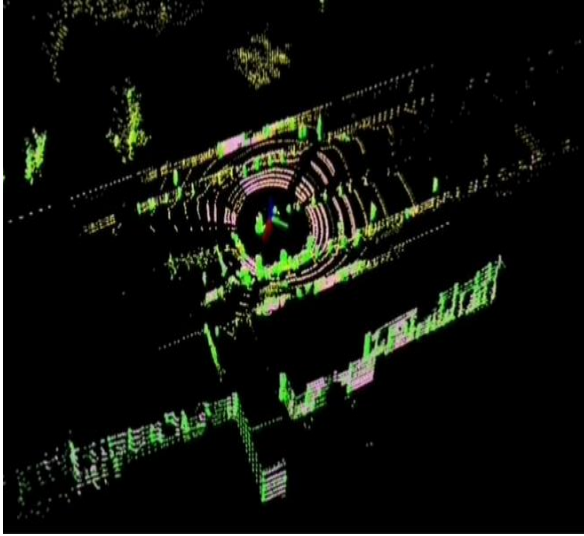


Figure 4. The start of the bridge mapping using the LiDAR SLAM system.

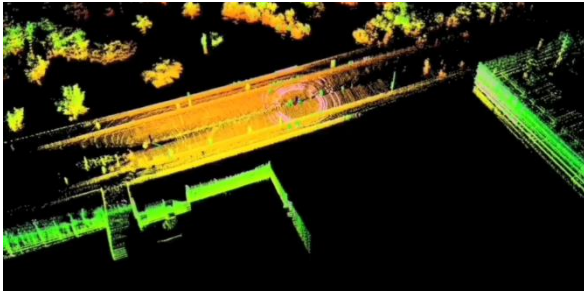


Figure 5. The middle of the bridge mapping using the LiDAR SLAM system

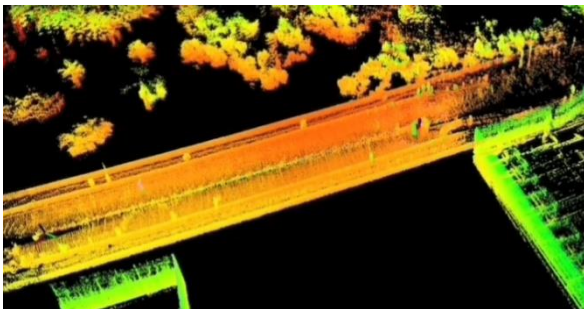


Figure 6. The end of the bridge mapping using the LiDAR SLAM system

One of the advantages of the proposed system is that it can fuse the GPR data with the 3D point cloud to create a comprehensive map of the indoor environment. The collected GPR data can be merged with the final 3D point cloud using the robot trajectory, as shown in Figure 7, and the distance measured indicator (DMI) data, which is connected to the GPR antenna to trigger the GPR data acquisition according to the distance. The DMI data provides the precise location of each GPR

scan along the robot path, which can be aligned with the corresponding 3D point cloud frame. The merged GPR and 3D point cloud data can provide rich information about indoor and outdoor environments' surface and subsurface features, such as pipes, wires, and defects.

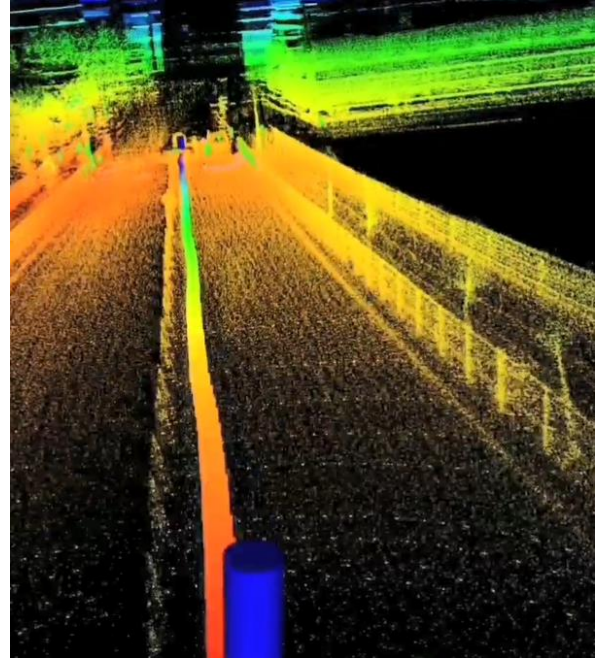


Figure 7. Zooming view of the bridge mapping using the LiDAR SLAM system.

5 Conclusion

In this paper, we presented a novel approach for surface mapping using an autonomous mobile robot equipped with a ground penetrating radar (GPR) antenna. The objective of our study was to evaluate the feasibility and performance of this method for detecting buried utilities and creating a 3D subsurface map. We designed and built a GPR system that can be mounted on a specially designed trailer that can be towed by the robot. We applied our method to an experimental site with an 80-meter-long bridge section, where we collected data using GPR and a 3D LiDAR sensor. Our method also demonstrated its potential for long-term mapping in outdoor and indoor environments where GPS signals are unavailable or unreliable. Our work opens up new possibilities for using GPR as a tool for robot navigation applications in various domains such as mining, construction, archaeology, and environmental monitoring limited by the environment accessibility to enable the ground wheeled robot for navigation.

Acknowledgement

The authors gratefully acknowledge the support from the Smart Traffic Fund (STF) under grant number PSRI/14/2109/RA.

References

- [1] Kouros, G., Psarras, C., Kostavelis, I., Giakoumis, D. and Tzouvaras, D., *Surface/subsurface mapping with an integrated rover-GPR system: A simulation approach*. in *IEEE International Conference on Simulation, Modeling, and Programming for Autonomous Robots (SIMPAR)*, pp. 15-22. IEEE, 2018.
- [2] Ékes, C. and B. Neduczka. *Robot mounted GPR for pipe inspection*. in *14th International Conference on Ground Penetrating Radar (GPR)*, pp. 160-164. IEEE, 2012.
- [3] Herman, H., *Robotic subsurface mapping using ground penetrating radar*. Doctoral dissertation, Carnegie Mellon University, 1997.
- [4] Hoxha, E., Feng, J., Sanakov, D. and Xiao, J., *Robotic Inspection and Subsurface Defect Mapping Using Impact-echo and Ground Penetrating Radar*. IEEE Robotics and Automation Letters, 2023.
- [5] Feng, J., Yang, L., Hoxha, E., Jiang, B. and Xiao, J., *Robotic inspection of underground Utilities for Construction Survey Using a ground penetrating radar*. Journal of Computing in Civil Engineering, 37(1): pp. 04022049, 2023.
- [6] Li, H., Chou, C., Fan, L., Li, B., Wang, D. and Song, D., *Robotic subsurface pipeline mapping with a ground-penetrating radar and a camera*. in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3145-3150. IEEE, 2018.
- [7] Kouros, G., Kostavelis, I., Skartados, E., Giakoumis, D., Tzouvaras, D., Simi, A. and Manacorda, G., *3d underground mapping with a mobile robot and a gpr antenna*. in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3218-3224. IEEE, 2018.
- [8] Le, T., Gibb, S., Pham, N., La, H.M., Falk, L. and Berendsen, T., *Autonomous robotic system using non-destructive evaluation methods for bridge deck inspection*. in *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3672-3677. IEEE, 2017.
- [9] Gibb, S., Le, T., La, H.M., Schmid, R. and Berendsen, T., *A multi-functional inspection robot for civil infrastructure evaluation and maintenance*. in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2672-2677. IEEE, 2017.
- [10] Kaur, P., Dana, K.J., Romero, F.A. and Gucunski, N., *Automated GPR rebar analysis for robotic bridge deck evaluation*. IEEE transactions on cybernetics, 46(10): pp. 2265-2276, 2015.
- [11] Lever, J.H., Delaney, A.J., Ray, L.E., Trautmann, E., Barna, L.A. and Burzynski, A.M., *Autonomous gpr surveys using the polar rover yeti*. Journal of Field Robotics, 30(2): p. 194-215, 2013.
- [12] Williams, R.M., L.E. Ray, and J.H. Lever. *Autonomous robotic ground penetrating radar surveys of ice sheets; using machine learning to identify hidden crevasses*. in *IEEE International Conference on Imaging Systems and Techniques Proceedings*, pp. 7-12. IEEE, 2012.
- [13] Williams, R.M., L.E. Ray, and J. Lever. *An autonomous robotic platform for ground penetrating radar surveys*. in *IEEE International Geoscience and Remote Sensing Symposium*, pp. 3174-3177. IEEE, 2012.
- [14] Trautmann, E., L. Ray, and J. Lever. *Development of an autonomous robot for ground penetrating radar surveys of polar ice*. in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1685-1690. IEEE, 2009.
- [15] Shan, T. and B. Englot. *Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain*. in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4758-4765. IEEE, 2018.
- [16] Khatib, O., *Real-time obstacle avoidance for manipulators and mobile robots*, in *Autonomous robot vehicles*, Springer. p. 396-404, 1986.
- [17] Rösmann, C., F. Hoffmann, and T. Bertram. *Planning of multiple robot trajectories in distinctive topologies*. in *European Conference on Mobile Robots (ECMR)*, pp. 1-6. IEEE, 2015.
- [18] Rösmann, C., Feiten, W., Wösch, T., Hoffmann, F. and Bertram, T., *Efficient trajectory optimization using a sparse model*. in *European Conference on Mobile Robots*, pp. 138-143. IEEE, 2013.
- [19] Rösmann, C., Feiten, W., Wösch, T., Hoffmann, F. and Bertram, T., *Trajectory modification considering dynamic constraints of autonomous robots*. in *ROBOTIK; 7th*

German Conference on Robotics, pp. 1-6. VDE,
2012.

Double-Pendulum Tower Crane Trolley Trajectory Planning: A Parameter Based Time-Polynomial Optimization

Shichen Sun^{1,2}, Juhao Su^{1,2}, Yu Hin Ng², Ching-Wei Chang²

¹Academy of Interdisciplinary Studies, The Hong Kong University of Science and Technology, Hong Kong S.A.R.

²Hong Kong Center for Construction Robotics, Hong Kong S.A.R.

ssunbb@connect.ust.hk, jsuan@connect.ust.hk, sampsonng@ust.hk, ccw@ust.hk

Abstract -

Modeling the tower crane as a double pendulum dynamic system introduces complexities in control considerations. To tackle this challenge, this paper presents a time-polynomial-based trajectory generation method. This method enables the reconstruction of direct commands and employs high-order fitting to align with various control constraints. The differential solutions of the swing angles, obtained from the linearized dynamic equations, can be minimized as the trolley completes its movement. Additionally, the trajectory is optimized based on time considerations to ensure the most efficient path while adhering to the safety limitations of the tower crane. With the proposed method, the trajectory curve of the trolley is a high-order polynomial with all the coefficients related to the system parameters, which makes the trajectory applicable against the change in the system parameters. Based on the trolley actuator output, the function of the swing angles could be derived as the feedback reference line to make the proposed control method robust against external disturbance. The efficacy of the proposed method is validated through real-scale simulations and compared to existing approaches, including linear quadratic regulator (LQR) and another published CTP method, demonstrating its good control performance.

Keywords -

Tower crane; Double pendulum; Time-based polynomial; Trajectory planning

1 Introduction

With the ongoing transformation and advancement of the construction industry, the utilization of Modular Integrated Construction (MiC) is steadily growing. MiC's factory-based prefabrication has streamlined on-site construction processes, enhanced construction efficiency and promoted standardized construction practices, thereby ensuring superior construction quality [1]. However, the installation sequence of prefabricated modules is predetermined [2], and their size and weight bring heightened transportation requirements at the construction site, with its expenses remaining consistently elevated [3].

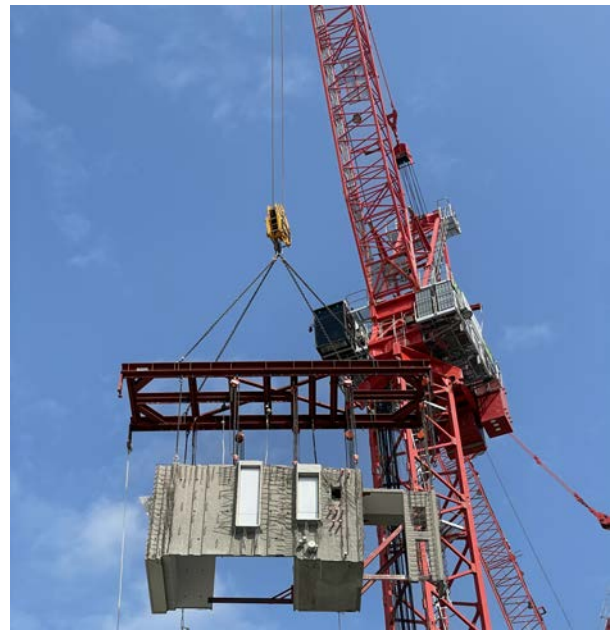


Figure 1. MiC and tower crane.

Tower cranes are widely used in construction sites for transporting large-sized objects, whose slight vibrations could be capable of causing significant accidents, thereby emphasizing the importance of advanced control methods. To achieve more precise control in such scenarios, it is necessary to establish an elaborate numerical model that represents a double pendulum structure. However, the complexity of the double pendulum structure poses greater demands on the design of the controller [4].

Over the past decades, numerous control theories have been proposed to regulate tower crane transportation and reduce angle swing. Feedback control systems, such as model predictive control [5], [6], sliding mode control [4], [7], [8], adaptive control [9], [10], neural network control [11], [12], and neural network control have shown effectiveness in mitigating external disturbances. However, the double pendulum tower crane system could be approximately modeled as the combination of two unidirectional systems, characterized as two single input multiple output

models. However, it is challenging to achieve coupling control effects across multiple outputs. Especially in real-scale models, the magnitude of displacement command changes is significantly greater than that of angle changes. Since the feedback compensations are calculated based on the differences between the desired states and actual states, the proportions of compensation are supposed to be small for positioning and large for angles. However, when the position of the trolley is close to the desired position, the small gain for the trolley error makes the positioning difficult.

Comparatively, open-loop control systems, including techniques like smooth shaping [13], input shaping [14], and trajectory planning [15], may exhibit lower robustness to external disturbances. However, these control methods consider the motor's output limitations and aim to achieve the desired state through relatively smooth control signals. By incorporating the output constraints of the motor, these control systems aim to ensure the system operates within the allowed limits and reaches the desired state in a controlled manner. In addition, open-loop control methods are not dependent on feedback signals from sensors, making them immune to delays in signal reception. This is particularly advantageous for systems with multiple control considerations. The primary concern only lies in the complexity of calculating the output based on the physical model.

Recently, Li et.al [16] published a time-polynomial-based trajectory planning method for double pendulum tower cranes. After mathematical transformation, the coefficients of the trajectory polynomial are in constant form according to a scaled-down model. However, the calculation basis is the linearized equations, which would cause the superposition of error. Besides, the trajectory proposed in the study is a polynomial curve with all the coefficients as constant. Besides, the simulation only contains the situations in which the lengths and masses are little changed. The proposed polynomial trajectory is not applicable to control the real-scale system.

While previous research on tower crane control has been discussed, there are still some limitations that need to be addressed: (a) Many existing models in the literature fail to address the control of real-size tower crane double pendulums adequately. This is primarily due to the difficulty in achieving a balance with single input feedback compensation control for larger-sized models. (b) The approximate derivation of mathematics from linearized equations is less accurate and may not be applicable when there are significant changes in the system characteristics.

This article proposes a tower crane control method based on polynomial solving of differential equations. Firstly, the linearized differential equation system for controlling the swing angle of the tower crane is derived based

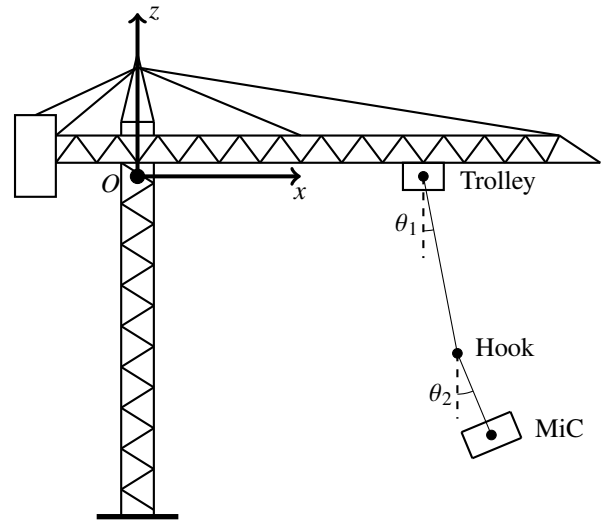


Figure 2. Depicted diagram for the model coordinate

on the physical model. Subsequently, the mathematical characteristics of the differential equation system are utilized to directly fit a polynomial of the corresponding order to reduce the errors, obtaining an appropriate motor output curve to achieve the desired control objective. Finally, through a true-size simulation model that contains saturation on trolley acceleration and velocity, the control effects of the proposed scheme on key elements are compared with those of other traditional controllers. Therefore, the following advantages are summarized:

- (a) The controller exhibits good control performance for the desired endpoint state, resulting in minimal simulation error.
- (b) This control method reduces the influence caused by the disparity in control weights for single-input multiple-output systems and achieves synchronous convergence control on multiple modules towards the endpoint state with the shortest time usage.
- (c) This control method imposes fewer performance requirements on the motor, ensuring smooth changes in jerk, acceleration, and speed throughout the entire process, which facilitates motor control and tracking.

2 Methodology

In this research, the trajectory control derivation is started from a single-degree-of-freedom (1 DOF) tower crane. With all the objects given coordinates, the dynamics model is derived and linearized, and the trajectory of the trolley is calculated to fulfill all the control constraints.

Table 1. Tower crane parameters

Parameter	Parameter explanation	Units
M_1	Mass of the hook	kg
M_2	Mass of the payload	kg
L_1	Length of the hoisting rope	m
L_2	Length of the rigging cable	m
R	The position of the trolley	m
g	Gravity constant	m/s^2
θ_1	Hook swing angle	rad
θ_2	Payload swing angle	rad

2.1 Dynamic modeling

In Figure 2, the depicted diagram showcases the ideal setup of a 1 DOF tower crane and its associated mathematical coordinate system. Table 1 provides an overview of the pertinent parameters that define the tower crane model. The Lagrangian method is used to derive the dynamics equation from the tower crane model. As shown in Figure 2, the position vectors of the two objects are:

$$d_1 = (R + L_1 \sin \theta_1) i_x - (L_1 \cos \theta_1) i_z; \quad (1)$$

$$d_2 = (R + L_1 \sin \theta_1 + L_2 \sin \theta_2) i_x - (L_1 \cos \theta_1 + L_2 \cos \theta_2) i_z. \quad (2)$$

The Lagrangian value (L) could be derived by the difference between kinetic energy (K_E) and potential energy (P_E), where K_E and P_E are computed as follows:

$$K_E = \frac{1}{2} M_1 < \dot{d}_1, \dot{d}_1 > + \frac{1}{2} M_2 < \dot{d}_2, \dot{d}_2 >; \quad (3)$$

$$P_E = -M_1 g L_1 \cos \theta_1 - M_2 g (L_1 \cos \theta_1 + L_2 \cos \theta_2). \quad (4)$$

The Lagrangian equations based on the two swing angles are:

$$\frac{d}{dt} \left(\frac{\delta L}{\delta \dot{\theta}_1} \right) - \frac{\delta L}{\delta \theta_1} = 0; \quad (5)$$

$$\frac{d}{dt} \left(\frac{\delta L}{\delta \dot{\theta}_2} \right) - \frac{\delta L}{\delta \theta_2} = 0. \quad (6)$$

The simplified equations are:

$$0 = L_1 (M_1 + M_2) \ddot{\theta}_1 + (M_1 + M_2) \cos \theta_1 \ddot{R} - (M_1 + M_2) g \sin \theta_1 + L_2 M_2 \sin(\theta_1 - \theta_2) \dot{\theta}_2^2 + L_2 M_2 \cos(\theta_1 - \theta_2) \ddot{\theta}_2; \quad (7)$$

$$0 = L_2 \ddot{\theta}_2 + \cos \theta_2 \ddot{R} - g \sin \theta_2 - L_1 \sin(\theta_1 - \theta_2) \dot{\theta}_1^2 + L_1 \cos(\theta_1 - \theta_2) \ddot{\theta}_1. \quad (8)$$

The linearized dynamics equations are:

$$0 = L_1 (M_1 + M_2) \ddot{\theta}_1 + (M_1 + M_2) \ddot{R} - (M_1 + M_2) g \theta_1 + L_2 M_2 \ddot{\theta}_2; \quad (9)$$

$$0 = L_2 \ddot{\theta}_2 + \ddot{R} - g \theta_2 + L_1 \ddot{\theta}_1. \quad (10)$$

2.2 Control objective

When following the command of a movement, the trolley is proposed to be controlled to the desired position. To derive the trajectory of the trolley, the initial and final target state of the trolley should satisfy the following constraints related to its jerks, accelerations, velocities, and positions:

$$R(0) = R_0; \dot{R}(0) = 0; \ddot{R}(0) = 0; \ddot{R}(0) = 0. \quad (11)$$

$$R(T_f) = R_d; \dot{R}(T_f) = 0; \ddot{R}(T_f) = 0; \ddot{R}(T_f) = 0. \quad (12)$$

where T_f is the transportation time, and R_d is the target position.

In addition, the swing angles at the end state are proposed to be controlled. The target state of the orientation of the hook and MiC should be as follows:

$$\theta_1(T_f) = 0; \dot{\theta}_1(T_f) = 0; \theta_2(T_f) = 0; \dot{\theta}_2(T_f) = 0. \quad (13)$$

Since the motor output has its physical limitations, during the motion, the jerk, acceleration, velocity, and position are limited under the safety values:

$$|\ddot{R}(t)| \leq J_{max}; |\ddot{R}(t)| \leq A_{max}; |\dot{R}(t)| \leq V_{max}. \quad (14)$$

The two angles are also limited to safety values:

$$|\theta_1(t)| \leq \theta_{1max}; |\theta_2(t)| \leq \theta_{2max}. \quad (15)$$

2.3 Polynomial planning

The initial orientation and its derivative of the hook and MiC are listed in equations (16). The trajectory of the system is formulated as a time-polynomial with 16 coefficients as illustrated in (17), such that it has a unique solution to the 16 equations of constraint mentioned in (11), (12), (13), and (16).

$$\theta_1(0) = 0; \dot{\theta}_1(0) = 0; \theta_2(0) = 0; \dot{\theta}_2(0) = 0. \quad (16)$$

$$R(t) = \sum_{i=0}^{15} k_{ri} \cdot \tau^i \cdot (R_d - R_0) + R_0. \quad (17)$$

where $k_{r0}, k_{r1}, \dots, k_{r15}$ are the coefficients to be derived by the control requirements. $\tau = t/t_f$ to find the most efficient transportation time.

Substituting (11) into equation (17), we get:

$$k_{r0} = k_{r1} = k_{r2} = k_{r3} = 0. \quad (18)$$

For the identities (9) and (10) to hold for any time t , the coefficients of the same order of t should sum up to zero for all t orders. According to the set-up form (17), the second derivative of R has up to 13_{th} order of t , thus θ_1

and θ_2 should be formulated as time-polynomials of 13^{th} order as well as illustrated in (19) and (20), such that they can act as counter terms to eliminate the 13^{th} order terms from \ddot{R} .

$$\theta_1(t) = \sum_{i=0}^{13} \alpha_{ri} \cdot \tau^i, \quad (19)$$

$$\theta_2(t) = \sum_{i=0}^{13} \beta_{ri} \cdot \tau^i, \quad (20)$$

where $\alpha_0, \alpha_1, \dots, \alpha_{13}$ and $\beta_0, \beta_1, \dots, \beta_{13}$ are the coefficients to be derived from (9) and (10).

By comparing the coefficient of the 0^{th} to 13^{th} order terms in the identities (9) and (10), 14 equations are obtained from each identity, and a total of 28 equations are set for the 28 variables α and β . By rearranging the equations in terms of α and β , the 28 equations can be expressed in matrix forms as follows:

$$A \cdot [\alpha_0, \alpha_1, \dots, \alpha_{13}, \beta_0, \beta_1, \dots, \beta_{13}]^T - C = 0. \quad (21)$$

where $A \in \mathbb{R}^{28 \times 28}$ is the coefficients matrix of α and β , and $C \in \mathbb{R}^{28 \times 1}$ is the constant terms from the equations.

From equation (21), the vector of variables α and β could be derived by the inverse of matrix A multiplies matrix C . $\theta_1(t)$ and $\theta_2(t)$ could thus be expressed only by the known parameters and the function $R(t)$.

After obtaining the expression of $\theta_1(t)$ and $\theta_2(t)$, 12 equations can be obtained by substituting the expressions (17) and (18) into the constraints (12), (13) and (16). The 12 equations can be expressed in matrix form as follows:

$$D(T_f, M_1, M_2, L_1, L_2, g) \cdot K = F. \quad (22)$$

Where $D \in \mathbb{R}^{12 \times 12}$ is the coefficients matrix for k_r , K is the vector of k_{r4} to k_{r15} , and $F = [1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]^T$.

Similarly, matrix K could be derived by the inverse of matrix D multiplies F , such that $k_{r4}, k_{r5}, \dots, k_{r15}$ are expressed only by known constant parameters and the variable T_f . The derived trajectory function $R(t)$ is simplified as:

$$R(t) = \sum_{i=4}^{15} K_{ri} \cdot \tau^i \cdot (R_d - R_0) + R_0. \quad (23)$$

where:

$$K_{r4} = \frac{10810800M_1L_1L_2}{(M_1+M_2)g^2T_f^4}; K_{r5} = -\frac{121080960M_1L_1L_2}{(M_1+M_2)g^2T_f^4};$$

$$K_{r6} = \frac{544864320M_1L_1L_2}{(M_1+M_2)g^2T_f^4} - \frac{360360(L_1+L_2)}{gT_f^2};$$

$$K_{r7} = -\frac{1297296000M_1L_1L_2}{(M_1+M_2)g^2T_f^4} + \frac{2882880(L_1+L_2)}{gT_f^2};$$

$$K_{r8} = 6435 + \frac{1783782000M_1L_1L_2}{(M_1+M_2)g^2T_f^4} - \frac{9729720(L_1+L_2)}{gT_f^2};$$

$$K_{r9} = -40040 - \frac{1427025600M_1L_1L_2}{(M_1+M_2)g^2T_f^4} + \frac{18018000(L_1+L_2)}{gT_f^2};$$

$$K_{r10} = 108108 + \frac{618377760M_1L_1L_2}{(M_1+M_2)g^2T_f^4} - \frac{19819800(L_1+L_2)}{gT_f^2};$$

$$K_{r11} = -163800 - \frac{112432320M_1L_1L_2}{(M_1+M_2)g^2T_f^4} + \frac{12972960(L_1+L_2)}{gT_f^2};$$

$$K_{r12} = 150150 - \frac{4684680(L_1+L_2)}{gT_f^2};$$

$$K_{r13} = -83160 + \frac{720720(L_1+L_2)}{gT_f^2};$$

$$K_{r14} = 25740; K_{r15} = -3432.$$

2.4 Time optimization

The minimum T_f satisfying (14) and (15) is found by the bisection method which is similar to [15]. After T_f is found, all unknowns except the time variable (t) in the trajectory function (23) are solved, and the trajectory of the trolley is fully determined while satisfying all constraints and safety limits mentioned. The polynomial function of the swing angles could also be derived.

3 Simulation

This section is conducted using Simulink in MATLAB. The section comprises two parts aimed at verifying the effectiveness of the proposed method. Section 3.1 demonstrates the accuracy of the proposed method under several situations. Section 3.2 shows the robustness of the proposed method against external disturbance. The below simulation is run with the gravitational constant $g = -9.81m/s^2$ and a control frequency of 20Hz. The system parameters and the constraints are set as in Table 2.

Table 2. Simulation setup values

Parameter	Parameter explanation	Values
M_1	Mass of the hook	200kg
M_2	Mass of the payload	1000kg
L_1	Length of the hoisting rope	30m
L_2	Length of the rigging cable	5m
J_{max}	Upper limit of jerk	1m/s ³
A_{max}	Upper limit of acceleration	0.1m/s ²
V_{max}	Upper limit of velocity	0.5m/s
θ_{1max}	The upper limit of θ_1	1deg
θ_{2max}	The upper limit of θ_2	1deg

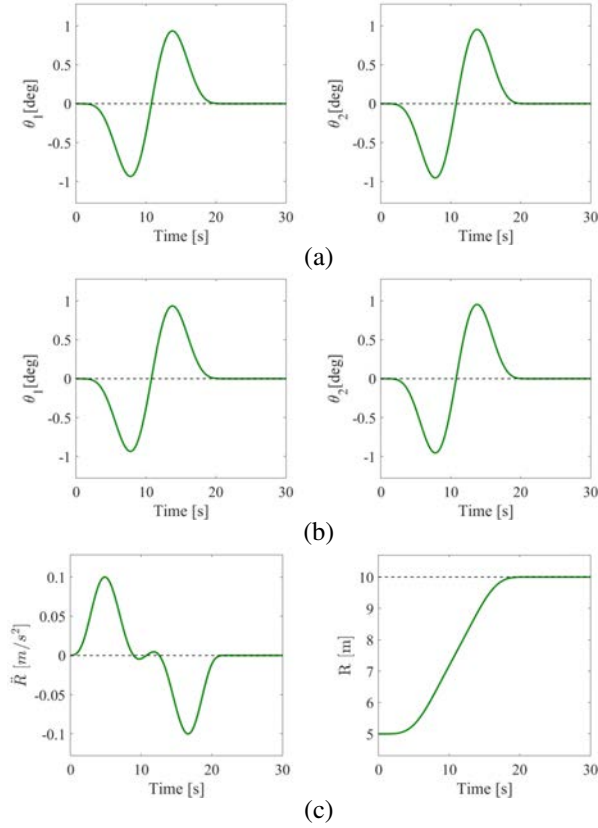


Figure 3. (a) Simulation results for the swing angle of hook (θ_1) and the payload swing angle (θ_2); (b) Theoretical values of the hook swing angle (θ_1) and the payload swing angle (θ_2); (c) The target acceleration and position of the trolley.

3.1 Accuracy verification

To prove the accuracy of the polynomial functions for trolley and swing angles, a movement command is set that the trolley moves from the position 5m to 10m. The corresponding simulation results are shown in Figure 3.

Based on the figure, it can be observed that the proposed method could execute the position-changing command. The swing angles produced by the simulation environment and the ones calculated from the polynomial functions are quite similar.

3.2 Comparison controller simulation

The proposed method is also tested in the simulation environment with external disturbance. The polynomial function of the trolley and swing angles are used as feedback reference lines, and the linear quadratic regulator (LQR) is used to compensate the system errors. It is compared to the same LQR, which follows the reference lines that the angles are supposed to be zero, and composite trajectory planning (CTP) [17] methods to show the per-

formance. The below part shows the simple derivation of LQR and the output of CTP.

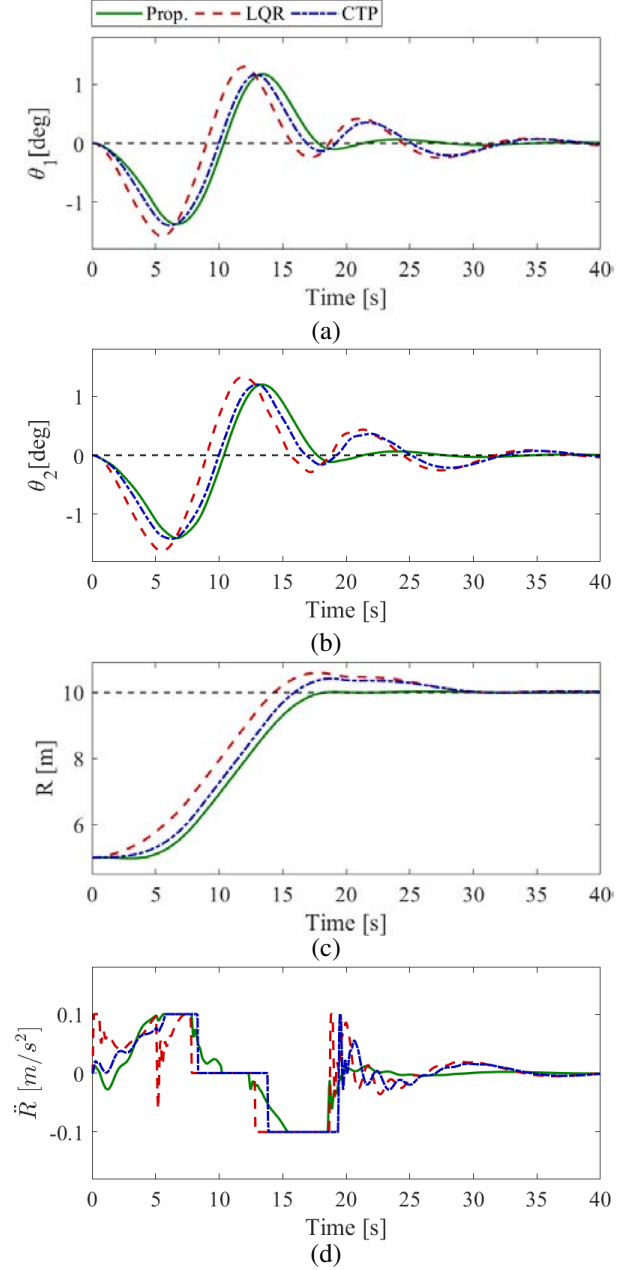


Figure 4. Simulation results for (a) the hook swing angle (θ_1), (b) the payload swing angle (θ_2), (c) the trolley position R , and (d) the trolley acceleration \ddot{R} of different controllers.

Based on (9) and (10), the linearized expressions of the two angles could be derived. The state matrix of LQR is thus:

$$s = [\theta_1; \dot{\theta}_1; \theta_2; \dot{\theta}_2; R; \dot{R}]; \quad (24)$$

The cost function is:

$$J = \int_0^\infty ((s - s_{des})^T Q (s - s_{des}) + W u^2). \quad (25)$$

Where $u = \ddot{R}$, $Q = \text{diag}[240, 48000, 40, 8000, 5, 100]$ and $W = 1$. The gain matrix of K is calculated by MATLAB to be:

$$K = [-414.85, -21.95, 279.72, 17.89, 2.24, 12.89].$$

An existing CTP controller gives the output of the trolley as [17]:

$$\ddot{R} = \begin{cases} 2\pi(R_d - R_0)(\sin(2\pi t/T_s))/T_s^2 \\ +k_r((M_1 + M_2)L_1\dot{\theta}_1 + M_2L_2\dot{\theta}_2), & 0 < t < T_s. \\ +k_r((M_1 + M_2)L_1\dot{\theta}_1 + M_2L_2\dot{\theta}_2), & t \geq T_s. \end{cases}$$

Where T_s is calculated to be 17.7s from the acceleration limitation.

The proposed method will use the above LQR controller to follow the polynomial lines. The LQR comparison will follow the bang-coast-bang acceleration command and zero angle line. The feedback loop of the CTP method is also enhanced with the LQR controller above. To make the acceleration output accessible for the actuator, the simulation is run with saturation of trolley acceleration and velocity.

As shown in Figure 4, the corresponding simulation results under external disturbance are presented for comparison, which contains the swing angles, trolley position, and the corresponding accelerations. In addition, the key control indicators are listed in Table 3, where T_f is the time after the trolley position error is less than 0.02m, and E_{θ_1} and E_{θ_2} are the largest angle errors after 25s.

Table 3. Quantified comparable results

Method	T_f	θ_{1max}	θ_{2max}	E_{θ_1}	E_{θ_2}
Prop.	28.7	1.37	1.37	0.06	0.06
LQR	39.8	1.58	1.58	0.25	0.25
CTP	40.0	1.40	1.40	0.21	0.21

Based on the figure and table, it can be observed that all three methods are capable of executing position-changing commands. The proposed method uses the shortest time to achieve the desired trolley position and the angle reduction. The maximum value of the angle during the whole progress is also the smallest. In trolley positioning, both LQR and CTP suffer from overshoot and require additional time to settle at the desired position. Besides, LQR and CTP have lower angle reduction than the proposed method. At the same time, the angle errors of the proposed method

at a certain time are smaller than the other two methods. In addition, the proposed acceleration of the motor is much smoother than LQR and CTP. The output curve would be easier to follow.

4 Conclusion

This paper introduces a trajectory planning method based on a position change request. A double pendulum system model is developed using the Lagrangian method, and a time-polynomial-based trajectory is derived. This trajectory minimizes the swing angles at the end state while simultaneously satisfying the actuator constraints (trolley jerk, acceleration, and velocity) for the entire progress. The controller exhibits remarkable performance in different simulated environments. Future research directions involve enhancing the degrees of freedom in the tower crane model to simulate real-world conditions better. It is important to note that incorporating trolley motion with jib rotation introduces additional nonlinear superposition terms, which may introduce inaccuracies during the linearization process. Additionally, to use 3D models in our future work, the integration of polynomial path planning for simultaneous control and obstacle avoidance could be explored. This comprehensive approach would offer a holistic solution for managing complex operational scenarios.

References

- [1] Jongchul Song, Walter R Fagerlund, Carl T Haas, Clyde B Tatum, and Jorge A Vanegas. Considering prework on industrial projects. *Journal of construction engineering and management*, 131(6):723–733, 2005.
- [2] Pei-Yuan Hsu, Panagiotis Angeloudis, and Marco Aurisicchio. Optimal logistics planning for modular construction using two-stage stochastic programming. *Automation in construction*, 94:47–61, 2018.
- [3] Bon-Gang Hwang, Ming Shan, and Kit-Ying Looi. Key constraints and mitigation strategies for prefabricated prefinished volumetric construction. *Journal of cleaner production*, 183:183–193, 2018.
- [4] Bingqing Zhao, Huimin Ouyang, and Makoto Iwasaki. Motion trajectory tracking and sway reduction for double-pendulum overhead cranes using improved adaptive control without velocity feedback. *IEEE/ASME Transactions on Mechatronics*, 27(5):3648–3659, 2021.
- [5] Šandor Ileš, Jadranko Matuško, and Mircea Lazar. Piece-wise ellipsoidal set-based model predictive

- control of linear parameter varying systems with application to a tower crane. *Asian Journal of Control*, 23(3):1324–1339, 2021.
- [6] Martin Böck and Andreas Kugi. Real-time nonlinear model predictive path-following control of a laboratory tower crane. *IEEE Transactions on Control Systems Technology*, 22(4):1461–1473, 2013.
- [7] Qingxiang Wu, Xiaokai Wang, Lin Hua, and Minghui Xia. Modeling and nonlinear sliding mode controls of double pendulum cranes considering distributed mass beams, varying roped length and external disturbances. *Mechanical Systems and Signal Processing*, 158:107756, 2021.
- [8] Menghua Zhang and Xingjian Jing. Model-free saturated pd-smc method for 4-dof tower crane systems. *IEEE Transactions on Industrial Electronics*, 69(10):10270–10280, 2022.
- [9] Ning Sun, Yongchun Fang, He Chen, Biao Lu, and Yiming Fu. Slew/translation positioning and swing suppression for 4-dof tower cranes with parametric uncertainties: Design and hardware experimentation. *IEEE Transactions on Industrial Electronics*, 63(10):6407–6418, 2016.
- [10] Zheng Tian, Lili Yu, Huimin Ouyang, and Guangming Zhang. Sway and disturbance rejection control for varying rope tower cranes suffering from friction and unknown payload mass. *Nonlinear Dynamics*, 105(4):3149–3165, 2021.
- [11] Menghua Zhang, Xingjian Jing, Weijie Huang, Quanli Lu, Jiyang Chen, and Zhenxue Chen. Robust fault accommodation approach for double-pendulum tower cranes via adaptive neural network-triggered control. *Nonlinear Dynamics*, pages 1–21, 2023.
- [12] Menghua Zhang and Xingjian Jing. Adaptive neural network tracking control for double-pendulum tower crane systems with nonideal inputs. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 52(4):2514–2530, 2021.
- [13] Jiaohui Peng, Jie Huang, and William Singhose. Payload twisting dynamics and oscillation suppression of tower cranes during slewing motions. *Nonlinear Dynamics*, 98:1041–1048, 2019.
- [14] Huimin Ouyang, Jinxin Hu, Guangming Zhang, Lei Mei, and Xin Deng. Decoupled linear model and s-shaped curve motion trajectory for load sway reduction control in overhead cranes with double-pendulum effect. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 233(10):3678–3689, 2019.
- [15] Zhuoqing Liu, Tong Yang, Ning Sun, and Yongchun Fang. An antiswing trajectory planning method with state constraints for 4-dof tower cranes: design and experiments. *IEEE Access*, 7:62142–62151, 2019.
- [16] Gang Li, Xin Ma, Zhi Li, and Yibin Li. Time-polynomial-based optimal trajectory planning for double-pendulum tower crane with full-state constraints and obstacle avoidance. *IEEE/ASME Transactions on Mechatronics*, 28(2):919–932, 2022.
- [17] Huimin Ouyang, Zheng Tian, Lili Yu, and Guangming Zhang. Motion planning approach for payload swing reduction in tower cranes with double-pendulum effect. *Journal of the Franklin Institute*, 357(13):8299–8320, 2020.

Advancements and Future Visions in Earthmoving Swarm Technology

Mikko Hiltunen¹ and Rauno Heikkilä²

¹University of Oulu, Finland

Mikko.Hiltunen@oulu.fi, Rauno.Heikkila@oulu.fi

Abstract -

This study examines autonomous earthmoving systems with collaboration among diverse machine types. It focuses on their application in a practical use case involving excavator loading, dumper transport, bulldozer spreading, and roller compacting for road construction. The study focuses on directly observed systems that primarily or partially align with the prescribed use cases. Drawing insights from evaluated systems, the research addresses implementation considerations such as the degree of human intervention, compatibility with diverse machine types, support for multiple simultaneous operations, adaptability to dynamic changes, integration of InfraBIM, reported productivity gains, and qualitative advantages. The study highlights the pioneering role of the earthmoving sector in adopting sensing and information technologies to reduce operation costs, enhance productivity, and improve automation and safety standards. The findings and use case formulation provide a foundation for the development of remote-controlled and autonomous earthmoving swarms, marking progress toward a future where complex construction tasks can be efficiently executed with minimal manual intervention.

Keywords -

Autonomous; Earthmoving; Robot Swarm;

1 Introduction

In recent years, the development of autonomous earthmoving has advanced to commercial solutions, with remote control as the current state-of-the-art technology [1]. This could potentially transform the labor-intensive and on-site nature of earthmoving into a desk job, changing the way we approach construction and excavation projects. The integration of remote-controlled and autonomous systems not only enhances the safety of earthmoving operations but also introduces the concept of supervising multiple machines simultaneously.

One interesting aspect of currently commissioned autonomous earthmoving solutions is their capability to record detailed operational routes and methods employed during a task. This feature allows repeating tasks with precision, minimizing human intervention and maximiz-

ing the overall efficiency of the operation. The ability to store and reproduce complex maneuvers opens new possibilities for synchronizing and optimizing earthmoving processes and achieving a higher level of consistency.[2]

The current development of radio technology has also played a major role in advancing autonomous and remote control capabilities of earthmoving. Reliable and broad radio communication has enabled remote control across great distances, transcending geographical boundaries [3]. The introduction of Gbps-level cellular networks enables live streaming of various sensor and telemetry data or ultra-high-resolution video feeds and remote controlling with tolerable latency, ensuring almost real-time responsiveness between input and feedback when operating machinery located even continents apart. This global reach not only enhances the flexibility of earthmoving operations but also transforms the traditional constraints of on-site management.

Swarm intelligence is described as a collective behavior of self-organized systems, which can operate without relying on any external or centralized control [4]. Examples of swarm intelligence can be found in many natural systems such as insect colonies or bird flocks.

A swarm of machines could be determined as multiple simultaneously moving and collaborating machines. It is often composed of homogeneous robots working towards a common goal hard to accomplish by a single robot but can also be formed from different types of machines that have tasks requiring interaction. [4]

Swarm is sometimes misleadingly used as a synonym for machine fleet [4]. The main difference between the two is that autonomous swarm machines can collect and share information to function more intelligently and coordinate joint tasks. The interaction may be accomplished with direct communications or indirectly through the environment.

This paper aims to determine the status of recent developments of cooperating machines in remote-controlled and autonomous earthmoving to underpin the beginning of the national Autonomous Swarm project led by the Digital Construction and Mining Research Group from the University of Oulu. Solutions in the construction and mining sector from around the world are briefed and their

challenges and potentials are discussed.

1.1 Literature review

In construction, the earthmoving sector has been at the forefront of embracing sensing and information technologies [5]. This adoption aims to lower operational costs, boost productivity, and elevate standards in automation and safety practices within the industry [5]. The literature on autonomous earthmoving equipment covers technical advances, optimization of earthworks planning, and fleet monitoring with a focus on safety.

Several review papers highlight innovations in earthmoving automation [5, 6, 7]. They discuss advances in fleet tracking, safety management, and machine control technology, categorizing research into equipment tracking, safety, pose estimation, and remote control. Future opportunities, especially in remote control and autonomous operation, are identified as underdeveloped areas.

Studies on fleet size optimization emphasize the integration of field data into simulation modeling. They propose methodologies, including multimodal-process data mining, to capture operational knowledge and create realistic simulation models [8]. Optimization efforts focus on minimizing both cost and emissions, considering factors such as equipment availability and project indirect costs [9, 10, 11, 12].

Fleet monitoring and safety papers explore decision-making tools for daily field management, web-accessible simulation solutions [13], and the use of motion data for real-time activity identification [14]. The importance of Multi-Agent Systems (MAS) for fleet-level coordination and safety is highlighted [15], offering proactive and reactive measures to prevent equipment-related collisions.

These papers collectively contribute to advancing autonomous earthmoving technology, addressing challenges in optimization, monitoring, and safety for more efficient and sustainable construction practices.

2 Methods

2.1 Evaluation of observed systems

Data for the evaluation is collected by observing a series of experiments and real-world demonstrations, utilizing the autonomous earthmoving machine swarm to execute various tasks in the construction and mining sector around the world.

The evaluated systems have been chosen from those that have been within the authors' direct observation in recent years. The study focuses on systems that primarily or partially align with the defined use cases.

The results obtained from the evaluation process will contribute valuable insights into the capabilities and limitations of the autonomous earthmoving system, enabling

informed decisions for its utilization in infrastructure construction machine swarms.

The observed remote-controlled and autonomous earthmoving systems are evaluated based on the following criteria.

Operational Principles:

1. **Human Intervention:** Quantifying the degree of human intervention required for the system to operate effectively.
2. **Supported Machine Types:** Identifying and assessing the machine types that have been tested and are supported by the system.
3. **Simultaneously Working Machines:** Evaluating the system's capacity to handle and coordinate multiple machines simultaneously.
4. **Adaptability to Changes:** Investigating how the system adapts to dynamic changes in the environment or the operational requirements of the construction site.
5. **Incorporation of InfraBIM:** Examining whether the system integrates with InfraBIM (Infrastructure Building Information Modeling) for enhanced project management and coordination.

Benefits:

1. **Productivity Gain:** Quantifying the reported increase in productivity achieved through the implementation of the autonomous earthmoving system.
2. **Qualitative Advantages:** Identifying and qualitatively describing the advantages offered by the system in terms of safety, precision, and overall operational improvements.



Figure 1. Use case vehicles in simulation.[16]

2.2 Use Case Formulation

To systematically organize the development of the autonomous earthmoving machine swarm, a practical use case has been formulated. The use case revolves around the construction of a designed road, encompassing four distinct types of machines involved in following tasks:

1. **Excavator Loading:** Navigation of a truck to the loading spot and excavation of soil from a designated area or pile to the dump body.
2. **Dump Truck Transport:** The transportation of excavated soil to predetermined locations.

3. **Bulldozer Spreading:** Focuses on using bulldozers to spread the transported soil evenly across the construction area.
4. **Roller Compacting:** Involves the use of rollers to compact the soil layers, ensuring a stable foundation for the road.

The formulated use case serves as a benchmark, allowing for a comprehensive understanding of the autonomous earthmoving system's capabilities in comparison to conventional operation. The use case will be tested both in simulation and real-world environments 1.

3 Evaluation

The first system (Figure 2) was demonstrated with a small wheel loader and 3 crawler dumps working together to transport a gravel pile [17]. The demonstration took place at Seikei University in Japan. The wheel loader featured a continuous curvature path planning algorithm and PurePursuit path tracking algorithm with dead time compensation based on Smith predictor. The path planning for crawler dump trucks featured a map creation algorithm narrowing down drivable areas. The shortest path was then generated with a grid-based search and smoothed to drivable form.



Figure 2. Vehicles used in autonomous system to transport gravel pile.[18]

As the second subject (Figure 3), an autonomous dam construction system is evaluated [19]. The demonstration took place at Naruse dam site in Japan. The system is used to build the cemented sand and gravel body of a trapezoidal-shaped dam. The control method utilizes work simulations based on operational data gathered from skilled operators. The construction process of multiple worksites around the country can be monitored and controlled from a central location. The quarrying excavator was conventionally operated, and construction included a highly laborious phase of laying steel reinforcement plates by hand.

The third system (Figure 4) was demonstrated in an isolated area at Ouluzone Motorsport Center where one operator remotely controlled a large wheel loader to ex-



Figure 3. Overlook of autonomous construction of Naruse Dam in Japan.[20]

cavate and load a dump truck [21]. The automatic transition routes were carefully recorded beforehand, and the machine repeated them during the demonstration. The operator took over with the remote control when adaption was required: for example, to excavate material from the pile and drop it into the dump body.



Figure 4. Control booth of retrofit remote control system.[22]

As the fourth subject (Figure 5), we evaluate a couple of autonomous haulage systems[23, 24]. The observed demonstration took place in Norway. Autonomous driving for dump trucks is attracting interest to cut costs of quarrying. Systems often include centralized traffic control, and each truck is equipped with GNSS receivers for localization and lidars and short-range radars for obstacle detection. Machines controlled by human operators have similar equipment to notify their position and for example specifying a loading position for the autonomous truck.

4 Discussion

While the demonstrations of autonomous earthmoving systems provide valuable insights into their capabilities, it is essential to acknowledge the potential bias introduced by the lack of detailed descriptions regarding the setup workload. Understanding the complexity of system setup

Table 1. Evaluation of observed remote-controlled and autonomous earthmoving systems

Principles:	1. AI swarm	2. Automated dam construction	3. Retrofit remote control	4. Autonomous haulage
Human Intervention	Demonstrated as a fully autonomous system requiring only supervision.	Conventional excavators on the material haulage phase. Autonomous construction phases require only monitoring.	Vehicles are mainly remote-controlled, can repeat actions previously recorded by the operator	Dump trucks are autonomous requiring only supervision, other machine types are remote-controlled.
Supported Machine Types	Wheel loader and dump truck (crawler)	Rollers, dozers, and dump trucks	Any, Universal retrofit	excavators, wheel loaders, dozers, and dump trucks
Simultaneously Working Machines	1 wheel loader and 3 dump trucks	4 operators control over 20 machines	2 machines per operator	not specified
Adaptability to Changes	Adaptive path planning with obstacle avoidance	Obstacle detection	No evidence of any machine vision components, relies on incessant operator supervision	Obstacle detection and automatic stopping
Incorporation of InfraBIM	No	Digital construction plans of dam	No	No
Benefits:				
Promoted Productivity Gains	Not specified	Consistent quality and speed	Enhanced operational visibility through video feed and statistical analytics	Longer operational times, lower fuel usage, extended machine life, consistent speed, reduced HR costs
Qualitative Advantages/ Disadvantages and other notes	Operational principles were discussed outright.	Special use case. Operational principles are not clear.	Adds a burden to the operator's cognitive capacity because of the lack of comprehensive sensory.	Wide range of advanced sensory implemented in retrofit installations.



Figure 5. Autonomous dump truck at loading point.[25]

is crucial for a comprehensive evaluation, as it directly impacts the practicality and feasibility of deploying these systems in research applications.

The setup workload encompasses tasks such as calibrating sensors, configuring communication networks, and in some applications creating digital maps of the construction site. Additionally, programming machine behaviors, defining operational parameters, and integrating with existing infrastructure are critical aspects that significantly contribute to the overall complexity of systems deployment.

A notable observation in the evaluation is the discrepancy in the documentation of methodologies among commercial solutions. While the first evaluated swarm system provides a comprehensive description of the utilized methodologies, commercial solutions often lack transparency in this regard. The absence of detailed information on the underlying methodologies in commercial systems hinders a thorough understanding of their operational principles, limiting the ability to assess their potential and limitations.

The evaluation shows varying levels of human intervention. Only the first system is demonstrated as fully autonomous and requires only supervision but in commercially deployed systems all or some of the machines are controlled manually or remotely. While the commercial technologies already show promising strides in improving efficiency and reducing labor-intensive tasks, challenges persist in achieving an unambiguously beneficial autonomy level, particularly in the more complex construction phases.

Across the commercial autonomous systems there is a common theme of recording and similarly repeating work processes. For example, the third and fourth evaluated systems seemingly lack integration of a more advanced path-planning interface and adaptive autonomy. The systems still demand the majority of the operator's time to monitor and intervene in case of changes in the environ-

ment. The first evaluated system was based on algorithms capable of adapting to the current situation, but during the demonstration two of the dump trucks were stopped due to algorithm malfunction.

The development of an intuitive interface for planning and monitoring is principal to the successful integration of autonomous earthmoving systems. Regardless of the control paradigm, it is necessary to provide operators with user-friendly tools that allow for efficient task planning, real-time monitoring, and intervention when necessary. This interface should have a balance between comprehensive oversight and machine-level control, ensuring that operators can easily understand and influence the swarm's behavior.

One prevailing trend observed across the evaluated autonomous earthmoving systems is the reliance on centralized control architectures. While this approach facilitates coordination and monitoring, especially in scenarios involving multiple machines and complex tasks, it raises questions about scalability and adaptability. In a conventional construction environment, centralized control aligns with the hierarchical organization of tasks and the need for coordinated actions. However, the paradigm shifts when considering the dynamics of a robot swarm, where decentralized control mechanisms may offer distinct advantages.

A notable revelation is the absence of any form of InfraBIM incorporation within most of the evaluated systems. Only the second evaluated system had digital plans of automatically constructed material layers. The benefits of InfraBIM, such as enhanced communication and streamlined data transfer, seem to be overlooked. Incorporating InfraBIM principles could pave the way for more intelligent and intuitive interactions between the modeled design and the autonomous systems, gaining a new level of efficiency and precision in construction operations.

The efficient handling of data poses a challenge in the context of autonomous earthmoving systems, with centralized systems often facing higher data processing and communication costs. As the scale of construction projects increases, the challenge of data scalability becomes more pronounced. Decentralized control systems, by distributing data processing among individual units, may offer a more scalable solution, potentially lowering the costs associated with the handling and transmission of large datasets.

5 Conclusion

The formulated use case, covering excavator loading, dumper transport, bulldozer spreading, and roller compacting, serves as a foundational framework for developing methods and concepts for autonomous earthmoving systems. Valuable insights are gleaned from evaluated systems and existing literature. The initial system offers tangible methods and functions for individual machines,

while others present refined user interfaces and comprehensive hardware solutions.

The study's evaluation predominantly relies on qualitative assessments due to the limited availability of quantitative data, particularly for commercial solutions where such information may not be publicly accessible. Furthermore, the study's emphasis on formulated use cases inherently lends toward subjective evaluations, as the aim is to identify systems or components adaptable to own research and development endeavours.

While the study's criteria ensure a comprehensive assessment of Earthmoving swarm technologies, it's essential to acknowledge the limitations of qualitative evaluation in fully capturing aspects like economical feasibility, operator skills, and potential time and cost savings. Future research needs to address these areas to delve into the quantitative aspects and provide a more holistic understanding of autonomous earthmoving systems' effectiveness and practical implications. As autonomous swarm technologies are advancing in the construction sector and data becomes more available, there is also a need for systematic analysis of the social and environmental aspects of using autonomous Earth-moving equipment, particularly in terms of benefits and impact evaluation.

The ongoing phase of the swarm project aims to set up the use case machines with the necessary equipment and control systems. The subsequent step involves establishing a secure yet dynamic connection across all robots and monitoring systems, allowing authorized units to seamlessly join the swarm.

The future objective is to demonstrate the use cases first individually and then collectively. The aspiration is to eliminate all manual intervention, with a transitional phase focusing on creating a user interface concept. This concept should empower a minimal number of human operators to instruct multiple machines, allow designing detailed tasks swiftly and intuitively, and need manual control only for exceptional circumstances.

The ultimate goal for InfraBIM-based automation development is to achieve machine readability for the models, enabling autonomous machines to directly comprehend their intended tasks from modeled designs without intermediary software[26]. With appropriate infrastructure, AI-powered methods could be implemented to take control of task execution, as already demonstrated for some machine types, especially in the transportation field.

References

- [1] J. S. Lee, Y. Ham, H. Park, and J. Kim. Challenges, tasks, and opportunities in teleoperation of excavator toward human-in-the-loop construction automation. *Automation in Construction*, 135:104119, 2022. doi:10.1016/j.autcon.2021.104119.
- [2] S. N. Naghshbandi, L. Varga, and Y. Hu. Technologies for safe and resilient earthmoving operations: A systematic literature review. *Automation in Construction*, 125:103632, 2021. doi:10.1016/j.autcon.2021.103632.
- [3] S. Ali, W. Saad, N. Rajatheva, K. Chang, D. Steinbach, B. Sliwa, C. Wietfeld, K. Mei, H. Shiri, H-J. Zepernick, et al. 6g white paper on machine learning in wireless communication networks. *arXiv preprint arXiv:2004.13875*, 2020. doi:10.48550/arXiv.2004.13875.
- [4] M. Rath, A. Darwish, B. Pati, B.K. Pattanayak, and C.R. Panigrahi. Swarm intelligence for resource management in internet of things. pages 21–45, 2020. doi:10.1016/B978-0-12-818287-1.00005-X.
- [5] E. R. Azar and V. R. Kamat. Earthmoving equipment automation: A review of technical advances and future outlook. *Journal of Information Technology in Construction (ITcon)*, 22(13):247–265, 2017. ISSN 1874-4753.
- [6] P. G.P.S. Fernandes, E. F. N. Júnior, and B. A. Prata. Optimization of earthworks planning: a systematic mapping study. *Canadian Journal of Civil Engineering*, 49(12):1781–1795, 2022. doi:10.1139/cjce-2022-0185.
- [7] H. A.D. Nguyen and Q. P. Ha. Robotic autonomous systems for earthmoving equipment operating in volatile conditions and teaming capacity: a survey. *Robotica*, 41(2):486–510, 2023. doi:10.1017/S0263574722000339.
- [8] R. Akhavian and A. H. Behzadan. Knowledge-based simulation modeling of construction fleet operations using multimodal-process data mining. *Journal of Construction Engineering and Management*, 139(11):04013021, 2013. doi:10.1061/(ASCE)CO.1943-7862.0000775.
- [9] S. A.R. Kaboli, M. Bahaaddini, and S. M. Kaboli. Optimal earthmoving fleet size for minimising emissions and cost. *Journal of Mining and Environment*, 11(4):949–965, 2020. doi:10.22044/jme.2020.9910.1918.
- [10] D. Morley, T. Joseph, and M. Lu. In search of the ideal truck-excavator combination. In *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*, volume 30, page 1. IAARC Publications, 2013. doi:10.22260/ISARC2013/0045.

- [11] C. Yi and M. Lu. A simulation-based earthmoving fleet optimization platform (sefop) for truck/excavator selection in rough grading project. In *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*, volume 35, pages 1–7. IAARC Publications, 2018. doi:10.22260/ISARC2018/0133.
- [12] M. Marzouk and O. Moselhi. Multiobjective optimization of earthmoving operations. *Journal of construction Engineering and Management*, 130(1):105–113, 2004. doi:10.1061/(ASCE)0733-9364(2004)130:1(105).
- [13] Y. Mohamed and M. Ali. A simplified online solution for simulation-based optimization of earthmoving operations. In *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*, volume 30, page 1. IAARC Publications, 2013. doi:10.22260/ISARC2013/0040.
- [14] K. M. Rashid and J. Louis. Automated activity identification for construction equipment using motion data from articulated members. *Frontiers in Built Environment*, 5:144, 2020. doi:10.3389/fbuil.2019.00144.
- [15] F. Vahdatikhaki, S. M. Langari, A. Taher, K. El Amari, and A. Hammad. Enhancing coordination and safety of earthwork equipment operations using multi-agent system. *Automation in construction*, 81: 267–285, 2017. doi:10.1016/j.autcon.2017.04.008.
- [16] M. Hiltunen. Test environment for autonomously operating swarm developed at the university of oulu. agx dynamics for unity -simulation software, 2023. URL <https://www.algoryx.se/agx-unity/>.
- [17] T. Takei, K. Ichikawa, K. Okawa, S. Sarata, T. Tsubouchi, and A. Torige. Path planning of wheel loader type robot for scooping and loading operation by genetic algorithm. In *2013 13th International Conference on Control, Automation and Systems (ICCAS 2013)*, pages 1494–1499, 2013. doi:10.1109/ICCAS.2013.6704123.
- [18] M. Hiltunen. Autonomously working system of small equipment excavating and transporting gravel in japan, 2023.
- [19] A4csel construction production system. On-line: https://www.kajima.co.jp/english/tech/c_a4csel/index.html, Accessed: 07/12/2023.
- [20] M. Hiltunen. Autonomously working dam construction cite at naruse dam in japan., 2022.
- [21] Teleo’s full-stack solution. On-line: <https://www.teleo.ai/technology/>, Accessed: 07/12/2023.
- [22] R. Heikkilä. Remote control booth of retrofit system for wheel loader manufactured by teleo, 2023. URL <https://www.teleo.ai/>.
- [23] T. Hamada and S. Saito. Autonomous haulage system for mining rationalization. *Hitachi Rev*, 67(1):87–92, 2018.
- [24] Steer autonomous. On-line: <https://www.steer.no/>, Accessed: 07/12/2023.
- [25] R. Heikkilä. Autonomously driving dump truck at loading point. retrofit control system manufactured by steer, 2023. URL <https://www.steer.no/>.
- [26] M. Hiltunen, R. Heikkilä, I. Niskanen, and M. Immonen. Open infrabim for remote and autonomous excavation. *Automation in Construction*, 156, 2023. doi:10.1016/j.autcon.2023.105148.

A construction robot path planning method based on safe space and worker trajectory prediction

Xiaotian Ye¹, Hongling Guo^{1,*}, Ziyang Guo¹, Zhubang Luo¹

¹Department of Construction Management, Tsinghua University, Beijing, China
yxt19@mails.tsinghua.edu.cn, hlguo@tsinghua.edu.cn, ziyangguo@mail.tsinghua.edu.cn,
lzb2017@mail.tsinghua.edu.cn

Abstract –

As the key to intelligent construction, construction robots can perform complex and dangerous tasks instead of workers. Construction robot path planning (CRPP) is a prerequisite for executing tasks. However, dynamic environments and moving workers at sites create significant difficulties for CRPP. To solve this issue, this research proposes a construction robot path planning method based on safe space and worker trajectory prediction. Firstly, a grid map with a target point is automatically established based on BIM (Building Information Modeling) and construction schedule. Secondly, an improved A* algorithm with a dynamic weight of the heuristic function is developed for global path planning. Thirdly, the worker and robot safe space are defined, and worker trajectory is predicted to improve the DWA (Dynamic Window Approach) for local path planning. Furthermore, a decision model is developed to deal with the path conflict based on the potential collision zone (PCZ). Finally, an experiment is designed and conducted to validate the proposed method. It is found that the method can effectively achieve the optimal path and resolve path conflict to ensure worker and robot safety.

Keywords –

Construction robot path planning; Safe space; Worker trajectory prediction; Decision model.

1 Introduction

The construction industry contributes to economic growth worldwide and boosts numerous countries' gross domestic product. However, it is confronted with global challenges, such as labour shortages, stagnant productivity, and frequent safety incidents [1]. An MGI (McKinsey Global Institute) survey indicates that over the past 20 years, the annual growth rate of labour productivity in the construction industry has been a mere 2%, significantly lagging behind other industries [2]. Construction robots and automation technologies are

recognized as the key to enhancing construction efficiency and safety [3]. Construction robots have demonstrated their advantages in various fields, including main structure and decoration works [4], building cleaning and waste recycling [5], and structural inspection.

In construction environment, safe and efficient path planning is not only the fundamental to the execution of construction tasks by robots but also a critical factor in enhancing their operation efficiency [6]. The continuous dynamic changes at construction sites, especially the uncertainty caused by workers and machinery, pose enormous challenges to their collaboration safety while accomplishing their tasks [7]. Therefore, it is urgent to improve construction robots' path-planning capability at construction sites.

Related advancements have been made in robotic path planning within construction scenarios. Pinto et al. developed a Vision-Guided Path-Planning System (V-GPP). The system employs RGB-D cameras to acquire real-time environmental data, integrating A* algorithm and 3D grid maps to plan safe routes for cable-driven robots [8]. Do et al. [9] utilized depth cameras to collect information about onsite environment and obstacles, refining the path planning of sTetro robots for stair-cleaning tasks through grid optimization and heat conduction analysis. However, current research mainly involves dynamic obstacles whose information is known and indoor scenarios, and ignores path conflict resolution strategies for random dynamic entities on site, especially in effectively integrating worker trajectory prediction with robotic path planning.

Therefore, this research proposes a dynamic path-planning method considering safe space and predicted worker trajectory to ensure the safety of worker-robot collaboration. A literature review is first made in Section 2. Then, the path planning method is illustrated in Section 3 and an experiment is conducted to test its performance in Section 4. In the end, a conclusion is drawn in Section 5.

2 Literature review

2.1 Path planning in dynamic environments

Path planning-related research primarily focuses on global path planning based on known information and local path planning based on real-time sensor feedback.

Global path planning can find the optimal route by fully exploiting the existing scene knowledge in static environments. Graph-search-based algorithms such as A* [10] and D* [11] have been extensively applied in path planning. Furthermore, biomimetic heuristic algorithms such as GA (Genetic Algorithm) [12], ACO (Ant Colony Optimization) [13], and neural network algorithms [5] have provided new perspectives for robotic path planning. By introducing a heuristic function, A* algorithm has significant advantages in enhancing search efficiency and is considered the best-first algorithm [14].

Local path planning is crucial in guiding robots to avoid dynamic obstacles. The most used local path planning algorithms include Artificial Potential Field (APF) [15], Dynamic Window Approach (DWA) [16], and reinforcement learning-based method [17,18]. For instance, Anirudh et al. [17] trained a Coverage Path Planning (CPP) model for a tile-laying robot using reinforcement learning, maximizing the area coverage of tile laying while minimizing energy consumption.

To address path planning problems in dynamic environments, a hybrid strategy that integrates global and local path planning has received significant attention. The hybrid approach initially generates an optimal route based on global path planning and then makes flexible adjustments using local path planning [19]. However, the method overlooks trajectory prediction for dynamic obstacles and fails to consider the safe space of workers and robots. Most studies expand on obstacles but lack a comprehensive and effective solution for path conflict.

2.2 Worker trajectory prediction

The dynamic and random environment of construction sites leads to frequent safety incidents [20] and poses significant challenges to construction robot path planning. The predicted trajectories of workers and machinery can enhance the efficiency and safety of path planning in such environments [21]. In current studies, trajectory prediction methods are primarily categorized into physics-based approaches (Sense - Predict) and learning-based approaches (Sense - Learn - Predict) [22].

Physics-based approaches are based on dynamic or kinematic models, using mathematical formulas to depict target motion. Zhu et al. utilized a Kalman filter to integrate parameters such as position, velocity, and acceleration for predicting the locations of workers and mobile equipment [23]. Considering the randomness of

worker movement, some studies have adopted Markov Models to forecast workers' potential trajectories and statuses [24,25]. The parameters of these models are derived from historical data and do not require training data, which are suited for short-term predictions.

Learning-based approaches can extract dynamic models and statistical behaviour patterns from vast training data, thus allowing long-term predictions in complex and dynamic environments. Long Short-Term Memory (LSTM) networks have been commonly used for motion trajectory prediction. Tang et al. [26,27] developed an LSTM encoder-decoder and combined it with a Mixture Density Network (MDN) to model the uncertainty in predictions, achieving entity trajectory forecasting up to 2 seconds. Cai et al. [28] proposed a context-aware LSTM-based method, extracting abundant contextual information (e.g., neighbour position, the relationship with the neighbour, and the distance from the destination) and feeding these data into the LSTM model for precise worker trajectory prediction.

Learning-based approaches rely on extensive historical data while acquiring high-quality data at construction sites is challenging. This limits the accuracy of the models. On the other hand, physics-based approaches do not rely on massive historical data, which is more suitable for construction environments.

3 Path planning method

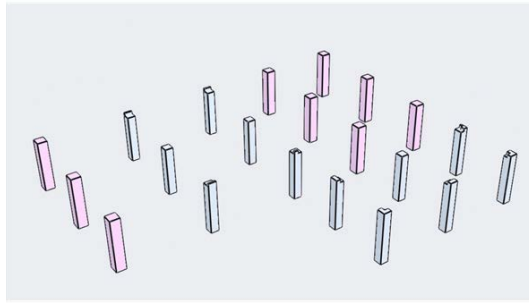
This method involves map establishment, global path planning, local path planning, and decision-making. This research establishes a grid map based on the BIM (Building Information Modeling) model and construction schedule, develops an improved A* algorithm for global path planning in the static map and an improved DWA for local path planning based on safe space and predicted worker trajectory, and builds a decision model to reduce unnecessary local planning.

3.1 BIM-based map establishment

BIM includes plenty of building information, which can provide a basis for establishing construction maps. By exporting a BIM model to an IFC (Industry Foundation Classes) file, basic information such as component type, component location, and geometric dimensions, can be extracted to generate a preliminary map. The IFC has a standard syntax to obtain component-related information. For example, IFCCARTESIANPOINT in Figure. 1(b) represents the position of a column, and IFCCOLUMNTYPE contains the component's global ID (e.g. the column's global ID is '1I2u_TieX4cg\$DL01NNHAJ') and its attribute information such as the component type and dimensions (e.g. the component type is COLUMN and its cross-section size is 500x500mm). With the location and cross-

section information, each component can be drawn in the grid map. Meanwhile, the construction schedule is automatically generated based on previous research by the authors [29] (e.g., Figure. 1(c)). Both completed construction tasks and those to be completed can be obtained from the schedule to update the grid map (e.g.,

Figure. 1(d)). Each gray rectangle not only represents an obstacle but also has semantic information (e.g., component type and size) through the global ID. The gray rectangles represent the columns that have been constructed, while the red one represents the column to be done, which is the target for the construction robot.



(a) BIM model

```
#933- IFCSTYLIDETEM(#932, (#260, $));
#936- IFCSHAPEPRESENTATION(#182, 'Body', 'Brep', (#932));
#938- IFCARTESIANPOINT((4830.9641400045, -293.303414205191, 0.));
#940- IFCAXIS2PLACEMENT3D(#6, $, $);
#941- IFCREPRESENTATIONMAP(#940, #936);
#942- IFCLOCALTYPE('112u1Tc4xGsd0n1IWH4', #41, '3G-500 x 500mm-0', $, $, $, (#41, '367932', '3G-500 x 500mm-0', 'COLUW.));
#944- IFCMAPDETITEM(#41, #921);
#946- IFCSHAPEPRESENTATION(#182, 'Body', 'MappedRepresentation', (#944));
#948- IFCPRODUCTDEFINITIONSHAPE($, (#946));
#950- IFCARTESIANPOINT((4830.9641400045, -293.303414205191, 0.));
#952- IFCLOCALPLACEMENT3D(#950, $);
#953- IFCLOCALPLACEMENT3D(#122, #952);
#954- IFCLOCALTYPE('112u1Tc4xGsd0n1IWH4', #41, '\X2(60F751D0571FX0\ - \X2(77E95F62\X0 - \X2(67F1X0\3G-500mm-0:373729', '3G-500 x 500mm-0', #953, #948, '373729');
#956- IFCPROPERTYSET('31c1Y1Zb739HW721c80', #41, 'Pset_ColumnCommon', $, (#314, #315, #316));
#959- IFCREDEFINISYBPROPERTIES('31juc6k1a3xxvZ8pC1L19', #41, $, $, (#954, #957));
#963- IFCLOCALPLACEMENT3D(#1808, $);
#11195- IFCREDEFINISYBTYPE('31VjKmlnL803apZ0m53e', #41, $, $, (#5663, #5651);
#9904- IFCARTESIANPOINT((500., 500., 0.));
#9906- IFCARTESIANPOINT((500., 400., 0.));
```

(b) IFC file

項目		年度																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
年度	項目	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373	2374	2375	2376	2377	2378	2379	2380	2381	2382	2383	2384	2385	2386	2387	2388	2389	2390	2391	2392	2393	2394	2395	2396	2397	2398	2399	2400	2401	2402	2403	2404	2405	2406	2407	2408	2409	2410	2411	2412	2413	2414	2415	2416	2417	2418	2419	2420	2421	2422	2423	2424	2425	2426	2427	2428	2429	2430	2431	2432	2433	2434	2435	2436	2437	2438	2439	2440	2441	2442	2443	2444	2445	2446	2447	2448	2449	2450	2451	2452	2453	2454	2455	2456	2457	2458	2459	2460	2461	2462	2463	2464	2465	2466	2467	2468	2469	2470	2471	2472	2473	2474	2475	2476	2477	2478	2479	2480	2481	2482	2483	2484	2485	2486	2487	2488	2489	2490	2491	2492	2493	2494	2495	2496	2497	2498	2499	2500	2501	2502	2503	2504	2505	2506	2507	2508	2509	2510	2511	2512	2513	2514	2515	2516	2517	2518	2519	2520	2521	2522	2523	2524	2525	2526	2527	2528	2529	2530	2531	2532	2533	2534	2535	2536	2537	2538	2539	2540	2541	2542	2543	2544	2545	2546	2547	2548	2549	2550	2551	2552	2553	2554	2555	2556	2557	2558	2559	2560	2561	2562	2563	2564	2565	2566	2567	2568	2569	2570	2571	2572	2573	2574	2575	2576	2577	2578	2579	2580	2581	2582	2583	2584	2585	2586	2587	2588	2589	2590	2591	2592	2593	2594	2595	2596	2597	2598	2599	2600	2601	2602	2603	2604	2605	2606	2607	2608	2609	2610	2611	2612	2613	2614	2615	2616	2617	2618	2619	2620	2621	2622	2623	2624	2625	2626	2627	2628	2629	2630	2631	2632	2633	2634	2635	2636	2637	2638	2639	2640	2641	2642	2643	2644	2645	2646	2647	2648	2649	2650	2651	2652	2653	2654	2655	2656	2657	2658	2659	2660	2661	2662	2663	2664	2665	2666	2667	2668	2669	2670	2671	2672	2673	2674	2675	2676	2677	2678	2679	2680	2681	2682	2683	2684	2685	2686	2687	2688	2689	2690	2691	2692	2693	2694	2695	2696	2697	2698	2699	2700	2701	2702	2703	2704	2705	2706	2707	2708	2709	2710	2711	2712	2713	2714	2715	2716	2717	2718	2719	2720	2721	2722	2723	2724	2725	2726	2727	2728	2729	2730	2731	2732	2733	2734	2735	2736	2737	2738	2739	2740	2741	2742	2743	2744	2745	2746	2747	2748	2749	2750	2751	2752	2753	2754	2755	2756	2757	2758	2759	2760	2761	2762	2763	2764	2765	2766	2767	2768	2769	2770	2771	2772	2773	2774	2775	2776	2777	2778	2779	2780	2781	2782	2783	2784	2785	2786	2787	2788	2789	2790	2791	2792	2793	2794	2795	2796	2797	2798	2799	2800	2801	2802	2803	2804	2805	2806	2807	2808	2809	2810	2811	2812	2813	2814	2815	2816	2817	2818	2819	2820	2821	2822	2823	2824	2825	2826	2827	2828	2829	2830	2831	2832	2833	2834	2835	2836	2837	2838	2839	2840	2841	2842	2843	2844	2845	2846	2847	2848	2849	2850	2851	2852	2853	2854	2855	2856	2857	2858	2859	2860	2861	2862	2863	2864	2865	2866	2867	2868	2869	2870	2871	2872	2873	2874	2875	2876	2877	2878	2879	2880	2881	2882	2883	2884	2885	2886	2887	2888	2889	2890	2891	2892	2893	2894	2895	2896	2897	2898	2899	2900	2901	2902	2903	2904	2905	2906	2907	2908	2909	2910	2911	2912	2913	2914	2915	2916	2917	2918	2919	2920	2921	2922	2923	2924	2925	2926	2927	2928	2929	2930	2931	2932	2933	2934	2935	2936	2937	2938	2939	2940	2941	2942	2943	2944	2945	2946	2947	2948	2949	2950	2951	2952	2953	2954	2955	2956	2957	2958	2959	2960	2961	2962	2963	2964	2965	2966	2967	2968	2969	2970	2971	2972	2973	2974	2975	2976	2977	2978	2979	2980	2981	2982	2983	2984	2985	2986	2987	2988	2989	2990	2991	2992	2993	2994	2995	2996	2997	2998	2999	3000

(c) Construction schedule

(d) Grid map with a schedule

Figure 1. The establishment of a grid map

3.2 Improved A* for global path planning

A* algorithm is a heuristic search algorithm suitable for global path planning in a static environment with known information. The basic idea is to sort the cost of the optional nodes around a current node, select the least-cost node, and repeat the process until it reaches the target point. In a 2D grid map, A* algorithm utilizes the cost function $f(n)$ to evaluate the path length as Equation (1).

$$f(n) = g(n) + h(n) \quad (1)$$

Where, $g(n)$ is the distance from the start node to the current node, and $h(n)$ is the heuristic function representing the distance from the current node to the destination.

Traditional A* algorithms use the same weights for $g(n)$ and $h(n)$, which leads to more search nodes and low efficiency. This research adds dynamic weight w to $h(n)$ (see Equation (2)), increasing w when $g(n)$ is less than $h(n)$ and decreasing w when $g(n)$ is greater

than $h(n)$. The dynamic weight can consider the characteristics of the construction schedule to reduce the search space and increase the search speed.

$$f(n) = g(n) + w \cdot h(n) \quad (2)$$

3.3 Improved DWA for local path planning

In traditional DWA, robots' speed limit only considers motor performance and braking distance. Besides, the distance and velocity functions are built based solely on the simple expansion of an obstacle. The dynamic properties of robots and workers and the future state of workers are not considered. Therefore, a safe space is defined and created for both workers and robots, and worker trajectory is predicted to modify DWA to ensure the safety of the path.

3.3.1 Safe space definition of workers and robots

(1) Worker safe space

This research defines worker safe space (WSS) as the worker operation space (WOS) and worker movement space (WMS), as shown in Figure. 2. WOS refers to the maximum space occupied by a worker's operation. This research adopts the maximum arm span of a standing worker as WOS, which is a circle with radius r_w . According to Chinese National Standard GB/T 13549-92 "Human dimensions in workspace" [30], the arm span of an adult male is 1.78m (90th percentile). WMS is the distance s_w traveled by a worker during the period of time when a robot detects the worker and reacts until the robot comes to a complete stop. WOS is a static region, while WMS is dynamic and related to worker speed.

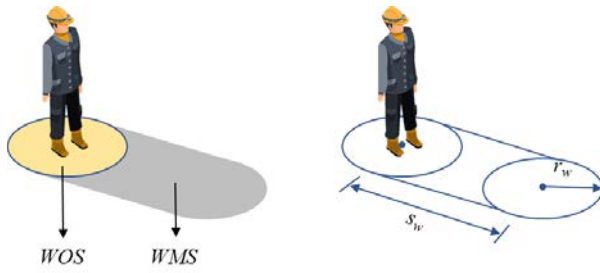


Figure 2. Worker safe space

Referring to ISO 15066 [31], s_w is calculated as Equation (3).

$$s_w = \int_{t_0}^{t_0+T_r+T_s} v_w(t) dt \quad (3)$$

Where, t_0 , T_r , T_s represent the present time, the reaction time of a construction robot, and the stopping time of the robot; v_w is the worker's speed in the direction of the robot. ISO 13855 [32] recommends a value of 1.6 m/s for the speed of human motion, thus Equation (3) can be simplified to Equation (4).

$$s_w = v_w(T_r + T_s) = 1.6 \times (T_r + T_s) \quad (4)$$

(2) Construction robot safe space

This research defines construction robot safe space (CRSS), which consists of self-occupied space (SS), safety distance (SD), and braking space (BS) (see Figure. 3). SS is the smallest outer circle occupied by a robot. SD refers to the minimum distance l_{cr} that should be maintained between a worker and the robot in a collaborative situation. According to ISO10128 [33], l_{cr} takes the value of 0.5m. BS means the distance s_{cr} traveled by the robot from the time point to detect the worker to that to stop.

As in Equations (5) and (6), s_{cr} consists of s_{cr1} and s_{cr2} . s_{cr1} represents the distance traveled during the reaction time of the robot, while s_{cr2} represents the distance travelled during the robot's stopping time.

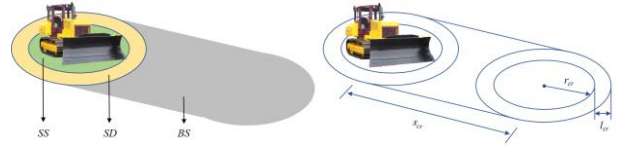


Figure 3. Construction robot safe space

$$s_{cr1} = \int_{t_0}^{t_0+T_r} v_r(t) dt \quad (5)$$

$$s_{cr2} = \int_{t_0+T_r}^{t_0+T_r+T_s} v_s(t) dt \quad (6)$$

Where, v_r is the directed speed of the robot in the direction of the worker. v_s is the speed of the robot in the course of stopping. In this research, the robot's motion during the reaction time is simplified as a uniform velocity motion, and the motion during the stopping time is simplified as a constant deceleration motion. Then, Equations (5) and (6) are respectively simplified to Equations (7) and (8).

$$s_{cr1} = v_r(t_0) \times T_r \quad (7)$$

$$s_{cr2} = \frac{1}{2} v_r(t_0) T_s \quad (8)$$

With the safe space of workers and robots, the objective functions (e.g., distance and velocity) can be optimized in DWA algorithm.

3.3.2 Worker trajectory prediction

In the context of intelligent construction, construction robots are disturbed mainly by workers and other robots on site. This paper assumes that construction robots onsite can interact with each other to avoid collisions, which can reduce some unpredictable interference. Construction workers are the most common and complex dynamic obstacles at construction sites. Therefore, this study focuses on the worker as a dynamic obstacle for localized collision avoidance. Predicting worker trajectory can provide abundant information for the local path planning of robots and improve their path safety. Differing from the traditional prediction that assumes an obstacle to move in a uniform linear motion, this research considers worker turning and establishes a nonlinear model of workers (i.e., constant linear velocity v and angular velocity ω). Then, worker trajectory is predicted based on CKF (Cubature Kalman Filter). The status vector ψ is defined as Equations (9) and (10).

$$\psi = \begin{bmatrix} x \\ y \\ v \\ \theta \\ \omega \end{bmatrix} \quad (9)$$

$$\psi_t = \psi_{t-1} + \int_{t-1}^t \begin{bmatrix} x'(t) \\ y'(t) \\ v'(t) \\ \theta'(t) \\ \omega'(t) \end{bmatrix} dt \begin{bmatrix} x_{t-1} + \frac{v}{\omega} \sin\theta_{t-1} \cos(\omega * \Delta t) + \frac{v}{\omega} \cos\theta_{t-1} \sin(\omega * \Delta t) - \frac{v}{\omega} \sin\theta_{t-1} \\ y_{t-1} - \frac{v}{\omega} \cos\theta_{t-1} \cos(\omega * \Delta t) + \frac{v}{\omega} \sin\theta_{t-1} \sin(\omega * \Delta t) + \frac{v}{\omega} \cos\theta_{t-1} \\ \theta_{t-1} + \omega * \Delta t \\ \omega \end{bmatrix} \quad (10)$$

Where, x , y , v , θ and ω denote the horizontal coordinates, vertical coordinates, velocity, angle, and the deflection angular velocity of a worker at time t , respectively.

The status and observation equations are defined as Equations (11) and (12), respectively.

$$\psi_t = F(\psi_{t-1}) + \varphi_t \quad (11)$$

$$Z_t = H * \psi_t + \lambda_t \quad (12)$$

Where, F , φ_t , Z_t , H , and λ_t is the nonlinear status transition matrix, the system noise vector, the observation vector, the observation matrix, and the observation noise vector, respectively. The observation vector Z can be obtained by collecting worker coordinates from sensors. The observation matrix H in Equation (14) shows the transformation from the state vector ψ_t to the observation vector Z_t .

$$Z = \begin{bmatrix} x \\ y \end{bmatrix} \quad (13)$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad (14)$$

φ_t and λ_t are assumed to satisfy positive definite, symmetric and uncorrelated, zero mean Gaussian white noise vector, e.g., $\varphi_t \sim N(0, Q_t)$, $\lambda_t \sim N(0, R_t)$ in Equation (15) and (16).

$$Q = \begin{bmatrix} \sigma_x & 0 & 0 & 0 & 0 \\ 0 & \sigma_y & 0 & 0 & 0 \\ 0 & 0 & \sigma_v & 0 & 0 \\ 0 & 0 & 0 & \sigma_\theta & 0 \\ 0 & 0 & 0 & 0 & \sigma_\omega \end{bmatrix} \quad (15)$$

$$R = \begin{bmatrix} \sigma_{x'} & 0 \\ 0 & \sigma_{y'} \end{bmatrix} \quad (16)$$

Where, σ_x , σ_y , σ_v , σ_θ , and σ_ω represent the variances of x , y , v , θ , and ω respectively. $\sigma_{x'}$ and $\sigma_{y'}$ respectively represent the variances of the observation vector. The status vector ψ includes 5 parameters belonging to a high dimensional system. For nonlinear state estimation, Arasaratnam and Haykin [34] proposed the CKF, which derives a third-degree spherical-radial cubature rule. The aforementioned kinematic model belongs to the nonlinear state model, solved using CKF. Then, the predicted trajectory can optimize the DWA's distance function and velocity range. Meanwhile, considering some unpredictable disturbances onsite, we shorten the exploration time (i.e., the search step of DWA) to minimize collisions in the search space.

3.4 A decision model for path conflict

Path conflict is the core of the dynamic planning problem. However, a path conflict does not mean that a collision will occur. Collisions happen when both spatial and temporal conflicts occur. Therefore, to reduce unnecessary local planning and make path planning more efficient, a decision model is proposed. Based on the global path and worker trajectory prediction, the potential collision zone (PCZ) between the paths can be obtained to determine if spatial conflicts occur (see Figure. 4). The PCZ is generated and updated in real time since the worker trajectory is predicted in real-time. If there are no PCZs between the two paths, there will be no spatial conflicts between the worker and the robot. Therefore, collisions do not occur, and the robot can follow the globally optimal path.

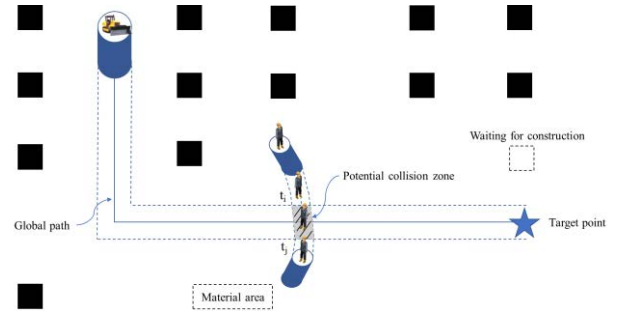


Figure 4. Potential collision zone (PCZ)

When a PCZ exists (i.e., a spatial conflict exists), the decision model will further determine the temporal conflict to select a decision. The time for workers to enter and leave the PCZ is calculated as t_i and t_j respectively. Then, the time for the construction robot to enter and leave the PCZ is calculated as t_p and t_q . Depending on the time when the robot enters the PCZ, the following robot decisions are made: 1) Maintain the original speed and path, 2) Maintain the original path and reduce speed, and 3) Change paths. The decision model is presented as follows:

Case 1: When $t_q < t_i$, the worker has crossed the PCZ, while the robot has not yet entered the PCZ. Therefore, the robot can travel at its original path and speed.

Case 2: When $t_p > t_j$, the worker has not yet entered the PCZ, while the robot has crossed the PCZ. Therefore, the robot can travel on its original route and speed.

Case 3: When $t_p < t_i < t_q < t_j$ and $t_p < t_i < t_j < t_q$,

the robot and the worker will meet in the PCZ, with the robot entering the PCZ first. Therefore, the robot needs to change a path locally.

Case 4: When $t_i < t_p < t_q < t_j$ and $t_i < t_p < t_j < t_q$, the robot and the worker will meet in the PCZ, with the worker entering the PCZ first. Therefore, the robot needs to reduce speed or change paths. If the robot cannot avoid the collision by decelerating, the robot needs to change path locally.

A sudden turn by the worker can lead to a failed trajectory prediction, in which case the decision model will control the robot to stop immediately (e.g. decision (2)). The improved DWA will plan a local path for these cases, where changing paths is required. The decision model reduces robot replanning and unnecessary local planning, allowing the robot to follow the optimal path as much as possible.

4 Experiment and test

4.1 Experiment

An experiment was designed to validate the feasibility of the proposed method. For global path planning, the experiments tested the improved A* algorithm on a grid map with a schedule (see Figure 5). In the grid map, the left side represents completed construction areas, and the right side represents areas that have not yet started. For local path planning, the test used a tracked vehicle to represent a construction robot and a humanoid robot to represent a worker, as shown in Figure. 6. Meanwhile, columns and materials were placed in the scene to simulate a construction site better.

A depth camera was utilized to acquire the position of the worker and robot, further calculating their safe spaces, and predicting worker trajectories. Then, the improved A* algorithm was utilized to plan the global path. As a dynamic obstacle, the worker moved at different speeds within a low-speed range and conflicted with the global path. Based on the decision model, the construction robot would decide whether to perform local path planning.

4.2 Results

The improved A* can successfully generate the global path as shown in Figure. 5. Meanwhile, table 1 compares the planning time of traditional and improved A* algorithms under various starting and target points. In the local path planning scenario, the global path (the purple line) and the predicted worker trajectory (the blue band) are shown in Figure. 6. A PCZ exists between the

global path and the predicted trajectory. In Figure. 6, since the worker would leave before the robot reached the PCZ, the robot followed its original path and speed without local path planning according to the decision model. In Figure. 7, the robot and the worker would meet in the PCZ, where the path needs to be locally planned using the improved DWA algorithm. Figure. 7 (a) shows the local path (the blue line) considering the predicted worker trajectory. The robot found a path behind the worker, preventing the secondary collision from going around in front of the worker. Compared with the local path without the prediction in Figure. 7 (b), the robot passed in front of the worker, which might cause a secondary collision.

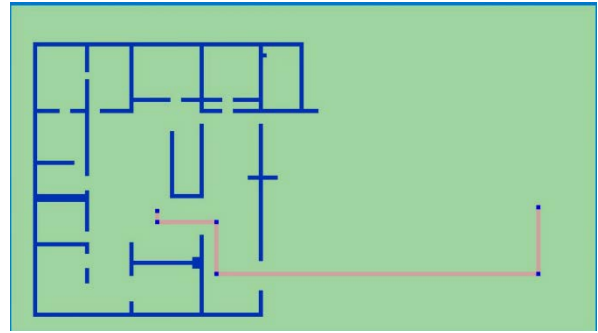


Figure 5. The experimental scenario for global path planning

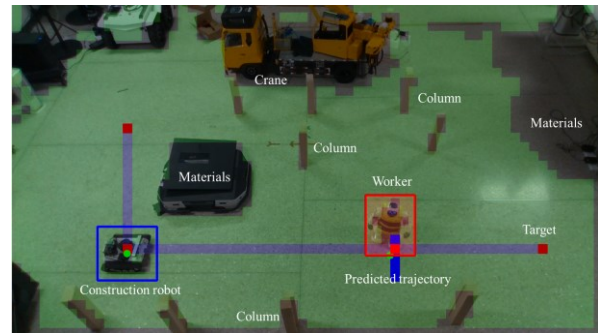
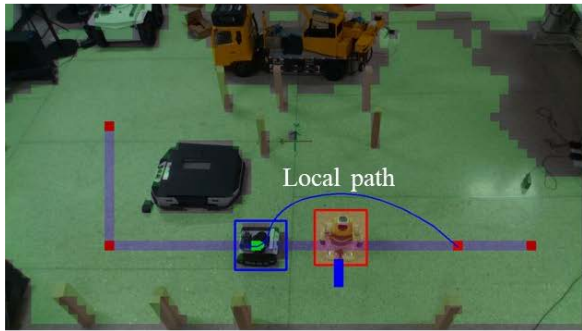
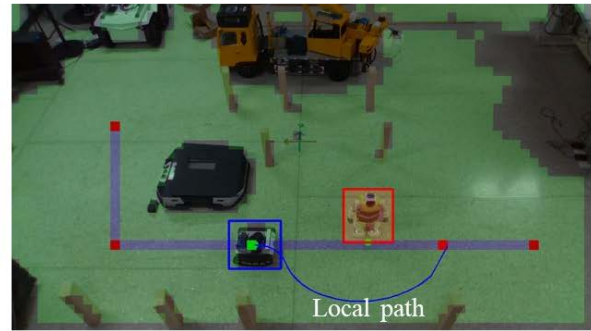


Figure 6. The experimental scenario for local path planning

Table 2 presents the comparison of prediction results for the worker and decisions for the robot with and without a sudden turn of the worker. Table 3 summarizes three groups of experiment results, including the path length. Path length means the length from the starting point to the target point, including locally planned paths.



(a) Dynamic planning with the predicted trajectory



(b) Dynamic planning without the predicted trajectory

Figure 7. The improved DWA

Table 1. Comparison of traditional and improved A*

Experiment	1	2	3
Time for traditional A* (s)	38.27	30.46	31.32
Time for improved A* (s)	13.05	12.96	11.42

Table 2. Comparison of prediction results for the worker and decisions for the robot

Experiment		1	2	3
Without a sudden turn	Predicted results	Success	Success	Success
	Decision for the robot	(2)	(3)	(1)
With a sudden turn	Predicted results	Failure	Failure	Failure
	Decision for the robot	(2)	(2)	(2)

Table 3. Comparison of traditional and improved DWA

Experiment	Path length (m)		Reduced weight of paths (%)
	Traditional DWA	Improved DWA	
1	5.41	4.97	8.1
2	5.49	5.01	8.7
3	6.04	5.17	14.4

4.3 Discussion

The results indicate that the construction robot can reach the target point by avoiding the moving workers, ensuring path safety. According to Table 1, the improved A* can enhance the efficiency of global planning. After fusing the safe space and the predicted worker trajectory, the improved DWA algorithm can find the path behind the worker, avoiding multiple planning and secondary collisions with the worker (see Figure. 7). The improvement further ensures path safety and path planning efficiency. Table 2 illustrates that the method can successfully predict the worker's trajectory without a sudden turn, while a sudden turn by a worker may lead to

failed predictions. In the case of failed predictions, the decision model will make the robot decelerate to avoid the collision.

According to Table 3, the improved method can achieve a shorter path, 8% less than the traditional A* and DWA methods. Meanwhile, the decision model can avoid unnecessary local planning, making the planning process more efficient. For example, although the robot and worker paths overlap in Figure. 6, they arrive at the PCZ at different times, which means there is no collision between them and no need for local path replanning.

5 Conclusion

This research proposes a construction robot path planning method to cope with the dynamic environment at construction sites. Specifically, a grid map incorporating a construction schedule is established, an improved A* is developed for global path planning, the safe space and worker trajectory prediction are considered to improve the DWA algorithm for local path planning, and a decision model is developed to enhance efficiency in local path planning. It is shown from the experiment that 1) the proposed method can find a global path efficiently, 2) the improved DWA can generate a safe path, avoiding secondary collisions with the worker, and 3) the decision model reduces robot local path replanning and better maintains globally optimal paths. Thus, the proposed method can effectively support construction robot path planning at construction sites.

The limitations of this research are summarized as well. Firstly, the experiment tests are conducted in a laboratory environment, which ignores some unpredictable disturbances. Secondly, only one worker and one robot are considered in this research. Future research will conduct experiments in the field and consider other disturbances, such as multiple robots, faster worker speeds and temporary piles of material.

Acknowledgment

We would like to thank the National Natural Science Foundation of China (Grant No. 52278310, 51578318) for supporting this research.

References

- [1] Bock T. The future of construction automation: technological disruption and the upcoming ubiquity of robotics. *Automation in Construction*, 59: 113–121, 2015.
- [2] Barbosa F., Mischke J., Parsons M. Improving construction productivity. On-line: <https://www.mckinsey.com/capabilities/operations/our-insights/improving-construction-productivity>, Accessed: 18/07/2017.
- [3] Cai S., Ma Z., Skibniewski M.J. *et al.* Construction automation and robotics for high-rise buildings over the past decades: A comprehensive review. *Advanced Engineering Informatics*, 42: 100989, 2019.
- [4] Ding L., Jiang W., Zhou Y. *et al.* BIM-based task-level planning for robotic brick assembly through image-based 3D modeling. *Advanced Engineering Informatics*, 43: 100993, 2020.
- [5] Wang Z., Li H., Zhang X. Construction waste recycling robot for nails and screws: computer vision technology and neural network approach. *Automation in Construction*, 97: 220–228, 2019.
- [6] Soltani A.R., Tawfik H., Goulermas J.Y. *et al.* Path planning in construction sites: performance evaluation of the Dijkstra, A*, and GA search algorithms. *Advanced Engineering Informatics*, 16(4): 291–303, 2002.
- [7] Wang B., Liu Z., Li Q. *et al.* Mobile robot path planning in dynamic environments through globally guided reinforcement learning. In *IEEE Robotics and Automation Letters*, pages 6932–6939, 2020.
- [8] Pinto A.M., Moreira E., Lima J. *et al.* A cable-driven robot for architectural constructions: a visual-guided approach for motion control and path-planning. *Autonomous Robots*, 41(7): 1487–1499, 2017.
- [9] Do H., Veerajagadeshwar P., Sun F. *et al.* Combined grid and heat conduction optimization for staircase cleaning robot path planning. *Automation in Construction*, 141: 104447, 2022.
- [10] Chen Z., Chen K., Song C. *et al.* Global path planning based on BIM and physics engine for UGVs in indoor environments. *Automation in Construction*, 139: 104263, 2022.
- [11] Stentz A. Optimal and efficient path planning for partially-known environments. In *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*, pages 3310–3317, 2014.
- [12] Ali M.S.A.D., Babu N.R., Varghese K. Collision free path planning of cooperative crane manipulators using genetic algorithm. *Journal of Computing in Civil Engineering*, 19(2): 182–193, 2005.
- [13] Yu L., Huang M.M., Jiang S. *et al.* Unmanned aircraft path planning for construction safety inspections. *Automation in Construction*, 154: 105005, 2023.
- [14] Duchoň F., Babinec A., Kajan M. *et al.* Path planning with modified A star algorithm for a mobile robot. *Procedia Engineering*, 96: 59–69, 2014.
- [15] Chen Y., Yu J., Su X. *et al.* Path planning for multi-UAV formation. *Journal of Intelligent & Robotic Systems*, 77(1): 229–246, 2015.
- [16] Xu C., Liu J., Wu Z. *et al.* Automated steel reinforcement detailing in reinforced concrete frames using evolutionary optimization and artificial potential field. *Automation in Construction*, 138: 104224, 2022.
- [17] Krishna Lakshmanan A., Elara Mohan R., Ramalingam B. *et al.* Complete coverage path planning using reinforcement learning for Tetromino based cleaning and maintenance robot. *Automation in Construction*, 112: 103078, 2020.
- [18] Cai J., Du A., Liang X. *et al.* Prediction-based path

- planning for safe and efficient human–robot Collaboration in Construction via Deep Reinforcement Learning. *Journal of Computing in Civil Engineering*, 37(1): 04022046, 2023.
- [19] Zhang W., Wang N., Wu W. A hybrid path planning algorithm considering AUV dynamic constraints based on improved A* algorithm and APF algorithm. *Ocean Engineering*, 285: 115333, 2023.
- [20] Duan P., Zhou J., Goh Y.M. Spatial-temporal analysis of safety risks in trajectories of construction workers based on complex network theory. *Advanced Engineering Informatics*, 56: 101990, 2023.
- [21] Rao A.S., Radanovic M., Liu Y. *et al.* Real-time monitoring of construction sites: sensors, methods, and applications. *Automation in Construction*, 136: 104099, 2022.
- [22] Rudenko A., Palmieri L., Herman M. *et al.* Human motion trajectory prediction: a survey. *The International Journal of Robotics Research*, 39(8): 895–935, 2020.
- [23] Zhu Z., Park M.-W., Koch C. *et al.* Predicting movements of onsite workers and mobile equipment for enhancing construction site safety. *Automation in Construction*, 68: 95–101, 2016.
- [24] Rashid K.M., Behzadan A.H. Risk behavior-based trajectory prediction for construction site safety monitoring. *Journal of Construction Engineering and Management*, 144(2): 04017106, 2018.
- [25] Arslan M., Cruz C., Ginhac D. Semantic trajectory insights for worker safety in dynamic environments. *Automation in Construction*, 106: 102854, 2019.
- [26] Tang S., Golparvar-fard M., Naphade M. *et al.* Video-based activity forecasting for construction safety monitoring use cases. In *Computing in Civil Engineering 2019*, pages 204–210, 2019.
- [27] Tang S., Golparvar-Fard M., Naphade M. *et al.* Video-based motion trajectory forecasting method for proactive construction safety monitoring Systems. *Journal of Computing in Civil Engineering*, 34(6): 04020041, 2020.
- [28] Cai J., Zhang Y., Yang L. *et al.* A context-augmented deep learning approach for worker trajectory prediction on unstructured and dynamic construction sites. *Advanced Engineering Informatics*, 46: 101173, 2020.
- [29] Guo H., Ye X., Ren Q. *et al.* Automatic generation of construction schedules based on BIM and rule reasoning. *Journal of Tsinghua University*, 62(2): 189–198, 2022.
- [30] The State Bureau of Quality and Technical Supervision. GB/T 13547-92 Human dimensions in workspace, China, 1992.
- [31] International Organization for Standardization. ISO/TS 15066 Robots and robotic devices-collaborative robots, Switzerland, 2016.
- [32] International Organization for Standardization. ISO 13855 Safe of machinery – Positioning of safeguards with respect to the approach speeds of parts of human body, Switzerland, 2010.
- [33] International Organization for Standardization. ISO 10218-2 Robots and robotic devices-safety requirements for industrial robots-part 2: robots systems and integration, Switzerland, 2011.
- [34] Arasaratnam I., Haykin S. Cubature kalman filters. *IEEE Transactions on Automatic Control*, 54(6): 1254–1269, 2009.

Factors Leading to Reduced Construction Productivity in Unmanned Construction

Genki Yamauchi¹, Takeshi Hashimoto¹, Mitsuru Yamada¹ and Shinichi Yuta²

¹Public Works Research Institute, Japan

²Shibaura Institute of Technology, Japan

yamauchi-g573bs@pwri.go.jp, t-hashimoto@pwri.go.jp, yamada-m@ceri.go.jp, yuta@shibaura-it.ac.jp

Abstract -

This study investigates factors affecting construction productivity in unmanned construction, a method developed in Japan for safely conducting recovery operations post-disaster using remotely operated machinery. Despite its widespread adoption, unmanned construction is typically less productivity than conventional methods. This research aims to identify and quantify the factors of this reduced productivity. Experiments were conducted using a hydraulic excavator in various operation environments, from manned to remote, focusing on visual information, operation interface, sensory information, and image display. The results reveal that the main factors decreasing productivity are differences in visual information for situational awareness and operating interface, with the former having a greater impact. Sensory information and image display differences were found not to be major contributors. The findings are specific to tasks similar to the model task used in the study and suggest the need for further research under various construction conditions to enhance productivity in unmanned construction, which is vital for disaster response and regular construction site productivity.

Keywords -

unmanned construction; remote control; construction equipment; working productivity

1 Introduction

Japan has experienced numerous disasters, including landslides due to heavy rain, earthquakes, and volcanic disasters. Recovery construction work post-disaster often takes place in areas at risk of secondary disasters. To safely conduct these recovery operations, a construction method has been developed using remotely operated construction machinery, allowing operators to work from a safe distance as shown in Figure 1. This method, known as "unmanned construction," was uniquely developed in Japan[1]. It was first introduced during the emergency recovery work of the Joganji River in 1969 and significantly evolved following the "Unzen Fugen Volcano Unmanned Construction Test Work" started in 1993[2]. Since then, it has been es-

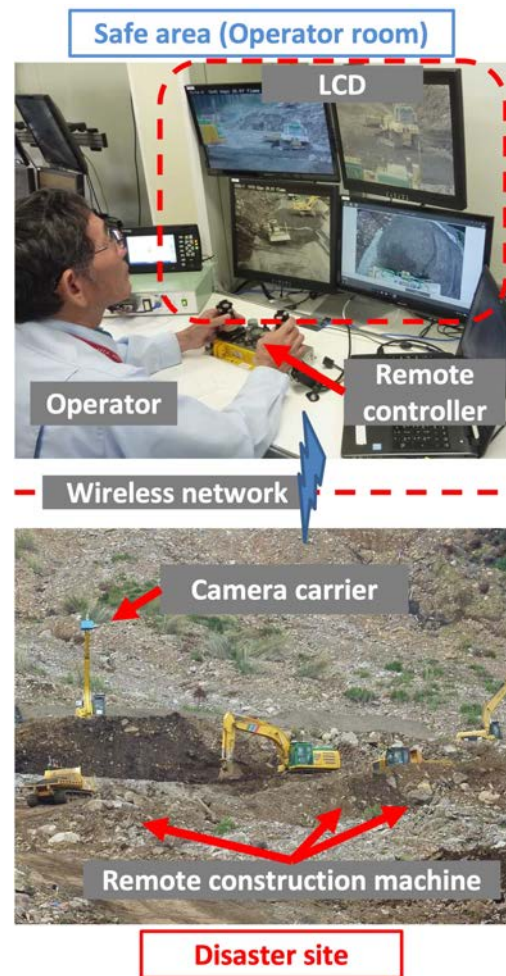


Figure 1. Overview of Remote Operation System

established as a practical construction method and adopted in various disaster sites, including the Usu Volcano eruption, the Niigata Chuetsu Earthquake, the Great East Japan Earthquake, the Kii Peninsula large-scale landslide disaster, and the Kumamoto Earthquake, with nearly 200 cases of implementation domestically[3][4].

However, it is generally said that unmanned construction is less productivity compared to conventional con-

Table 1. Examples of Factors Leading to Reduced Construction Productivity in Previous Studies

Factors Identified in Previous Studies	Classification in This Study
Delay in camera footage	1) Visual Difference
Insufficient resolution of camera footage	
Lack of depth perception in camera footage	
Insufficient camera angle	2) Operation I/F Difference
Short control levers	
Different response of control levers	
Different feedback in terms of sound and vibration	3) Sensory Information Difference

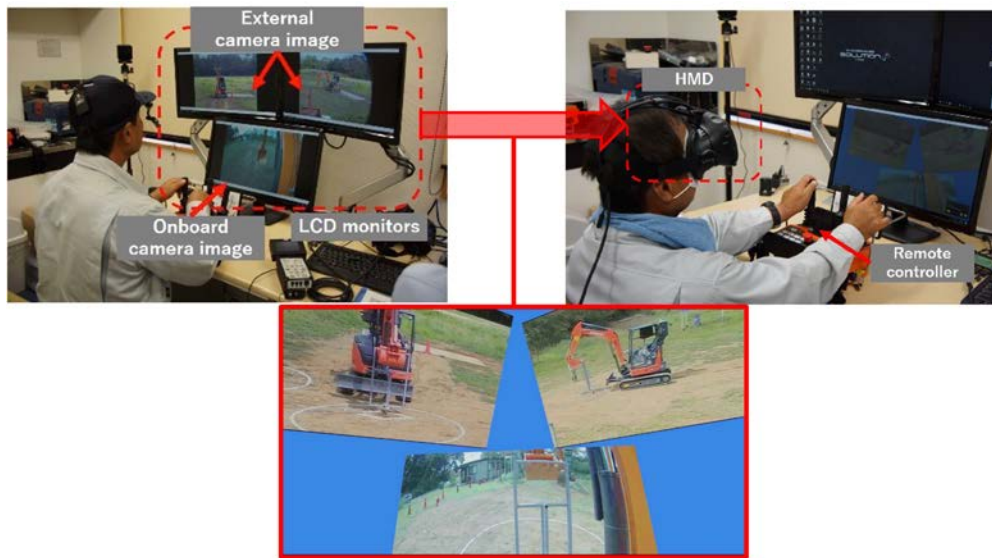


Figure 2. Visual System for Remote Operation. Left image shows conventional remote operation system using LCD, and Right image shows HMD based remote operation system

struction methods. Quantitative studies[5][6] have shown that the productivity of remote operation is about 45% of that during manned operation. To promote the adoption of unmanned construction and complete disaster recovery work safely and swiftly, it is necessary to improve the productivity of unmanned construction.

To improve construction productivity, it is essential to clearly identify the factors of productivity reduction in unmanned construction and focus resources on research and development to address these factors. However, while there are examples of identifying factors of productivity reduction through operator interviews[7], no study has quantitatively verified these factors. Therefore, this study aims to identify the factors of reduced productivity during unmanned construction by conducting experiments using actual remotely operated construction machinery.

2 Factors Analysis Based on Previous Studies

The reason for lower productivity in remote operation compared to manned operation is believed to be due to the

operator working in a different environment than during manned operation. The current standard remote operation in unmanned construction, as shown in Figure 1, is conducted inside a control room using a joystick controller for remote operation. The operator watches multiple displays showing footage from cameras mounted on the construction machinery (hereafter referred to as onboard cameras) and cameras installed at positions that provide an overall view of the work area (hereafter referred to as external cameras). Therefore, the differences between manned and remote operation can be broadly divided into the following three points:

1) Visual difference for situational awareness

In manned operations, situational awareness is obtained by the operator through direct visual observation from their seat inside the construction machinery. In remote operations, situational awareness is achieved by monitoring footage from both onboard and external cameras.

2) Difference in operation interface (hereafter referred to as operation I/F)

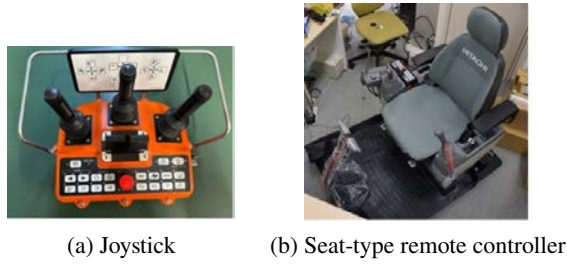


Figure 3. Operation Device for Remote Operation System

In manned operations, the operation is carried out using control levers located in the operator's seat within the construction machinery. In remote operations, the operation is conducted through a joystick controller designed for remote control.

3) Difference in sensory information

In manned operations, the operator can perceive sensory information such as the tilt of the machine, vibrations, and engine noise. In remote operations, this kind of sensory information is not available to the operator.

Previous studies[5][8][9][10][11] have identified various factors leading to reduced construction productivity. Representative factors are listed in Table 1 and classified into the above 1) to 3) categories.

In this experiment, we attempted to identify which of these differences contribute most to the reduction in construction productivity by gradually changing each of these differences from a manned operation environment to a remote operation environment and comparing the completion times of the same tasks under each condition.

3 Experiment Design

The experiment used a 12t hydraulic excavator capable of remote operation. The target environment for remote operation was set to be similar to the general unmanned construction environment shown in Figure 2.

To clearly identify the factors affecting construction productivity when changing the operation environment from manned to remote operation, it is important to minimize each change (for example, only change the operation I/F). Therefore, in this experiment, we utilized the HMD system developed in [12]) and the driver's seat-type remote controller developed in [13]) to minimize each change. HMD system is designed to replicate the visual setup of the three LCD monitors as shown in Figure 2. The driver's seat-type remote controller is composed of the same driver's seat and control levers as those of the hydraulic excavator, allowing for remote operation using these components as shown in

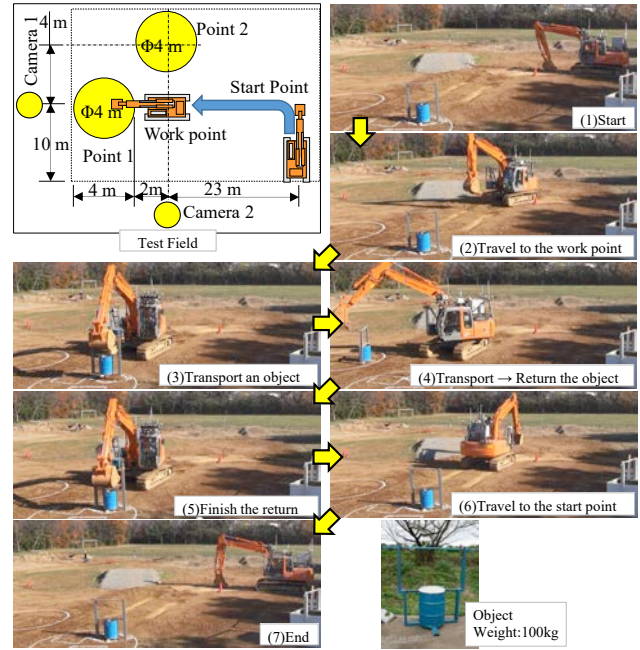


Figure 4. Outline of Model Task II[5]

Figure 3. Table 2 shows the patterns of changes in experimental condition. By utilizing the HMD and driver's seat-type remote controller, we created experimental condition changes in addition to the aforementioned 1) to 3) differences, including differences in image display.

Table 3 lists the equipment used in the experiment other than operation devices. The equipment and system used in the current standard unmanned construction are almost equivalent to those listed in Table 3, making it possible to understand the current factors of productivity reduction in unmanned construction through this experiment. However, as the HMD may be difficult to measure for some operators, as mentioned later, the experiment was conducted with caution.

The operators' age and work experience who participated in the experiment are listed in Table 4. The purpose of this study is to understand the factors of productivity reduction when operators with no experience in unmanned construction perform such operations for the first time. Therefore, we selected four operators who regularly operate hydraulic excavators but have no experience in unmanned construction. The lever operation patterns for manned operation, joystick and driver's seat-type controller were set to the commonly used patterns by all operators. The tasks performed under each pattern were based on the Model Task II[5] as shown in Figure 4, which involves a series of tasks using a hydraulic excavator, including traveling, moving objects, and replacing objects. Operators practiced this model task twice and performed it five times for a total of seven times under each pattern,

Table 2. Experiment Pattern






Pattern	Overview	Operation device	Vision device	Bodily sensation
a		Cabin Seat Controller	Direct View	Enable
b		Cabin Seat Controller	Footage by HMD	Enable
c		Seat-type Remote Controller	Footage by HMD	Disable
d		Seat-type Remote Controller	Footage by LCD	Disable
e		Joystick	Footage by LCD	Disable

Table 3. Equipments

Device	Model
Hydraulic excavator	Hitachi ZX120
Onboard camera	AXIS Q1615-MKII
External camera	AXIS Q6155-E
LCD Monitor	iiyama X2382HS
HMD	HTC Vive

Table 4. Operator Information

Operator	Age	Experience (year)
A	49	24
B	64	30
C	65	20
D	64	44

and the completion time of the task (hereafter referred to as cycle-time) was measured.

4 Experiment Result

Table 5 shows the average of five cycle times under each experimental pattern. Since there are differences in cycle times among operators, the data is presented as a percentage of the average cycle time of five runs under

experimental pattern a (manned operation), referred to as the cycle time ratio, shown in Figures 5. The cycle time ratio was normalized with the results of pattern A set as 100. For Operator B's pattern b, data could not be obtained due to strong discomfort when wearing the HMD. Similarly, for Operator D's pattern b, only one run could be conducted due to the same reason, so only one data point is presented. T-test was conducted using the data of five cycle times to determine if the difference in cycle time ratio due to the difference in experimental patterns was significant. A significance probability p less than 0.05 was considered significant, and a star mark ★ was added to the graph. However, for Operator D's pattern a and b, and between b and c, as mentioned above, since there was only one data point for pattern b, a t-test was not conducted, and a star mark was added to the graph. According to Figures 5, the average cycle time ratio under experimental pattern e is 220.6, and the construction productivity compared to experimental pattern a (manned operation) is about 45%. This is consistent with the results of previous studies. Furthermore, the differences in cycle time ratios between experimental patterns are shown in Table 6. Positive values indicate an increase in cycle

time, i.e., a reduction in construction productivity, while negative values indicate an improvement in construction productivity. Similar to the graphs, environments with significant differences are marked with a star ★.

5 Discussion

Based on Table 6, we examined how the differences in each operation environment affected the reduction in construction productivity in this experiment.

- **Difference in Visual Information (Experimental pattern a→b)**

According to Table 6, among the three operators for whom data was available, two showed a significant difference in cycle time ratio due to the difference in visual information, with an increase in the value (approximately 33-43%). For the remaining operator, the difference in cycle time ratio also showed a significant increase (approximately 94%). This suggests that the difference in visual information is a major factor in reduced construction productivity.

- **Difference in Sensory Information (Experimental pattern b→c)**

According to Table 6, among the three operators for whom data was available, two did not show a significant difference in cycle time ratio due to the difference in sensory information. For the remaining operator, the cycle time ratio decreased, indicating an improvement in construction productivity, and the difference was very small (approximately 2%). This suggests that the difference in sensory information is not a major factor in reduced construction productivity.

- **Difference in Image Display (Experimental pattern c→d)**

According to Table 6, none of the operators showed a significant difference in cycle time ratio due to the difference in image display. This suggests that the difference in image display is not a major factor in reduced construction productivity.

- **Difference in Operation I/F (Experimental pattern d→e)**

According to Table 6, among the four operators, one did not show a significant difference in cycle time ratio due to the difference in operation I/F. However, for all other operators, there was a significant difference in cycle time ratio due to the difference in operation I/F, with an increase in the value (approximately 25-105%). This suggests that the difference in operation I/F can be a major factor in reduced construction productivity for some operators.

From these results, it can be said that the major factors in the decrease of construction productivity in unmanned

construction are differences in visual information and operation I/F.

6 Conclusion

To examine the factors leading to a decrease in construction productivity in unmanned construction, remote operation experiments were conducted using the model task. From these experimental results, the following points were clarified:

- 1) The major factors in the decrease of construction productivity during unmanned construction are "differences in visual information for situational awareness" and "differences in operating interface".
- 2) "Differences in sensory information" and "differences in image display" are not major factors in the decrease of construction productivity.

The above results 1) and 2) are specific to tasks like the model task involving excavation work with hydraulic excavators, and for other tasks such as off-road driving or breaking work with breakers, sensory information like machine tilt, sound, and vibration may become important. Further examination under various construction patterns is desired in the future. Furthermore, we have incorporated 'cycle time' as a criterion for evaluation within the framework of our proposed model in Task II, undertaking an analysis of the elements that diminish construction efficiency. However, 'cycle time' permits only the appraisal of the entire task, with assessments of more granular details within the task being constrained. For precise evaluations, future research should explore methodologies that consider the operator's visual perspective, movements of the control lever, and hydraulic dynamics.

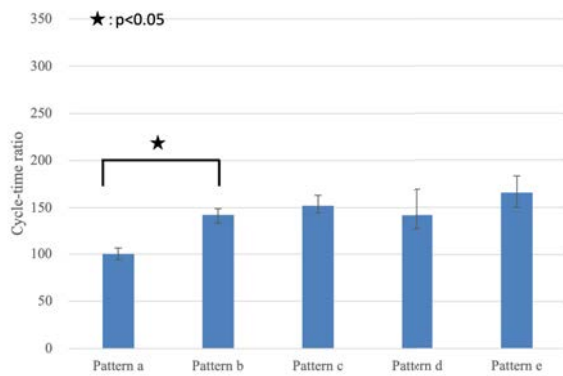
Unmanned construction is expected to be a technology that contributes not only to disaster response but also to improving productivity in regular construction sites through work style reform for operators and increased productivity through day and night work. Therefore, improving construction productivity is essential, and further research is desired.

Table 5. Cycle-time ratio comparison between each pattern [s]

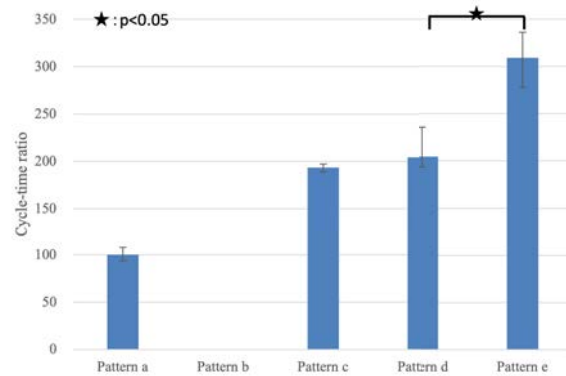
Operator	Pattern a	Pattern b	Pattern c	Pattern d	Pattern e
A	166	237	253	236	274
B	161	-	310	328	497
C	203	270	267	279	331
D	184	358	355	336	453

Table 6. Cycle-time ratio comparison between each pattern

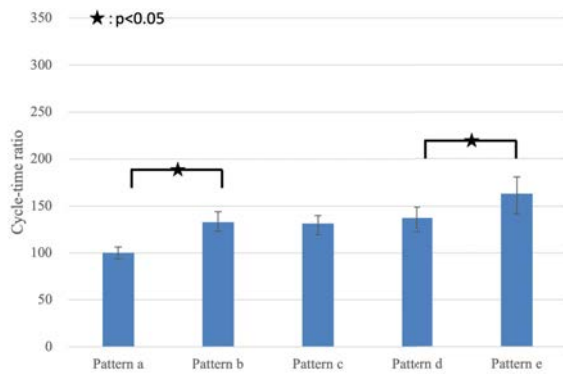
Comparison Factor	Pattern a→b	Pattern b→c	Pattern c→d	Pattern d→e
	Vision difference	Body sensation difference	Visual device difference	I/F difference
Operator A	42.6★	9.7	-10.1	23.1
Operator B	-	-	11.2	104.8★
Operator C	32.8★	-1.5	6.2	25.3★
Operator D	94.1	-1.7	-10.3	63.5★



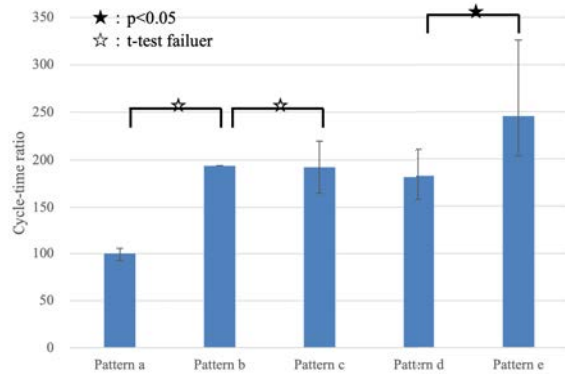
(a) Operator A's result



(b) Operator B's result



(c) Operator C's result



(d) Operator D's result

Figure 5. Experimental results of cycle-time ratios by each operator

References

- [1] Kazuhiro Chayama, Akira Fujioka, Kenji Kawashima, Hiroshi Yamamoto, Yasushi Nitta, Chikao Ueki, Atsushi Yamashita, and Hajime Asama. Technology of unmanned construction system in japan. Journal of Robotics and Mechatronics, 26(4):403–417, 2014.
- [2] Yuji Hiramatsu, Takashi Aono, and Masami Nishio. Disaster restoration work for the eruption of mt usuzan using an unmanned construction system. Advanced Robotics, 16(6):505–508, 2002.
- [3] Eiji Egawa, Kensuke Kawamura, Masaharu Ikuta, and Takayuki Eguchi. Use of construction machinery in earthquake recovery work. Hitachi Review, 62(2): 136–141, 2013.
- [4] Takeshi Hashimoto. Unmanned construction system for disaster response. In JAPAN-AMERICA Frontiers of Engineering Symposium, 2014.
- [5] Masamaru Moteki, Shinichi Yuta, and Kenichi Fujino. The proposal of the model task for efficiency evaluation of the construction work by remote control of a hydraulic excavator. Journal of JCMA, 66 (8):71–79, 2014 (in Japanese).
- [6] Takeshi Hashimoto, Kenichi Fujino, and Shinichi Yuta. A study of the working efficiency of a remote control type hydraulic excavator. In Proceedings of International Society for Terrain-Vehicle Systems (ISTVS), 2017.
- [7] Takeshi Yamaguchi., Tadashi Yoshida., and Yutaka Ishimatsu. Field survey on man-machine interfaces in remote operations (in japanese). Proceedings of the 59th Annual Meeting on Japan Society of Civil Engineers (2004), pp. 373–374, 2004.
- [8] Sadanori Ito, Yuichi Sakano, Kenichi Fujino, and Hiroshi Ando. Effects of remote operation visual environment on work efficiency in unmanned construction (in japanese). Japan Society of Civil Engineers, 73(1):15–24, 2017.
- [9] Hiroshi Furuya, Nobuo Kuriu, and Chiharu Shimizu. Development of next-generation remote-controlled machinery system. Report of obayashi corporation technical research institute, (76):1–10, 2012.
- [10] Takeshi Hashimoto, Genki Yamauchi, Kenichi Fujino, Shinichi Yuta, and Kazuyoshi Tateyama. Study of operator’s line of sight in unmanned construction systems. In 2018 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR), pages 1–6, 2018.
- [11] Masaru ITO, Toshio TSUJI, Yuichi KURITA, Seiji SAIKI, and Yoichiro YAMAZAKI. Hydraulic excavator remote operation system using seat vibration feedback. The Robotics and Mechatronics Conference, (0):2A2–B03, 2018.
- [12] Genki Yamauchi, Takeshi Hashimoto, and Shinichi Yuta. Assessment of work efficiency of hmd viewing system for unmanned construction work. In ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction, volume 36, pages 824–830. IAARC Publications, 2019.
- [13] Masaharu Moteki, Akihiko Nishiyama, Takeshi Hashimoto, Kenichi Fujino, and Shinichi Yuta. Improvement of work efficiency due to differences in visual and operational interfaces in remote operation of hydraulic excavators (in japanese). The 16th Symposium on Construction Robotics in Japan, O6-4, 2016.

A Framework and Cyber-Physical System Architecture for Cloud-Based Construction Monitoring with Autonomous Quadrapeds

Aras Maqsoodi and Javier Irizarry

School of Building Construction, Georgia Institute of Technology, United States
 aras.maqsoodi@gatech.edu, javier.irizarry@gatech.edu

Abstract –

Construction progress monitoring is a crucial aspect of project management, ensuring project completion on time, within budget, and meeting quality standards. Current data collection and processing methods are time-consuming and labor-intensive, and modern technologies, including robots and computer vision, offer a potential solution to this challenge. While past research has improved data collection and processing independently, a significant manual effort is still required to classify and transform collected data for processing. This research introduces the foundation of a novel framework to bridge this gap, leveraging an autonomous four-legged robot using the Software Development Kit (SDK) and the Robot Operation System (ROS). Additionally, Simultaneous Localization and Mapping (SLAM) for navigating, and cloud computing platforms for data processing, are employed to establish a repository for raw data from images captured by the robot, facilitating simultaneous processing stages. The research proposes a cloud-based machine-learning service that hosts a deep-learning algorithm for material identification and quantification. The objective of the proposed framework and system architecture is to capture images of the indoor environment of construction projects, classify and process data to create quantified material lists, and compare them to the full material list on the cloud, construction schedules can potentially be updated in real-time.

Keywords –

Computer Vision; Cloud-Based Construction Monitoring; Autonomous Quadraped; Framework; Cyber-Physical System Architecture

1 Introduction

The construction industry faces challenges in adopting novel technologies, rendering it unproductive

and inefficient [1, 2]. Despite the improvement rates observed in other sectors, such as manufacturing and information technology, over the past decades, the construction industry's productivity remains stagnant [3, 4]. Technology-driven systematic approaches, such as construction automation and robotics, have been introduced to enhance efficiency, constructability, planning, and overall productivity [5].

Limited real-time data and analysis from construction environments result in a gap between ongoing project stages and decisions made by top managers [6, 7]. While traditional methods such as human interactions and paper-based task recording have long been relied upon for many project monitoring tasks, there is a growing trend among project participants to adopt technology and modern tools [8, 9]. Considering the industry's current status, efficiency in the data-collecting process is crucial to support managers in making decisions grounded in accurate data and up-to-date information [10, 11].

Effective and regular inspections, along with quantifying completed and ongoing construction tasks, are essential and require innovative strategies. Technology-driven systematic approaches, such as construction automation and robotics, have been introduced to enhance efficiency, constructability, planning, and overall productivity [5]. Optimizing the data-collecting process is crucial to assist managers in making decisions based on accurate and updated information [10], and data-collection automation can facilitate construction monitoring [12].

To evaluate the effectiveness of the proposed application of robots, specifically utilizing quadrapeds for live data collection and updating construction schedules, a survey was conducted among academic researchers and professionals in the construction industry as part of another research study [13]. There is data available to verify the benefits of the proposed utilization of the robot and its ability to perform required tasks (Figure 1). Current applications of quadrapeds in this domain mostly require manual navigation for tasks such as collecting data or remote inspection. However, the

scope of this paper, as a prerequisite for a larger research project, is to present the foundation and hypotheses for the autonomous operation of the robot.

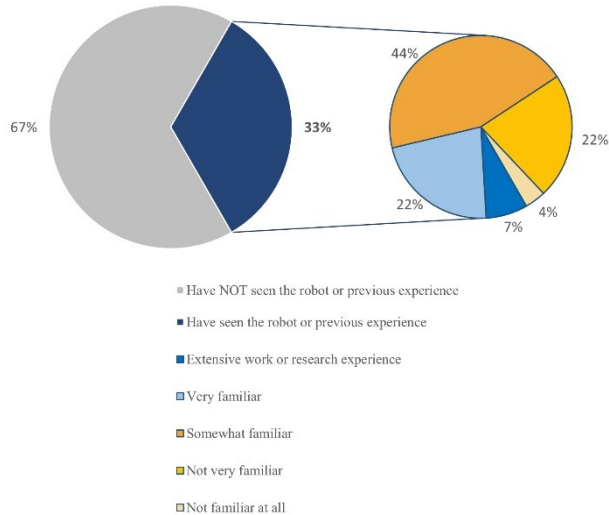


Figure 1. Survey results on utilizing the robot for the assigned task

While several studies have explored the use of collected data from the construction environment for construction monitoring, a gap still persists between data collection and the processing and updating of construction schedules by an autonomous robot. This paper contributes to the existing body of knowledge by proposing a framework and a cyber-physical system architecture to leverage Spot, a four-legged autonomous robot, for real-time construction schedule updates. A preliminary workflow was introduced in a prior study [4], and this paper presents the required developed framework along with the necessary system architecture.

2 Research Significance

The primary objective of this research is to propose the fundamental components of an automated method for collecting, classifying, and processing data to update construction schedules in real-time. The system is designed to automatically update construction schedules, with a specific emphasis on enhancing efficiency and accuracy within indoor construction environments and eliminating the lag between 'actual' and 'reported' progress.

Despite advancements in data collection and processing within indoor construction environments, a notable gap persists, requiring substantial manual effort to classify and prepare collected data for further processing. This research contributes to the existing body of knowledge by proposing the initial steps of addressing this critical gap, creating the foundation of a platform for

streamlining the process, and facilitating real-time updates to construction schedules and well-detailed information, preventing possible delays and cost overruns as the final goal of future studies [10].

3 Methodology

A comprehensive review of existing literature, involving the search, selection, and analysis of relevant sources, was conducted to understand the functionality of the robot. Since the implementation of the idea of using a quadruped for regular inspection and real-time updating of construction schedules is complex, a meticulous selection of system components and architecture is necessary. Part of the selection process was based on the authors' experience in other research, while the remaining parts were informed by literature review.

Additionally, as the study focuses on a specific type of autonomous robot for construction monitoring and planning, technical specifications, manuals, and manufacturer's websites were consulted. The application of autonomous robots as the selected model was thoroughly investigated, taking into account both challenges and benefits.

To establish a feasible goal for the study, the previously proposed workflow [4] underwent updates through further research and optimization. Furthermore, aligned with the required system performance and anticipated outcomes, a cyber-physical system architecture was designed to provide a platform for the next step of the research, which involves developing a Construction Material Library (CML) using Convolutional Neural Networks (CNNs).

4 Literature Review

The use of modern technologies and the improvement in data collection methods have enabled project managers to reduce the gap between ongoing project stages and crucial decisions [14]. Given the inefficiency, labor-intensive, and error potential of traditional data collection methods [15], innovative solutions, along with regular and effective inspections, play a pivotal role in identifying schedule deviations [10]. This approach facilitates resource allocation to prevent delays and cost overruns in construction projects [1, 11, 12].

Computer vision has emerged as a promising alternative solution in recent years, providing automated and continuous monitoring capabilities through image and video capture. This approach offers rich insights into project entities, behaviors, and site conditions, facilitating a more comprehensive understanding of complex construction tasks [16]. Machine learning algorithms have streamlined the progress monitoring process, employing the analysis of as-built images for

accurate calculations [17].

Fard & Peña-Mora [18] utilized captured images from construction environments for camera matching and material detection to assess and identify physical progress in projects. While this approach simplifies procedures, reducing both time and costs, it necessitates the use of handheld cameras and personnel for image classification and alignment with planned models. Recognizing the limitations of this manual process, there is an emerging need to incorporate advanced equipment to minimize human intervention throughout the entire process [19-21].

In another research, Golparvar-Fard et al. [10] used captured images to create an as-built 4D BIM model and proposed to compare as-planned and as-is models to recognize construction progress. Vähä et al. [22] emphasized that the integration of cutting-edge tools, such as easy-to-use robots in construction automation, would deliver benefits in terms of value, safety, quality, productivity, and overall performance.

Petersen et al. [23] explored various categories of construction robots, classifying them based on equipment, performance, and outcomes. Moreover, Asadi et al. [14, 24] demonstrated how a mobile wheel-enabled robot can identify obstacles, deliver real-time vision-based data in particular conditions, and be integrated with other tools like unmanned aerial vehicles in hard-to-reach areas. A system architecture for a Multi-Agent Robotic System is proposed for path planning in indoor environments [25]. It demonstrated its capabilities by assisting in automated path planning and navigation for collecting 2D image and 3D point cloud data.

An autonomous mobile robot equipped with a navigation and drift correction algorithm was employed to collect data from an indoor environment and incorporate it into the decision-making process [26]. While ground mobile wheel-enabled robots exhibit proficiency in collecting data from indoor environments and elements situated on the same building story, the primary challenge lies in autonomous travel between levels without human intervention [23, 27]. Taking into account mobility and diverse capabilities, legged robots offer advantages compared to wheeled or other ground mobile robot types. Multi-legged systems play a crucial role in attaining superior mobility performance, energy efficiency, speed [27, 28], and obstacle negotiation capabilities [29].

This research proposes a workflow and a cyber-physical system architecture to leverage Spot, an autonomous quadruped robot, for collecting and processing data to update real-time construction schedules in real-time. The robot has been employed in construction-related research, demonstrating its potential and capabilities in indoor environments [30-32].

5 Framework

Nearly half of the construction cost is associated with building materials, and effective material monitoring can facilitate the accomplishment of project goals [32, 33]. This research proposes a framework that employs Spot as the central component of a system designed to recognize materials, assess physical progress by comparing the implemented materials list with the full materials list, and update construction schedules in real-time. The robot's task is to autonomously navigate across different stories, inspect designated areas on each floor, and collect data for real-time processing. Given the system's complexity with various subsystems and components, this paper focuses on the foundational aspects of the system, particularly recognizing components and their interactions to achieve the required performance.



Figure 2. Boston Dynamic's SPOT with 360° camera and a LiDAR payload.

The robot can navigate diverse terrains while simultaneously sensing and analyzing its environment [34]. Equipped with sensor modules on its front, rear, and sides, including stereo cameras, a wide-angle camera, and a texture projector, the robot utilizes these sensors to facilitate SLAM navigation and wayfinding, enabling autonomous movement across various building stories and areas [35]. As an autonomous quadruped, Spot is equipped with ten cameras, comprising five optical cameras for operator view and five depth cameras for robot perception and obstacle avoidance (Figure 3).



Figure 3. Front, side, and rear sensors and wide-angle cameras

This configuration provides Spot with a comprehensive optical Field of View (FoV) and a depth

camera range of approximately 2 meters [35]. However, since the robot's cameras do not offer sufficient quality for the purposes of this research, and considering issues with lighting and shadows in captured images [36, 37], a powerful inspection tool such as the Spot CAM+ is required. Featuring a payload 360° spherical camera, it ensures accurate and high-resolution images (Figure 4), facilitating the material recognition process [38].



Figure 4. Spot CAM+ payload, and spot Core I/O and Velodyne VLP-16 LiDAR

Utilizing LiDAR as additional sensing detectors (Figure 4), the robot can create 3D point clouds of environments, which are beneficial for obstacle detection and avoidance, as well as secure and efficient pathfinding [37, 39]. While research has explored using dynamic point cloud scans, the high level of noise and the extensive computational processes required for the robots [40] make this method inefficient for our goal, which is to update construction schedules in real-time during progress monitoring. Due to data size and computational limitations, LiDAR can only serve as a navigation tool. Instead, semantic data extracted from 360° camera RGB images, along with their locations, can identify implemented materials in specified locations and compare them with the total amount to calculate overall progress.

The current study represents the initial phase of a larger project aimed at collecting data from indoor construction environments, identifying physical progress, and updating construction schedules in real-time. The proposed method involves directing the robot along predefined paths to regularly visit designated locations, capturing images and collecting data. The framework utilizes cloud services and CNNs to process the data and update the construction schedule promptly (Figure 5).

Autonomous quadrupeds, exemplified by Boston Dynamics Spot, have demonstrated their capabilities in construction environments [41, 42]. In our approach, Spot is considered to perform two parallel sets of actions: data collection and data processing. During data collection, the quadruped utilizes SLAM and the robot's

camera, Spot CAM+, to navigate specified areas, capture images, and upload them to the cloud.

Amazon SageMaker, as a cloud-based platform hosting the CNN, is utilized for material identification and quantification through a CML. The quadruped is tasked with regularly visiting specified areas to compare the used materials with the comprehensive material list, thereby identifying physical progress and updating construction schedules in real-time.

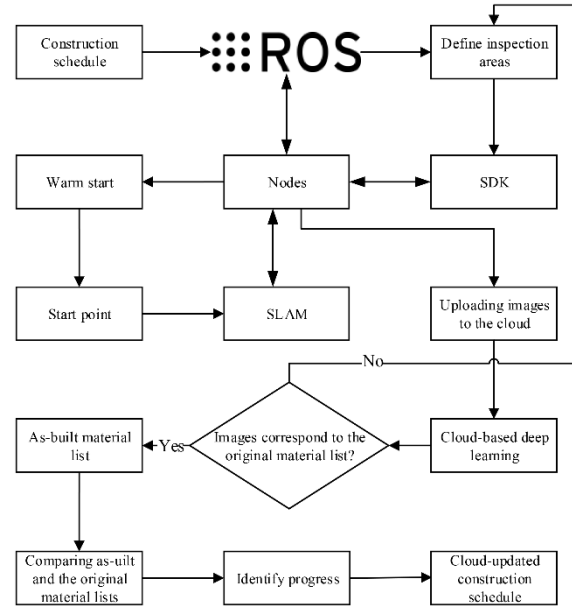


Figure 5. Spot's Real-Time Construction Schedule Updates Workflow

In the proposed framework, ROS Noetic, running on Ubuntu 20.04, serves as the primary framework for the robot, facilitating communication among various components. It ensures coordination of nodes, such as sensing cameras and LiDAR [43], for traveling between locations and collecting accurate information from specified areas or uploading captured data to a cloud computing platform like Amazon Web Services (AWS). Users have the capability to program and interface with the robot using the SPOT SDK [37, 43].

Autonomous robots can navigate indoor environments using either a pre-built map or live sensor data. During the warm start, the quadruped employs SLAM in the initial step and navigates through the designated area, either manually or autonomously, to construct an offline map that includes all physical components in the environment, such as obstacles and available paths [44]. Spot uses Velodyne VLP-16 LiDAR for localization [37], allowing the robot to determine its position and employ grid-based path planning to create a path. Subsequently, the robot utilizes live sensor data and an autonomous navigation and mapping library, such as

RTABMap, for further localization and navigation through specified requested areas [45, 46].

The robot can visit requested areas, capture images, leverage a cloud computing platform such as AWS, and upload the collected images to the cloud [47]. This framework relies on computer vision to extract semantic data or utilize filters to cluster images based on their physical features, enabling material classification through a CNN [48]. The algorithm hosted by SageMaker can be trained for material classification, enabling the robot to recognize and quantify materials based on their locations [49]. This facilitates the subsequent step of comparing implemented material lists with full material lists, considering start and finish dates, locations, and recognizing construction progress [48].

The proposed cyber-physical system architecture for this research implements the designed framework and integrates the required sensors' performance of Spot into ROS (Figure 6). It establishes a closed loop for acquiring data from nodes, which include the VLP-16 LiDAR for wayfinding and message transmission in ROSCORE, and the CAM+ 360° camera for capturing RGB images.

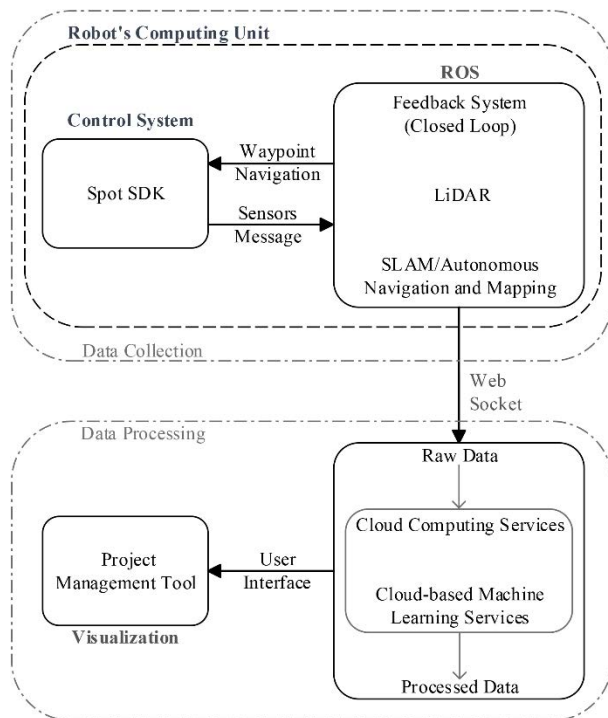


Figure 6. Spot's Cyber-Physical Systems Architecture for Real-Time Construction Schedule Updates

In terms of autonomous wayfinding, SLAM libraries like RTAB-Map enable Spot to map and navigate its environment effectively. Given that LiDAR is designated for wayfinding in this framework, the RTAB-Map library processes 3D scans generated by the VLP-16 [50].

During the initialization phase, ROS undergoes registration, and publishers are instantiated. As the process unfolds, the program enters a loop dedicated to SDK calls for data acquisition. The collected data undergoes processing, is packaged into ROSCore messages, and subsequently published [37]. Since the proposed CNN for the next phase of this research should consider a wide range of Hue-Saturation-Value (HSV) color values for construction materials due to different types of shadows and lighting conditions, a large dataset is needed [48]. AWS and SageMaker provide an appropriate platform for our cloud-based model training for the future step of the study [49].

The system architecture comprises nodes responsible for uploading captured images to the cloud and utilizes a cloud-based machine learning service like SageMaker for image organization. A deep learning algorithm hosted on the cloud is employed to train the dataset, classify, and quantify materials within images captured in construction environments [51]. With data collected and processed in the cloud, outputs such as implemented material lists can be generated. This enables the updating of construction schedules across various interfaces, such as Primavera or MS Project.

6 Discussion and Limitation

In recent years, novel technologies and equipment have been employed for data collection in construction environments. Despite this, a substantial portion of the necessary data for construction monitoring and progress assessment still relies on personnel. Through routine inspections and comparisons between actual progress and planned objectives, construction schedules can be updated, guided by inspection sequences and the availability of construction workers [10]. However, the collected data must be organized and classified for various applications, including construction monitoring. The manual nature of these processes highlights the need for innovative solutions to minimize human intervention during both the data collection and processing stages [19].

Regular inspections and schedule updates can identify schedule deviations early, enabling managers to allocate resources effectively, enhance overall performance, and prevent potential delays and cost overruns [10]. Computer vision has emerged as an innovative and successful solution for updating construction schedules, offering automation in regular construction monitoring through the use of captured images and videos from construction environments for further processing [16].

Utilizing autonomous robots for data collection and processing captured images to identify construction progress improves project efficiency by enabling appropriate workforce allocation [12, 22]. Automated

data collection and processing have the potential to enhance construction labor productivity and add value to project performance [52]. Additionally, automated construction monitoring eliminates the gap between task completion, reporting, and approval. It has the potential to streamline financial transactions for invoices and bridge gaps [53].

The primary focus of this study centered on an in-depth review of existing literature and the authors' knowledge regarding the integration of autonomous robots in construction monitoring, including their experience with similar experiments. However, due to the limited body of research available, the efficacy of the suggested framework requires further investigation through upcoming experiments within construction projects. Additionally, the successful deployment of quadruped robots in indoor settings necessitates careful consideration of existing limitations, technology acquisition costs, operational challenges, ambient environmental factors, and safety considerations associated with human-robot collaboration.

Evaluating the long-term sustainability of benefits in adopting a quadruped robot necessitates project stakeholders to comprehensively assess cost factors, covering initial investment, operational expenses, and maintenance costs over its entire lifespan, rather than merely throughout a specific project [54]. In addition, by ensuring ambient lighting in the construction environment [35, 36, 42, 55], the robot can be utilized during sparsely populated night hours [56], improving safety and efficiency and adding value to the project's daily performance and outputs.

During the construction phase, the application of quadrupeds may encounter operational limitations in tasks involving elevated heights, such as inspecting steel framing, building facades, and false ceilings [35, 41]. Optimal lighting conditions are crucial for the robot's effective operation, as it may struggle to detect obstacles in shadowy areas and transparent elements like glass doors and windows [35, 36, 55, 56]. Consideration of these factors ensures safe and optimal performance in diverse environments and enables the construction industry to explore new technological advancements to improve construction monitoring processes.

7 Conclusion and Future Research

Construction robotics has experienced remarkable progress, marked by notable advancements. Despite this progress, persistent technical challenges include functional integration, localization, planning, guiding algorithms, sensor technology, and robot intelligence [57]. Nevertheless, the construction industry's unstructured, complex, and dynamic nature is being effectively addressed by recent advances in artificial

intelligence, incorporating machine learning and deep learning methods [58]. This research has proposed the initial step to bridge the gap between collecting data and processing and updating construction schedules, eliminating the lag between actual and reported progress in construction environments.

Additional steps are required to Implementing the proposed framework and system architecture for a robot in construction settings. Moreover, customizing the robot for a specific construction site would aid in assessing its functionality and operational efficiency for project monitoring and scheduling. Furthermore, investigating the integration of diverse unmanned vehicles alongside the robot, such as ground vehicles, drones, and swarm systems, either individually or in conjunction, can help overcome the limitations of quadrupeds.

References

- [1] Bock T. The future of construction automation: Technological disruption and the upcoming ubiquity of robotics. *Automation in construction*. 2015;59:113-21.
- [2] Construction Labor Productivity [Internet]. Office of Productivity and Technology. 2023 [cited Apr 5]. Available from: <https://www.bls.gov/productivity/highlights/construction-labor-productivity.htm>.
- [3] Abioye SO, Oyedele LO, Akanbi L, Ajayi A, Delgado JMD, Bilal M, et al. Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges. *Journal of Building Engineering*. 2021;44:103299.
- [4] Maqsoodi A, Irizarry J. Proceedings of the 46th AUBEA Conference: A conceptual workflow for live updating construction schedules by collecting and classifying data using a four-legged autonomous robot. *AUBEA 2023: Springer Nature*; 2024, [in press].
- [5] Sawhney A, Riley M, Irizarry J, Riley M. *Construction 4.0*. Sawhney, A, Riley, M, Irizarry, J, Eds. 2020.
- [6] Teizer J, Lao D, Sofer M, editors. Rapid automated monitoring of construction site activities using ultra-wideband. *Proceedings of the 24th International Symposium on Automation and Robotics in Construction*, Kochi, Kerala, India; 2007.
- [7] Teizer J. Status quo and open challenges in vision-based sensing and tracking of temporary resources on infrastructure construction sites. *Advanced Engineering Informatics*. 2015;29(2):225-38.
- [8] Woo J, Shin S, Asutosh AT, Li J, Kibert CJ, editors. An overview of state-of-the-art technologies for data-driven construction. *Proceedings of the 18th*

- International Conference on Computing in Civil and Building Engineering: ICCCB E 2020; 2021: Springer.
- [9] Martin M, Banijamali K, Kazemian A. Reality Capture Technologies for Automated Quality Control during Construction 3D Printing. *Computing in Civil Engineering* 2023. p. 361-70.
- [10] Golparvar-Fard M, Peña-Mora F, Savarese S. D4AR—a 4-dimensional augmented reality model for automating construction progress monitoring data collection, processing and communication. *Journal of information technology in construction*. 2009;14(13):129-53.
- [11] Navon R, Sacks R. Assessing research issues in automated project performance control (APPC). *Automation in construction*. 2007;16(4):474-84.
- [12] Navon R, Goldschmidt E. Monitoring labor inputs: automated-data-collection model and enabling technologies. *Automation in construction*. 2003;12(2):185-99.
- [13] Maqsoodi A. Autonomous Robot-Enabled Data Collection, Classification, and Processing Method for Real-Time Construction Schedule Updating [Unpublished master's thesis]: Georgia Institute of Technology; 2023.
- [14] Asadi K, Suresh AK, Ender A, Gotad S, Maniyar S, Anand S, et al. An integrated UGV-UAV system for construction site data collection. *Automation in Construction*. 2020;112:103068.
- [15] Lin JJ, Golparvar-Fard M. Visual and virtual progress monitoring in Construction 4.0. *Construction 40*: Routledge; 2020. p. 240-63.
- [16] Seo J, Han S, Lee S, Kim H. Computer vision techniques for construction safety and health monitoring. *Advanced Engineering Informatics*. 2015;29(2):239-51.
- [17] Baduge SK, Thilakarathna S, Perera JS, Arashpour M, Sharafi P, Teodosio B, et al. Artificial intelligence and smart vision for building and construction 4.0: Machine and deep learning methods and applications. *Automation in Construction*. 2022;141:104440.
- [18] Fard MG, Peña-Mora F. Application of visualization techniques for construction progress monitoring. *Computing in Civil Engineering* (2007)2007. p. 216-23.
- [19] 19. Kopsida M, Brilakis I, Vela PA, editors. A review of automated construction progress monitoring and inspection methods. *Proc of the 32nd CIB W78 Conference* 2015; 2015.
- [20] Paneru S, Jeelani I. Computer vision applications in construction: Current state, opportunities & challenges. *Automation in Construction*. 2021;132:103940.
- [21] Banijamali K, Vosoughi P, Arce G, Noorvand H, Hassan M, Kazemian A, editors. Early-Age Strength Monitoring of Sensor-Embedded 3D Printed Structures. *Construction Research Congress* 2024; 2024.
- [22] Vähä P, Heikkilä T, Kilpeläinen P, Järviluoma M, Gambao E. Extending automation of building construction—Survey on potential sensor technologies and robotic applications. *Automation in construction*. 2013;36:168-78.
- [23] Petersen KH, Napp N, Stuart-Smith R, Rus D, Kovac M. A review of collective robotic construction. *Science Robotics*. 2019;4(28):eaau8479.
- [24] Asadi K, Ramshankar H, Pullagurla H, Bhandare A, Shanbhag S, Mehta P, et al. Building an integrated mobile robotic system for real-time applications in construction. *arXiv preprint arXiv:180301745*. 2018.
- [25] Hu D, Gan VJ, Wang T, Ma L. Multi-agent robotic system (MARS) for UAV-UGV path planning and automatic sensory data collection in cluttered environments. *Building and Environment*. 2022;221:109349.
- [26] Mantha BR, Menassa CC, Kamat VR. Robotic data collection and simulation for evaluation of building retrofit performance. *Automation in Construction*. 2018;92:88-102.
- [27] Lee JH, Park J-H, Jang B-T, editors. Design of Robot based Work Progress Monitoring System for the Building Construction Site. 2018 International Conference on Information and Communication Technology Convergence (ICTC); 2018: IEEE.
- [28] Hwangbo J, Lee J, Dosovitskiy A, Bellicoso D, Tsounis V, Koltun V, et al. Learning agile and dynamic motor skills for legged robots. *Science Robotics*. 2019;4(26):eaau5872.
- [29] Raibert M, Blankespoor K, Nelson G, Playter R. Bigdog, the rough-terrain quadruped robot. *IFAC Proceedings Volumes*. 2008;41(2):10822-5.
- [30] Hutter M, Gehring C, Lauber A, Gunther F, Bellicoso CD, Tsounis V, et al. Anymal-toward legged robots for harsh environments. *Advanced Robotics*. 2017;31(17):918-31.
- [31] Halder S, Afsari K. Real-time Construction Inspection in an Immersive Environment with an Inspector Assistant Robot. *EPiC Series in Built Environment*. 2022;3:389-97.
- [32] Navon R. Automated project performance control of construction projects. *Automation in construction*. 2005;14(4):467-76.
- [33] Kini DU. Materials management: The key to successful project management. *Journal of management in engineering*. 1999;15(1):30-4.
- [34] Guizzo E. By leaps and bounds: An exclusive look at how boston dynamics is redefining robot agility.

- IEEE Spectrum. 2019;56(12):34-9.
- [35] Dynamics B. Robot Specifications 2022 [updated Oct 26, 2022. Available from: <https://support.bostondynamics.com/s/article/Robot-specifications>.
- [36] Wetzel E, Liu J, Leathem T, Sattineni A, editors. The Use of Boston Dynamics SPOT in Support of LiDAR Scanning on Active Construction Sites. ISARC Proceedings of the International Symposium on Automation and Robotics in Construction; 2022: IAARC Publications.
- [37] Nilsson O. Building dense reconstructions with SLAM and Spot: MSc Thesis TFRT-6158, Department of Automatic Control, Faculty of ...; 2022.
- [38] Dynamics. B. Spot CAM setup and usage 2023 [updated Jun 7, 2023. Available from: <https://support.bostondynamics.com/s/article/Spot-CAM-Spot-CAM-Spot-CAM-IR>.
- [39] Wang J, Sun W, Shou W, Wang X, Wu C, Chong H-Y, et al. Integrating BIM and LiDAR for real-time construction quality control. *Journal of Intelligent & Robotic Systems*. 2015;79:417-32.
- [40] Kim P, Chen J, Cho YK. SLAM-driven robotic mapping and registration of 3D point clouds. *Automation in Construction*. 2018;89:38-48.
- [41] Halder S, Afsari K, Serdakowski J, DeVito S, Ensafi M, Thabet W. Real-Time and Remote Construction Progress Monitoring with a Quadruped Robot Using Augmented Reality. *Buildings*. 2022;12(11):2027.
- [42] Halder S, Afsari K, Chiou E, Patrick R, Hamed KA. Construction inspection & monitoring with quadruped robots in future human-robot teaming: A preliminary study. *Journal of Building Engineering*. 2023;65:105814.
- [43] Dynamics. B. Concepts [Available from: <https://dev.bostondynamics.com/docs/concepts/real-dme>.
- [44] Xiao X, Liu B, Warnell G, Stone P. Motion planning and control for mobile robot navigation using machine learning: a survey. *Autonomous Robots*. 2022;46(5):569-97.
- [45] Koval A, Karlsson S, Nikolakopoulos G. Experimental evaluation of autonomous map-based spot navigation in confined environments. *Biomimetic Intelligence and Robotics*. 2022;2(1):100035.
- [46] Labbé M, Michaud F. RTAB - Map as an open - source lidar and visual simultaneous localization and mapping library for large - scale and long - term online operation. *Journal of field robotics*. 2019;36(2):416-46.
- [47] Kehoe B, Patil S, Abbeel P, Goldberg K. A survey of research on cloud robotics and automation. *IEEE Transactions on automation science and engineering*. 2015;12(2):398-409.
- [48] Dimitrov A, Golparvar-Fard M. Vision-based material recognition for automated monitoring of construction progress and generating building information modeling from unordered site image collections. *Advanced Engineering Informatics*. 2014;28(1):37-49.
- [49] Lu W, Chen J, Xue F. Using computer vision to recognize composition of construction waste mixtures: A semantic segmentation approach. *Resources, Conservation and Recycling*. 2022;178:106022.
- [50] Zhang G, Zhisheng Z, Zhijie X, Min D, Meng P, Cen J, editors. Implementation and research on indoor mobile robot mapping and navigation based on RTAB-Map. 2022 28th International Conference on Mechatronics and Machine Vision in Practice (M2VIP); 2022: IEEE.
- [51] Wang H, Niu D, Li B, editors. Distributed machine learning with a serverless architecture. *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*; 2019: IEEE.
- [52] Lin JJ, Golparvar-Fard M. Construction progress monitoring using cyber-physical systems. *Cyber-physical systems in the built environment*. 2020:63-87.
- [53] Hamledari H, Fischer M. Construction payment automation using blockchain-enabled smart contracts and robotic reality capture technologies. *Automation in Construction*. 2021;132:103926.
- [54] Bademosi F, Issa RR. Factors influencing adoption and integration of construction robotics and automation technology in the US. *Journal of Construction Engineering and Management*. 2021;147(8):04021075.
- [55] Ekanayake B, Wong JK-W, Fini AAF, Smith P. Computer vision-based interior construction progress monitoring: A literature review and future research directions. *Automation in construction*. 2021;127:103705.
- [56] Halder S, Afsari K, Chiou E, Patrick R, Hamed KA. Construction inspection & monitoring with quadruped robots in future human-robot teaming: A preliminary study. *Journal of Building Engineering*. 2023:105814.
- [57] Delgado JMD, Oyedele L, Ajayi A, Akanbi L, Akinade O, Bilal M, et al. Robotics and automated systems in construction: Understanding industry-specific challenges for adoption. *Journal of building engineering*. 2019;26:100868.
- [58] You K, Zhou C, Ding L. Deep learning technology for construction machinery and robotics. *Automation in Construction*. 2023;150:104852.

Integrating Automation into Construction Site: A System Approach for the Brick Cutting Use Case

Cinzia Slongo¹, Vincenzo Orlando¹, Dietmar Siegele¹ and Dominik T. Matt^{1,2}

¹Fraunhofer Italia Research, Italy

²Free University of Bolzano-Bozen, Italy

([cinzia.slongo](mailto:cinzia.slongo@fraunhofer.it), [vincenzo.orlando](mailto:vincenzo.orlando@fraunhofer.it), [dietmar.siegele](mailto:dietmar.siegele@fraunhofer.it))@fraunhofer.it, dominik.matt@unibz.it

Abstract –

The integration of robotics, automation and digital technologies into construction processes offers opportunities for increased efficiency, adaptability and sustainability. However, a comprehensive integration of these technologies is still missing. This paper proposes an integrated systems approach using collaborative construction robotics, a modular construction container and an advanced tasks application connected to a multi-agent control system to address challenges in efficiency, sustainability and labour shortage in construction site to optimise construction workflows and processes. The focus is on establishing on-site factories and using mobile manufacturing technologies to revolutionise traditional construction methods. The methodology includes modular construction containers, collaborative construction robotics and advanced task applications to streamline construction operations. A case study of automated brick cutting illustrates the feasibility and potential benefits of the proposed approach. The multi-agent control system coordinates human and robotic agents to increase the efficiency and resilience of construction processes. The paper concludes with insights into future research directions, emphasising safety, efficiency, adaptability and scalability in addressing the dynamic nature of construction environments. Through this integrated approach, the construction industry can achieve improved productivity, safety and sustainability, paving the way for transformative changes in construction practices.

Keywords – Construction container; Construction processes Automation; Collaborative Construction Robotics; Advance Task Planning; Multi-agent-based system

1 Introduction

Robotics and automation are transforming the construction industry, but implementation challenges

remain due to the sector's complexity and fragmented logistics. The deployment of automated technologies and prefabrication is hindered by various site management and logistics issues, which negatively impact productivity in terms of time, cost, and safety. Construction automation is challenging in the construction industry due to the need for on-site operations and the absence of a fixed workspace like most manufacturing sites have. Moreover, the high cost of automation systems and the unstructured nature of construction sites require specialized personnel. However, robotics offers a promising solution to address issues relating to labour shortages, safety concerns, and efficiency. Customizing construction elements, such as bricks or insulation panels, or replacing parts of existing structures, poses a high level of risk on-site, as it often requires of saws, cutters or drills and is carried out frequently throughout the working day. One possible solution is to automate established design processes, employing programmable machinery and robotic equipment to load machinery and transport components to the assembly location. This would improve working conditions and enhance procedural accuracy by partially relieving workers of hazardous, monotonous, and physically demanding tasks.

This paper introduces an integrated systems approach based on an initial prototype employing robotics and multi-agent-based strategy aimed at automating construction tasks. The prototype integrates human, production equipment, and robotic agents to streamline material preparation and tool management. Employing an advanced task planning application alongside a modular construction container, the system aims to bolster on-site production, fostering enhanced efficiency and resilience. This early prototype serves as a crucial step in identifying and addressing open issues within construction automation, setting the stage for more targeted and specific research in the field.

2 State of the art

The integration of robotics in construction has

received considerable attention in recent years, with numerous studies investigating different aspects of robotic systems in construction applications. The current state of automation and robotics in construction, coupled with BIM-based task planning and multi-agent control platforms, reflects a transformative shift towards efficiency, precision, and adaptability in the industry. Research has explored the use of multi-agent collaboration for building construction [1], BIM-integrated construction robot task planning and simulation [2], and human-robot collaboration in construction [3]. In addition, the research has addressed the current task management process and state of control systems in construction sites and associated digital tools [4] to understand how using Industry 4.0 enablers and interconnection on a construction site can improve productivity [5]. Task planning and control activities can meet the information needs of workers and preparation of materials, thus optimising the time for installation work [6].

The use of multi-agent systems has been explored in various domains, including supply chain risk management [7] and robot scheduling [8]. [9] advocate for a platform-based approach to automation in construction, emphasizing the need for integrated systems to optimize resource allocation and project coordination highlighting the potential for automation to streamline on-site construction processes. Some studies have explored BIM-integrated collaborative robotics for application in building construction and maintenance [10, 11] and BIM-based semantic building world modelling for robot task planning and execution [12] so that the use of automated or robotic tools combined with BIM, which provides positioning data for automated or semi-automated assembly, can speed up the construction process [13].

In addition, the concept of digital twins in construction has been explored [14], [15], together with its integration with collaborative robotics for construction tasks in generating realistic simulations to prototype robotic solutions suitable for specific tasks [16]. Recognizing the recent advancements related to applications of machine learning and digitalization, [17] propose to open the perspective from the discussion of single construction activities (e.g. additive manufacturing, automated installation, assembly or bricklaying) towards an integrated robotized construction site including innovative materials, improved robotic hardware and streamlined construction workflows.

In particular, component assembly requires workers to repeatedly measure and calibrate, resulting in low construction efficiency. Using on-site factories and automated tasks can create opportunities for improved productivity, profitability, and sustainability compared to manual labour and traditional off-site prefabrication,

unrealistic and expensive. [18] discuss the challenges and opportunities for digital in situ fabrication in architecture and construction. Their research underscores the transformative potential of robotic technologies in redefining traditional fabrication methods. This approach aims to mitigate waste, predominantly resulting from transport, rework, and storage, by maximising outputs from enhanced inputs [19, 20]. [19] advocate for the establishment of on-site factories to support lean principles and industrialized construction practices. Their research highlights modularization and prefabrication as strategies for enhancing project scalability and adaptability. Implementing a process of low to high-level fabrication and assembly of components on a site using fixed or mobile automated machines/robots, instead of traditional construction methods, has the potential to automate construction processes [21]. [22] present mobile additive manufacturing, a robotic system for cooperative on-site construction exploring the potential of additive manufacturing technologies to enable on-site fabrication and construction for greater flexibility and adaptability in building processes.

The construction industry is characterised by greater uncertainty and variability of operations and requires modern organisational models based on lean principles for small, flexible and scalable production units that can be moved directly to the construction site [23], reducing lead times and waste in the construction processes, minimising the use of materials to ensure excellent mobility and on-time delivery. The reviewed on-site manufacturing models, even those based on BIM, focus on safety stocks for material and time predictability, but suffer from a lack of planning of workers and resources, where space is also a resource [24]. In particular, work preparation has not been targeted and there is limited literature on the topic of modular on-site fabrication [17, 25, 26]. However, the literature lacks a holistic concept that integrates collaborative robotics into construction processes, covering the entire process from material transport and preparation to placement and use.

Despite recognizing the potential synergies between BIM-based task planning and multi-agent control platforms, existing research reveals significant gaps in integrating robotics and automation into construction processes comprehensively. These gaps encompass various aspects, including task planning, control, collaboration, and the absence of holistic methodologies that address the entire construction lifecycle from planning to execution. Addressing the identified gaps in the literature and advancing the understanding of the integration of robotics, automation and digital technologies in construction processes requires answering critical research questions. These questions include how on-site factories and mobile manufacturing

technologies can revolutionise traditional construction methods, enabling greater flexibility, adaptability and sustainability in construction processes. Additionally, methods and strategies to improve material preparation, workforce planning and resource allocation within construction projects must be explored. An integrated systems approach using robotics, automation, BIM-based task planning and multi-agent control platforms would result in streamlined construction workflows and processes to optimise material preparation and resource utilisation. A new systems approach, together with the implementation of flexible production units, would enable agile and responsive construction operations, leading to improved project scalability, adaptability and sustainability.

3 Implementation

The scope of this research is limited to providing and testing a system approach and methodology for handling heterogeneous and dynamic processes and data in the construction industry, specifically for some common and hazardous tasks on construction sites.

3.1 Modular approach

The integration of collaborative construction robotics, modular construction containers, and advanced task applications linked to multi-agent control systems an attempt at a paradigm shift in construction methodologies and processes. This section describes the methodology adopted to realize this integrated approach, aimed at generating a seamless workflow that enhances construction efficiency and adaptability. At the core of this integrated approach lies the modular construction container as a flexible and adaptable system designed to facilitate material preparation, storage, and transportation.

This modular approach supports the idea that on-site factories can be used on different projects and be scalable according to the project needs. Modules can be attached or removed according to the project requirements, enabling material preparation tasks such as cutting bricks or insulation panels. The prepared materials, resulting from a single assembly line with a low-mid level of automation, are then packed in the correct sequence for efficient use on the construction site (Figure 1).

By using a task planning application specifically designed for allocating tasks to both human and robotic agents, we can use the digital twin concept and real-time data from the construction site to allocate tasks intelligently. This process considers factors such as agent availability, skills, and the urgency of tasks. For instance, an application could generate brick laying patterns based on BIM model and construction site data, optimizing material utilization. The same application could optimize task sequences and schedules, taking into account task

interdependence and the availability of resources like tools, equipment, and workers.

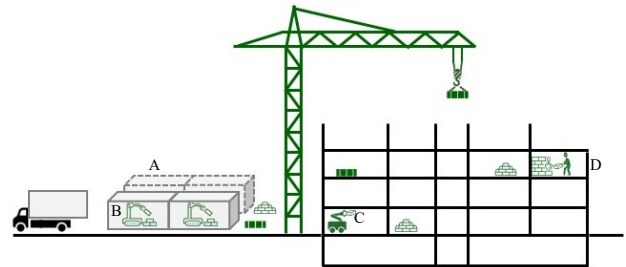


Figure 1. Modular construction container and multi-agent system

The methodology consists of several steps. A pivotal phase is included, focusing on software development and integration to digitally manage equipment, including robots and machinery, to fulfil the identified use cases of the on-site factory. The first step was the implementation of the on-site factory and configuring the container and its equipment. Subsequently, an automated and programmable system is developed to equip the machines for semi-automatic operation. Next, focus on software development and integration to digitally manage equipment (robots and machinery) to fulfil the tasks for the on-site factory's identified use cases. Achieving integration between the various technologies required to streamline a system and an adaptable approach and process flow that could, in the future, with specific research, meet the standards of rapid reconfigurability, accuracy and performance required by the industry.

3.2 Case study

The construction industry has historically underinvested in technology and appears to be resistant to change. The use of technology in planning, delivering, and operating construction projects carries great potential for saving costs, enhancing quality, and increasing speed.

Brick cutting is a relevant use case because modifications are frequently necessary. Brick cutting is a common task on a construction site, as walls are often not multiples of brick sizes. This requires adjusting the dimensions of one or more bricks to achieve the correct size. Additionally, changes to the dimensions of openings or other structural elements during construction also necessitate brick cutting. Unfortunately, this task exposes workers to significant health risks. The implementation of programmable machines, along with robotic tools for loading and unloading the work surface, offers a high degree of flexibility and automation. As a result, it enables customisation and automation of a significant proportion of the work executed on-site, which is

currently performed by on-site workers, frequently outdoors and near the building site.

From the BIM model data, precise element dimensions and placement details essential for construction can be obtained. This information is crucial as it informs the cutting parameters for the machine and coordinates the final delivery of the finished part on the robotic platform. Creating an isolated and controlled area, monitored by the workers, for this activity that utilizes high-risk equipment will enhance the safety and working conditions of the workers by enabling better control. By relocating this task to a secure and closely monitored facility and enabling robotic machinery to carry out the task instead of or in support of the worker, the aim is to achieve greater precision and efficiency with a focus on safety. The handling of the bricks cut-out by machines can be passed on to robotic platforms, which precisely deliver the parts to the operator on-site when and where they are required. This process saves time by enabling the worker to stay on site instead of interrupting operations to cut and transport bricks, while also enhancing safety by reducing the number of workers working at the site.

3.3 Proposed framework

3.3.1 Definition of the requirements

The challenges of the system approach based on a construction container for the automation of construction tasks and integrated agents' management include a restricted working area, complications in handling various brick shapes and weights, unpredictability regarding agents' collaboration in the operational environment, and task variability, which requires operations and activities to be performed through motion and interaction control. The definition of functional and technical requirements of the systems involved interpreting the needs of the end users, so they were organically derived from the use cases in terms of reliability, consistency, productivity safety and accuracy.

- Configuration: provide a system, software and control framework related to the configuration required by the use case.
- Operation: provide a system for planning and executing the use case.
- Safety: provide a comprehensive safety management system for the use case.
- Planning and monitoring: provide a system for planning and monitoring the execution of the tasks through an intuitive GUI linked to the digital twin of the construction container.
- Accuracy: the accuracy needed by the operation has been defined by improving the quality of the component information provided by the BIM model.

3.3.2 Agents definition

The multi-agent control system coordinates the actions of the human and robotic agents involved in the construction tasks. This approach ensures effective allocation of tasks, taking into account agent skills, workload, and priority. As a result, the construction process becomes more efficient. The construction container, represented as agent A, customizes and stores bricks. Agent B, a separate robot, sets up the brick cutter (agent C) and separates usable bricks from rejects. The transportation and construction tasks, specifically constructing a brick wall within a storey, can be efficiently performed by robots - agent D (Figure 2).

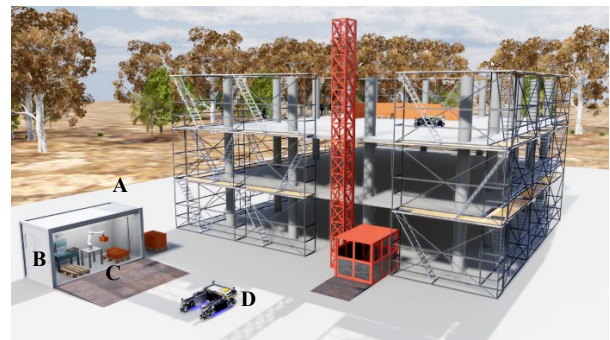


Figure 2. The multi-agent control system

This increases the resilience of the construction process by replacing a faulty robot or resolving a bottleneck in carrying out a given task with an additional agent for a limited time. The data-driven approach gathers information from the task planning application, enabling robots to receive updates about upcoming tasks on their control systems while humans can receive tasks through mobile or smart devices.

3.3.3 Hardware Architecture

The construction container consists of a 4.8x2.4m on-site container housing the automated production system. The development of the automated production system involved compiling a list of components and tools based on defined requirements and desired features.

The completed setup included a 6-axis collaborative robot arm (Universal Robots UR10e) with a two-finger gripper (Onrobot RG6) and a brick-cutting machine (IMER M400 Smart). The machine has a water-cooled saw with a diameter of 400 mm and a pusher that moves independently via a linear actuator. In addition, the machine is enclosed in a Plexiglas structure to prevent splashes that could damage electrical components. The elements communicate via a private Ethernet connection within the container, with each component assigned a unique IP address. This communication is facilitated by a central switch (D-Link DGS-108).

The gripper was integrated using the MODBUS

communication protocol. The saw was chosen for its compatibility with the automation, which was achieved by incorporating a DC motor moving a linear guide connected to the saw carriage by an Oldham joint and controlled by a DC brick (Tinkerforge).

3.3.4 Software Architecture

The construction container's software architecture comprises three main elements: one for robot motion planning, another for controlling the brick-cutting machine, and a third for the User Interface (UI). These components are connected via an MQTT broker, enabling message exchange in JSON format, as illustrated Figure 3.

The robot's trajectory has been planned for brick handling using the ROS Moveit framework. The Pilz Motion Planner performs the planning using a Python program for simulation in RViz to provide visualisation and validation.

The robot is capable of performing four distinct pick and place tasks, depending on whether brick cutting is necessary. These tasks include transferring items from the source pallet to the machine, transferring items from the source pallet to the destination pallet, transferring items from the machine to the destination pallet to unload the desired part, and transferring items from the machine to the warehouse to unload the remaining part. The machine can move both forwards and backwards. An error management protocol has been integrated to handle any failures. The user interface enables communication with the backend using the three.js JavaScript library.

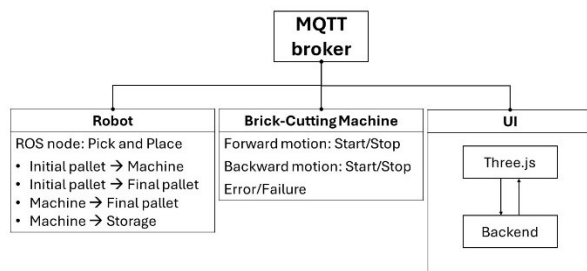


Figure 3. Software architecture

Figure 4 presents an example of agent communication. The scenario illustrates the UI indicating that one brick requires cutting, while another brick needs to be placed directly on the final pallet as a whole brick. The UI initiates communication by sending a message that includes the dimensions of the bricks, the specific tasks assigned to each agent, and the predetermined order in which these tasks are to be performed. The manipulator loads the brick onto the machine. Once the task is complete, it sends a message to load a brick directly onto the final pallet and another message to the machine to start cutting. After the cutting process, the machine sends

a message to the robot to unload the two parts, which then informs the UI that the task is complete.

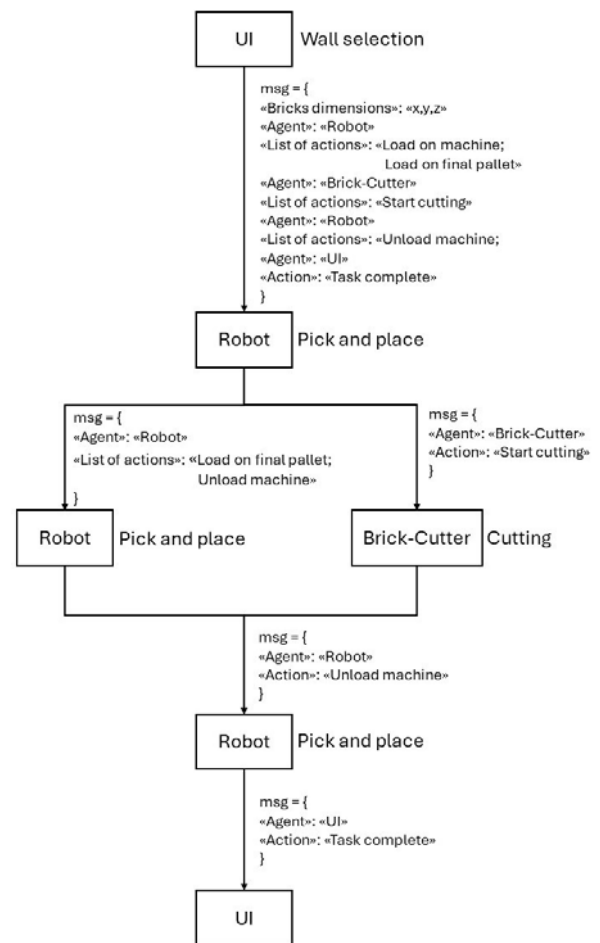


Figure 4. Agent communication

3.3.5 Task planning application and User Interface

Effective communication among the agents involved in the on-site construction phase is a crucial factor for project success. Planning and control actions that facilitate construction work include tasks such as organising production areas, coordinating material transportation, and setting up tools.

The implementation of process automation and robot operations has led to the creation of a task planning application. This tool effectively manages the necessary information sources for executing tasks and the available actors for completing them. It generates multiple tasks from the building's BIM file and assigns them to the most appropriate and available actor (either a robotic arm, a human worker, or manufacturing equipment) via an intuitive graphical interface (Figure 5). The application integrates machines and robots to manage and prioritize tasks, cooperatively handling more complex tasks, thus

facilitating on-site coordination and communication.

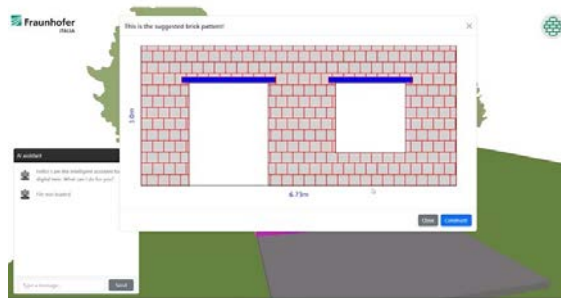


Figure 5. Task Planning Application User Interface

The system consists of a high-level planner which identifies tasks and robot base positions, and a basic planner, which computes a collision-free trajectory.

4 Results

The feasibility of the first prototype for the proposed integrated system approach for automating construction tasks using collaborative construction robotics and multi-agent control systems has been tested. The results indicate that it could be adopted as an alternative workflow, with the potential to revolutionise the construction industry and contribute to a more efficient, resilient, and sustainable built environment if the remaining issues are addressed by specific research. The study suggests that this approach has the potential to be a game-changer in the construction industry, but measurable benefits have yet to be demonstrated. The presented work is currently undergoing further development since its initial presentation at the DiSC - Digital and Sustainable Construction Conference 2023 in Bolzano. A dedicated showcase session was held to gather feedback from participants. The test carried out suggests that it may be possible to automate labour-intensive and time-consuming tasks, such as brick cutting, material customisation, and placement. This could lead to improved working conditions on construction sites, as workers can focus on more complex and value-added activities, potentially resulting in improved safety, productivity, efficiency, and accuracy.

4.1 The construction container - Automated on-site prefabrication

Due to the interdisciplinary nature of construction engineering and research, and the innate challenges of conducting studies in real-life settings, validating construction domains poses various challenges.



Figure 6. The construction container - Automated on-site prefabrication

The construction container (shown in Figure 6) acted as a mobile factory shell, serving as an external laboratory for testing and verifying the feasibility of the proposed approach and its applications in real-world environments such as construction sites. The feedback provided reliable information on the consistency of the method and process, as well as immediate feedback on critical issues. It also has the potential for resolution or future development. Automating operations that involve exposure to hazardous equipment or materials can create a safer working environment for human workers and reduce the potential for workplace accidents. The presented system approach has been shown to be effective in achieving these goals.

The multi-agent control system and modular construction container allow for adaptation to changing task requirements and unforeseen challenges. The task planning system can allocate tasks dynamically to human or robotic agents located in different containers based on their availability and skill set. This ensures the construction process continues even in the face of unexpected disruptions. This flexibility could help to make the construction process more adaptable to uncertainties and fluctuations in the workforce on site.

4.2 Automated brick cutting system

Due to the high degree of customisation in the construction industry, it is challenging to apply standardisation and prefabrication techniques off-site without requiring frequent revisits and modifications on site. Therefore, the possibility of establishing mobile factories on site has emerged to meet the demand for customised special pieces, including bricks and insulation panels. An automated system for cutting bricks has been developed and tested with the aim of producing tailor-made building materials.

Robot tasks involve (Figure 7a/b):

1. Brick grasping from their initial position.
2. Loading the machine to execute the task

and to match brick dimensions.

3. Unloading the cut parts.

4. Loading entire bricks onto the final pallet.

The waste can be reused in the construction process if its dimensions match the need for other special pieces to build walls on the site, as derived from the BIM model. By optimising material consumption and placement, this method would minimise waste generation and would help to reduce overall construction costs.



Figure 7 a/b: Automated brick cutting system

The integrated system approach consisting of the modular construction container concept and the task planning application could be easily adapted and scaled to meet the needs of other tasks or to accommodate new technologies, tools and construction methods. The test carried out showed that the approach remains relevant and worthy of further development in the face of ongoing technological advances.

5 Conclusion & Outlook

The aim of this paper is to outline the proposed system approach for addressing automation and collaborative robotics in an integrated workflow to manage performance and safety issues related to specific construction site tasks.

The modular system approach focused on the on demonstrating the feasibility of using a fully automated workflow to perform tasks on site, using task planning and execution BIM-based and collaborative robotic arms. The results indicate that automated brick cutting workflow is faster than the manual process, where workers have to go to the point, measure, and return to the cutting machine to execute the task. However, so far, no measurable data has been collected in terms of final accuracy and overall performance.

Research indicates that the path is feasible, but it can be further improved to fulfil the requirements for safety and efficiency, and for demonstrating a high level of adaptability and versatility in response to the constantly changing and dynamic nature of the construction site.

Issues regarding speed and performance must be addressed while prioritising safety. Enhanced grasping capabilities can be achieved through the implementation of advanced vision object recognition algorithms based on deep learning and the development of custom machines and tools. This will lead to increased speed and greater robot efficiency.

Future research should also consider the complete process, including the delivery of bricks to the construction site. This would address the challenges of handling and transporting pallets loaded with bricks.

Additional insights should therefore concentrate on fulfilling container space requirements that meet market demands, coordination of multiple on-site factories, and testing the automated system's speed limits to optimise productivity.

References

- [1] Ankit K., Tony L., Jana, S. and Ghose D.: Multi-Agent Collaboration for Building Construction, 2020.
- [2] Kim S., Peavy M., Huang P.-C. and Kim, K. Development of BIM-integrated construction robot task planning and simulation system. *Automation in Construction*, 127, 103720 2021.
- [3] Liang, C.-J., Wang X., Kamat V.R., Menassa C.C. Human-Robot Collaboration in Construction. Classification and Research Trends. *J. Constr. Eng. Manage*, 147, 03121006, 2021.
- [4] Schöberl M., Bartmann D., Kessler S., Fottner J. Towards a Construction Site Control System? Task Management in Construction Operations and Intralogistics. Presented at the *38th International Symposium on Automation and Robotics in Construction*, Dubai, UAE November 2, 2021.
- [5] Turner C.J., Oyekan J., Lampros S. et al. (1 more author) Utilizing industry 4.0 on the construction site: challenges and opportunities. *IEEE Transactions on Industrial Informatics*, 17 (2). pp. 746-756. ISSN

- 1551-3203, 2021
- [6] Görsch C., Seppänen O., Peltokorpi A. and Lavikka R. Task planning and control in construction: revealing workers as early and late planners, *Construction Management and Economics*. 2023
- [7] Giannakis M. and Louis M. A multi-agent based framework for supply chain risk management. *Journal of Purchasing and Supply Management*, 17, 23–31, 2011.
- [8] Huang Y. and Zhu J. Modeling and Design of Multi - Agent Logistics Robot Scheduling Simulation Platform. Presented at the 2017 *5th International Conference on Mechatronics, Materials, Chemistry and Computer Engineering (ICMMCCE 2017)* September, 2017.
- [9] Masters K. and Johnston J. Automated construction: boosting on-site productivity using a platform-based approach. *Proceedings of the Institution of Civil Engineers – Civil Engineering* 172(6): 23–28, 2019
- [10] Follini C., Magnago V., Freitag K., Terzer M., Marcher C., Riedl M., Giusti A. and Matt D.T. BIM-Integrated Collaborative Robotics for Application in Building Construction and Maintenance. *Robotics*. 10, 2, 2020.
- [11] Follini C., Terzer M., Marcher C., Giusti A. and Matt D.T. Combining the Robot Operating System with Building Information Modeling for Robotic Applications in Construction Logistics. In: Zegloul, S., Laribi, M.A., and Sandoval Arevalo, J.S. (eds.) *Advances in Service and Industrial Robotics*, pp. 245–253, Springer International Publishing, Cham, 2020.
- [12] Kim K. and Peavy M. BIM-based semantic building world modeling for robot task planning and execution in built environments. *Automation in Construction*. 138, 104247, 2022.
- [13] Cai S., Ma Z., Skibniewski M. J., Bao S., and Wang H. Construction Automation and Robotics for High-Rise Buildings: Development Priorities and Key Challenges. *Journal of Construction Engineering and Management*, 146(8), 04020096, 2020
- [14] Sepasgozar S.M.E., Khan A.A., Smith K., Romero J.G., Shen X., Shirowzhan S., Li, H. and Tahmasebinia, F. BIM and Digital Twin for Developing Convergence Technologies as Future of Digital Construction. *Buildings*, 13, 441, 2023.
- [15] Nour El-Din M., Pereira P.F., Poças Martins J. and Ramos N.M.M. Digital Twins for Construction Assets Using BIM Standard Specifications. *Buildings*, 12, 2155, 2022.
- [16] C. Brosque E. Galbally O. Khatib and M. Fischer. Human-Robot Collaboration in Construction: Opportunities and Challenges. *International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)* pp. 1-8, Ankara, Turkey. 2020
- [17] Gharbia M., Chang-Richards A., Lu Y., Zhong R.Y. and Li H. Robotic technologies for on-site building construction: A systematic review. *Journal of Building Engineering*, 32, 101584, 2020.
- [18] Buchli J., Giftthaler M., Kumar N., Lussi M., Sandy T., Dörfler K. and Hack N. Digital in situ fabrication - Challenges and opportunities for robotic in situ fabrication in architecture, construction, and beyond. *Cement and Concrete Research*. 112. 10.1016/j.cemconres.2018.05.013, 2018.
- [19] Rosarius A. and García de Soto B. On-site factories to support lean principles and industrialized construction. *Organization Technology and Management in Construction*, An International Journal. 13. 2353-2366. 10.2478/otmcj-2021-0004, 2021
- [20] Calvetti D., Méda P., Chichorro Gonçalves M. and Sousa H. Worker 4.0: The Future of Sensored Construction Sites. *Buildings*. 2020
- [21] Zhu A., Pauwels P., and De Vries B.. Smart component-oriented method of construction robot coordination for prefabricated housing. *Autom. Constr.*, vol. 129, p. 103778, 2021
- [22] Dielemans, G., and K. Dörfler. Mobile Additive Manufacturing: A robotic system for cooperative on-site construction. *Proceedings of the International Conference of Intelligent Robots and Systems (IROS)*, Workshop Robotic Fabrication: Sensing in Additive Construction, Prague, Czech Republic. Vol. 27. 2021
- [23] Rauch E., Matt D. T. and Dallasega P. Mobile On-site Factories — Scalable and distributed manufacturing systems for the construction industry. In *2015 International Conference on Industrial Engineering and Operations Management (IEOM)*, Dubai: IEEE, Mar. 2015, pp. 1–10, 2015
- [24] Magill L.J., Jafarifar N., Watson A. and Omotayo T. 4D BIM integrated construction supply chain logistics to optimise on-site production. *International Journal of Construction Management*, 22, 2325–2334, 2022.
- [25] Brusselaers N., Fufa S.M. and Mommens K. A Sustainability Assessment Framework for On-Site and Off-Site Construction Logistics. *Sustainability*, 14, 8573, 2022.
- [26] Wei H-H., Zhang Y., Sun X., et al. Intelligent robots and human-robot collaboration in the construction industry: A review. *Journal of Intelligent Construction*, 1(1): 9180002, 2023.

Case study on the applicability of a construction site layout planning system

Jiyu Shin¹ Jongwoo Cho² and Tae Wan Kim³

¹ Researcher, Division of Architecture & Urban Design, Incheon National University, South Korea

² Researcher, Ph.D., Division of Architecture & Urban Design, Incheon National University, South Korea

³ Associate Professor, Ph.D., Division of Architecture & Urban Design, Incheon National University, South Korea
jiyuSHIN@inu.ac.kr, jjwoo@inu.ac.kr, taewkim@inu.ac.kr

Abstract

This study presents the development and application of a construction site layout planning (CSLP) system, emphasizing the importance of efficient, safe, and cost-effective planning in construction management. Recognizing the limitations of current algorithmic and simulation-based approaches, this research advocates for a CSLP system that prioritizes communication speed, ease of use, and adaptability over precision, integrating user experience and intuition into the planning process. A prototype CSLP system is developed based on these principles. It incorporates actual business process and decision support functionalities to address the dynamic nature of construction sites, supporting user-dependent planning. The system's development was informed by a comprehensive literature review that identifies gaps in existing CSLP tools and methodologies, leading to a new framework that aligns with practical field requirements. The effectiveness of the prototype is evaluated through a case study on an apartment complex construction site. This study assesses the system's capacity to deliver appropriate planning information and support plan review. The methodology includes a thorough analysis of the CSLP creation process, prototype development based on identified functional requirements, and a case study for validation. Findings indicate that the prototype enhances decision-making, improves safety, and optimizes the placement of temporary facilities. It facilitates collaboration and information sharing, thereby improving overall project management. This research contributes to construction project management theory by proposing a user-centred CSLP system that significantly improves efficiency, safety, and cost-effectiveness in construction projects.

Keywords –

Site layout; Temporary facilities; Plan assessment; Case study

1 Introduction

In the field of construction management, CSLP is an important consideration from the very beginning of a project, determining the size, number, and type of temporary facilities required for the project [1]. These facilities must be located within the project site appropriately and can include offices, guardhouses, restrooms, yards, workshops, shifts, water pumps, tower cranes, etc. [2]. The placement of temporary facilities has a significant impact on the cost, time, and safety of a project, and, thus, is critical for efficient project operations [3].

Many researchers have recognized the importance of CSLP and have conducted research on mathematical calculations and genetic algorithms for equipment and personnel movement, facility location optimization, safety, and transportation costs [4]–[7]. However, because CSLP is created in the early stages of a project, there is a high degree of variability in the plan, and the placement precision required in the field is lower for temporary facilities than for construction objects. Therefore, CSLP tools for the field should focus on communication speed and ease of use rather than placement precision. Algorithmic and simulation methods for identifying the best locations for temporary facilities often fall short in real-world applications. This is primarily due to their reliance on predetermined sizes and quantities, which may not align with practical situations. This misalignment hinders practitioners from implementing these approaches. To overcome these limitations, CSLP tools need to be integrated with business processes where users utilize them as decision support tools, and where human interaction with the tools leads to CSLP results. Therefore, in CSLP system development, it is important to determine how decision support functions can be effectively combined with existing CSLP creation processes through user experience. In addition, user experience and intuition are important factors in deployment planning decisions, as

they allow users to periodically update the required information in the plan to reflect changes in the field situation and improve decision-making by better understanding on the available resources (i.e., personnel, materials, and equipment). Therefore, the CSLP system should have a user-dependent structure.

In addition, the main purpose of a temporary facility layout plan in a construction project is to outline the route and potential hazard spaces before construction begins. Typically, this risk assessment is done in accordance with the legislation applicable to the construction phase and the company's manual. However, this review process is highly dependent on the subjective judgment of the planner. Therefore, there is a need for an auxiliary information provision function that enables more effective judgment of rough movements and potential hazard spaces.

In this study, a prototype of a CSLP system is developed, and a case study is conducted to evaluate how well the system meets the requirements. We develop a prototype of a CSLP system and reenact the CSLP process through a case study. In doing so, we examine the following two main aspects: 1) whether the CSLP output effectively conveys the appropriate level of planning information, and 2) whether the additional information provided is useful for plan review.

2 Method

The methodology of this study consists of three parts.

Analysis of the CSLP creation process: We thoroughly analysed each process required to create a CSLP. Through this analysis, we derived the functional requirements for each procedure, which is the basis for system development.

Prototype development: We developed a CSLP prototype system based on the functional requirements obtained by referring to previous studies. This prototype was designed to reflect the functional requirements.

Validation through case study: We conducted a case study to validate the effectiveness of the developed prototype. This process evaluated whether the CSLP prototype meets the functional requirements defined earlier. The feedback from the case study would play an important role in the further improvement of the prototype.

3 Literature Review

In a different direction from algorithmic research, there is research on developing tools to support material and equipment location planning to address safety, cost, and time issues on construction sites [8]–[12](Table 1).

In [8], a computerized system (ArcSite) for CSLP was developed. The system integrates a GIS and a DBMS to

automatically identify and place the best locations for temporary facilities. In [9], a genetic algorithm-based model (EvoSite) was developed to optimize CSLP. The main difference between the two early systems is the environment in which the site and temporary facilities are modelled. ArcSite is designed for layout planning modelling in GIS, i.e., continuous space environment, and EvoSite is designed for layout planning and result presentation in grid system environment, which is easy to apply genetic algorithm.

To address the complexity and uncertainty of CSLP, [10] proposed a simulation-based decision support tool (DST), which also uses rectangular facility modelling and Euclidean distance function for optimization of equipment, material, and worker placement in a grid system environment. Modelling in digital environments tends to focus on drawing methods in continuous space environments and prefer grid environments for optimization and plan review.

In [11], a new automation framework integrating BIM and optimization algorithms for prefabricated building CSLP is proposed. Recently, there have been attempts to combine BIM and parametric modelling support tools (e.g., Dynamo) into CSLP. However, these methods differ significantly from current practice and are limited in terms of user accessibility.

The work of [12] provides some insight into the functional requirements of a CSLP tool. In this study, several parameters essential for CSLP were extracted through social network analysis. These parameters are classified into four main categories, and in particular, three parameters (i.e., modelling a continuous site space, considering the site's surrounding environment, main and secondary traffic roads for project gates) belonging to the construction site category are used in this study to set the development direction of the CSLP system.

The CSLP support tools presented in the aforementioned studies have the disadvantage that they do not sufficiently consider integration with existing CSLP creation processes. This leads to a decrease in the accessibility of users to new tools for CSLP creation, so the system development was organized by simulating the work processes of 1) site surrounding investigation, 2) layout drawing, 3) layout evaluation and review, and 4) modification and maintenance.

In addition, the collaboration function facilitates the flow of information, helping all relevant parties to get the information they need in a timely manner. This contributes to the improved quality of decision making and reduced misunderstandings between them [13]. For proper CSLP consensus, the collaboration functions of the system are crucial. For this purpose, we have integrated a documentation task generation function into the system to help support information sharing and communication.

Table 1 Existing CSLP systems & tools

Reference	Research purpose	Approach	Utilized elements
[8]	Optimization of temporary facility layout in Construction site	Development of GIS and DBMS integrated system(Arcsite), use of heuristic algorithms and proximity index	Drawing method over continuous space environment
[9]	Development of an optimization model for CSLP	Development of the genetic algorithm-based Evosite system	Use of grid system for applying analysis algorithm
[10]	Development of a simulation-based decision support tool for construction site layout planning	Development of a simulation-based decision support tool (DST)	Use of grid system for applying analysis algorithm
[11]	Automation of layout planning for prefabricated construction sites	Integration of BIM and optimization algorithms, use of GA and PSO	Exclusion of parametric modelling support tools for user accessibility
[12]	Analysis and identification of research gaps in construction site layout planning models	Analysis of existing research models using social network analysis (SNA)	Utilization of parameters for CSLP tasks

4 Description of CLSP process incorporated in proposed System

The functions of the system were derived by considering the functions and information required by users to perform CSLP tasks, as described in [14]. The functions of the CSLP system are categorized based on the order in which the CSLP was created in practice and are defined as 1) Surrounding Situation Investigation, 2) Drawing Site layout Planning, 3) Layout Evaluation and Review, and 4) Modification and Maintenance. The reason for categorizing the functions according to the order of creation is to make it easier for potential system users to accept and quickly adapt to the new system in a similar way to the existing CSLP creation method.

The functions are placed in the order of the existing CSLP creation, as shown in Figure 1, and this order defines the modules from which they will be recruited in the system.

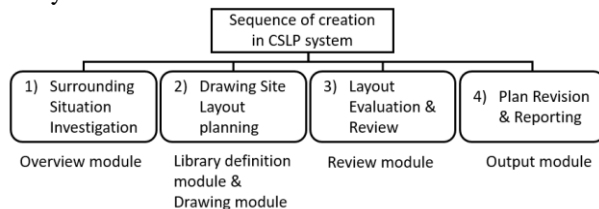


Figure 1. Hierarchy of CSLP system functions and system modules.

4.1 Functional Requirements for CSLP System Module

This section describes the CSLP authoring process and the authoring order considered. It identifies the

modules defined in the creation order and the functional requirements considered in each module. This is presented in Figure 1.

- 1) *Surrounding Situation Investigation*: The user needs to investigate the surrounding environment of the site where the project will be started. In this step, the surrounding factors, such as terrain, transportation accessibility, and neighbouring buildings, are checked, and information is collected. For example, by considering the presence or absence of major roads or residential complexes, the user can prevent complaints that may arise after the project starts and plan solutions to resolve them. Reporting on the surrounding conditions is also an important part of the existing CSLP planning report.
- 2) *Drawing Site Layout Planning*: The user designs the layout of the construction site based on the collected surrounding information. In this process, the space arrangement is established by considering buildings, equipment, materials, security, etc. The user considers the number, size, and type of temporary facilities required for the project and determines the location of the required facilities.
- 3) *Layout Evaluation and Review*: This is the stage of reviewing and evaluating the designed layout, considering the efficiency, safety, environmental impact, and schedule of the construction process. The CSLP is evaluated and revised if there are any inadequacies. It is also important to consider site management, material management, worker transportation, and safety.
- 4) *Plan Revision and Reporting*: In this step, the maintenance and management plan is developed,

and the plan is verified and adjusted. The user shares the plan with stakeholders and incorporates their feedback. Generating the reports on temporary facility deployment would increase work efficiency and facilitate information sharing with stakeholders. Information is delivered in the form of data files, which facilitates the exchange of views between the user and stakeholders.

As a result, the eight functional requirements (i.e., R1 to R8) considered in the system are described below.

- (R1) The system needs to be integrated with imagery that shows geographical characteristics so that users can understand their surroundings.
- (R2) The system should be able to help the user plan the CSL (construction site layout) without missing any temporary facilities that need to be placed.
- (R3) When starting a construction project, the system should be able to utilize the CAD files that are naturally available and reflect the accuracy of the site location and scale in the CSLP.
- (R4) The system should allow the user to have visibility into the planned CSL and identify hazardous areas.
- (R5) It should confirm that the paths between equipment and workers do not overlap for the safety of workers on site.
- (R6) It should allow for schedule conscious CSL, as schedules may require unnecessary temporary facilities to be moved or removed.
- (R7) The user should be able to derive the planned CSL in a report format to reduce the time spent on paperwork.
- (R8) The user should be able to generate data in flexible file formats that support communication with stakeholders, which are time- and location-independent.

4.2 Development of the interface design

In the development of the User Interface (UI) for a CSLP system, our strategy was firmly rooted in fostering an intuitive and efficient interaction between the user and the system. By diligently applying the four UI design principles established by [15], we anticipated a significant enhancement in the user experience, particularly in terms of communication with potential users.

- **Familiarity:** We adopted the ribbon interface framework, widely recognized from Microsoft applications, to minimize the learning curve and ensure a consistent user experience across platforms. This decision promotes intuitive engagement with the application by leveraging

familiar elements.

- **Clarity:** We emphasized clarity in our design for visual presentation, conceptualization, and terminology. Specifically, in the design of function icons, we aimed to reduce user errors and increase efficiency, making the application more intuitive and accessible.
- **Grouping:** We organized function modules logically to make it easier for users to quickly find and understand the information they need. This approach supports cognitive ease, aiding users in categorizing and recalling application functionalities, thus optimizing navigation and usability.
- **Compatibility:** We ensured our UI supports compatibility with multiple software file formats to accommodate the varied task and job requirements of our users. This is crucial for CSLP activities, allowing seamless integration of a wide array of file formats into their workflows, thereby broadening the application's applicability and flexibility.

5 CSLP System Application through a Case Study

To verify that the proposed CSLP meets the requirements, a case study was conducted in an apartment complex site. The existing temporary facility layout plan, CAD drawings, and project schedule information are utilized to derive the CSLP using the prototype system. The case study follows the sequence of system modules shown in Figure 1 in Chapter 4, which allows to evaluate the functional completeness of the aforementioned requirements.

The case study is based on an apartment site located in a city centre in Korea, and the construction work was carried out over a period of 28 months from September 2015 to January 2018. The site consists of five apartment buildings and is surrounded by a residential complex with a shopping centre, a school, and a subway station. In preparation for the case study, the data was reformatted for web use by converting the previously received DWG (CAD) files to DXF file format. During the case study, the type, location, and number of temporary facilities are determined by referring to the existing layout plan.

5.1 Overview Module

After accessing the CSLP system, a user can use the project creation function. In the project creation window, the user can search for the address of the apartment construction site, load a map image, and set the start and end dates of the project (Figure 2).

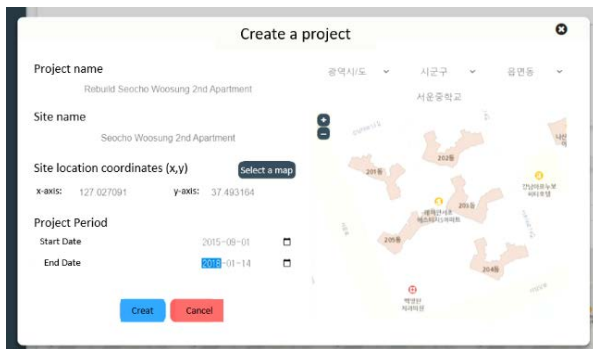


Figure 2. Overview module: Creating a project.

Through this map, the user can identify the surrounding area of the site and consider complaints from neighbouring residential complexes and traffic congestion on the main road (R1). After that, the user can add the prepared project schedule in the project schedule management window (Figure 3).

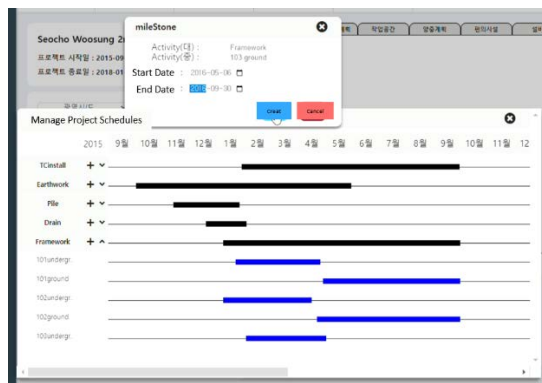


Figure 3. Overview module: Adding a schedule.

5.2 Library Definition Module

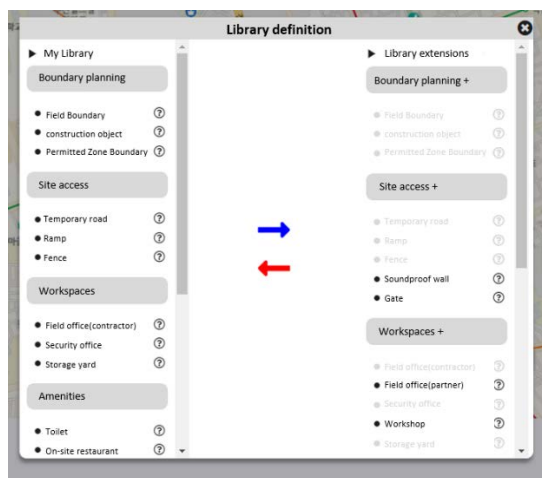


Figure 4. Library Definition module.

In this module, the user specifies the temporary facilities required for the case project. If the user use the “Library Extension” to add the some facilities to “My Library,” the added facilities are displayed on the left side of the system and can be viewed in the Drawing Module. The user can add, delete, and modify facilities to customize them for the case project. This module can be utilized as a checklist of facilities needed in the field (R2) (Figure 4).

5.3 Drawing Module

Before moving to the Drawing module of the system, the user should prepare the site map in the overview step and the list of temporary facilities for his or her project in the library definition step. The user can then load the CAD (DXF) file of the site into the system and place it on the map (① in Figure 5), and refer to the Library on the left to draw various temporary facilities on the map (R3). In this case project, there were 5 apartment buildings, 3 ramps, 4 site offices, 1 restroom, 13 lifts, and 3 T/Cs. Except for the apartments, the rest of the facilities were drawn using shapes (② in Figure 5), and the placement of apartment buildings utilized CAD files to improve the accuracy of the building geometry and enhance visual understanding. In the drawing stage, polygons and dots are placed on the map and coloured to distinguish between facility types intuitively. In this stage, opinions are shared with many stakeholders, and a note function is included to support this (③ in Figure 5). In addition, to allow multiple stakeholders to participate in the project, in this system, the edit mode and view mode are separated for quick plan checking (④ in Figure 5). In addition, the user can capture the screen as he or she creates the CSLP to keep track of the work.

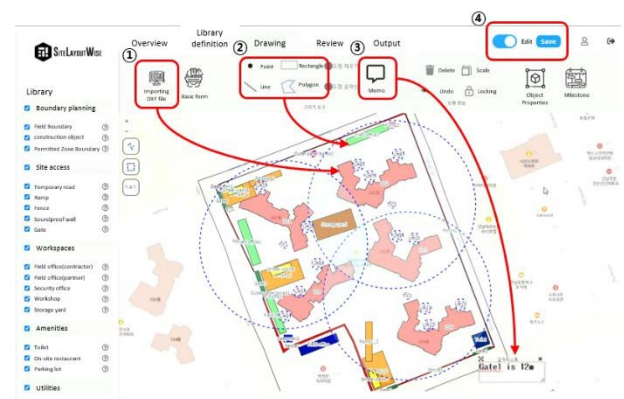


Figure 5. Drawing module.

5.4 Review Module

In this case study, the efficiency and safety of the designed CSLP were evaluated and reviewed. This module includes the ability to identify blind spots and hazardous areas. For example, the Figure 6 is an image of the blind spot identification function, which indicates low visibility locations sequentially from green to yellow to red. In this case, the field shows that the area at the edge of the field has a low visibility score and the centre has a high score. The Figure 7 shows a feature that shows the hazardous areas between temporary facilities and the target building in order from green to yellow to red. This feature shows the danger around lifts, T/Cs, and ramps in red. The results of the hazardous space check are based on the results of the blind spot check, and facilities located in blind spots are displayed as more dangerous (R4). In addition, the “Placement Constraint Check” function allows the user to define maximum separation values for different facilities to check whether overlaps between facilities occur (R5). This feature can be used to select routes that consider the safety of workers. These features help the user evaluate the appropriateness of your CSLP and consider modifications if necessary. In addition, temporary facilities can be scheduled according to the fluid CSLP situation on site, and the system displays the temporary facilities at a specific point in time (R6). In the Review module, the “View Modification Log” function is provided to allow the user to see who modified what, when, and by whom, increasing the transparency of planning and collaboration. The existing CSLP method relies on the individual's ability to review, but the system proposed in this study improves the review function by applying computational algorithms (Figures 6 and 7).

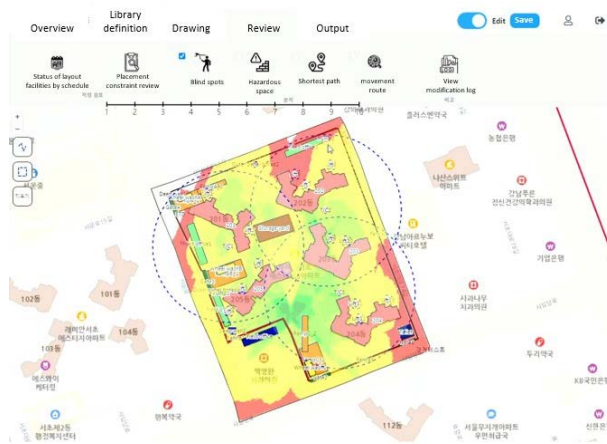


Figure 6. Review module: Checking blind spot.



Figure 7. Review module: Checking hazardous space.

5.5 Output Module

After the CSL is created in the Drawing module and reviewed and adjusted in the Review module, the Output module generates and outputs a CSLP report for the field use (R7). The report includes some details, such as the overall layout plan, redundancy plan, amenities, and workspace, and provides additionally results for hazardous spaces and blind spots identified through the Review module. On the report generation page, the user can make additions or notes, and organize and modify the list of deployed facilities using legends. For example, Figure 8 shows a report of a site access plan, with objects in that category highlighted in the image and labelled in the legend. This report is useful for reducing the amount of additional paperwork required after using the system. At the Output module stage, the user can save a DXF file of the CSLP and a PDF file of the report, which is a convenient file format to deliver to various stakeholders (R8).



Figure 8 Output Module

6 Conclusion

Previous studies on CSLP have pointed out the lack of integration with existing work processes, which leads to the problem of limited user accessibility of CSLP tools. Therefore, a new system developed in this study organizes functions by simulating work processes including site perimeter survey, layout drawing, layout evaluation and review, and modification and maintenance. The system is implemented according to the order of CSLP creation and provides analysis-based information that the existing method does not provide. In the study, a total of eight functional requirements were derived by analysing the four-step CSLP creation process, and the system was developed based on them. Through case studies of selected projects, it was confirmed that these eight requirements were satisfied.

The CSLP system enables efficient and systematic deployment planning compared to the traditional CSLP creation methods, which are typically time-consuming and limited to simple drawing functions. Through algorithmic evaluation, the system overcomes many of the drawbacks of traditional methods, especially time-consuming, the possibility of creating improper batches, and limitations in terms of safety and efficiency. By evaluating user-created CSLPs, the CSLP system helps users review their plans more carefully and determine the optimal placement. It shows significant performance improvements, especially in terms of speed and ease of use. In addition, the system provides the flexibility to incorporate feedback from various stakeholders in real time, which improves the user experience.

This study contributes to construction project management theory through the development of a user-centred construction site layout planning (CSLP) system. If this system is used by construction managers, it will help them to layout and manage temporary facilities more efficiently and precisely, contribute to the improved safety and cost-effectiveness of construction projects.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2021R1A2C1013188)

References

- [1] R. R. Rapp and B. L. Benhart, *Construction Site Planning and Logistical Operations: Site-Focused Management for Builders*. Purdue University Press, 2015.
- [2] F. Sadeghpour, O. Moselhi, F. Asce, and S. T. Alkass, "Computer-Aided Site Layout Planning," 2006, doi: 10.1061/ASCE0733-93642006132:2143.
- [3] H. M. Osman and M. E. Georgy, "Layout planning of construction sites considering multiple objectives: A goal-programming approach," in *Construction Research Congress 2005: Broadening Perspectives - Proceedings of the Congress*, 2005. doi: 10.1061/40754(183)130.
- [4] J. Xu, S. Zhao, Z. Li, and Z. Zeng, "Bilevel Construction Site Layout Optimization Based on Hazardous-Material Transportation," *Journal of Infrastructure Systems*, vol. 22, no. 3, Sep. 2016, doi: 10.1061/(asce)is.1943-555x.0000303.
- [5] L. C. Lien and M. Y. Cheng, "A hybrid swarm intelligence based particle-bee algorithm for construction site layout optimization," *Expert Syst Appl*, vol. 39, no. 10, pp. 9642–9650, Aug. 2012, doi: 10.1016/j.eswa.2012.02.134.
- [6] A. Kaveh, M. Khanzadi, M. Alipour, and M. Rastegar Moghaddam, "Construction site layout planning problem using two new meta-heuristic algorithms," *Iranian Journal of Science and Technology - Transactions of Civil Engineering*, vol. 40, no. 4, pp. 263–275, Dec. 2016, doi: 10.1007/s40996-016-0041-0.
- [7] S. RazaviAlavi and S. AbouRizk, "Site Layout and Construction Plan Optimization Using an Integrated Genetic Algorithm Simulation Framework," *Journal of Computing in Civil Engineering*, vol. 31, no. 4, Jul. 2017, doi: 10.1061/(asce)cp.1943-5487.0000653.
- [8] M.-Y. Cheng and J. T. O'Connor, "Site layout of construction temporary facilities using an enhanced-geographic information system (GIS)," *Autom Constr*, vol. 3, no. 1, pp. 11–19, 1994.
- [9] T. Hegazy and E. Elbeltagi, "EvoSite: Evolution-based model for site layout planning," *Journal of computing in civil engineering*, vol. 13, no. 3, pp. 198–206, 1999.
- [10] S. R. RazaviAlavi and S. AbouRizk, "Construction Site Layout Planning Using a Simulation-Based Decision Support Tool," *Logistics*, vol. 5, no. 4, Dec. 2021, doi: 10.3390/logistics5040065.
- [11] B. Yang, T. Fang, X. Luo, B. Liu, and M. Dong, "A BIM-Based Approach to Automated Prefabricated Building Construction Site Layout Planning," *KSCE Journal of Civil Engineering*, vol. 26, no. 4, pp. 1535–1552, Apr. 2022, doi: 10.1007/s12205-021-0746-x.
- [12] M. Salah, R. Khallaf, E. Elbeltagi, and H. Wefki, "Construction Site Layout Planning: A Social

- Network Analysis,” *Buildings*, vol. 13, no. 10. Multidisciplinary Digital Publishing Institute (MDPI), Oct. 01, 2023. doi: 10.3390/buildings13102637.
- [13] S. H. A. Rahman, I. R. Endut, N. Faisol, and S. Paydar, “The Importance of Collaboration in Construction Industry from Contractors’ Perspectives,” *Procedia Soc Behav Sci*, vol. 129, pp. 414–421, May 2014, doi: 10.1016/j.sbspro.2014.03.695.
- [14] J. Shin, K. Park, T.-W. Kim, and M. Kim, “Towards a Planning System for Construction Site Layout: Surveying Current Situation and Identifying Key Functions,” *Korean Journal of Computational Design and Engineering*, vol. 27, no. 4, pp. 392–401, Dec. 2022, doi: 10.7315/cde.2022.392.
- [15] N. U. Bhaskar, P. P. Naidu, R. C. Babu, and S. R. & Govindarajulu, “General Principles of User Interface Design and Websites,” 2011.

Discrete Event Simulation to Predict Construction Equipment Emissions on a Digital Twin Platform

Kepeng Hong, Alexandros Apostolidis, Jochen Teizer

Department for Civil and Mechanical Engineering, Technical University of Denmark
keho@dtu.dk, aapo@dtu.dk, teizerj@dtu.dk

Abstract

Emissions from machinery that is primarily fueled by Diesel represent a significant environmental concern in the construction sector. Traditional monitoring methods, including both Simplified and Portable Emissions Measurement Systems (SEMS and PEMS, respectively) encounter practical and financial constraints when deployed extensively across the diverse machinery types. This paper introduces a novel approach on predicting emissions and fuel consumption by leveraging a priori recorded emissions data from non-road mobile machinery (NRMM) in a Discrete Event Simulation (DES) as part of a Digital Twin Platform (DTP). Focusing on three types of construction machines (drilling rig, loading excavator, and hauling dump truck) the DES models their basic operations on a DTP purposed for earthwork and foundation activities for a high-rise building project in Denmark. With the input of different configurations (e.g., machine quantity and type, location), DES allows for the prediction of emissions and work output. Verification of the approach occurred in a field-realistic outdoor construction laboratory setting while the validation was demonstrated on a construction site. The results provide an efficient and economical avenue for monitoring emissions related to construction equipment operations. Beyond the environmental benefits, the proposed method generates knowledge that can supply construction managers with critical insights into performing proper resource leveling.

Keywords –

Construction equipment emissions, Digital Twin, Discrete Event Simulation, Site layout optimization, Non-road Mobile Machinery, Portable Emission Measurement System, Prediction and avoidance.

1 Introduction

Construction machinery is mainly powered by diesel fuel and contribute substantially to Greenhouse Gases (GHG) and thus, global warming [1]. Heavy construction

machinery, often referred to NRMM, accounts for about half of all CO₂ (Carbon Dioxide) emissions produced by the construction industry in Denmark in 2004 [2]. Besides CO₂, NRMM produces Nitrogen Oxides (NO_x), Carbon Monoxide (CO), Particulate Matter (PM), and Hydrocarbon (HC) pollutants that are dangerous to human health and the environment. Specifically, off-road diesel equipment is identified as the third largest contributor to NO_x emissions (14.5%) and the second largest contributor to PM emissions (24.3%) among mobile vehicles [2].

The European Union (EU) applies regulations for threshold levels of emissions from NRMM and to that end, various sensor technologies are set up to measure emission levels during the construction phase [3]. At a national level, the Danish Government has set a target for the green transformation of the built environment. It aims to achieve the goal of carbon neutrality by 2050 [4].

This is where Digital Twin (DT) concepts come into play that integrate, among other data, Building Information Modeling (BIM), construction schedule (4D component), Internet of Things (IoT, incl. data from sensors communicated wirelessly to a DTP and being processed for further analysis), and user interfaces (i.e., UI dashboards). Yet, a challenge of DTs in construction seems to be the reliable gathering of accurate field data and connecting the different data from the various sources for further reasoning in knowledge-based representations, for example, as part of high-fidelity information that is already available in 4D BIM models.

In essence, the further purpose of a DT is “a virtual representation of an object or system that spans its lifecycle, is updated from real-time data, and uses simulation, machine learning, and reasoning to help decision-making” [5]. While the virtual models of DTs can monitor real-time data from SEMS [6] and PEMS [7], DTs may also access data recorded in the past that assist in generating further insights from the data-driven simulations [8]. In construction simulations newly created knowledge benefit decision making, for example, reducing fuel consumption in construction logistics [9].

However, an open research question is: *How can these individual components (4D BIM, IoT/sensors, DTP*

incl. UI) be applied together in a meaningful way that can demonstrate the environment benefits of reducing construction equipment emissions while increasing output?

The following sections explain the background, the methodology, and preliminary findings how data from heavy construction equipment involved in creating foundation piles, loading and hauling the excavated earth material off site are used in predictive simulations for construction site operations emissions assessment.

2 Background

This section is subdivided in a few very brief reviews of some of the relevant topics an interested reader might want to be familiar with, for example, the key alternatives to (a) gather live construction equipment emissions data (2.1-2.3) and (b) conduct basic construction equipment emissions simulations (2.4).

2.1 Portable Emission Measurement Systems

Portable Emissions Measurement Systems (PEMS) have been used for several decades, gaining significant traction in the late 1990s and early 2000s. These systems are designed to measure and analyze emissions from various sources like road vehicles and NRMM. There is a diverse range of PEMS available in the market. Therefore, they vary in measurement capabilities, features, and compatibility with different machines or equipment. PEMS systems consist of sensors, analyzers, and data acquisition systems that can be transported and installed for on-site measurements. The adoption of this technology grew as researchers recognized the importance of obtaining accurate emissions data. They sought data from real-world field experiments, with machines working in their operational settings, going beyond traditional laboratory testing. PEMS find applications in research, emissions monitoring, and other activities related to emissions control and environmental studies [7, 9, 10]. Noteworthy, PEMS is supported by both the U.S. Environmental Protection Agency (EPA) and the European Environment Agency (EEA) for measuring emissions.

2.2 Telematics

Telematics provides an advanced monitoring method widely employed in the automotive industry to track assets like cars, trucks, and construction equipment. It utilizes Onboard Diagnostics (OBD), often in combination with Global Navigation Satellite Systems (GNSS) technology to record and map a vehicle's movement, providing valuable data for fleet tracking and management. End-user access to telematics data, if granted, further enables efficient monitoring of various vehicle aspects, such as speed, idling, fuel consumption,

tire pressure, etc. This is feasible through a sophisticated onboard computer in the vehicle, capable of capturing comprehensive information [11].

2.3 Simplified Emission Measurement Systems

Compared to PEMS, a Simplified Emission Measurement System (SEMS), as used in [2] and [6], comes at a much lower expense (a few thousand Euros). SEMS is a portable emission meter and can be applied to several means of transportation, including NRMM [12]. Due to its small size and the less time-consuming installation, usually less than 15 minutes, it is suitable for simplified measurements when equipment operations are not to be interrupted for a long time or inspections, for example, by regulatory authorities. Due to the lower complexity, fewer experts are required to mount the system than PEMS, simplifying its installation process and use overall [13]. Yet, [2, 6] pointed out that the latest released emission standards from the EEA for NRMM make it difficult if not impossible for some of the relatively inexpensive sensor components in SEMS to record the ever-smaller emission levels very accurately.

2.4 Discrete Event Simulations

Discrete Event Simulation (DES) involves modeling systems where the state variable undergoes changes at specific time points [14]. It has found widespread adoption as an effective technique for comprehending system behaviors and assessing different operational strategies. Since the inception of the CYCLic Operations Network (CYCLONE) [15] has been instrumental in crafting computer-based simulation models for construction projects, aimed at analyzing and optimizing their performance [16]. Following CYCLONE's introduction, a multitude of construction simulation systems emerged, including STROBOSCOPE [17] and RiSim [18], and SDESA [19]. DES was also implemented for estimating construction emissions by investigating load factors and based on various equipment activities [20]. Besides other tools, SimPy emerged as an open-source library that includes components enabling DES workflows efficiently [21]. These simulation systems offer valuable tools for project managers to replicate the dynamic interactions between resources and activities, facilitating comprehensive performance evaluation and the generation of insights.

3 Research Method

The employed Digital Twin Platform (DTP) plays a pivotal role in gathering and streaming the raw machinery data via IoT-sensors (Internet of Things) to an accessible cloud storage space. There, a pre-processing module filters the data first for erroneous elements and

then processes it online for emissions reasoning by applying simplified mathematical formulas. Finally, the resulting information is transformed as input values into a DES simulation model. This model finds out how environmental and economical the construction operations process can be. Subsequently, a User Interface (UI) presents the results to the end user in a visual and accessible dashboard format in commercially available web browsers. As shown in Figure 1, the designed technical components in this workflow seamlessly incorporate some manual user input, fostering a collaborative and interactive working environment for easily refining the required parameters to run the DES, which predicts previously not available outcomes.

Multiple steps are involved in the workflow of data recording, communication, processing and visualizing: (1) Recording raw equipment emissions and location data (latter, if available), (2) wireless communication to cloud storage space via mobile networks, (3) processing raw data streams online by applying simplified mathematical conversions, (4) reasoning by applying basic thresholds to detect emission breaches that violate regulations, (5) linking the results to parameters of already existing elements of 4D BIM models, (6) taking this automatically generated information and further manual input as values to run the DES, (7) visualizing the results as part of a dashboard in the DTP.

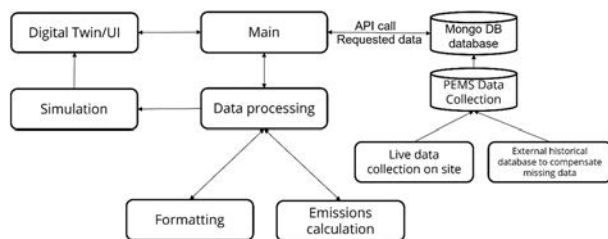


Figure 1. Workflow of acquiring, communicating, processing, and populating emission information to the Digital Twin Platform (DTP).

For demonstration, the method is applied to a realistic but highly simplified use case involving two alternative construction plans of up to two drilling rigs purposed to create reinforced concrete piles for a foundation wall, one excavator loading the excavated earth material from the drilled piles into at least one dump truck that hauls the material off the building site.

4 Implementation

4.1 Live and online emission datasets gathering

Emissions data for 3 different NRMM were collected. The drilling rig, Bauer BG55, collects the emissions data via PEMS on the (final demonstration) high-rise building

construction site for 1 full workday (7 effective work hours). Mobilization of PEMS took several hours. Due to several unknown false data readings of NO_x emissions, the dataset was manually filtered and cleaned before used in the further analysis. More information on the recording of emissions data of the drilling rig (2 hrs.) is here [3].

For the excavator, Caterpillar 325F, a structured experiment took place at another location, a realistic outdoor construction laboratory setting. This was necessary since only one PEMS was available for testing at the final demonstrator site. However, the same PEMS device (AVL492) was utilized for recording 1.5 effective working hours of the excavator's emissions, like NO_x and CO₂. The excavator operated in different working modes including loading and idling, during cold and hot starts.

All machines operated on Diesel fuel. The utilized PEMS in all tests (drill rig and excavator) logged data at high frequency (1 Hz) via the CAN-bus system. Such data can also easily be exported as CSV (Comma Separated Values) file for potential other post processing.

Since no dump truck was readily available at either location in Denmark, the PEMS data used for a dump truck came from an online data repository. Chosen was a 10-ton dump truck, Isuzu FTR850 AMT. The PEMS device in the study that collected the emission data [22] was a SEMTECH[®]DS unit recording multiple trips on urban routes in South Africa at a rate of 1 Hz. While some risk is potentially involved in procuring data over the internet for our research purposes, the data set was previously published in a peer-reviewed academic outlet and although the recording location was not identical to the climate in Denmark, still deemed trustworthy.

4.2 Datasets processing into averaging heat maps

After dataset gathering, we proceeded with assessing the impact of engine speed and engine load on the emissions within the context of the machine being in several modes, incl. direct work, supporting work, and other works, including idling.

First, we removed data points where the engine speed was below 800 rpm, as it is unlikely for engines to be working at such speeds in practical scenarios. Based on observation and preliminary analysis by checking the distribution of engine speed value when the machine is idling, we defined the idling mode as when engine speed is between 800 rpm and 900 rpm. We calculated the average value for the emission rate (g/s) of CO₂ and NO_x and the consumption rate (g/s) of fuel at the idling mode, which is then used in DES for simulating the emission and consumption of idling machine as shown in Table 1.

For other working modes, the observation heat map of engine speed and engine load is created. The approach focused on creating informative heatmaps for detecting the machine's operating mode automatically based on, specifically, its engine speed, engine load, and emissions

values. The first step included extracting the relevant data from the dataset and then dividing its engine speed and engine load into sections to construct a more detailed grid for the heat map. The process proceeded otherwise with calculating the average NO_x emissions within each grid element by applying certain filtering conditions to the dataset. The same process was used for the CO₂-values and for the fuel consumption. Additionally, we visualized the resulting information data as a heat map, using a color scheme to represent average emissions, and included annotations for clarity. Axis labels and a title provide context, and a color bar assists in the interpretation. Given the space restrictions in this paper, Figure 2 presents the calculated heat maps for the drill rig only. Figure 2 (a) shows the percentage of engine load and speed of the drilling rig when it is working in drilling mode. Figure 2 (b-d) shows the CO₂ and NO_x emission rate as well as Fuel consumption rate when the drilling rig is working at different sections of engine load and speed. Figure 2 (b-d) is created from the observation on all working modes and reflects the correlation between emissions and engine speed and load, which are thus also used for the simulation of other working modes.

Table 1. Average NO_x and CO₂ emission and FUEL consumption for idling equipment.

Equipment	NO _x (10 ⁻³ g/s)	CO ₂ (g/s)	Fuel (g/s)
Drilling rig Bauer BG 355G	13.80	5.38	1.71
Excavator CAT 325F	3.77	2.20	0.50
Truck Isuzu FTR850 AMT	0.87	1.21	0.38

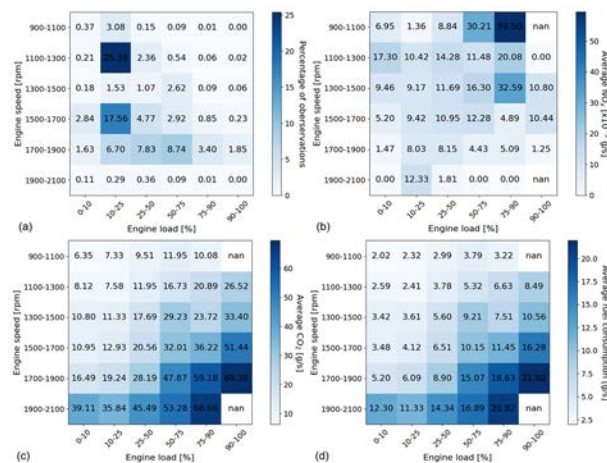


Figure 2. Heat maps for Bauer 355G drilling rig when it is in the drilling mode for 3 hours of observed work.

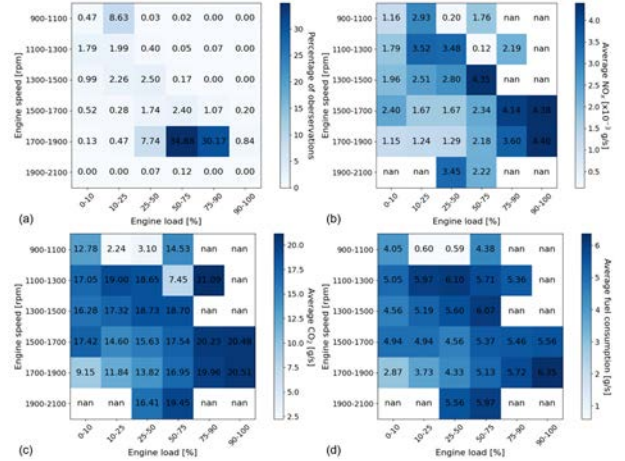


Figure 3. Heat maps for CAT 325F excavator when it is in the loading mode for 3 hours of observed work.

Overall, few deviations in the engine load and the speed are observed, and therefore, the emissions can be associated automatically when engine parameters fall within operating modes as outlined above. Likewise, it is a useful source of input data for building a simplified DES for equipment emissions (explained later).

For the loading activity as shown in Figure 3, the excavator operates in higher engine load for the majority of the time (50%-90%) and between 1700 to 1900 rpm. It is a more intensive activity and it is logical to observe the differentiation in engine load and engine speed between the loading and idling activity. In the same manner the rate of fuel consumption and CO₂ emission are much higher than in the idling activity, whereas the NO_x emission shows a significant monotonic correlation with engine speed and engine load. This is explained due to the catalyst existing in the machine which activates when it reaches temperatures above 120 °C.

4.3 Emissions calculating from simulation

Based on the live recorded PEMS emissions data two arrays are generated: a probability array and a value array. The probability array encapsulates the likelihood of engine speed and load of the machine when the machine is in different working mode. Simultaneously, the value array associates real average values with each combination of engine speed and load. The core of our method involves the calculation of cumulative probabilities. After flattening the probability array, we calculated the cumulative probabilities of that array. The next step is the generation of random numbers distributed uniformly between 0 and 1. Subsequently, our approach identified the corresponding combination of engine load and speed for each randomly generated number. From the simulation we receive the time of each machine's operation in the unit of seconds. For each second, we

estimate the emission and consumption calculation on the basis of the specified probabilities and associated values. The final result is a sum of emission and consumption values sampled from the distribution.

4.4 Simulation engine, logic and parameters

We used the SimPy simulation engine to perform the DES. SimPy is an open-source process-based discrete-event simulation framework for Python. It provides the necessary components for modeling and simulating complex systems with discrete events, such as the advancement of time and interactions between different entities in the simulation (Table 2). The simulation (Figure 3) starts with a human user entering numerous input parameters it needs to run (see Table 3).

Here, the vehicle agents of the excavator and trucks are modeled as resources (Figure 4). Conditional activities represent a task that starts when the resources are available in the queues. For instance, the loading operation starts when an excavator and a dump truck are available. A normal activity starts whenever an instance of any preceding activity ends. Therefore, the transportation operation takes place when the loading is finished. The queue nodes represent events that can occur in line ups, for example, when any of the machines turn to idling state. Additional information regarding the aforementioned terminology can be found in [19].

Figure 4 shows dump truck/s stationed at an idling stage, forming a queue until an available loading spot becomes vacant for filling. The number of available loading spaces mirrors the count of excavators on site [1]. Upon reaching a loading spot, a truck awaits while the assigned excavator commences the filling operation, subsequently returning to the idling stage. The filling operation relates to the filling time. Once a dump truck is loaded, it departs from the construction site to haul the excavated soil to a temporary dump location. In cases where another truck awaits its turn, it promptly occupies the newly vacated loading space, ensuring seamless loader utilization. Conversely, should all loading spots be temporarily occupied, the excavator transitions into an idling mode until its services are requested. The loading process adheres to predefined parameters, specifically the truck's capacity, the excavator's bucket size, considering bank and loose states, which collectively determine the filling efficiency. Notably, the drilling rig operates independently from the excavator-dump truck system. It executes drilling operations and can enter very short idling modes when encountering major changes in soil type classifications, with the duration governed by distribution patterns, all are more detailed in Table 3. This structured approach intends to optimize construction site efficiency, ensuring resource allocation aligns with operational demands.

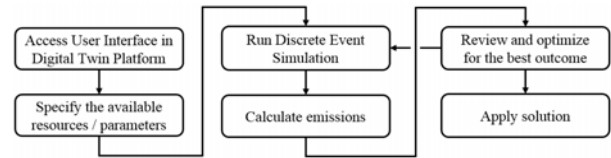


Figure 3. Simulation logic to retrieve desired solution.

Table 2. Time distributions of activities for construction machines used in the DES.

Equipment	Working modes	Time distributions
Drilling rig	Drilling	Uniform (10,14) [min]
	Idling	Normal (30,10) [sec]
Excavator	Loading	Normal (25,5) [sec]
	Idling	
Truck	Driving	Uniform (5,8) [min]
	Idling	

Table 3. Simulation parameters and preselected units.

Parameter	Explanation	Units
Simulation time	The corresponding work time the simulation runs	min [480 min = 1 workday]
Excavator	Loading material on dump truck, idling at times	No. [1]
Dump truck/s	Hauling off material, idling at times	No. [1..n]
Drilling rig/s	Boring piles in uncertain soil conditions, unloading material, idling at times	No. [1-2]
Distance to travel	Hauling distance from construction to dump site	km [1..30]
Initial soil amount	The amount of soil left over from previous drills	m ³ [1..n]
Soil type variations	Possibility of slightly varying excavated soil	Factor [1-3]
Excavator bucket size	Volume for one swing of the excavator arm	m ³
Truck capacity	Truck bed size, incl. bank and loose states	m ³ [5-10]

4.5 Realistic simulation scenario

The logic explained above was applied to simulate activities, inspired by a real-life construction operation (a high-rise building in Aarhus, Denmark). On this project, a drilling rig produced foundation walls and, in the process, an excavator loaded the bored soil material on a dump truck that hauled it away to a nearby temporary storage location (Figure 5). A site manager interested in the construction equipment emissions and fuel consumption values provided the simulation parameters according to Table 3. The simulation was run 100 times to improve the confidence level of the results. Mean values were returned to the user for review (Table 4).

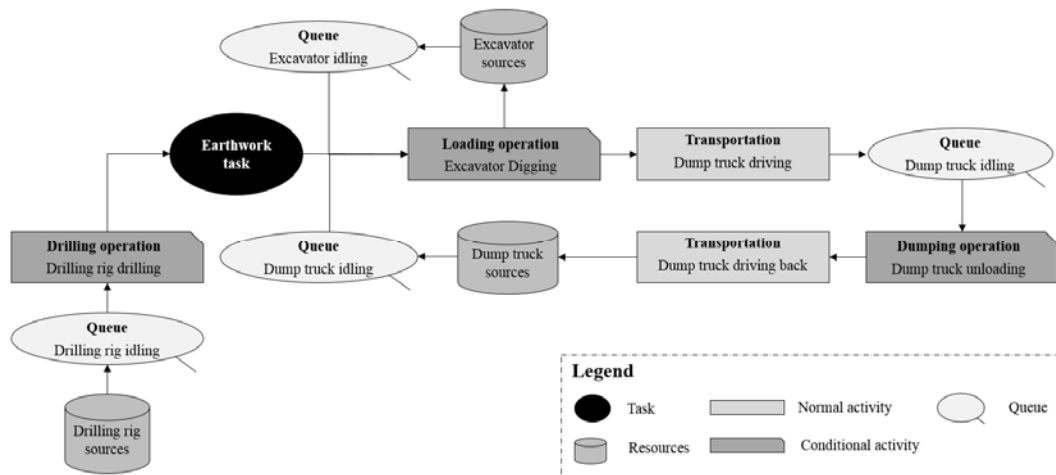


Figure 4. Discrete Event Simulation (DES) flowchart of the construction site operation.

The simulated activities replicate a real high-rise building construction site in Aarhus, Denmark. A drilling rig produced foundation walls and the excavator loaded a dump truck to haul excessive earth material away to a nearby temporary dumping location (Figure 5).

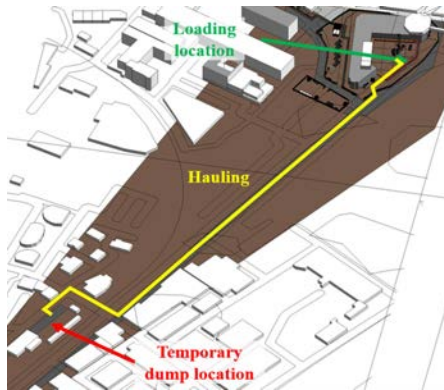


Figure 5. BIM-based construction site layout visible in the online user interface on the Digital Twin Platform

Our DTP contains a BIM-based construction site layout plan that is linked via the Autodesk Platform Services (APS) to our DES. This is accessible online in a user interface in a web browser. The simulated scenarios allow to change any of the parameters given in Table 2, for example, the hauling distance can change from the nearby site (1 km) to further away (30 km) increasing the transportation time, creating more emissions and higher idling times for the truck, unless the end user (a project manager) decides to run the DES to optimize the system for both environmental and economic goals. Detailed results from the simulation can be visualized in the dashboard of the DTP (note: not shown) and are accessible in a .CSV-file for further processing using tools construction site practitioners are more familiar with (e.g., Excel).

5 Results and Discussion

For demonstration, we simulated two construction alternatives based on the realistic scenario. One serves as a baseline scenario (Alternative A) where one drilling rig, one truck and one loader are employed for soil digging and hauling for one full day. Alternative B employs one additional drilling rig of identical machine parameters.

In the comparison between the two alternative construction plans (Table 4: A and B), it is evident that the utilization of two drilling rigs instead of using one is more productive. In the same work time Alternative B dug out nearly twice (199.04%) the amount of soil as Alternative A (158.12 m³ vs. 79.44 m³, respectfully, using the same number of trucks and loaders). The trucks also moved twice the amount of soil (153.33 m³ vs. 76.07 m³, respectfully). However, the sum of all equivalent emission values (NO_x and CO₂) and the total fuel consumption of all machines increased by only 57.65%, 70.05%, and 74.39%, respectfully.

The system's productivity of the two drill rigs, excavator and truck doubled compared to the scenario with only one drilling rig. From the decomposition of emission and fuel consumption, one can see that two drill rigs create nearly twice as much the amount of emissions. However, since the utilization rate of both loader and truck increased (note: less idle times), emissions drop sharply by using Alternative B instead of A. Moreover, the excavator loading the excavated material emits more CO₂ emissions and consumes more fuel in Alternative B as it idles less and performs more loading cycles.

Another key finding is that the alternative construction plan B doubled the amount of soil transferal in the same work time, resulting in higher productivity and less emissions per unit of work output as shown in Table 4. As anticipated, the drilling rig consumes a substantial amount of Diesel fuel, resulting in higher emissions. The loader follows as the second-largest

emitter, trailed by the dump truck.

In brief, CO₂ emissions exhibit a strong correlation with fuel consumption across all the machines. These and more insights can offer a comprehensive overview of emissions across the machinery fleet that is used in

construction projects, aiding eventually in environmentally conscious decision-making and optimization efforts.

Table 4. Results to a simplified Discrete Event Simulation (DES) comparing two alternative systems (A and B, while B consists of one additional drilling rig) (note: the choice of selecting alternative construction plans depends on the environmental and economic targets of the construction project and might be site and situational dependent).

Machine	Time [mins]		NO _x [g]		CO ₂ [kg]		Fuel [kg]		Soil moved [m ³]	
Alternative	A	B	A	B	A	B	A	B	A	B
Drilling rig 1	477.48	475.86	149.21	148.80	191.82	191.15	60.79	60.58		
Drilling	398.30	396.80	83.64	83.33	166.28	165.65	52.66	52.46		
Idling	79.18	79.06	65.56	65.46	25.54	25.50	8.13	8.12		
Drilling rig 2		475.98		148.79		191.21		60.59	79.44	158.12
Drilling		396.90		83.31		165.70		52.48		
Idling		79.08		65.48		25.51		8.12		
Loader	460.59	469.49	102.78	103.26	67.99	76.68	16.25	18.91		
Idling	446.30	440.67	101.07	99.79	58.81	58.06	13.48	13.31		
Loading	14.29	28.82	1.71	3.46	9.19	18.61	2.77	5.60		
Truck	477.37	477.11	7.25	7.84	25.03	25.34	7.84	7.94		
Idling	395.03	381.46	2.96	2.86	19.04	18.38	5.97	5.76		
Loading	33.76	32.79	1.76	1.71	2.46	2.39	0.77	0.75		
Hauling	15.24	30.73	0.79	1.60	1.11	2.24	0.35	0.70	76.07	153.33
Unloading	33.34	32.13	1.74	1.67	2.43	2.34	0.76	0.73		
Soil left [m ³]									3.38	4.78
Sum	480	480	259.24	408.69	284.84	484.38	84.88	148.02		
Savings according to Sum (2*A-B)/A [%]			42.35%		29.95%		25.61%			
Rates for emission or fuel consumption by work output (def. as soil moved by truck)			3.41 g/m ³	2.67 g/m ³	3.71 kg/m ³	3.15 kg/m ³	1.16 kg/m ³	0.97 kg/m ³		

6 Conclusion and Outlook

This paper presented a novel approach where a Digital Twin Platform came into play to integrate, among other important construction data, a BIM-based construction site layout model, emissions from heavy equipment coming from IoT-sensors and being further processed, and a DES. While the simulation presented one noteworthy case of comparing alternative construction plans (resources and schedules), initial value was generated when assessing for environmental vs. economic project objectives.

While the overall workflow and method have proven to work successfully, further work is necessary exploring more complex construction scenarios. This includes some statistical analysis of the results and the implications on environmental policies. Future work can also explore agent-based (simulation) modeling (ABM). Existing ABM primarily replicate how individuals within an organization or across various organizations interact in a synthetic environment, where agents make decisions and engage in communication [23]. Yet, several studies have explored the use of ABM to enhance efficiency in construction operations. With respect to their work, ABM has proven valuable particularly in earthmoving operations, due to its ability to accommodate diverse

equipment specifications and provide more accurate time and cost estimates compared to simulation methods like Discrete Event Simulation (DES) [24]. Aside from researchers employing ABM to simulate earthmoving operations by equipping equipment agents with state charts and static and dynamic properties, ABM can be utilized to evaluate how off-site congestion and traffic flow of equipment agents impact the earthmoving efficiency [25]. Future work can also concentrate on the coordination between the earthmoving equipment agents at the project level and the provided safety measures to prevent collisions [26] and noise emissions [27].

Acknowledgements

The authors gratefully acknowledge the Ministry of Environment of Denmark for their financial support of the project “Green Construction Site of the Future” (MUDP Sags nr.: 2020-37291).

References

- [1] Jassim H.S.H., Lu W., Olofsson T. Assessing energy consumption and carbon dioxide emissions of off-highway trucks in earthwork operations: An

- artificial neural network model. *Cleaner Production*, 198: 364-380, 2018. <https://doi.org/10.1016/j.jclepro.2018.07.002>
- [2] Winther K. Sensor based control device – For fast measurement of NOX and particles from internal combustion engines. DTI report, 2020.
 - [3] Andrade L.M., Teizer J. Monitoring climate forcers from heavy construction equipment emissions in a digital twin framework. *CIBW099W123*, 409-420, 2023. <https://doi.org/10.24840/978-972-752-309-2>
 - [4] Dansk Byggeri. Recommendations to the Danish Government from the Climate Partnership of the Constr Industry. <https://www.frinet.dk/media/3174/climate-partnership-construction-report-march-2020-bat-kartellet.pdf> (Accessed 20.06.2022)
 - [5] Trauer J. et al. What is a Digital Twin? Definitions and insights from an industrial case study in technical product development. In *Design Society*, 757-766, 2020. <https://doi.org/10.1017/dsd.2020.15>
 - [6] Teizer J., Wandahl S. Simplified Emissions Measurement System for Construction Equipment, *Construction Research Congress*, 474-482, 2022. <https://doi.org/10.1061/9780784483961.050>
 - [7] Cao T. et al. Evaluations of in-use emission factors from off-road construction equipment. *Atmospheric Environment*, 147: 234-245, 2016. <https://doi.org/10.1016/j.atmosenv.2016.09.042>
 - [8] Jiang Y. et al. Digital twin-enabled real-time synchronization for planning, scheduling, and execution in precast on-site assembly. *Autom. Constr.*, 141: 104397, 2022. <https://doi.org/10.1016/j.autcon.2022.104397>
 - [9] Abolhasani S., Fey C. Real-world in-use activity, fuel use, and emissions for nonroad construction vehicles: a case study for excavators. *Air & Waste Mgmt. Assoc.*, 58(8): 1033-1046, 2012. <https://doi.org/10.3155/1047-3289.58.8.1033>
 - [10] Lewis P. et al. Requirements and incentives for reducing construction vehicle emissions and comparison of nonroad diesel engine emissions data sources. *Constr. Eng. Mgmt.*, 135(5): 341-351, 2009. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000008](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000008)
 - [11] Monnot J.M., Williams R.C.. Construction Equipment Telematics. *Constr. Eng. Mgmt.* 137(10). [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000281](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000281)
 - [12] Sato et al. International and Sectoral Variation in Industrial Energy Prices 1995–2015. *Energy Economics*, 78: 235-258, 2019.
 - [13] Ligterink N.E., et al. Supporting Analysis on Real-world Light-duty Vehicle CO2 Emissions. TNO report R10419v32016, TU Delft, 2016.
 - [14] Brentnall A. Discrete-Event System Simulation (International Edition). *J. Simulation*, 1(3): 223-223, 2007. <https://doi.org/10.1057/palgrave.jos.4250022>
 - [15] Halpin D.W., Riggs L.S. *Planning and analysis of construction operations*. Wiley, 1992. ISBN: 9780471555100.
 - [16] AbouRizk S. et al. Research in Modeling and Simulation for Improving Construction Engineering Operations. *Constr. Eng. Mgmt.*, 137(10), 2011. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000288](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000288)
 - [17] Martinez J.C., Ioannou P.G. General purpose simulation with Stroboscope. *Winter Simulation Conference*, 1159-1166, 1994. <https://doi.org/10.1109/WSC.1994.717503>
 - [18] Chua D.K.H., Li G.M. RISim: Resource-interacted simulation modeling in construction. *Constr. Eng. Mgmt.*, 128(3): 195-202, 2002. [https://doi.org/10.1061/\(ASCE\)0733-9364\(2002\)128:3\(195\)](https://doi.org/10.1061/(ASCE)0733-9364(2002)128:3(195))
 - [19] Lu M. Simplified discrete-event simulation approach for construction simulation, *Journal of Constr. Eng. Mgmt.*, 129(5): 537-546, 2003. [https://doi.org/10.1061/\(ASCE\)0733-9364\(2003\)129:5\(537\)](https://doi.org/10.1061/(ASCE)0733-9364(2003)129:5(537))
 - [20] Zhang H. Discrete-Event Simulation for Estimating Emissions from Construction Processes. *Mgmt. Eng.*, 31(5): 2015. [https://doi.org/10.1061/\(asce\)me.1943-5479.0000236](https://doi.org/10.1061/(asce)me.1943-5479.0000236)
 - [21] Zinoviev D. Discrete Event Simulation. It's Easy with SimPy! Suffolk University: Boston, MA, USA, 20182018. (Accessed: 16.05.2023)
 - [22] Joubert J.W., Gräbe R.J. Real driving emissions data: Isuzu FTR850 AM. *Data in Brief*, 41, 107975, 2022. <https://doi.org/10.1016/j.dib.2022.107975>
 - [23] Fioretti G. Agent-Based simulation models in organization science. *Org. Res. Met.* 16(2):227-242, 2013. <https://doi.org/10.1177/1094428112470006>
 - [24] Jabri A., Zayed, T. Agent-based modeling and simulation of earthmoving operations. *Autom. Constr.*, 81: 210-223, 2017. <https://doi.org/10.1016/j.autcon.2017.06.017>
 - [25] Kim K., Kim K.J. Multi-agent-based simulation system for construction operations with congested flow. *Autom. Constr.*, 19(7): 867-874. <https://doi.org/10.1016/j.autcon.2010.05.005>
 - [26] Vahdatikhaki F., et al. Enhancing coordination and safety of earthwork equipment operations using Multi-Agent System. *Autom. Constr.*, 81: 267–285, 2017. <https://doi.org/10.1016/j.autcon.2017.04.008>
 - [27] Babazadeh N., Teizer J., Bargstädt H.-J., Melzner J. Predictive simulation of construction site noise emissions from heavy equipment.” *Smart and Sustainable Built Environment*, 2024, <https://doi.org/10.1108/SASBE-08-2023-0226>

Performance Evaluation of Genetic Algorithm and Particle Swarm Optimization in Off-Site Construction Scheduling

Mizanoor Rahman¹, Sang Hyeok Han^{1,2}

¹ Department of Building, Civil and Environmental Engineering, Concordia University, Montréal, QC, Canada

² Centre for Innovation in Construction and Infrastructure Engineering and Management (CICIEM), Gina Cody School of Engineering and Computer Science, Concordia University, Montréal, QC H3G 1M8, Canada

mizanoor.rahman@mail.concordia.ca, sanghyeok.han@concordia.ca

Abstract –

Off-site construction (OSC) is gaining significant attention due to its promising benefits, including reduced time, cost, and waste, along with improved quality, productivity, and safety. However, the dynamic nature of the production process (i.e., non-typical process time) introduces challenges in OSC production line, such as: (i) bottlenecks; (ii) workstation idle time; and (iii) identification of an optimal production sequence. To leverage the full benefits of OSC, a superior production planning and scheduling optimization method become imperative. Therefore, this paper aims to compare the computational performance of the Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) for optimizing OSC production schedule. The methodology consists of the three key steps, including: (i) data analysis; (ii) development of GA and PSO algorithms; (iii) implementation of both GA and PSO in a real-life wall panel production line in Edmonton, Canada. The results reveal that GA outperforms PSO in minimizing project completion time (PCT). Specifically, for 160 wall panels, the PCT using GA is 6112 min, whereas with PSO, it is 6122 min. Conversely, PSO produces results more quickly than GA. For the same set of 160 wall panels, the model runtime is 17.97 sec for GA and 6.0 sec for PSO. The findings of this study offer valuable insights for production managers in selecting the most effective algorithm for optimizing production schedules.

Keywords –

Off-site construction; Genetic algorithm; Particle swarm optimization; Manufacturing; Schedule

1 Introduction

Off-site construction (OSC) is a process in which the building components (e.g., wall, floor, roof) are fabricated in a controlled environment (i.e., factory) and

then transport it to the site for assembly. The adoption of OSC is growing, as it is reducing construction time, waste, environmental impact while increase productivity, safety, and quality [1]. However, it is essential to achieve optimal efficiency in the production line, as most of the activities are performed within a factory environment. In practice, the process time for same type of building component (e.g., wall) within the same production line dynamically varies. For instance, in a wall production line, the process time for each wall panel is not uniform due to varying design parameters, such as wall length, height, thickness, number of studs, and the presence of doors and windows. Due to this non-typical process time, the production line encounters various challenges, such as: (i) struggling to identify the optimal production sequence; (ii) facing significant bottlenecks; and (iii) dealing with excessive idle time at workstations. To address these challenges, an optimal production schedule is crucial. Therefore, it is paramount to identify the best optimization algorithm specifically tailored for OSC.

In this regard, limited research has been conducted to identify a superior schedule optimization algorithm, specifically tailored for OSC. For example, Altaf et al. [2] compared the optimization performance of PSO and simulated annealing (SA) in the event of OSC production line and found that PSO outperformed SA. Yazdani et al. [3] combined three metaheuristic algorithms, namely differential evolution (DE), imperialist competitive algorithm (ICA), and GA, to simultaneously minimize the duration and cost of precast production processes. They found that DE provides better result compared to ICA and GA. Lee and Hyun [4] used GA and simulated annealing (SA) to create an optimal production schedule for multiple projects. However, the literature reveals a gap in research regarding the identification of the superior optimization algorithm between GA and PSO, specifically tailored for OSC production lines.

Therefore, the primary objective of this paper is to determine the superior optimization algorithm between GA and PSO for minimizing project completion time in

OSC. This research conducted through three key steps, such as: (i) data analysis; (ii) develop GA and PSO model for minimizing project completion time (PCT) and obtain optimal production sequence; and (iii) compare the computational performance of GA and PSO to determine best optimization algorithm for OSC.

2 Literature Review

This section presents a literature review on GA and PSO, as the primary focus of this paper revolves around these two algorithms.

2.1 Genetic algorithm

Genetic Algorithm (GA) draws inspiration from biological evolution as a search and optimization method. It initiates a pool of potential solutions, assesses their fitness for a specific problem, selects the top performers, and merges their characteristics through crossover, occasionally incorporating mutations. This cyclic procedure persists across numerous generations, striving to approach an optimal or nearly optimal solution. GA is renowned for its adaptability, proving effective in addressing intricate issues with extensive solution spaces, and finds widespread application in the domains of scheduling and optimization. In the literature, extensive research has been conducted on using GA for OSC schedule optimization. For example, Ko and Wang [5], Ko and Wang [6] developed a GA-based multi-objective optimization model to address the flow shop-sequencing issue in the manufacture of precast components (PC) while considering the buffer sizes between production stations. Nassar [7] used GA to develop an optimal schedule for reducing project duration and interruption durations in a linear project. However, this model may not be suited for OSC because the working process of linear projects (e.g., road projects) differs significantly from that of OSC projects (e.g., wood-based wall manufacturing). Fan et al. [8] used GA to find the optimal schedules for repetitive projects. They introduced a soft logic strategy (i.e., sequencing) to minimize project costs. Agrama [9] developed a multi-objective GA to minimize project duration, number of interruptions, and resource levelling for a repetitive project (i.e., a multi-story building). Yazdani et al. [3] combined three metaheuristic algorithms, namely differential evolution (DE), imperialist competitive algorithm (ICA), and GA, to simultaneously minimize the duration and cost of precast production processes. Hyun et al. [10] developed a multi-objective optimization model using NSGA-II to reduce production time and labour costs for a continuous modular unit production line. Zhang et al. [11] developed an NSGA-II model to solve the multi-objective optimization problem in off-site construction (i.e., precast production) schedules by considering the impact

of disturbance events such as machine malfunctions, order modifications, and unexpected order insertions.

2.2 Particle Swarm Optimization (PSO)

The Particle Swarm Optimization (PSO) is an optimization algorithm inspired by the coordinated movement of bird flocks in nature. In PSO, a group of particles traverses a search space to identify an optimal solution for a given problem. Each particle adapts its position by considering both its individual experience and the collective experience of its peers, all with the goal of discovering the most favourable solution. Several researches have been conducted on PSO for optimizing OSC and flow shop schedule. For example, Tasgetiren et al. [12] used PSO algorithm for minimizing the makespan and total flow time for the permutation flow shop sequencing problems. Guo et al. [13] developed a modified PSO algorithm for obtaining the optimal production schedule by sequencing the manufacturing process. Koulinas et al. [14] created a PSO-based hyper-heuristic algorithm to solve the resource-constrained project scheduling problem (RCPSPP). This hyper-heuristic serves as an upper-level controller for multiple low-level heuristics that navigate the solution space. Altaf et al. [15] combined a DES model with optimization using PSO to generate a more realistic schedule that captures the dynamics of the panel prefabrication process. However, this model may not be ideal for comparing its performance with other algorithms (e.g., GA) because it generates mean PCT. Zhang and Yu [16] developed a planning technique utilizing PSO algorithm to optimize the PC transport process. Hayat et al. [17] introduces the hybridization of the particle swarm optimization with variable neighbourhood search and simulated annealing to tackle permutation flow-shop scheduling problems.

In summary, optimizing a production schedule in OSC is crucial for minimizing overall production costs and ensuring faster project delivery. However, previous studies primarily used either GA or PSO to optimize schedules in OSC. Despite the proficiency of both GA and PSO as optimization algorithms, a comprehensive investigation into determining a superior optimization algorithm specifically for OSC production lines remains unexplored. Accordingly, this paper aims to identify a best performing algorithm between GA and PSO explicitly for OSC.

3 Methodology

To fulfill the objectives of this paper, a research framework is summarized in Figure 1. The framework primarily consists of three procedures, such as: (i)

estimating process times from historical data; (ii) developing optimization models, including GA and PSO, specifically suited for OSC; and (iii) implementing and comparing the results of GA and PSO using a real-life wall frame manufacturing factory located in Edmonton, Canada. The process time for each component (e.g., wall panel) at every workstation, the number of workstations, and the number of panels to be produced in a given project are used as input. The criteria of this proposed framework involve one-panel flow (i.e., each workstation can only perform its tasks on one panel at a time) and the sequence (i.e., order) of workstations. In this respect, the optimization models are implemented using Python version 3.11.3. Moreover, the models are run on an Intel® Core™ i7 CPU with a processing speed of 3.40 GHz. Simultaneously, their computational performance in terms of PCT and runtime is recorded.

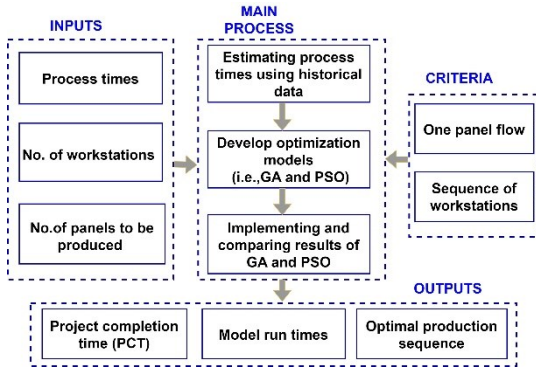


Figure 1. Proposed research framework

3.1 Estimate process times

The overall process to estimate the expected process times for each panel at each workstation using empirical data is shown in Figure 2. In this respect, the process mainly consist of three phases, such as: (i) calculating labor productivity using Equation (1) and (2); (ii) developing histogram to obtain optimistic, most likely and pessimistic productivity because histograms provide a powerful and intuitive way to analyze and interpret the characteristics of a dataset, aiding in decision-making and drawing meaningful conclusions from the data [18]; (iii) calculating weighted average productivity using Equation (6) to incorporate probabilistic process time. It is important to note that before creating the histogram, the data were preprocessed to exclude outliers because outliers significantly influence the determination of realistic statistical parameters, such as mean, upper bound, lower bound[18]. Outliers refer to individual data points that deviate significantly from the overall dataset. Usually represented as individual points beyond the

whiskers in a plot. In essence, if any data point is above the upper boundary or below the lower boundary, it is considered an outlier. It can be calculated using Equation (3)-(5). Finally, the process time for each panel at each workstation is estimated in accordance with Equation (2).

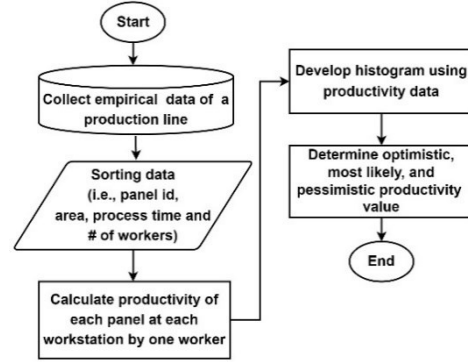


Figure 2. Flow chart for data analysis

$$T = t * N \quad (1)$$

where, T = process time by one worker; t = process time by N number of workers; and N = number of workers.

$$P = A/T \quad (2)$$

where, P = labour productivity; and A = panel area.

$$IQR = Q3 - Q1 \quad (3)$$

where, IQR = interquartile range; $Q1$ = first quartile (i.e., median of lower half of data); and $Q3$ = third quartile (i.e., median of upper half of data).

$$\text{Upper bound of data set} = Q3 + 1.5 * IQR \quad (4)$$

$$\text{Lower bound of data set} = Q1 - 1.5 * IQR \quad (5)$$

The weighted average productivity for each workstation is calculated using the optimistic, most likely, and pessimistic value of productivity using Equation (6).

$$P_{avg} = \frac{O + 4m + P}{6} \quad (6)$$

where P_{avg} = weighted average productivity; O = optimistic productivity; m = most likely productivity; and P = pessimistic productivity.

3.2 Optimization model formulation

The objective of this research is to find the best optimization algorithm between GA and PSO in terms of minimize the project completion time (PCT) and run time

in an OSC production line. The objective function is representing in Equation (7).

$$\text{Min } C(P_i, S_j) = \text{Min} [\max\{C(P_{i-1}, S_j), C(P_i, S_{j-1})\} + T_{i,j}] \quad (7)$$

where $C(P_i, S_j)$ = PCT of i^{th} panel at workstation j ; $C(P_{i-1}, S_j)$ = completion time of $(i-1)^{\text{th}}$ panel at workstation j ; $C(P_i, S_{j-1})$ = completion time of i^{th} panel at workstation $(j-1)$; and $T_{i,j}$ = process time of i^{th} panel at station j . The wall panel production sequence is considered as the decision variables. Moreover, two types of constraints are considered, such as: (i) the wall panel must follow the sequence of workstation; (ii) each workstation should work on a single wall panel (i.e., one panel flow).

3.2.1 GA model

The process flow of minimizing PCT using GA is outlined in Figure 3. The process begins with the initialization of an initial population and each population is evaluated for its fitness based on Equation (7). After that GA generate optimum solution through selection, crossover, and mutation. In the selection process, population with superior fitness are chosen to form a mating pool.

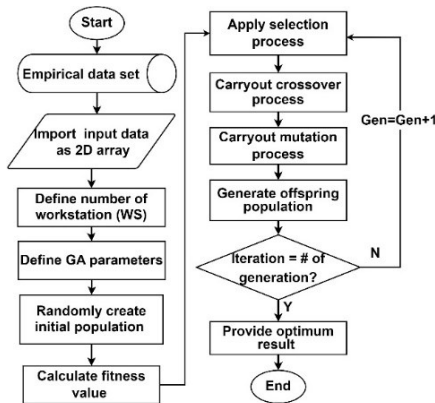


Figure 3. Process flow of GA

In the crossover phase, a partially mapped crossover (PMX) strategy is utilized. This involves generating offspring solutions by selecting a sub-set of gene from one parent and replacing it with another parent. Through this genetic recombination process, offspring can get advantageous attributes from parents, potentially leading to improved solutions. As shown in Figure 4, the random mapping of genes J3-J6 in chromosome 1 (i.e., parent 1) interchange with the genes J6-J1 in chromosome 2 (i.e., parent 2) for crossover, resulting in the generation of offspring 1 (i.e., child 1) and offspring 2 (i.e., child 2).

Step 1: Randomly select crossover region (i.e., partially mapped)

Parent 1	J1	J2	J3	J4	J5	J6	J7	J8	J9
Parent 2	J2	J9	J6	J4	J5	J3	J7	J8	J1

Step 2: Generate offspring by exchanging genetic information between parents

Child 1	J1	J2	J6	J4	J5	J3	J7	J8	J9
Child 2	J2	J9	J3	J4	J5	J6	J7	J8	J1

Figure 4. Illustration of crossover process

Furthermore, to preserve population diversity and avoid premature convergence, the mutation operation is employed as an effective strategy. In this study, a two-point swap mutation technique is utilized, where a random pair of genes (i.e., panels) within the chromosome is chosen, and their positions are swapped to generate offspring, forming the foundation for the subsequent generation. As shown in Figure 5, child 1 exchanges the positions of genes J2 and J3, while child 2 swaps the positions of genes J3 and J8, leading to the creation of new offspring.

Child 1	J1	J2	J6	J4	J5	J3	J7	J8	J9
Child 2	J2	J9	J3	J4	J5	J6	J7	J8	J1

Before mutation

Child 1	J1	J3	J6	J4	J5	J2	J7	J8	J9
Child 2	J2	J9	J8	J4	J5	J6	J7	J3	J1

After mutation

Figure 5. Illustration of mutation process

3.2.2 PSO model

The process flow of PSO is depicted in Figure 6, encompassing essentially six steps, including: (i) randomly generate initial no of particles where each particle represents a solution. For OSC scheduling problem, particle is a list of sequential panels; (ii) calculate the fitness value (i.e., PCT) for each particle; (iii) find local best position (i.e., chose a job sequence that provide minimum PCT between current and previous iteration) and global best position (i.e., chose the job sequence that provides minimum PCT among all the particle in current iteration); (iv) the iteration continue if it does not reached maximum number; (v) update the velocity of the particles as per Equation (8); and (vi) update the position (i.e., sequence of panels) for next iteration.

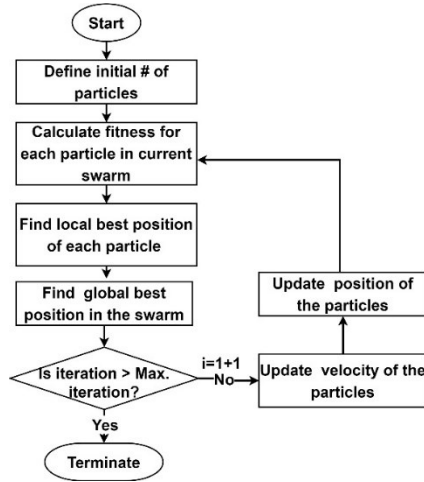


Figure 6. Process flow of the PSO algorithm

$$V_i(t+1) = \omega v_i(t) + c_1 r_1 (p(t)(i, lb) - x_i(t)) + c_2 r_2 (p(t)(gb) - x_i(t)) \quad (8)$$

where $v_i^{(t+1)}$ = velocity of i^{th} particle at $(t+1)$ iteration; ω = inertia weight; $v_i^{(t)}$ = velocity of i^{th} particle at current (i.e., t) iteration; c_1 and c_2 = acceleration coefficient; r_1 and r_2 = random numbers value between 0 and 1; $p^{(t)}(i, lb)$ = local best of i^{th} particle at current iteration (i.e., t); $p^{(t)}(gb)$ = global best position.

4 Implementation and results

The proposed scheduling method is applied to a light gauge steel (LGS) wall panel production line located in Edmonton, Canada, dedicated to manufacturing light gauge steel (LGS) wall panels for both commercial and residential building. The production line primarily consists of three workstations, such as: (i) assembly station where the required number of studs and tracks are prepared based on the shop drawings; (ii) framing station, where the framing components are securely fastened together to form wall panels; and (iii) sheathing station where the drywall is installed on framed components. In this paper, the process time for 166 wall panels is collected and analysed. However, for illustrative purposes, the collected data for 10 panels are shown in Table 2. Subsequently, Equations (1) and (2) are employed to calculate labour productivity for each panel. For example, the area of panel E831 in Table 2 is 94 ft², it takes 41 minutes to complete by 2 workers at assembly workstation. Therefore, the process time by one worker is 82 minutes, and labor productivity is calculated as $94/82 = 1.15$ ft²/min. Similarly, productivity for panel E831 at the framing and sheathing workstations is 0.94

ft²/min and 1.57 ft²/min, respectively.

From the data analysis, it is evident that productivity is not consistent, meaning it varies from panel to panel. To address this variability in productivity, a weighted average productivity is adopted to determine the expected productivity for each workstation. To derive the optimistic, most likely, and pessimistic productivity, histograms are constructed for each workstation using 166 empirical data points. As illustrates in Figure 7 the framing station's diverse productivity metrics, with optimistic, most likely (i.e., median), and pessimistic values of 0.16 ft²/min, 1.15 ft²/min, and 3.19 ft²/min, respectively. Similarly, Figure 8 shows the sheathing station's productivity statistics, showcasing optimistic, most likely, and pessimistic values of 0.14 ft²/min, 0.79 ft²/min, and 2.29 ft²/min, respectively. As shown in Figure 9, the nailing station's productivity spanning 0.14 ft²/min (optimistic), 0.71 ft²/min (average), and 1.91 ft²/min (pessimistic) The weighted average productivity for the assembly, framing, and sheathing workstations is calculated as 1.33 ft²/min, 0.93 ft²/min, and 0.82 ft²/min, respectively. These productivity values are employed to calculate the process time for each panel at each workstation, as presented in Table 1

Table 1. Estimated process time at each workstation

Panel Id	Panel area (ft ²)	Process time (min)		
		Assembly	Framing	Sheathing
E 1099	160	60	57	65
E 819	111	42	40	45
E 561	132	50	47	54
E 767	130	49	46	53
E 779	130	49	46	53
E 861	130	49	46	53
E 1140	136	51	49	56
E 807	136	51	49	55
E 1129	92	35	33	37

Table 2. Sample empirical data for 10 panels

Panel Id	Panel area (ft ²)	Assembly station			Framing station			Sheathing station		
		Process time (min)	No of workers	Productivity (ft ² /min)	Process time (min)	No of workers	Productivity (ft ² /min)	Process time (min)	No of workers	Productivity (ft ² /min)
E 831	94	41	2	1.15	25	4	0.94	20	3	1.57
E 789	151	65	1	2.32	110	3	0.46	100	3	0.50
E 780	132	48	1	2.75	40	3	1.10	85	3	0.52
E 809	56	38	2	0.73	49	3	0.38	24	3	0.78
E 814	109	65	2	0.84	60	3	0.61	72	3	0.50
E 820	62	26	2	1.19	32	3	0.64	54	3	0.38
E 790	22	15	2	0.74	11	2	1.00	13	3	0.57
E 651	136	35	2	1.94	55	2	1.23	47	2	1.44
E 622	69	15	2	2.31	59	2	0.59	33	2	1.05
E 630	45	20	2	1.12	29	2	0.77	40	2	0.56

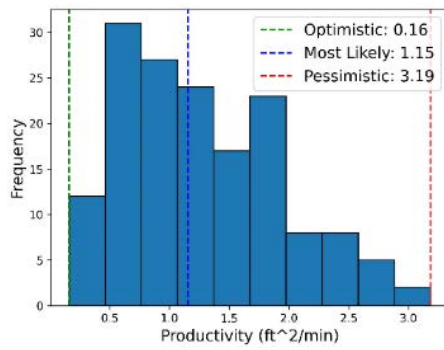


Figure 7. Histogram for assembly station

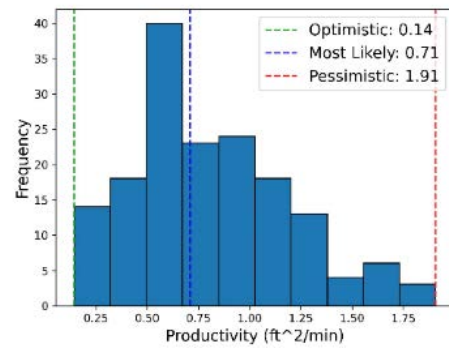


Figure 9. Histogram for sheathing station

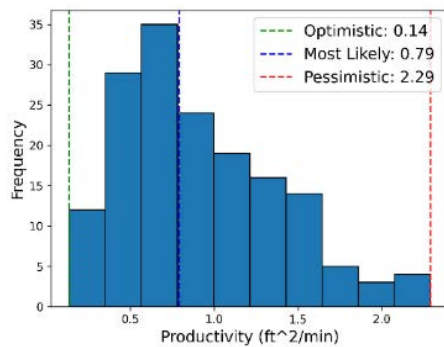


Figure 8. Histogram for framing station

4.1 Application of GA

To automate the implementation of GA and enhance the efficiency of minimizing PCT through panel sequencing, a Python-based script has been meticulously developed. The adopted GA parameters are follows: (i) population size 20; (ii) mutation rate 0.1; (iii) crossover rate 0.8; (iv) number of generations 2500; and (v) number of panels 50. As shown in Figure 10, following 1600 iterations, the calculated optimum PCT is 2196 minutes.

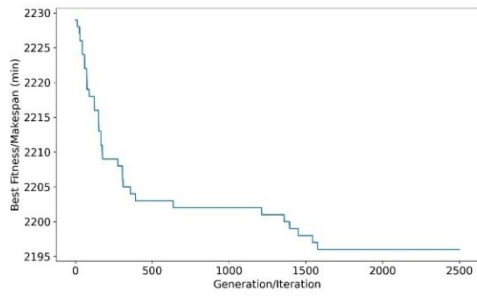


Figure 10. PCT during each iteration using GA

4.2 Implementation of PSO

The objective of this PSO-based optimization model is also to minimize the PCT by sequencing the wall panel production. The selected parameters for running the PSO model are as follows: (i) number of wall panels - 50; (ii) number of particles - 20; (iii) number of iterations - 500; (iv) inertial weight - 0.5; (v) cognitive weight - 2.0; and (vi) social weight - 2.0. As shown in Figure 11, the convergence achieved at 150 iterations and, the PCT reduced from 2237 min to 2226 min.

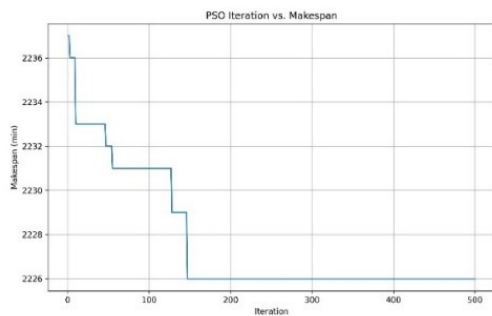


Figure 11. PCT during each iteration using PSO

4.3 Comparison of results

To identify the best-performing algorithm between GA and PSO for minimizing PCT in OSC, 13 sets of wall panels are selected, such as number of wall panel 10,20,30,40,50,60,70,80,90,100,120,140 and 160. As shown in Figure 12, the performance of GA for minimizing PCT is outperformed PSO. For instance, the minimum PCT for 160 panels is 6112 min using GA, whereas with PSO, it is 6122 min. Moreover, the run time at each iteration is record for both GA and PSO. As shown in Figure 13, the run time for GA is relatively higher than PSO. For example, the runtime for 160 panels

using GA is recorded as 17.97 sec, whereas with PSO, it is 6.0 sec.

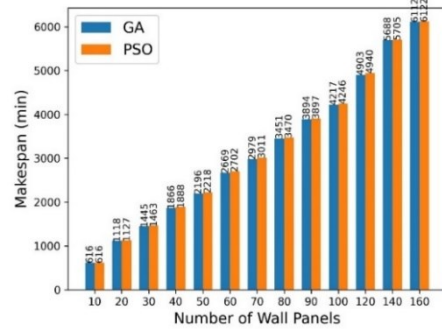


Figure 12. Comparative PCT for GA and PSO

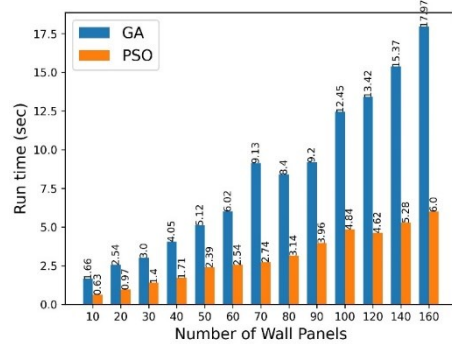


Figure 13. Comparative run time for GA and PSO

5 Conclusions and future works

In this study the performance of GA and PSO for minimizing PCT in OSC production schedule is evaluated. The proposed methodology employed GA, and PSO to optimize OSC production schedule in terms of PCT and run time. In the data analysis stage, a histogram is developed to estimate expected productivity for each workstation. The proposed research framework is implemented on a light gauge steel (LGS) wall panel production line at Edmonton in Canada. Each model (i.e., GA and PSO) is implement in 13 sub-sets of wall panels to minimize the PCT. The results demonstrate that GA provides better results than PSO for minimizing PCT. For instance, the PCT using GA is 6112 min, while it is 6122 min using PSO for 160 wall panels. In contrast, PSO generates output relatively faster than GA for the same set of 160 wall panels. For example, the model run time for GA is around 17.97 sec, while it takes around 6 sec

for PSO.

The original contribution of this research assists planners in choosing the best optimization algorithm, eliminating the need for trial and error with multiple algorithms. While this study yields satisfactory outcomes, it is limited to the comparison between two algorithms. To address this limitation, in the future, the proposed methodology can be further expanded by comparing these two algorithms with other search algorithms, including Simulated Annealing, Ant Colony, and Tabu Search algorithms, to find the best performing algorithm for this type of optimization problem, specifically in the context of OSC.

Acknowledgements

The authors gratefully acknowledge the financial support provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) to conduct this research. Additionally, they extend their heartfelt gratitude to Angat Pal Singh Bhatia for generously providing the production line data.

References

- [1] Rahman, M. and H.R. Sobuz. Comparative study of IPS & PPVC precast system—A case study of public housing buildings project in Singapore. in Proceedings of the 4th International Conference on Civil Engineering for Sustainable Development (ICCESD 2018), KUET, Khulna, Bangladesh. 2018.
- [2] Altaf, M.S., et al., Integrated production planning and control system for a panelized home prefabrication facility using simulation and RFID. *Automation in construction*, 2018. 85: p. 369-383.
- [3] Yazdani, M., et al., Production scheduling of off-site prefabricated construction components considering sequence dependent due dates. *Environmental Science and Pollution Research*, 2021: p. 1-17.
- [4] Lee, J. and H. Hyun, Multiple modular building construction project scheduling using genetic algorithms. *Journal of Construction Engineering and Management*, 2019. 145(1): p. 04018116.
- [5] Ko, C.-H. and S.-F. Wang, Precast production scheduling using multi-objective genetic algorithms. *Expert Systems with Applications*, 2011. 38(7): p. 8293-8302.
- [6] Ko, C.-H. and S.-F. Wang, GA-based decision support systems for precast production planning. *Automation in construction*, 2010. 19(7): p. 907-916.
- [7] Nassar, K., Evolutionary optimization of resource allocation in repetitive construction schedules. *Journal of Information Technology in Construction (ITcon)*, 2011. 10(18): p. 265-273.
- [8] Fan, S.-L., K.-S. Sun, and Y.-R. Wang, GA optimization model for repetitive projects with soft logic. *Automation in Construction*, 2012. 21: p. 253-261.
- [9] Agrama, F.A. Multi-objective genetic optimization for scheduling a multi-storey building. in *The International Conference on Civil and Architecture Engineering*. 2014. Military Technical College.
- [10] Hyun, H., et al., Multiobjective optimization for modular unit production lines focusing on crew allocation and production performance. *Automation in Construction*, 2021. 125: p. 103581.
- [11] Zhang, R., et al., Green optimization for precast production rescheduling based on disruption management. *Journal of Cleaner Production*, 2023. 420: p. 138406.
- [12] Tasgetiren, M.F., et al., A particle swarm optimization algorithm for makespan and total flowtime minimization in the permutation flowshop sequencing problem. *European journal of operational research*, 2007. 177(3): p. 1930-1947.
- [13] Guo, Y., et al., Applications of particle swarm optimisation in integrated process planning and scheduling. *Robotics and Computer-Integrated Manufacturing*, 2009. 25(2): p. 280-288.
- [14] Koulinas, G., L. Kotsikas, and K. Anagnostopoulos, A particle swarm optimization based hyper-heuristic algorithm for the classic resource constrained project scheduling problem. *Information Sciences*, 2014. 277: p. 680-693.
- [15] Altaf, M.S., M. Al-Hussein, and H. Yu. Wood-frame wall panel sequencing based on discrete-event simulation and particle swarm optimization. in *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*. 2014. IAARC Publications.
- [16] Zhang, H. and L. Yu, Dynamic transportation planning for prefabricated component supply chain. *Engineering, Construction and Architectural Management*, 2020. 27(9): p. 2553-2576.
- [17] Hayat, I., et al., Hybridization of Particle Swarm Optimization with Variable Neighborhood Search and Simulated Annealing for Improved Handling of the Permutation Flow-Shop Scheduling Problem. *Systems*, 2023. 11(5): p. 221.
- [18] Sheng, Y., et al., Outlier identification in radiation therapy knowledge-based planning: a study of pelvic cases. *Medical physics*, 2017. 44(11): p. 5617-5626.

Developing a novel application to digitalize and optimize construction operations using low-code technology

Eder Martinez^{1,2}, Louis Pfister², and Felix Stauch²

¹School of Architecture, Civil Engineering and Geomatics, University of Applied Sciences and Arts Northwestern Switzerland (FHNW), Switzerland

²Division Civil Engineering, Implenia AG, Switzerland
eder.martinez@fhnw.ch, louis.pfister@implenia.com, felix.stauch@implenia.com

Abstract

Low-code is an emerging technology allowing subject experts with no information technology background to develop their own digital solutions to operational challenges. Low-code has shown potential to support digitalization in different business scenarios. Nevertheless, there is a lack of scientific literature exploring how this technology could be leveraged in the construction industry. This paper addresses this gap using case study research to analyze the development and implementation of a novel low-code application to monitor production in tunneling construction. The outcomes reveal that the involvement of people closer to operation in the low-code development process supports productive development practices enabling the delivery of an application meeting project requirements and supporting on-site adoption. The introduction of the application resulted in an increase in productivity during the reporting process. The innovative use of low-code technology in the construction industry also reveals opportunities for future research.

Keywords –

Low-code, Citizen developer, Low-code development platform, Lean construction, Building Information Modeling, Construction industry, Digitalization.

1 Introduction

Low-code development is an emerging technology aiming to enable people without coding or programming expertise to develop applications and digital solutions without the support of Information Technology (IT) professionals. Richardson et al. [1] coined the term low-code in 2014 arguing that it emerged as a response to rapid changing business environments and requirements for faster and cheaper technology development [2, 3].

In comparison to traditional programming methodologies, low-code seeks to reduce manual coding

efforts and lower technical entry barriers in technology development, thereby fostering a more efficient and productive development. Low-code abstracts complex programming into an application modeler equipped with an intuitive graphical interface containing pre-defined components/templates to support the application development process. Users can drag and drop elements onto the graphical interface or connect directly to related databases without the need of writing code [4]. Embracing this programming approach empowers a distinct group of individuals, commonly identified as "citizen developers", to actively create their own digital solutions for addressing operational challenges [5, 6].

Low-code is usually seen as an evolution of principles within Computer-Aided Software Engineering (CASE) [7]. For instance, Bucaioni et al. [8] argue that low-code should be regarded as a set of tools and techniques within the domains of Model-Driven Engineering (MDE). Some authors even contend that the rise in popularity of the low-code trend may not in itself offer significant technical innovation, suggesting that its rise may be due to a rebranding of related CASE/MDE concepts [9].

Low-code offers an alternative to the ever-increasing demand of digitalization, particularly in light of the IT skills shortage [4, 10, 11]. This considering that in the next 5 years, the digitalization of different industries will demand the development of more than 500 million applications and digital solutions, exceeding the delivering capacity of IT organizations and developers [12, 13]. In this regard, market reports are very adventurous forecasting that by 2025 nearly 70% of business applications will be developed using some sort of low-code technology [14], with the number of "citizen developers" largely surpassing the number of professional developers [15]. This translates into an expected market growth from USD 13.89 billion in 2021 to USD 94.75 billion by 2028 [16].

If these projections materialize, low-code will influence not only digitalisation of various industries but also the way projects are managed. In fact, the Project Management Institute (PMI) has proactively initiated

educational programs aiming to train project managers to leverage the use of this technology [17].

Low-code is also considered in international research consortiums, such the Horizon 2020 Lowcomote project [18]. Lowcomote aims to equip the next generation of professionals, empowering them to emerge as leaders in the future engineering of low-code development platforms. There are also emerging competence centres such as LowCodeLab@OST [19] in Switzerland seeking to focus on building up strategic citizen development and low-code skills.

Since the introduction of the concept in 2014, low-code technology has been used to support digitalization in a variety of business scenarios. Nevertheless, there is a lack of literature delving into low-code use in the context of processes and operations in the construction industry. This paper contributes to filling this gap by describing and analyzing a case study where low-code was used to develop an application that supports specific construction tasks in the context of tunneling operations.

2 Related literature

Prinz et al. [20], Bucaionni et al. [8], and Pinho et al. [21] offer recent literature reviews about low-code. These studies highlight the limited literature on low-code as well as underline the growing interest of the academic community to fill the gap and assimilate the knowledge generated by the industry. Prinz et al. [20] point out that most of research focuses on the technical aspects of low-code platforms. However, social aspects related to roles and responsibilities, alignment of processes, and communication related to the deployment of low-code in business environment has received little attention. Bucaionni et al. [8] reveal that almost half of sources correspond to grey literature, with the first scientific publication appearing in 2018, four years after the introduction of the low-code term. The authors recommend further investigating the characteristics and nuances of applications created through low-code development. This aims to assess the current state of low-code development practices and analyze broader topics such as scalability and performance metrics. Pinho et al. [21] focus the analysis into advantages and disadvantages in low-code development with particular emphasis on platforms' usability. The authors make notes of caution related to the prominent presence of industry and low-code vendors in literature. This considering that low-code vendors have a strong incentive to enhance their products, making them potentially influential in driving research in this field. Although recognizing them as catalysts for advancing research in this domain can be advantageous, it is crucial to prioritize scientific rigor in peer-reviewed literature, and the primary motivation for their involvement should not solely revolve around business

success.

Based on a systematic review, Martinez and Pfister [22] provide an overview of relevant low-code literature and classify articles using thematic analysis. The articles were clustered in 5 themes: technical, interfaces with contemporary topics, implications to software development, literature reviews, and business applications. This analysis a) confirms the growing scientific interest on low-code, b) confirms that most of the literature focuses on technical aspects of low-code development platforms, c) consolidates a variety of use cases describing low-code use in different business scenarios, and d) reveals the lack of scientific work exploring low-code in the context of the construction industry.

Only a handful of studies address the implications of low-code to business operations, and overall digital transformation. Waszkowski [3] describes the use of a low-code platform for automating business processes in a manufacturing environment. The author emphasizes that one of the main benefits of low-code is the reduction of time in the transfer of requirements from the end-user to the IT developers. This is because with low-code, the citizen developer can play the dual developer/end-user role. In the same direction, Wolff [23] states that manufacturing is very suited for low-code because most of the engineers running business processes are familiar with programming language, thus low-code environment is not fully foreign to them. The author also describes popular low-code manufacturing applications in supply chain, inventory tracking, logistics, and sales. From a broader perspective, Sanchis et al. [24] explore the suitability of low-code to enable digitalization in manufacturing. The authors analyze the status of research and benchmark different low-code platforms to identify challenges and opportunities.

The use of low-code in business scenarios has enabled the identification of different benefits. Practitioners identify faster development, ease of use, and low development cost as the main benefits of low-code [25–27]. On top of reducing hand-coding, the assisted low-code process facilitates data integration, provides space for testing and experimenting new ideas, as well as becomes a single control point for configuration and application maintenance [1]. Richardson and Rymer [28] state that low-code development process can be 5 to 10 times faster than traditional approaches. Industry surveys report that software developers are embracing low-code to accelerate the digital transformation, increase response to changing business requirements, and reduce reliance on hard-to-hire IT developers [29].

One of the main limitations of low-code platforms relates to restricted customization options [25, 30]. The lack of customization derives from the fact that low-code

development is naturally less powerful than traditional programming, and users are limited to the options offered in the low-code platform. Tisi et al. [31] also identify scalability and fragmentation as low-code shortcomings. Scalability refers to the ability to expand the use of a given application to larger systems, projects, or organizations. Fragmentation relates to limited interoperability among low-code platforms and supporting databases. Some organizations hesitate to fully embark into low-code because of the high cost per user [25], and concerns about potential vendor lock-in [29]. Although low-code platforms are intended to be intuitive, users and even IT professionals also require investing time in learning and getting familiarized with a platform's user interface [26]. In more complex scenarios, a citizen developer may be unable to meet demanding requirements, eventually requiring support from software developers [29].

Literature on low-code use in the construction industry context is particularly limited. Martinez et al. [32] proposed the idea of using low-code to support the delivery of construction project in a book chapter involving lean and construction industry 4.0. Expanding on this idea, Martinez and Pfister [22] introduced the first academic work exploring the use of low-code in the context of construction base analyzing the outcomes of multiple case studies. Martinez and Cisterna [33] provide another use case integrating low-code with artificial intelligence.

This study contributes to the emerging body of knowledge by extending the analysis and lessons learned from a prior case study, now considering information available after 18-months of application use.

3 Research methodology

This study uses Case Study research design and adheres to the methodologies suggested by Yin [34]. Case study research is an empirical inquiry that investigates a contemporary phenomenon within a real-life context leveraging the use of multiple sources of data and evidence [34]. In this context, the authors acted as citizen developers supporting the project team in the collection of requirements and developing the application using a low-code platform. The authors carried out a comprehensive study of the interactions that took place throughout the entire development process. This involved analysing various project documents, including meeting minutes and paper data collection forms. They also undertook a comparative assessment, contrasting the current manual process with the improved process incorporating the application developed by the project team.

4 Case study

4.1 Context

The case study involves a tunnelling project using different underground construction techniques. The initial construction phase of the project considers earthmoving operations to remove soft soil in front of the planned tunnel section opening. In a second phase, the project considers 500 m. of tunnelling excavation using the umbrella pipes technique. The final phase, considering more stable geology, involves 3.500 m. of tunnelling construction using drill and blast technique.

Traditionally, the reporting of activities and resources used in different operations is done using paper forms. The information included in these forms considers the different activities and corresponding execution times, as well as data about different resources used (e.g., machines and personnel hours, key materials, and subcontractors). This data is then manually transferred to a spreadsheet where different analyses can be performed.

This manual process to capture and process data has the typical shortcomings of a manual paper-based approach. Manual data processing is time-consuming and error-prone. In addition, it often results in delays in data processing, preventing managers and production teams from accessing real-time project performance information that is critical for timely and fact-based decision making. In this project, three daily shifts are tasked with managing the different paper forms.

4.2 Application development process

The project team aimed to tackle the shortcomings of paper-based project reporting by digitalizing and automating the process of data capturing and handling through a customized mobile application. For this purpose, the project team conducted a series of workshops to understand the data capturing process, identify relevant data inputs, and define the requirements of a digital solution. Along with reducing manual effort for data processing, the team aimed to have real time information about activities and resource utilization which facilitates the identification of improvement opportunities to increase productivity.

The authors acted as "citizen developers" supporting the team in the elicitation of requirements and the materialization of the application in the project digital collaboration environment. In this exercise participated the project manager, site manager, and several key project members. From the workshops, the team was able to consolidate requirements to replace the paper forms used in the project and derive all relevant relationships to build a data model that enables real time generation of project performance reports. The team also defined iterative loops to review prototypes to verify the data

model and improve the application's user interface.

4.3 System architecture

The architecture of the system is depicted in Figure 1. The digital solution is built on Microsoft Power Platform. Microsoft Power Platform is a set of cloud applications that allows automation of business processes and create application with low-code [35]. The data capturing and processing concept developed in this case study combines different components of this platform including PowerApps, Dataverse, and PowerBI. The application development process did not require manual code writing; rather, it involved integrating various functions to establish connections between data and facilitate interactions among pre-defined components.

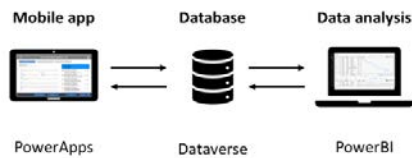


Figure 1: System architecture

In Figure 2, primary users report activities and resources through a Graphical User Interface (GUI) built in PowerApps. PowerApps is a suite of apps, services, connectors, and a data platform offering a rapid application development environment to realize custom applications tailored to specific business needs [36]. PowerApps allows connecting data stored Microsoft Dataverse (or in related online data sources such as SharePoint, Dynamics 365, and SQL Server).

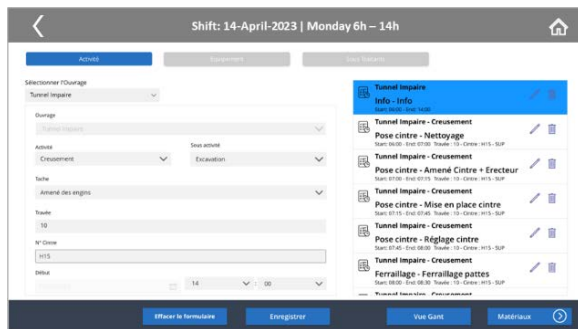


Figure 2: Graphical User Interface

The GUI is connected to the Microsoft Dataverse via integrated data connectors predefined in the Power Platform. From the mobile application data inputs are therefore sent and stored in a Dataverse. Dataverse allows to structure data in a set of cloud-based tables including relationships. In Figure 3, the data model includes two groups of data: master and transactional. Master tables contain predefined project data such as

materials, type of machines, shifts identification number, personal, etc. This data can only be edited by certain key users at the project site and must be updated on a regular basis according to project needs. For example, a new machine arriving to the project must be added to the corresponding table to appear in the GUI.

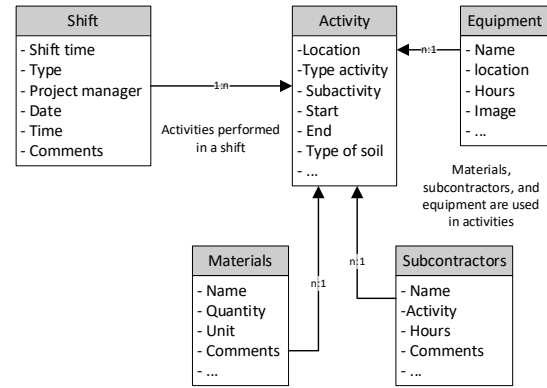


Figure 3: Data model

Transactional data complements master data and corresponds to all process inputs captured onsite through the mobile application. For instance, an application user selects a machine stored in the master data and complements this information with the number of hours it was used during the shift. Transactional data is captured through a series of views in the mobile application and sent to the database where relationships among different data entries are defined via unique identification numbers.

The different views included in PowerApps GUI emulate the paper forms intended to be replaced. The GUI is designed to walk the user through a logical workflow to input data in an organized way. Users begin by entering data about the shift (e.g., shift number, start and end time, location, project manager, etc.). In the background the creation of a shift generates a unique ID number where all the other inputs are attached. The user then can add further transactional data connected to the shift entry by entering the details in the corresponding GUI screen. This allows establishing several one-to-many relationships between the shift and activity/resources tables in the data model.

PowerBI is used for report generation, data analysis, and overall business intelligence. Through interconnected queries and established data relationships, PowerBI retrieves and combines data to analyze and generate the visual insights needed by the project team.

4.4 Outcomes and findings

The project team, working part-time over three months, successfully delivered a functional application through iterative development loops. The authors

highlight efficient requirement elicitation and meeting end-user needs as the primary benefits of low-code. Citizen developers with operational knowledge played a key role, articulating requirements, optimizing development, and bridging exchange between the project team and operations, resulting in faster validation loops.

In terms of process optimization, the materialization of the application removed manual steps from the process, enabling significant saving in personnel time. Table 1 presents estimated times to perform different tasks in the data capturing, controlling, and analysis process. These times were obtained through semi-structured interviews with users on-site. Site engineers, who are responsible for reporting data onsite, save approximately 35 min per shift. That is, 105 min of time savings per day (3 shifts / day). The project manager performs the task of controlling, analyzing data, and reporting daily. For these tasks, the project manager saves approximately 110 min per day. Approximately, the project saves a total of 215 min per day. That translated into an estimate of 235.425 min (3.924 hours) for the 3 years of project duration (215 min/day \times 365 days \times 3 years). Furthermore, there is also the implicit value of having project data processed in real time to make informed decisions about the course of the project. For instance, by allowing the team to promptly define corrective actions when detecting important deviation in the execution of activities and use of related key resources.

Table 1: Estimated time savings

Personnel	Savings (min/day)	Description
Foreman		
Report filling & data processing	105	No need to write on paper and manually digitalize data
Project Manager		
Controlling data	20	Control quality of data inputs
Extracting relevant data	60	Machine, materials, data is readily available
Report creation	30	Processing of data to generate report
Total	215	

In terms of challenges, a limitation associated with low-code development is the restricted ability for customization. The development process is restricted to the predefined options and structures available in a specific platform, potentially limiting the flexibility of development. In this case study, the project's requirements were consistently evaluated in relation to the capabilities of the low-code platform and the expertise of citizen developers. This approach was

explicitly adopted to avoid the need for customizing applications through manual coding, which would require professional IT support. The requirements for digitizing the data capture process in this case did not necessitate customizations beyond what the platform could offer, which may not be the case in other projects.

In terms of scalability, the project team initially considered the development of the data model through basic online lists (SharePoint lists). However, as the data model's complexity grew, a shift to a relational database became essential to streamline the management of relationships among tables. This challenge led to subsequent development iterations and additional licensing costs. This reveals the necessity for careful consideration during the development process, addressing not only user interface requirements but also the comprehensive system architecture defined for the digitalization initiative.

5 Discussion

The authors explore broader industry implications of low-code, specifically its role in supporting the digitalization in the construction sector.

5.1 One-of-a-kind nature of construction

Low-code development offers opportunities to facilitate digitalization of the construction industry processes since it has the potential to cope with the so-called one-of-a-kind nature of construction business. Processes in the construction industry lack standardization because they are usually influenced by site conditions. As a result, information technology tools are unable to cope with a variety of requirements to match every business case, affecting their applicability on site and broader use in other projects. In this aspect, low-code has the potential to support the development of tailored applications with low effort and advanced engagement of people in the field. This could enable different degrees of digitalization of specific construction operations. Related literature highlights the potential of low-code to enable a bottom-up innovation process [37].

5.2 Digitalization and innovation

Several processes on construction sites are still planned and controlled using pen and paper, thus low-code could facilitate their digitalization. Traditionally, the data generated in a project is not properly used and learnings are hardly transferred from one project to the other. In this aspect, low-code offers an alternative to capture key process data and use it to analyze and improve operations. At a project level, teams can have real time access to project performance metrics, identify bottlenecks, and define actions to improve overall

productivity of key processes. At an enterprise level, construction firms could consolidate data coming from different projects and leverage the use of this information to improve performance in other phases of the project life cycle. For example, by consolidating and analyzing data from similar projects, an organization can generate a solid performance database and productivity indicators to improve accuracy when bidding on new projects. In this direction, the authors suggest exploring opportunities for further research on the impact of low-code in areas like Lean Construction, Integrated Project Delivery, Building Information Modeling (BIM), and Construction 4.0. For instance, it would be interesting to analyze how low-code platforms can interact with the Industry Foundation Classes in BIM.

5.3 Computing literacy

Although low-code platforms are intended to be intuitive, the citizen developer still requires some level of technology literacy and basic data modelling knowledge to deliver a working digital solution. This seems to be contradictory with low-code platform vendors and market research statements implying that basically anyone can build applications without software expertise. It is certainly true that any individual could use a low-code platform to build digital solutions. However, the realization of the value of a particular solution relies on how it fits the organizational context, processes, and people interacting with the technology [38]. The role of the citizen developer is therefore key for this purpose and discussions about training programs to equip them with the required skills already started in modern industrial engineering and project management education [17, 39]. In this aspect, exploring what are the necessary skills and competences of a citizen developer to leverage low-code in the context of the construction industry offers an interesting area for further research.

5.4 Traditional versus low-code programming

Low-code development is commonly positioned as a faster and more cost-effective alternative to conventional coding practices. While it is undeniable that low-code offers apparent advantages when compared to traditional approaches, such as the expedited elicitation of requirements leading to potential reductions in both development time and costs, it is crucial to acknowledge the limited body of literature establishing robust foundations for these statements. Estimating software development time and cost is challenging considering the many variables to consider [40], and traditional approaches may not suit the emerging low-code development paradigm [41]. Conducting a comparative analysis between traditional and low-code programming, with a focus on predefined parameters such as time and

cost, presents a great opportunity for future research.

6 Limitations

The study's findings are derived from a single case using a specific low-code platform. To generalize these results, additional studies with similar contexts are needed. In the case study, the project team streamlined the digitalization process by already having a low-code platform integrated into their internal organization's IT system, eliminating the need for evaluation and introduction.

7 Conclusions

The paper contributes to the emerging body knowledge related to low-code use in the construction context by documenting the development and implementation of a novel low-code based application to track project progress.

By lowering the technical barriers to programming, low-code enabled the involvement of people closer to operation in the development process. This enables productive development practices and agility to derive and application that meets the requirements of the project user and facilitates the adoption of the technology on-site. The implementation of the low-code based application allowed significant reduction of the manual work required to capture and process data in the reporting process. Promising results emerge in leveraging low-code for construction operations digitalization. However, it's crucial to address and explore limitations in customization, scalability, and organizational aspects associated with low-code implementation. The novelty of this topic also allows the identification of several areas for further research. For instance, overlaps with Lean Construction, Integrated Project delivery, Building Information Modeling, and Construction 4.0. are interesting areas for future research.

8 Acknowledgments

The authors would like to thank the project team members that participated in the development process of the application discussed in this study. Their contributions to the digitalization initiative, as well as their inputs to the whole process, significantly enriched the findings of this study.

9 References

- [1] Richardson, C., Rymer, J. R., Mines, C., Cullen, A., & Whittaker, D. (2014). *New development platforms emerge for customer-facing applications*.

- On-line: <https://www.forrester.com/report/New-Development-Platforms-Emerge-For-CustomerFacing-Applications/RES113411>
- [2] Fryling, M. (2019). Low code app development. *Journal of Computing Sciences in Colleges*, 34(6), 119. <https://dl.acm.org/doi/abs/10.5555/3344051.3344061>
- [3] Waszkowski, R. (2019). Low-code platform for automating business processes in manufacturing. *IFAC-PapersOnLine*, 52(10), 376–381. <https://doi.org/10.1016/j.ifacol.2019.10.060>
- [4] Sahay, A., Indamutsa, A., Ruscio, D. D., & Pierantonio, A. (2020). Supporting the understanding and comparison of low-code development platforms. In *46th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)* (pp. 171–178). <https://doi.org/10.1109/SEAA51224.2020.00036>
- [5] Oltrogge, M., Derr, E., Stranksy, C., Acar, Y., Fahl, S., Rossow, C., ... Backes, M. (2018). The Rise of the Citizen Developer: Assessing the Security Impact of Online App Generators. In *IEEE Symposium on Security and Privacy (SP)* (pp. 634–647). <https://doi.org/10.1109/SP.2018.00005>
- [6] Everhard, J. (2019). The Pros And Cons Of Citizen Development. Retrieved August 3, 2021, On-line: <https://www.forbes.com/sites/johnneverhard/2019/01/22/the-pros-and-cons-of-citizen-development/?sh=4d04dfae84fd>
- [7] Bock, A. C., & Frank, U. (2021). Low-Code Platform. *Business & information systems engineering*, 63(6), 733–740. <https://doi.org/10.1007/s12599-021-00726-8>
- [8] Bucaioni, A., Cicchetti, A., & Ciccozzi, F. (2022). Modelling in low-code development: a multi-vocal systematic review. *Software and systems modeling*, 21, 1959–1981. <https://doi.org/10.1007/s10270-021-00964-0>
- [9] Cabot, J. (2020). Positioning of the low-code movement within the field of model-driven engineering. In *Proceedings of the 23rd ACM/IEEE International Conference on Model Driven Engineering Languages and Systems: Companion Proceedings* (pp. 1–3). <https://doi.org/10.1145/3417990.3420210>
- [10] Breaux, T., & Moritz, J. (2021). The 2021 software developer shortage is coming. *Communications of the ACM*, 64(7), 39–41. <https://doi.org/10.1145/3440753>
- [11] Carroll, N., & Maher, M. (2023). How Shell Fueled Digital Transformation by Establishing DIY Software Development. *MIS Quarterly Executive*, 22(2). On-line: <https://aisel.aisnet.org/misqe/vol22/iss2/3>
- [12] Moore, S. (2015, December). How to Deliver Enterprise Mobile Apps Faster. On-line: <https://www.gartner.com/smarterwithgartner/how-to-deliver-enterprise-mobile-apps-faster>
- [13] FutureScape, I. D. C. (2019, October). IDC FutureScape: Worldwide IT Industry 2020 Predictions. On-line: <https://www.idc.com/getdoc.jsp?containerId=US45599219>
- [14] Wong, J., Iijima, K., Leow, A., Jain, A., & Vincent, P. (2021). Magic Quadrant for Enterprise Low-Code Application Platforms. *Gartner*. Retrieved August 3, 2023, On-line: <https://www.gartner.com/en/documents/4005939>
- [15] Wong, J., Driver, M., & Ray, S. (2019). *The Future of Apps Must Include Citizen Development*. Gartner. On-line: <https://www.gartner.com/en/documents/3970067>
- [16] Fortune Business Insights. (2022). Low Code Development Platform Market. Retrieved August 29, 2023, On-line: <https://www.fortunebusinessinsights.com/low-code-development-platform-market-102972>
- [17] Project Management Institute. (2021). Introducing PMI Citizen Developer. Retrieved August 3, 2021, On-line: <https://www.pmi.org/citizen-developer>
- [18] Lowcomote. (2023). Lowcomote. Retrieved August 21, 2023, On-line: <https://www.lowcomote.eu/>
- [19] 1LowCodeLab. (2023). LowCodeLab@OST: Kompetenzzentrum für innovatives Low-Coding und strategisches Citizen Development. Retrieved August 21, 2023, On-line: <https://www.ost.ch/de/forschung-und-dienstleistungen/interdisziplinaere-themen/themencluster/lowcodelab>
- [20] Prinz, N., Rentrop, C., & Huber, M. (2021). Low-Code Development Platforms - A Literature Review Completed Research. In *Proceedings of the 27th annual Americas Conference on Information Systems (AMCIS)*. On-line: https://aisel.aisnet.org/amcis2021/adv_info_systems_general_track/adv_info_systems_general_track/2
- [21] Pinho, D., Aguiar, A., & Amaral, V. (2023). What about the usability in low-code platforms? A systematic literature review. *Journal of Computer Languages*, 74, 101185. <https://doi.org/10.1016/j.cola.2022.101185>
- [22] Martinez, E., & Pfister, L. (2023). Benefits and limitations of using low-code development to support digitalization in the construction industry. *Automation in Construction*, 152, 104909. <https://doi.org/10.1016/j.autcon.2023.104909>
- [23] Wolff, I. (2019). Making In-House Apps with Low-Code, No-Code Platforms, 2021(Nov.). On-line: <https://www.sme.org/technologies/articles/2019/oc>

- tober/software-making-in-house-apps-with-low-code-no-code-platforms/
- [24] Sanchis, R., García-Perales, Ó., Fraile, F., & Poler, R. (2020). Low-code as enabler of digital transformation in manufacturing industry. *Applied Sciences*, 10(1), 12.
- [25] Luo, Y., Liang, P., Wang, C., Shahin, M., & Zhan, J. (2021). Characteristics and Challenges of Low-Code Development: The Practitioners' Perspective. In *Proceedings of the 15th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)* (pp. 1–11). <https://doi.org/10.1145/3475716.3475782>
- [26] Alsaadi, H. A., Radain, D. T., Alzahrani, M. M., Alshammari, W. F., Alahmadi, D., & Fakieh, B. (2021). Factors that affect the utilization of low-code development platforms: survey study. *Romanian journal of information technology and automatic control-revista romana de informatica si automatica*, 31(3), 123–140. <https://doi.org/10.33436/v31i3y202110>
- [27] Sahinaslan, E., Sahinaslan, O., & Sabancioglu, M. (2021). Low-Code Application Platform in Meeting Increasing Software Demands Quickly: SetXRM. In *Fourth International Conference Of Mathematical Sciences (ICMS)* (Vol. 2334, p. (070007) 1-4). <https://doi.org/10.1063/5.0042213>
- [28] Richardson, C., & Rymer, J. R. (2016). *Vendor landscape: The fractured, fertile terrain of low-code application platforms*. Forrester. On-line: <https://www.forrester.com/report/Vendor-Landscape-The-Fractured-Fertile-Terrain-Of-LowCode-Application-Platforms/RES122549>
- [29] Outsystems. (2019). *The State of Application Development Is IT Ready for Disruption?* On-line: <https://rb.gy/stqzxlh>
- [30] Alamin, M. A. A., Malakar, S., Uddin, G., Afroz, S., Haider, T. B., & Iqbal, A. (2021). An Empirical Study of Developer Discussions on Low-Code Software Development Challenges. In *IEEE/ACM 18th International Conference On Mining Software Repositories (MSR)* (pp. 46–57). <https://doi.org/10.1109/MSR52588.2021.00018>
- [31] Tisi, M., Mottu, J.-M., Kolovos, D., Lara, J. D., Guerra, E., Ruscio, D. D., ... Wimmer, M. (2019). Lowcomote: Training the next generation of experts in scalable low-code engineering platforms. In *STAF Co-Located Events Joint Proceedings*. <https://doi.org/10.5281/zenodo.4314656>
- [32] Martinez, E., Ezzeddine, A., & García de Soto, B. (2022). Integrating project delivery and Information Technology: challenges and opportunities. In V. González, F. Hamzeh, & L. F. Alarcón (Eds.), *Lean Construction 4.0: Driving a Digital Revolution of Production Management in the AEC Industry* (1st ed., pp. 275–287). Routledge.
- [33] Martinez, E., & Cisterna, D. (2023). Using Low-Code and Artificial Intelligence to Support Continuous Improvement in the Construction Industry. In *Proceedings of the 31st Annual Conference of the International Group for Lean Construction (IGLC31)* (pp. 197–207). Presented at the 31st Annual Conference of the International Group for Lean Construction (IGLC 31), Lille, France. <https://doi.org/10.24928/2023/0236>
- [34] Yin, R. K. (2014). *Case study research: design and methods* (5th Edition.). Los Angeles, CA, USA: SAGE Publications.
- [35] Microsoft. (2022). Microsoft Power Platform. Retrieved August 15, 2022, On-line: <https://powerplatform.microsoft.com/en-us/>
- [36] Microsoft. (2022). What is Power Apps. Retrieved August 11, 2022, On-line: <https://docs.microsoft.com/en-us/powerapps/powerapps-overview>
- [37] Elshan, E., Germann, D., Dickhaut, E., & Li, M. (2023). Faster, Cheaper, Better? Analyzing How Low Code Development Platforms Drive Bottom-Up Innovation. *ECIS 2023 Research-in-Progress Papers*. On-line: https://aisel.aisnet.org/ecis2023_rip/82
- [38] Peansupap, V., & Walker, D. H. T. (2006). Information communication technology (ICT) implementation constraints. *Engineering, Construction and Architectural Management*, 13(4), 364–379. <https://doi.org/10.1108/09699980610680171>
- [39] Adrian, B., Hinrichsen, S., & Nikolenko, A. (2020). App Development via Low-Code Programming as Part of Modern Industrial Engineering Education. In I. L. Nunes (Ed.), *Advances in Human Factors and Systems Interaction* (pp. 45–51). https://doi.org/10.1007/978-3-030-51369-6_7
- [40] Boehm, B., Abts, C., & Chulani, S. (2000). Software development cost estimation approaches — A survey. *Annals of Software Engineering*, 10(1), 177–205. <https://doi.org/10.1023/A:1018991717352>
- [41] Butler, R. R. (2020). *Estimating Effort for Low-code Applications* (PhD Thesis). *ProQuest Dissertations and Theses*. On-line: <https://www.proquest.com/dissertations-theses/>

GPT-based Logic Reasoning for Hazard Identification in Construction Site using CCTV Data

Dai Quoc Tran¹, Yuntae Jeon², Minsoo Park³, and Seunghee Park⁴

¹Global Engineering Institute for Ultimate Society, Sungkyunkwan University, South Korea

²Department of Global Smart City, Sungkyunkwan University, South Korea

³Sungkyun AI Research Institute, Sungkyunkwan University, South Korea

⁴School of Civil, Architectural Engineering and Landscape Architecture, Sungkyunkwan University, South Korea

daitran@skku.edu, jyt0131@g.skku.edu, pms5343@skku.edu, shparkpc@skku.edu

Abstract -

The applications of deep learning-based robust surveillance are vital for improving safety at construction sites, with closed-circuit television (CCTV) systems serving as a pivotal tool in achieving this goal. Despite the recent progress in state-of-the-art deep learning models, the task of hazard identification remains a persistent difficulty due to the complexity of the working environment. This paper presents a novel end-to-end pipeline termed “*Image-to-Hazard*” that aims to address the disparity between individual single-model predictions. The pipeline incorporates multimodal inputs and uses logical reasoning to establish connections. The pipeline integrates a model based on GPT architecture from the OpenAI API, encompassing various tasks such as detection, depth estimation, danger identification, and logical reasoning. Firstly, an actual video dataset was obtained from construction sites and annotated. Subsequently, customized object detection models were trained and optimized. Afterward, a thorough extraction of visual features was conducted by utilizing pre-trained models for tasks such as semantic segmentation and depth estimation. Subsequently, prompt engineering was conducted to seamlessly include the input of visual feature information, and these structures were integrated into OpenAI GPT-based models to enhance their capacity for logical reasoning. As a result, a proposed approach showed its robustness in integrating the GPT-based model and vision model for automated hazard identification and management at construction sites.

Keywords -

GPT; Logic Reasoning; Hazard Identification; CCTV; Safety Management

1 Introduction

Construction sites have constantly been identified as one of the most hazardous work environments worldwide [1]. According to the construction accident data from the Republic of Korea, there were an average of 538 deaths per year between 2010 and 2020 [2]. These fatalities

accounted for around 27.9% of all accidents across ten different industries, making it the industry with the highest accident rate [2]. In 2021, the Occupational Safety and Health Administration (OSHA) [3] recorded a fatality rate of 12.3 deaths per 100,000 full-time equivalent workers in the construction industry. It is essential to have a thorough grasp of the various dangers involved, including falls, electrocution, and exposure to hazardous materials, to improve worker safety. Closed-circuit television (CCTV) cameras are widely acknowledged as important tools for monitoring safety on construction sites. CCTV cameras facilitate the identification of possible dangers, surveillance of adherence to safety protocols, and examination of incidents by capturing visual evidence of actions. However, there is not much research on synchronizing the multi-vision model with NLP for logical reasoning. With the advancement of the generative pre-trained transformer (GPT) model currently, the GPT model acts as a strong tool for combining knowledge and providing insight into a given scenario [4]. Manual monitoring of large-scale CCTV channels in construction applications can result in the failure to identify potential hazards. The GPT model enables the identification of hazards not only through visual information but also from a linguistic perspective. Subsequently, the safety report can be generated automatically. This study aims to bridge the gap between separate single-model predictions using multimodal connective logical reasoning, whereby *Image-to-Hazard*, a novel end-to-end pipeline that uses a GPT-based model from the OpenAI API, is proposed for detection, depth estimation, hazard identification, and logic reasoning for safety monitoring at construction sites. The contributions of this study can be summarized as follows:

1. A novel end-to-end pipeline using a GPT-based model from the OpenAI API is proposed for detection, segmentation, depth estimation, hazard identification, and logical reasoning for safety monitoring at construction sites.
2. Detection models are trained on real-life CCTV

datasets.

2 Literature Review

Deep learning (DL) has been widely employed in diverse fields such as computer vision, natural language processing, and robotics. Furthermore, deep learning methods have demonstrated resilience in improving safety surveillance at construction sites [1]. Integrating deep learning algorithms into CCTV systems enhances the ability to identify and mitigate potential safety issues at building sites [5]. Deep learning algorithms provide a range of approaches to improve safety monitoring using CCTV systems. As stated by [6], 400,000 photos are taken throughout the construction stages. Furthermore, CCTV systems are built at nearly every construction site and are consistently employed to monitor the situation. Safety management, progress tracking, and quality inspections can benefit from the use of high-resolution photos, videos, and algorithms that rely on deep learning techniques. Object recognition and tracking have garnered considerable interest among various deep-learning approaches [7, 8]. Deep learning-based object detection can be categorized into two main types: one-stage and two-stage detectors [9]. Notable architectures in this field include Fast R-CNN (fast region-based convolutional neural network) [10] and YOLO (you-only-look-once) [11]. These two object detection methods are used based on unique research aims, taking into account the trade-off between accuracy and inference time. Real-time detection is essential in monitoring construction site safety using CCTV footage, requiring a careful balance between accuracy and inference speed. Some common applications of deep learning in construction site, such as: Nath *et al.* employed the YOLO architecture [11] to create three distinct models for identifying worker personal protective equipment (PPE). The constructed models were verified using a dataset that was created specifically for this purpose, called *Picto-v3*. The technique, designed using the YOLO architecture for one-stage detection, exhibited an inference speed of around 13 frames per second (FPS), which is close to real-time. Following that, several studies have endeavored to improve the efficacy of the YOLO framework in monitoring safety at building sites. Park and colleagues [12] enhanced the YOLO architecture by incorporating an attention mechanism, resulting in SOC-YOLO. This modification aimed to enhance the detection accuracy of small and overlapping worker images. SOC-YOLO is a technique that employs advanced methods such as distance intersection over union (DIoU), non-maximum suppression (NMS), weighted triplet attention, expansion feature levels, and a soft pool. It has proven to be effective in detecting small and overlapping targets (workers) at complex construction sites. When conducting studies with SOC-YOLO, the av-

erage precision (AP) for small items showed an increase from 67.52% to 73.88% mAP for minute objects, and from 74.56% to 77.57% for overlapping objects. This demonstrates the practicality of SOC-YOLO in safety monitoring. Currently, the research paradigm is shifting from single-model to multimodal development, highlighting the necessity of considering the context of a given image or video alongside object detection, as quantifying objects cannot rely solely on a single model. By incorporating various types of models, including detection, segmentation, and depth estimation, it is possible to develop a context-aware model that encompasses a broader understanding of the visual data. Chern *et al.* [13] proposed modularized context-aware safety monitoring for fall accidents. Using the CCTV far-field monitoring dataset YUD-COSAv2, the approach involves training detection models and subsequently combining them with segmentation and depth map models to create a context-aware model. This model takes the scenario into account rather than relying solely on detection results, enabling a more informed and comprehensive inference of results. This model was able to differentiate workers by height and could apply different PPE compliance rules, whereby average precisions of 78.50% and 86.22% were obtained, respectively. Overall, these studies focus on single-model development, and the improvement is derived from improved datasets, customized models, or ensemble models. However, a multimodal understanding of scenarios from input images or videos is still lacking. Chen *et al.* [14] developed a framework consisting of three modules for: (1) automated processing of regulatory rules and transformation of sentences into computable graph structure representations; (2) combining two object detection and pose estimation models to represent scene information; and (3) automated reasoning of hazard notification from the above two output graphs. The proposed framework, which extracts safety rules through feature engineering, was effective in identifying individuals operating a grinder. Khan *et al.* [15] identified mobile scaffolding and workers using Mask R-CNN. This study proposed a correlation-based approach for mobile scaffold safety monitoring and the detection of unsafe worker behaviors. Mask R-CNN was used to classify and segment worker tasks and an object correlation detection (OCD) module was used to detect unsafe behaviors. Subsequently, safety rules were used to determine whether the scenario was safe based on the detection results. The test results exhibited 85-97% precision and recall for class-1 (safe behavior) and 91-65% precision and recall for class-2 (unsafe behavior). An overall accuracy of 86-96% confirms the Mask R-CNN-based OCD module's applicability in the construction environment.

3 Proposed Approach

This section describes the deep learning and self-developed models used in the proposed approach. The proposed approach is illustrated in Figure 1 and comprises four major phases: data acquisition, model inference, synchronization, and logic reasoning. During the data acquisition phase, an actual CCTV dataset is acquired and labeled to train the detection model. The input image is then used for inference in multiple types of deep learning models, including object detection, and depth-map estimation. The information from the detected objects is then synchronized using multimodal synchronization modules. Finally, the GPT-based model is used for hazard-scenario logic reasoning using the developed auto-generating prompt structure. Each phase of the proposed approach is essential to contributing to the overall objective. The data-acquisition phase ensures that the model has access to a representative, high-quality dataset. The inference phase enables the model to recognize and categorize the objects within the input image. The synchronization phase ensures that the information regarding the detected objects is consistent across models. The reasoning phase enables the model to generate a logic-based hazard scenario that can be utilized to prevent catastrophic events. The proposed approach is a promising novel method for hazard detection and prevention with the potential to improve worker and property safety at construction sites.

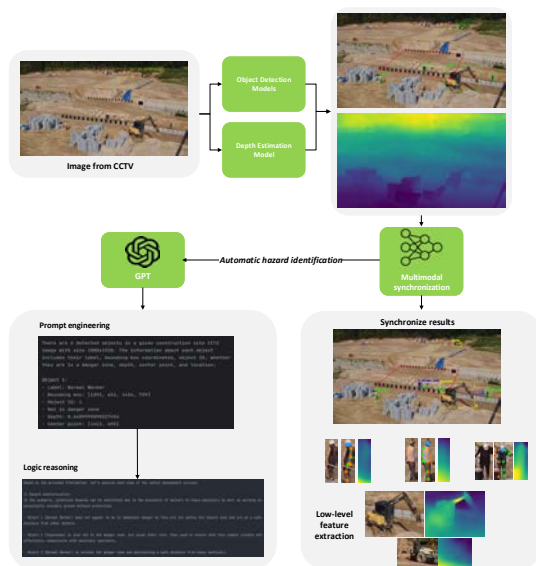


Figure 1. Image-to-Hazard.

3.1 Object Detection Modules

A construction site is a complex environment containing numerous objects that must be detected and analyzed.

The focus of this study is to classify and identify primary target objects on a construction site, including construction vehicles, workers, and signalman. The specific class names are listed below.

1. Construction vehicle with eight classes (*Excavator, Dozer, Forklift, Dump Truck, Mixer Truck, Cargo Truck, Scissor Lift, Crane*).
2. Signalman detection with two classes (*Worker, Signalman*).

The exploratory data analysis is presented in Figure 2. Numerous detection models can be used to detect these objects. In this study, the state-of-the-art one-stage object detector YOLOv8X was used to train and finetune the detection model. To enhance the diversity of the training samples, the default data for training all the models in this study were augmented using YOLOv8. The augmentation technique used was "mosaic," which combines four different training images into one in a mosaic-like pattern. This approach helps to improve the variation and representation of the training data. Subsequently, a unique augmentation technique specific to YOLOv5, known as 'CopyPaste', was employed to randomly select an object from one image and paste it onto another, thereby enhancing the complexity of the image data. The model then applied random affine transformations, including scaling, shearing, and rotation, to the images. The ranges of these transformations were predefined. Following these steps, another YOLOv5 specific technique called 'MixUp' was used to blend two images and their labels to generate a new, more complex image. Following 'MixUp,' the mask information was removed from the data, and then a series of augmentations from the Albumentations library were performed. These transformations include techniques such as Blur, MedianBlur, conversion to grayscale, and contrast-limited adaptive histogram equalization, each applied with a certain probability. The experiment was conducted using CentOS Linux 8 with two GTX A6000 graphics processor units, each with 48 GB of memory. The model was developed using PyTorch, based on the MMYolo [16] library. The mean Intersection over Union (mIoU) was used to evaluate the model. mIoU is a frequently used metric in computer vision for evaluating the efficacy of object detection and segmentation algorithms. This metric quantifies the amount of overlap between the predicted bounding boxes (or segmentation masks) and ground truth bounding boxes (or masks), indicating the accuracy of model prediction.

Table 1 displays model detection performance for construction vehicles and signalman datasets. In this study, the mIoUs of small, medium, and large objects are considered in the analysis. Comparative studies with other SOTA models were not considered because the objective

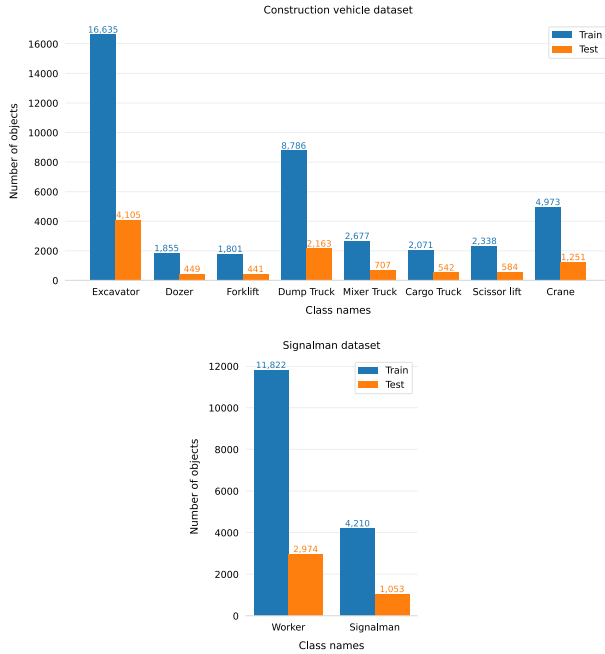


Figure 2. Dataset class distribution.

of this research was to develop end-to-end multimodal logic reasoning. Objects that occupy 0 to 1024 pixels (32×32 pixels) are considered 'small'; objects that occupy between 1024 and 9216 pixels (32×32 to 96×96 pixels) are considered 'medium'; and objects that occupy more than 9216 pixels (96×96 pixels or larger) are considered 'large.'

The construction vehicle detection model revealed a significant size-dependent performance, with smaller objects proving to be more challenging for the model to detect. The model exhibited a precision of 0.377 for small objects, which increased substantially to 0.763 for medium objects and peaked at 0.921 for large objects. This steep increase suggests that the model is particularly adept at detecting larger construction vehicles, possibly because of their distinct high-contrast features that are easier to discern at larger scales. The signalman detection model demonstrated comparatively consistent performance across object sizes, with precision measures ranging from 0.652 for small objects to 0.765 for medium objects, and then slightly increasing to 0.838 for large objects. Unlike the other models, this model did not show a steep size-dependent performance gradient, which may be attributed to the distinct characteristics of the signalmen, making them easier to detect irrespective of their size. The discrepancies in model performance across different sizes and categories indicate strengths and limitations. These models are currently more effective in detecting larger objects. Smaller objects prove to be a common challenge in

visualization, possibly because of their indistinguishable features at smaller scales. This highlights the need for further research to enhance the precision of object detection models, particularly for small- and medium-sized objects.

Table 1. Detection model performance.

Model	IoU=0.5:0.95↑		
	Small	Medium	Large
Construction vehicle	0.377	0.763	0.921
Signalman	0.652	0.765	0.838

3.2 Human Pose Estimation

The overlapping area of the upper body precludes the exact location of the worker from being extracted if only the bounding box from the object detection model is utilized. In numerous instances, the bounding box of the detected worker does not encompass the entire body; consequently, the ankle coordinates cannot be estimated. For estimating the ankle midpoint of the detected worker, the well-known pre-trained model HRNet [17] utilizes keypoint extraction. Figure 3 presents examples of the algorithm. The detected person bounding boxes are extracted first, followed by the application of a pose estimation model, to estimate all the key points of the body. In some instances, ankle points cannot be detected because of overlapping objects, and the ankle midpoint is estimated using an estimation ratio between upper body keypoint.

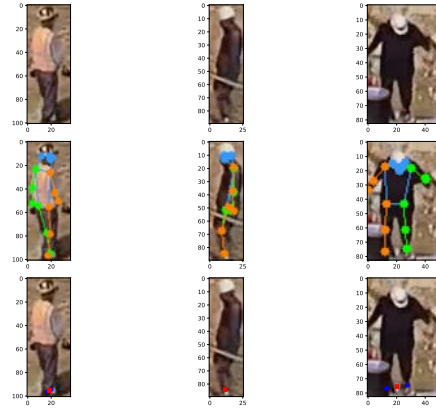


Figure 3. Actual dataset.

3.3 Depth Estimation Module

Objects detected by a detection model can be extracted from image coordinates without considering spatial information. A monocular CCTV depth-estimation map can contribute to spatial information analysis. Therefore, the trained MiDaS [18] model was used to extract depth data

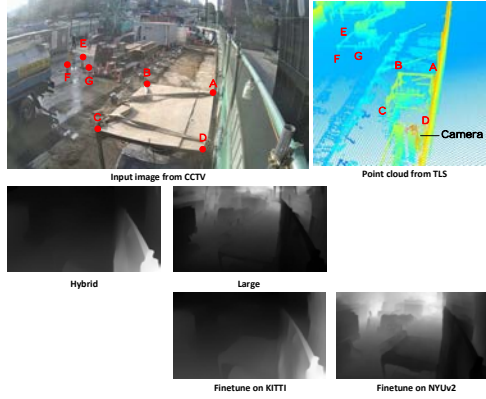


Figure 4. Comparison of depth map estimation results using point cloud.

from the CCTV video. Using the x and y coordinates in the image and depth z , the Euclidean distance in three dimensions can be estimated. As shown in Figure 4, the depth estimation results from monocular CCTV were then compared with the actual point cloud data from terrestrial laser scanning (TLS). The four trained models were used to compare the results. The MiDaS model finetuned on NYUv2 dataset robustly estimated depth using the monocular CCTV image. In this study, the trained MiDaS, finetuned on the NYUv2 dataset, was selected to estimate the depth map at the construction site.

3.4 Prompt Engineering

For logical reasoning, GPT is used. Currently, GPT-based models perform phenomenally in the natural language processing (NLP) domain [4]. GPT-based models are a type of large-language model (LLM) trained using a technique called generative pre-training. GPT-based models are typically trained on a massive dataset of text and can be used for a variety of NLP tasks, such as text generation, translation, and question-answering. With this large-scale dataset, the model can understand extremely complex text structures and provide meaningful answers. Therefore, in this research, GPT-based models were utilized for the analysis of hazard scenarios, using the extracted multimodal visual information.

An important technique is “*Prompt engineering*”, which is the process of designing and optimizing inputs (prompts) to a model, such as GPT, in a way that maximizes the quality and relevance of its outputs. This is a crucial part of using AI language models because the way a question or task is framed can significantly affect the model’s response. Object detection models identify and classify objects within images typically by providing a bounding box around the detected object and a label indicating the

object. This output can be used as an input prompt for a language model such as GPT for further analysis, reasoning, or narrative generation.

An ontology model can be used to define entities at a construction site (such as workers, vehicles, and equipment), their properties (location, PPE status, and moving direction), and the relationships between these entities. Once the ontology model is developed, it can be used to generate prompts for the GPT model.

Given:

- D : the set of detected objects.
- d_i : a detected object in D , which is a tuple $(label, bbox, object_id, in_danger_zone, depth, center_point)$.
- P : the set of object distances.
- p_i : a pair of objects in P , which is a tuple $(object1, object2, distance)$.

We define:

- A function $f(d_i)$ to generate the description for a detected object d_i .
- A function $g(p_i)$ to generate the description for a pair of objects p_i .

Then, the function $h(D, P)$ used to generate the sentence for the GPT model is defined as

$$h(D, P) = \bigcup_{d_i \in D} f(d_i) \cup \bigcup_{p_i \in P} g(p_i)$$

Algorithm 1 illustrates the conversion of visually detected features into a GPT-based model.

Using the aforementioned ontology models, the generated prompt is then input into the GPT-based model for logical reasoning.

4 Experiment and Discussion

This Section presents sample hazard scenarios and analyzes the input and output of logical reasoning from the GPT-based model. Currently, OpenAI supports API for GPT-based models up to GPT4.

As shown in Figure 5 and Figure 6, after extracting object information using detection models, such as draw a danger area, provide a reason for danger area, estimate the Euclidean distance in 3D between “Normal Worker” and “Vehicle,” extract depth and segmentation map. The GPT model infers the hazard scenarios in this input image based on three main questions, as follows:

1. Identify potential hazards between objects.
2. Detail the specific risks associated with these potential hazards.

Algorithm 1 Generating descriptive sentences for object detection

```

1: procedure FORMAT_OBJECT_INFO(obj)
2:   if not obj['in_danger_zone'] then
3:     danger_zone_status ← 'Not in'
4:   else
5:     danger_zone_status ← 'In'
6:   end if
7:   info ← format_info(obj, danger_zone_status)
8:   if obj['is_signalman'] then
9:     append_signalman_info(info, obj)
10:  end if
11:  return join_info(info)
12: end procedure
13: procedure FORMAT_DISTANCES(distances)
14:  descriptions ← format_pairwise_distances(distances)
15:  append_safe_distance_info(descriptions)
16:  return join_descriptions(descriptions)
17: end procedure
18: procedure GENERATE_SENTENCE(detected_object_info_danger,
19:   object_distances)
20:  num_objects ← count_objects(detected_object_info_danger)
21:  if object_distances is not empty then
22:    distances_text ← format_distances(object_distances)
23:    base_text ← format_base_text(num_objects, ob-
24:   jects_text, distances_text)
25:  else
26:    base_text ← format_base_text_no_distances(num_objects, objects_text)
27:  end if
28:  additional_text ← format_additional_text()
29:  return base_text + additional_text

```

3. Highlight the safety management approach for reducing risks.

Considering the visual information extraction input, the GPT4 model provides feedback regarding hazard scenarios in a given image, as follows:

- Potential hazards: Object 1 (normal worker) and Object 5 (excavator): The normalized distance between Object 1 and Object 5 is 0.159, which is less than the safe distance threshold of 0.2. This indicates that the worker is too close to the excavator. Object 4 (normal worker) and Object 5 (excavator): The normalized distance between Object 4 and Object 5 is 0.275, which passes the safe distance threshold, but is close enough to warrant attention.
- Analyze the "Risks associated with these potential hazards": The primary risk in both of these scenarios is injury due to the close proximity of workers to heavy machinery. These injuries can occur in a

variety of ways, such as: (1) Being struck by moving parts of the machinery (e.g., the bucket of the excavator). (2) Being caught between machinery and another object. (3) Slips, trips, or falls due to uneven or unstable ground near the machinery.

- Safety management approach for reducing risks.

Clearly, with multimodal inference, different aspects such as worker type, vehicle name, location, and 3D distance between objects, are used to extract visual information. Thus, the GPT model can perform logical reasoning much more effectively. The construction site is a complex environment. Without a tailored training model and considering all possible hazards, the pretrained single-detection model can meet the challenge of understanding context-aware scenarios.

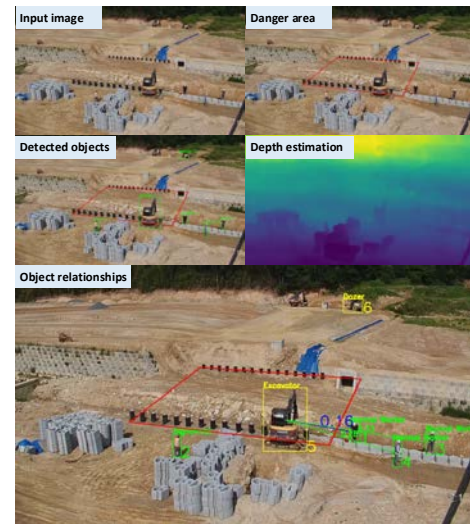


Figure 5. Step-by-step model detection.

In Figure 7 and 8, the report identifies potential hazards and risks in a construction site scenario, indicating close proximity of workers to heavy machinery and an unstable slope. Notably, Object 2 (normal worker) is dangerously close to Object 7 (excavator), with a normalized distance of 0.196, which is less than the safe limit of 0.2. Additionally, Object 3 (normal worker) is within the danger zone, being close to an unstable slope. These situations raise significant risks, including physical injury or fatality for Object 2, from possible contact with the excavator and potential slope collapse hazards for Object 3. Several safety strategies have been proposed to mitigate these risks. These strategies include enforcing safe distances between workers and machinery, particularly for Object 2; evacuating and restricting access to the danger zone around the unstable slope where Object 3 is located; continuous site monitoring with the CCTV system, possibly



Figure 6. GPT logic reasoning.

supplemented with automated alerts for danger zone intrusions; slope stabilization procedures; mandatory use of PPE for all workers; and regular safety training sessions. This integrated approach aims to enhance worker safety and reduce accident risk.

5 Conclusion

This study addresses the challenges of identifying hazards in complex and unpredictable construction site environments, with the support of advanced deep-learning models. To overcome these obstacles, we propose a novel end-to-end pipeline, *Image-to-Hazard*, designed to bridge the gap between separate single-model predictions using multimodal and logical reasoning. This pipeline integrates detection, depth estimation, hazard identification, and logical reasoning, by employing a GPT-based model from OpenAI API for safety monitoring at construction sites. A large-scale video dataset was obtained from actual construction sites and labeled. Subsequently, a large-scale dataset was utilized to train and optimize tailored object detection models. Pre-trained models for semantic segmentation and depth estimation were utilized to generate a comprehensive visual feature extraction dataset. Visual feature information was then integrated with prompt structures and input into OpenAI GPT-based models for logical reasoning. In conclusion, this study demonstrated that the *Image-to-Hazard-Scenarios* pipeline, which combines multimodal data for context-aware hazard identification, was successful in enhancing safety monitoring at con-

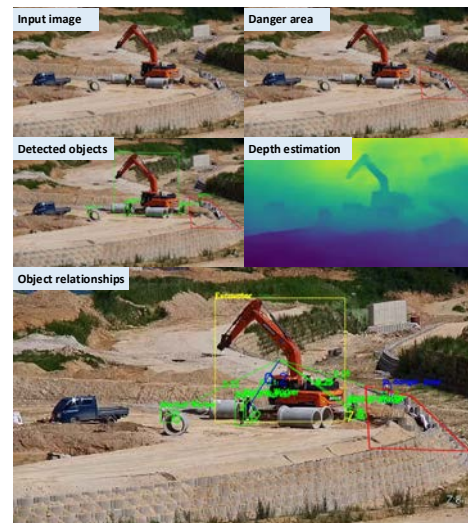


Figure 7. Step-by-step model detection.

struction sites. The method was validated using actual CCTV footage, quantifying its performance and leading to the development of a deployment procedure for its implementation in practical construction site settings. This finding underscores the significant potential of employing integrated multimodal approaches in improving the safety of complex working environments, such as construction sites. Future research should focus on refining this pipeline and exploring its applicability to other similarly complex settings using real-time inference.

Acknowledgements

This research was supported by a grant [2022-MOIS38-002 (RS-2022-ND630021)] from the Ministry of Interior and Safety (MOIS)'s project for proactive technology development safety accidents for vulnerable groups and facilities, and this research was supported by a grant from the Korean Government (MSIT) to the NRF [RS-2023-00250166]. This work is financially supported by Korea Ministry of Land, Infrastructure and Transport(MOLIT) as Innovative Talent Education Program for Smart City.

References

- [1] Fan Zhang, Hasan Fleyeh, Xinru Wang, and Minghui Lu. Construction site accident analysis using text mining and natural language processing techniques. *Automation in Construction*, 99:238–248, 2019.
- [2] Large Accident Case Book, 2022. [Online; accessed 2022-07-18].



Figure 8. GPT logic reasoning.

- [3] U.S. Bureau of Labor Statistics. <https://www.bls.gov>, 2023.
- [4] Hanmeng Liu, Ruoxi Ning, Zhiyang Teng, Jian Liu, Qiji Zhou, and Yue Zhang. Evaluating the logical reasoning ability of chatgpt and gpt-4. *arXiv preprint arXiv:2304.03439*, 2023.
- [5] JoonOh Seo, SangUk Han, SangHyun Lee, and Hyoungkwan Kim. Computer vision techniques for construction safety and health monitoring. *Advanced Engineering Informatics*, 29(2):239–251, 2015.
- [6] Kevin K Han and Mani Golparvar-Fard. Potential of big visual data and building information modeling for construction performance analytics: An exploratory study. *Automation in Construction*, 73: 184–198, 2017.
- [7] Dai Quoc Tran, Minsoo Park, Daekyo Jung, and Seunghee Park. Damage-map estimation using uav images and deep learning algorithms for disaster management system. *Remote Sensing*, 12(24):4169, 2020.
- [8] Dai Quoc Tran, Minsoo Park, Yuntae Jeon, Jinyeong Bak, and Seunghee Park. Forest-fire response system using deep-learning-based approaches with cctv images and weather data. *IEEE Access*, 10:66061–66071, 2022. doi:10.1109/ACCESS.2022.3184707.
- [9] Zhengxia Zou, Keyan Chen, Zhenwei Shi, Yuhong Guo, and Jieping Ye. Object detection in 20 years: A survey. *Proceedings of the IEEE*, 2023.
- [10] Ross Girshick. Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 1440–1448, 2015.
- [11] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.
- [12] Minsoo Park, Dai Quoc Tran, Jinyeong Bak, and Seunghee Park. Small and overlapping worker detection at construction sites. *Automation in Construction*, 151:104856, 2023. ISSN 0926-5805. doi:<https://doi.org/10.1016/j.autcon.2023.104856>. URL <https://www.sciencedirect.com/science/article/pii/S0926580523001164>.
- [13] Wei-Chih Chern, Jeongho Hyeon, Tam V Nguyen, Vijayan K Asari, and Hongjo Kim. Context-aware safety assessment system for far-field monitoring. *Automation in Construction*, 149:104779, 2023.
- [14] Shi Chen, Kazuyuki Demachi, and Feiyan Dong. Graph-based linguistic and visual information integration for on-site occupational hazards identification. *Automation in Construction*, 137:104191, 2022.
- [15] Numan Khan, Muhammad Rakeh Saleem, Doyeop Lee, Man-Woo Park, and Chansik Park. Utilizing safety rule correlation for mobile scaffolds monitoring leveraging deep convolution neural networks. *Computers in Industry*, 129:103448, 2021.
- [16] MMYOLO Contributors. MMYOLO: OpenMMLab YOLO series toolbox and benchmark. <https://github.com/open-mmlab/mmyolo>, 2022.
- [17] Jingdong Wang, Ke Sun, Tianheng Cheng, Borui Jiang, Chaorui Deng, Yang Zhao, Dong Liu, Yadong Mu, Mingkui Tan, Xinggang Wang, et al. Deep high-resolution representation learning for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, 43(10):3349–3364, 2020.
- [18] René Ranftl, Katrin Lasinger, David Hafner, Konrad Schindler, and Vladlen Koltun. Towards robust monocular depth estimation: Mixing datasets for zero-shot cross-dataset transfer. *IEEE transactions on pattern analysis and machine intelligence*, 44(3): 1623–1637, 2020.

Adopting Automation in Premanufacturing: A Two-Mode Network Analysis on Factors and Roles in Iran and North America's Construction Industry

Amirhossein Babaei Ravandi¹ and Tamer E. El-Diraby²

¹ & ² Centre for Information Systems in Infrastructure and Construction, Dept. of Civil & Mineral Engineering, University of Toronto, Ontario M5S 1A4, Canada

¹amir.babaei@mail.utoronto.ca, ²tamer.diraby@utoronto.ca

Abstract

This study embarks on an in-depth exploration of the integration of automation in premanufacturing within the construction industry, with a specific focus on contrasting the perspectives and adoption strategies in Iran and North America. Utilizing a two-mode network analysis, the research explores the complex interplay between various roles within the construction sector and a range of 27 key factors influencing the shift toward automated processes. The findings reveal distinct approaches between the two regions: Iran, at a pivotal stage of technological evolution, emphasizes foundational economic and operational efficiencies, with key factors like 'Profitability' scoring a Degree Centrality of 0.37. In contrast, North America adopts a mature, holistic approach that balances economic, operational, and environmental considerations, prioritizing 'Efficiency and Productivity' with a Degree Centrality of 0.33. This comparative analysis offers crucial insights into the differing stages of technological adoption in construction, shaped by regional market conditions and developmental phases. Furthermore, the study provides valuable guidance for policymakers and decision-makers in crafting informed strategies that accommodate specific regional needs and advancements in construction technology.

Keywords

Automation in construction; Premanufacturing, Network Analysis; Two-mode Networks; Modular Construction; Automate Premanufacturing

1 Introduction

The construction industry, an integral component of the global economy, plays a significant role in shaping Gross Domestic Product (GDP) and creating

employment opportunities across the globe. It also stands as a dominant force in energy and natural resources consumption and represents a considerable share of the world's total energy use [1]. This dual impact underscores the critical need for sustainable and efficient practices within the construction domain to balance its economic contributions with environmental stewardship.

Despite its potential, the industry struggles with resource optimization, waste reduction, and the adoption of new technologies, challenges that are critical to address in today's climate-conscious world [2]. Prefabrication emerges as a key technology to address these challenges by offering benefits such as enhanced quality, speed, and sustainability in construction, alongside cost and waste reduction. However, its integration into mainstream construction practices has been met with reluctance due to inherent industry characteristics, such as the uniqueness of each project and the variability in the design and stakeholder composition [3], [4]

In this study, we explore the potential of automating the prefabrication process in the construction sector, with a particular focus on the factors influencing the choice between manual and automated premanufacturing. Manual premanufacturing relies on human labor to construct building components or entire sections off-site in a controlled factory setting before transporting them to the construction site for assembly. It is valued for its flexibility, allowing for customization and adjustments based on specific project needs or when dealing with intricate designs that automated processes cannot easily accommodate [5]. On the other hand, automated premanufacturing leverages robotics and advanced technologies to increase efficiency, precision, and safety while addressing challenges such as labor shortages and rising construction costs [6]. We investigate the technical, business, economic, and environmental drivers that play an important role in automation decisions, aspects vital for both private-sector investment and

public policymaking. By conducting expert surveys in two markedly different markets, North America and Iran, and utilizing two-mode network analysis, our research not only aims to explore the dynamics among various factors but also goes beyond traditional methods to offer an in-depth understanding of their interconnectedness and relative importance within the broader system. Our approach also places a significant emphasis on the roles within the construction sector and their impact on the adoption of automation in premanufacturing. By examining the perspectives and influences of various professional roles we gain invaluable insights into the sector's diverse viewpoints on automation.

2 Background

The integration of automation in premanufacturing within the construction industry signifies a transformative shift propelled by the demands for increased efficiency, sustainability, and innovation. This review aims to synthesize the recent relevant studies in this field.

Central to this transformation is the implementation of automation throughout different construction stages, from the initial design phase to the robotic assembly. A notable example is automated modeling achieved through the integration of Building Information Modeling (BIM) with prefabrication for sustainable construction, and employing parametric design optimization to enhance design layouts for greater efficiency and lower environmental impact [7]. Additionally, the development of digital twin frameworks, as discussed in [8], [9], is a promising step toward the automation of various stages of prefabrication to enhance the adaptability of managing diverse premanufacturing projects and elements and facilitate their construction through automated processes.

The burgeoning role of Artificial Intelligence (AI) in refining the capabilities of automated robots is increasingly evident. [10] introduced a risk management system using AI algorithms to enhance the accuracy and speed of decision-making and improve the reliability and cost-effectiveness of prefabricated building projects. [11] further advanced the field by employing Deep Reinforcement Learning for automated assembly planning in robot-based construction to optimize the assembly processes through a Markov Decision Process model and advance the efficiency and safety of implementing automation in premanufacturing.

Automation in robotics assembly is another notable field of innovation and research in premanufacturing due to its potential for advancing efficiency and creative solutions. In this context, [12] demonstrated the potential of integrative robotic prefabrication and co-design methods. This research highlights the effective

combination of computational design with fabrication planning and marks a significant advancement in digital automation workflows. Similarly, [13] contributed to this field by developing robotic setups that combine subtractive and additive processes to enable large-scale spatial fabrications, thereby maintaining digital accuracy and minimizing waste. Also, [14] explored efficiency gains in automated production by using simulation to optimize layout, enhance workspace utilization, and boost production. This study demonstrates automation's practical advantages in prefabrication.

The existing literature extensively explores the implementation of prefabrication in construction, including barriers, challenges, and advantages, but few studies investigate its automation. Despite technological progress, fully automating prefabrication presents complex challenges. For instance, [15] examined the complexities in timber-frame prefabrication automation to pinpoint barriers related to technology adoption, machinery sophistication, and business scale. Also, [16] outlined seven dimensions that challenge automation in modular construction, with low standardization and individual customer requests emerging among the most important barriers. Similarly, [17] highlighted key strategies to encourage automation in the Nigerian construction industry, including funding and subsidies, mandatory policies, and incentives for adoption.

These studies emphasize the complexities and potential of automating prefabrication, highlighting the need for a comprehensive approach that encompasses technical, socio-economic, regulatory, and strategic dimensions. However, their reliance on traditional data analysis may overlook the intricate interplay of influencing factors, as [16] suggests a deeper exploration into the interrelationships and dynamics of the influential factors. This study contributes by applying network analysis for a more objective, in-depth examination of automation in prefabrication, utilizing mathematical algorithms to quantify the significance of interrelated factors [18], [19]

3 Methodology

In this study, we employed a comprehensive two-mode network analysis methodology to investigate the key factors and roles influencing automation in premanufacturing. An extensive literature review meticulously analyzed over 50 papers to discern the distinctions between manual and automated premanufacturing processes in the construction industry, initially identifying over 40 potential influential factors. This preliminary set was further refined through semi-structured interviews with eight industry and academic experts from North America and Iran. These interviews aimed to evaluate the factors' relevance, clarify

definitions, identify overlaps, and uncover any overlooked elements. The process ensured the factors were relevant across different construction industry contexts, particularly considering the distinctive practices between North America and Iran. This thorough examination and expert consultation ultimately streamlined the list to a focused set of 27 factors.

The complete list of final factors is as follows: (a) Factors impacting productivity and speed of the manufacturing process include complex and highly-integrated designs (F1), design flexibility and customization (F2), coordinated parallel production (F3), safety (F4), material availability (F5), faster delivery (F6), production time (F7), and smaller and modularized factories (F8); (b) Factors impacting productivity and speed of the manufacturing process encompass efficiency and productivity (F9), quality improvement (F10), the resilience of the manufacturing program (F11), long-term competitive power (F12), foreign markets (F13), profitability (F14), and unstable conditions (F15); (c) Social issues linked to or impacted by premanufacturing involve diversification of workforce (F16) and skilled workers (F17); (d) Factors related to the schedule of projects cover set-up time (F18) and long implementation lead-time (F19); (e) Overall project costs include high initial costs (F20), labor cost (F21), maintenance costs (F22), and space requirements (F23); and (f) Environmental impacts of projects comprise GHG emissions and energy saving (F24), more sustainable design (F25), noise pollution (F26), and waste reduction (F27) [20], [21], [22], [23], [24], [25], [26], [27], [6], [28], [29].

Subsequently, A survey, featuring both open and closed questions and available in English and Farsi, was distributed to firms in the construction sectors of Iran and North America. It aimed to assess perspectives on the importance and interrelations among various factors and included demographic details and familiarity with off-site construction. Considering the complex and wide-ranging nature of the factors under examination, we employed purposive sampling to engage a broad spectrum of industry stakeholders—from CEOs and university professors to general workers, to explore the relationships between each pair of factors. This scientific method allowed us to deliberately select individuals who could provide valuable insights into the relationships between each pair of factors. The survey, validated by a high Cronbach's Alpha (0.87) for reliability, also provided definitions and examples of factors to ensure clarity. To avoid neutral responses, a six-point Likert scale was utilized. Also, the questionnaire provided researchers' contact information to encourage participants to offer accurate feedback. Networks for each region were then modeled from the survey data, reflecting the diverse responses [30].

3.1 Two-mode Social Network Analysis

This study leverages a two-mode network analysis, ideal for investigating complex systems involving distinct entity categories, to examine the intricate relationships between 27 identified key factors and various stakeholders from Iran and North America. This approach offers critical insights into the systemic patterns and structures within these interactions [31].

To build the networks for our study, we first constructed adjacency matrices for role-factor relationships in each region. An adjacency matrix, a square matrix representing a finite graph, indicates the adjacency between pairs of nodes [32]—in our case, the roles and factors. In these matrices, a '1' is assigned to a cell when respondents of specific roles acknowledge a factor's importance, signifying a direct relationship, while a '0' indicates the absence of such recognition. These matrices not only provide detailed insights into the varying perceptions of different roles on key factors across regions but also lay the groundwork for constructing two-mode networks. In this bipartite network structure, nodes from one category (roles/stakeholders) are connected exclusively to nodes from the other category (factors), based on the stakeholders' affirmation of the factors' importance.

Upon constructing these networks, our analysis will leverage several centrality measures, namely degree, betweenness, and eigenvector centrality, to elucidate the influence and position of each node within the network, as explained in the following.

Degree Centrality (DC) reflects the number of ties a node has to other nodes, which is an indication of its activity or popularity within the network [33]. For roles, it represents how many factors they are connected to, while for factors, it shows the number of stakeholders recognizing their importance. This is crucial for identifying key factors that are widely acknowledged across different roles.

Betweenness Centrality (BC) indicates a node's capacity to act as a bridge within the network. It is calculated based on the number of shortest paths that pass through a node [33]. A high betweenness centrality for a role or factor suggests it plays a critical role in connecting various parts of the network, which can indicate stakeholders who are influential in bridging different factors or vice versa.

Eigenvector Centrality (EC) refines the basic premise of degree centrality by accounting for both the quantity and quality of a node's connections. It is based on the principle that connections to highly connected nodes contribute more to the score of the node in question [34]. Therefore, in this study, a factor with high eigenvector centrality is one that is recognized by stakeholders who themselves are widely connected or influential within the network.

4 Results and Discussion

Tables 1, 2, and 3 present the complete adjacency matrix and the results of centrality measures for selected factors (F) and roles (R) in Iran, respectively. Similarly, Tables 4 and 5 showcase the outcomes of network analysis for North America, detailing centrality measures for a subset of factors and roles. Also, Figures 1 and 2 display the two-mode networks of roles and factors for Iran and North America, respectively.

4.1 The Case of Iran

The integration of findings from the role-factor adjacency matrix and network analysis of Iran's construction industry reveals insightful trends in the perceived importance of different factors in the automation of premanufacturing processes. The adjacency matrix highlights a strong focus on key factors such as F14 (Profitability), F9 (Efficiency and Productivity), and F8 (Smaller and Modularized Factories). This pattern underscores a significant emphasis on financial viability, operational efficiency, and the adaptability of manufacturing processes within the sector.

The emphasis on profitability and efficiency in Iran's construction industry reflects a focused effort to optimize output while maintaining economic sustainability within the regional context. The focus on smaller and modularized factories, in particular, signifies a strategic shift toward more flexible, scalable production methods. This approach reflects a keen understanding of the economic and operational dynamics specific to Iran's evolving construction landscape, where cautious yet strategic steps are being taken toward modernization.

From the perspective of network measures, the equally high betweenness centrality of the mentioned factors suggests they act as key bridges in the network, connecting various aspects of the industry and facilitating the flow of influence and information. This indicates that investigation/shifts/advancements in these areas could have ripple effects, influencing a range of other connected factors and decisions within the sector.

The betweenness centrality of F15 (Unstable Conditions) and F20 (High Initial Costs) implies these factors, while not as central as the top three, are nonetheless critical in linking different areas of the industry, potentially acting as catalysts for change or areas of concern that require careful management. Their role in the network could be indicative of underlying challenges or barriers that the industry must navigate in its transition towards more automated practices.

In contrast, the lower betweenness centrality of F24 (GHG Emissions and Energy Saving) and F16 (Diversification of Workforce), coupled with their lower eigenvector centrality, suggests that these areas, while

Table 1. The Adjacency Matrix of the Role-factor Network

	F	1	10	11	12	13	14	15	16	17	18	19	2	20	21	22	23	24	25	26	27	3	4	5	6	7	8	9	Sum
R	8	0	1	1	1	1	1	1	0	1	0	0	1	1	1	1	0	0	0	0	1	0	1	0	0	1	1	1	16
	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0	0	0	1	0	1	1	0	1	1	1	19
	5	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	21
	9	1	1	0	1	1	1	1	0	1	1	0	0	1	1	1	0	1	1	0	1	0	1	0	1	1	1	1	19
	2	1	1	1	1	1	1	1	0	1	1	0	0	0	1	1	1	0	0	0	1	0	0	1	1	1	1	1	17
	3	1	1	0	1	0	1	1	1	1	1	0	1	1	1	1	0	0	1	0	1	1	0	1	0	1	1	1	19
	10	0	0	0	0	1	1	1	0	1	1	1	0	1	1	1	0	1	0	0	1	0	0	1	1	1	1	1	16
	7	0	1	0	0	0	1	1	0	1	1	1	0	1	1	1	0	1	0	0	0	0	0	1	1	1	1	1	16
	11	0	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	21
	12	0	1	0	1	1	1	1	0	1	0	0	1	1	0	1	0	0	1	0	0	0	0	0	1	1	1	1	13
	13	0	1	1	0	1	1	1	0	1	1	0	0	0	1	1	1	0	0	0	1	0	1	1	1	1	1	1	17
	14	0	1	0	0	1	1	0	0	1	1	0	1	1	0	1	0	0	0	1	0	1	1	0	0	1	1	1	13
	15	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	0	1	0	1	0	0	0	0	1	1	16
	6	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1	0	0	1	0	0	0	1	1	1	1	1	1	19
	4	0	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	0	1	0	0	0	0	0	1	1	1	1	18
Sum		3	14	9	10	12	15	14	4	14	12	3	9	12	13	15	3	4	7	2	10	3	8	11	9	14	15	15	

Table 2. Centrality Measures of Selected Factors in Iran (Ranked Based on DC)

Nodes (Factors)	DC	BC	EC
F14 (Profitability)	0.3659	0.0251	0.1899
F8 (Smaller and Modularized Factories)	0.3659	0.0251	0.1899
F9 (Efficiency and Productivity)	0.3659	0.0251	0.1899
F15 (Unstable Conditions)	0.3415	0.0200	0.1808
F20 (High Initial Costs)	0.2927	0.0157	0.1511
F24 (GHG Emissions and Energy Saving)	0.0976	0.0014	0.0512
F16 (Diversification of Workforce)	0.0976	0.0013	0.0545

Table 3. Centrality Measures of Selected Roles in Iran (Ranked Based on DC)

Nodes (Roles)	DC	BC	EC
R5 (Construction Manager)	0.5122	0.0684	0.2128
R11 (University Prof)	0.5122	0.0843	0.2040
R3 (Design Engineer)	0.4634	0.0616	0.1882
R6 (Superintendent)	0.4634	0.0329	0.2045
R12 (Real Estate Developer)	0.3171	0.0141	0.1435

Table 4. Centrality Measures of Selected Factors in North America (Ranked Based on DC)

Nodes (Factors)	DC	BC	EC
F9 (Efficiency and Productivity)	0.3333	0.0205	0.1891
F14 (Profitability)	0.3333	0.0205	0.1891
F17 (Skilled Workers)	0.3077	0.0171	0.1761
F8 (Smaller and Modularized Factories)	0.3077	0.0174	0.1741
F5 (Material Availability)	0.3077	0.0170	0.1754
F10 (Quality Improvement)	0.2821	0.0136	0.1641
F24 (GHG Emissions and Energy Saving)	0.2051	0.0070	0.1174
F25 (More Sustainable Design)	0.2051	0.0068	0.1198

Table 5. Centrality Measures of Selected Roles in North America (Ranked Based on DC)

Nodes (Roles)	DC	BC	EC
R1 (Project Managers)	0.5641	0.0701	0.2344
R5 (Construction managers)	0.5641	0.0806	0.2283
R17 (Product Manager)	0.5128	0.0690	0.2022
R20 (Safety Engineers)	0.4872	0.0571	0.2007
R16 (Business Specialist)	0.4103	0.0341	0.1735



Figure 1. Two-mode network (role – factor) of Iran



Figure 2. Two-mode network (role – factor) of North America

recognized, are not yet central to the industry's main discourse or influence chains. This reflects a sector still in the early stages of integrating environmental and social considerations into its core operational strategies. However, their presence in the network hints at a growing awareness and potential for these aspects to become more integral as the industry evolves, driven by global trends and local developments in sustainability and workforce management.

Network analysis result also reveals key roles instrumental in the adoption of automation in premanufacturing by particularly highlighting Construction Managers and University Professors for their high degree and betweenness centrality. This underscores their integral role in shaping industry practices, blending technical expertise with strategic oversight. In contrast, roles with a more focused selection of factors, such as R12 (Real Estate Developer), suggest a targeted approach, reflecting specific priorities or specialized requirements within their area of the construction industry. This might indicate a concentration on aspects like investment viability and project feasibility, central to their professional domain.

The R6 (Superintendent) role, with its notable eigenvector centrality, implies a strong connection within influential networks, possibly stemming from their on-ground experience and practical insights into project execution. These insights, coupled with the moderate level of industry experience and familiarity with premanufacturing among the Iranian respondents, point towards a sector at the cusp of technological evolution. The emphasis on modularization, efficiency, and profitability highlights a transition towards more innovative, cost-effective, and adaptable construction methodologies, aligning with the sector's growing inclination towards embracing technological advancements in an increasingly competitive market.

Overall, the network analysis indicates a construction industry at a pivotal juncture, with certain factors steering current practices, while others, currently on the periphery, hold the potential to shape future directions and priorities in the Iranian context.

4.2 The Case of North America

In North America, the analysis reflects a sector that prioritizes economic sustainability, operational excellence, quality enhancement, and skilled labor. The network measures reaffirm the significance of factors such as F9 (Efficiency and Productivity), F4 (Profitability), and F17 (Skilled Workers), each demonstrating a high degree centrality. This underscores their central role in the discourse around automation which is probably influenced by the region's advanced practices in premanufacturing and a workforce relatively familiar with these technologies. The high betweenness

centrality of these factors indicates they are not just central but also serve as critical junctions within the network. Also, their significant eigenvector centrality suggests their strong interconnectedness within a network of influential factors and emphasizes their impact on the sector's decision-making processes.

Additionally, F8 (Smaller and Modularized Factories) and F5 (Material Availability) demonstrate notable betweenness centrality, revealing their capacity to link distinct parts of the network. This focus reflects an understanding of the dynamic nature of construction projects in North America and the need for responsive production methods. The attention to F10 (Quality Improvement) aligns with the region's proficiency in premanufacturing, indicating an ongoing commitment to high standards and superior outcomes.

Interestingly, F24 (GHG Emissions and Energy Saving) and F25 (More Sustainable Design), while having lower centrality measures, still reflect a growing awareness and integration of environmental sustainability within the sector. This interconnectedness indicates a sector that is progressively weaving environmental sustainability into its operational fabric, recognizing that long-term efficiency and productivity must harmoniously coexist with ecological responsibility.

These network measures paint a comprehensive picture of a region where strategic economic, operational, environmental, and human resource considerations are intricately interwoven into the fabric of the construction sector's approach to automation. This blend of priorities, underpinned by solid technical understanding and broader societal implications, reflects a mature, forward-looking industry ready to leverage the benefits of advanced manufacturing techniques.

Regarding the roles, the network analysis reveals the prominent roles of R1 (Project Managers), R5 (Construction Managers), R17 (Product Managers), and R16 (Business Specialists) in shaping the move towards automation in premanufacturing. Their significant centrality measures indicate a deep engagement with and understanding of the industry's diverse aspects, from operational to environmental considerations.

Meanwhile, R20 (Safety Engineers), with its specific centrality measures, highlights the focus on technical and safety aspects in automation processes. This diversity in roles reflects a multifaceted approach within the industry, combining strategic management with specialized expertise. The centrality of these roles, particularly in betweenness and eigenvector measures, suggests their importance in relaying and contextualizing the industry's collective view on critical factors of automation.

4.3 Comparative Analysis of Prefabrication Automation Integration in Iran and North America

When comparing the network analyses of Iran and North America, several key insights emerge about their respective approaches to automation in premanufacturing. In Iran, the analysis reflects a sector strategically navigating its initial stages of technological integration, primarily focusing on foundational economic and operational efficiencies. This approach is characterized by a cautious yet strategic adaptation to modern methodologies, hinting at an industry gradually embracing change while considering local market dynamics and challenges.

Conversely, North America's analysis exhibits a more advanced stage of technological adoption, characterized by a holistic and integrated approach. This region demonstrates a well-rounded consideration of various factors, not just limited to immediate economic gains but also encompassing long-term sustainability and workforce development. The interplay of centrality measures in North America suggests a more interconnected and mature industry, where diverse professional roles contribute to a comprehensive understanding of automation's broader implications.

These contrasting scenarios highlight differing priorities and stages of development in the two regions. Iran's focus on key foundational elements suggests a phase of building toward more complex integrations of automation, whereas North America's balanced approach indicates an existing advanced implementation stage, where the industry is fine-tuning and expanding its existing automated processes. This comparison offers valuable insights into how different regions adapt to technological changes in construction, shaped by their unique market conditions and developmental stages.

5 Conclusion

The comparative study of Iran and North America's construction sectors provides valuable insights into the diverse approaches and stages of adoption of automation in premanufacturing. The network analysis reveals Iran's focused strategy on economic and operational aspects, indicative of an industry preparing for more complex technological integrations. In contrast, North America's construction sector showcases an advanced stage of adoption, characterized by a well-rounded integration of automation, considering long-term sustainability alongside immediate economic benefits. This study highlights the importance of understanding regional differences in technological adoption, offering guidance for policymakers and industry stakeholders in shaping future strategies. It underscores the need for a holistic

approach to integrating new technologies in construction, tailored to the unique challenges and opportunities presented by each market's stage of development and priorities.

References

- [1] McKinsey, "The Next Normal in Construction," McKinsey & Company, Jun. 2020. Accessed: May 06, 2023. [Online]. Available: https://www.mckinsey.com/~media/mckinsey/industries/capital%20projects%20and%20infrastructure/our%20insights/the%20next%20normal%20in%20construction/executive-summary_the-next-normal-in-construction.pdf
- [2] N. H. E. Omar and A. H. E. Omar, "Towards applying the global roadmap for technology development for zero energy projects," *Int. J. Adv. Eng. Bus. Sci.*, vol. 4, no. 1, 2023.
- [3] D. Hořínková, "Advantages and Disadvantages of Modular Construction, including Environmental Impacts," in *IOP Conference Series: Materials Science and Engineering*, 2021, p. 3. Accessed: May 08, 2023. [Online]. Available: <https://iopscience.iop.org/article/10.1088/1757-899X/1203/3/032002/meta>
- [4] S. Navaratnam, A. Satheeskumar, G. Zhang, K. Nguyen, S. Venkatesan, and K. Poologanathan, "The challenges confronting the growth of sustainable prefabricated building construction in Australia: Construction industry views," *J. Build. Eng.*, vol. 48, p. 103935, 2022.
- [5] W. Lu, K. Chen, F. Xue, and W. Pan, "Searching for an optimal level of prefabrication in construction: An analytical framework," *J. Clean. Prod.*, vol. 201, pp. 236–245, Nov. 2018.
- [6] K. Orlowski, "Automated manufacturing for timber-based panelised wall systems," *Autom. Constr.*, vol. 109, p. 102988, 2020.
- [7] S. K. Yevu, E. K. Owusu, A. P. C. Chan, K. Oti-Sarpong, I. Y. Wuni, and M. O. Tetteh, "Systematic review on the integration of building information modelling and prefabrication construction for low-carbon building delivery," *Build. Res. Inf.*, vol. 51, no. 3, pp. 279–300, Apr. 2023, doi: 10.1080/09613218.2022.2131504.
- [8] B. Kaiser, D. Littfinski, and A. Verl, "Automatic Generation of Digital Twin Models for Simulation of Reconfigurable Robotic Fabrication Systems for Timber Prefabrication," in *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*, IAARC Publications, 2021, pp. 717–724.
- [9] S. Kosse, O. Vogt, M. Wolf, M. König, and D. Gerhard, "Digital twin framework for enabling

- serial construction,” *Front. Built Environ.*, vol. 8, p. 864722, 2022.
- [10] H. Liu, Y. He, Q. Hu, J. Guo, and L. Luo, “Risk management system and intelligent decision-making for prefabricated building project under deep learning modified teaching-learning-based optimization,” *PLoS One*, vol. 15, no. 7, p. e0235980, 2020.
 - [11] A. Zhu, T. Dai, G. Xu, P. Pauwels, B. de Vries, and M. Fang, “Deep reinforcement learning for real-time assembly planning in robot-based prefabricated construction,” *IEEE Trans. Autom. Sci. Eng.*, 2023, Accessed: Dec. 04, 2023. [Online].
 - [12] H. J. Wagner, M. Alvarez, A. Groenewolt, and A. Menges, “Towards digital automation flexibility in large-scale timber construction: integrative robotic prefabrication and co-design of the BUGA Wood Pavilion,” *Constr. Robot.*, vol. 4, no. 3–4, pp. 187–204, 2020.
 - [13] P. Eversmann, F. Gramazio, and M. Kohler, “Robotic prefabrication of timber structures: towards automated large-scale spatial assembly,” *Constr. Robot.*, vol. 1, no. 1–4, pp. 49–60, 2017.
 - [14] E. Lachance, N. Lehoux, and P. Blanchet, “A Simulation Model to Analyze Different Automation Scenarios in a Mixed-Assembly Manufacturing Line: Timber-Frame Prefabrication Industry,” *J. Constr. Eng. Manag.*, vol. 149, no. 10, p. 04023091, Oct. 2023.
 - [15] E. Lachance, N. Lehoux, and P. Blanchet, “Automated and robotized processes in the timber-frame prefabrication construction industry: A state of the art,” in *2022 IEEE 6th International Conference on Logistics Operations Management (GOL)*, IEEE, 2022, pp. 1–10.
 - [16] F. G. Feldmann, “Towards Lean Automation in Construction—Exploring Barriers to Implementing Automation in Prefabrication,” *Sustainability*, vol. 14, no. 19, p. 12944, 2022.
 - [17] A. E. Oke, J. Aliu, P. Fadamiro, P. Akanni, P. S. Jamir Singh, and M. Shaharudin Samsurijan, “Unpacking the strategies to promote the implementation of automation techniques in the construction industry,” *Constr. Innov.*, 2023.
 - [18] D. Hevey, “Network analysis: a brief overview and tutorial,” *Health Psychol. Behav. Med.*, vol. 6, no. 1, pp. 301–328, Jan. 2018, doi: 10.1080/21642850.2018.1521283.
 - [19] S. Letouche and B. Wille, “Connecting the Dots: Exploring Psychological Network Analysis as a Tool for Analyzing Organizational Survey Data,” *Front. Psychol.*, vol. 13, p. 838093, 2022.
 - [20] C. P. Chea, Y. Bai, X. Pan, M. Arashpour, and Y. Xie, “An integrated review of automation and robotic technologies for structural prefabrication and construction,” *Transp. Saf. Environ.*, vol. 2, no. 2, pp. 81–96, 2020.
 - [21] E. Dahlgren, C. Göçmen, K. Lackner, and G. Van Ryzin, “Small modular infrastructure,” *Eng. Econ.*, vol. 58, no. 4, pp. 231–264, 2013.
 - [22] D. P. Butler, E. M. Malstrom, and S. C. Parker, “A model for the economic evaluation of automated manufacturing systems,” *Cost Eng.*, vol. 38, no. 6, p. 25, 1996.
 - [23] A. Gunasekaran, A. R. Korukonda, I. Virtanen, and P. Yli-Olli, “Improving productivity and quality in manufacturing organizations,” *Int. J. Prod. Econ.*, vol. 36, no. 2, pp. 169–183, 1994.
 - [24] R. Kangari and T. Yoshida, “Automation in construction,” *Robot. Auton. Syst.*, vol. 6, no. 4, pp. 327–335, 1990.
 - [25] A. Mathur, G. S. Dangayach, M. L. Mittal, and M. K. Sharma, “Performance measurement in automated manufacturing,” *Meas. Bus. Excell.*, vol. 15, no. 1, pp. 77–91, 2011.
 - [26] N. Viswanadham and T. L. Johnson, “Fault detection and diagnosis of automated manufacturing systems,” *IFAC Proc. Vol.*, vol. 21, no. 15, pp. 95–102, 1988.
 - [27] A. B. Smith Jr, “Boycotts of prefabricated building products and the regulation of technological change on construction jobsites,” *ILR Rev.*, vol. 25, no. 2, pp. 186–199, 1972.
 - [28] D. Smith, T. R. Heathman, A. Klarer, C. LeBlon, Y. Tada, and B. Hampson, “Towards automated manufacturing for cell therapies,” *Curr. Hematol. Malig. Rep.*, vol. 14, pp. 278–285, 2019.
 - [29] C. Rosa, F. J. G. Silva, and L. P. Ferreira, “Improving the quality and productivity of steel wire-rope assembly lines for the automotive industry,” *Procedia Manuf.*, vol. 11, pp. 1035–1042, 2017.
 - [30] S. Y. Chyung, K. Roberts, I. Swanson, and A. Hankinson, “Evidence-based survey design: The use of a midpoint on the Likert scale,” *Perform. Improv.*, vol. 56, no. 10, pp. 15–23, 2017.
 - [31] T. Opsahl, “Triadic closure in two-mode networks: Redefining the global and local clustering coefficients,” *Soc. Netw.*, vol. 35, no. 2, pp. 159–167, 2013.
 - [32] M. Newman, *Networks*. Oxford university press, 2018.
 - [33] L. C. Freeman, “Centrality in social networks: Conceptual clarification,” *Soc. Netw. Crit. Concepts Sociol. Lond. Routledge*, vol. 1, pp. 238–263, 2002.
 - [34] P. Bonacich, “Power and centrality: A family of measures,” *Am. J. Sociol.*, vol. 92, no. 5, pp. 1170–1182, 1987.

Assessing the Financial Dynamics of Modular Offsite Building Projects

Hisham Said¹, Tarek Salama², and Anthony Gude³

¹ School of Engineering, Santa Clara University, Santa Clara, California, USA

² Department of Construction Management, California State University Sacramento, Sacramento, California, USA

³ Gude Capital management

hsaid@scu.edu, salama@csus.edu, agude@gudecapital.com

Abstract –

Current project development financial modeling approaches have been utilized for traditional stick-built construction. This paper presents a novel system dynamics simulation model to analyze the financial dynamics of modular offsite building projects. The model captures the complex interactions between construction activities and cash flows, allowing for the evaluation of various financial parameters of developers. The developed model was applied to a real-world case study, demonstrating alignment with traditional financial models. Sensitivity analysis revealed the significance of specific parameters in influencing project profitability, such as factory deposit ratio, equity capital ratio, and interest rates. These findings provide valuable insights and strategies to enhance the financial attractiveness of modular construction projects. The paper contributes to the body of knowledge by offering a comprehensive tool for studying the financial implications of modular construction projects. This tool can assist stakeholders in making informed decisions and optimizing project outcomes, thus promoting the growth and adoption of modular construction as a sustainable and efficient building approach.

Keywords – Modular Construction; Offsite; System Dynamics; Simulation; Financial; Modeling.

1 Introduction

Smart manufacturing, automation, and industrial production (prefabrication, and offsite manufacturing) are main technologies involved in the Industry 4.0 (I4.0) revolution [1]. I4.0 is defined as an ‘integration of technologies that reshapes the way things are made’ [1]. Construction 4.0 represents the Industry 4.0 version for the construction industry by digitization using the following three transformational trends: 1) Industrial production and construction (offsite manufacturing, 3D

printing, and automation); 2) Cyber-physical systems (Internet of things (IoT) and sensors); 3) Digital technologies (Building information modeling (BIM), artificial intelligence (AI), big data, Blockchain, augmented reality, cloud computing, and laser scanning, etc. [1]. Modular and offsite construction (MOC) systems are classified based on size of prefabricated components into prefabricated and processed materials, panelized (2D modules), modular (3D modules), and hybrid construction (combination of 2D and 3D modules) [2,3]. Modular and offsite construction (MOC) provides many advantages due to its competency in delivering better quality, lower cost, shorter schedules, and higher safety [4,5]. However, modular and offsite construction is known by its unique risks and uncertainties, which can be different from those of traditional construction [6]. Hence, the modular and offsite construction industry is facing many obstacles and barriers and not all modular construction projects were delivered successfully [7]. However, according to the modular building institute (MBI), the market share for permanent modular construction as a percentage of the total construction industry increased from 2.14% in 2015 to 6.03% in 2022 [8]. Many studies investigated the barriers to increase the market share for modular and offsite construction [3,6,9,10]. Salama et al. [3] investigated five main barriers determined by industry professionals using a questionnaire to collect data for: 1) the negative stigma for modular and offsite construction; 2) the lack of examples of past success in this industry; 3) the lack of standards and regulations. 4) the unclear procurement strategies utilized in this industry; and 5) project financing obstacles. While other studies assessed the risks associated with this industry [11,12]. Li et al. [11] identified modular risks such as engineering, occupational and cultural, socio-economic, and financial risks and a fuzzy analytic hierarchy process (AHP) was utilized to rank these risks; then simulation techniques were used to assess risks of modular projects. Abdul Nabi and El-adaway [12] developed a new approach to forecast the cost performance of modular construction

projects after identifying 50 modular risks that include risks for material supply, transportation, manufacturing installation, contracts and disputes. This study also included financial risks such as the high initial (capital) costs and the inconsistent cash flow. Moreover, many offsite construction companies are providing good examples of success [8], however, in certain markets, this industry is still immature and encounters financial obstacles that might lead to disputes in modular construction projects [13,14]. Hence, this research is investigating the effect of different financial parameters that can be changing based on the current economic challenges existing in the market using a newly developed system dynamics model to simulate different scenarios for these financial parameters.

2 Literature review

Cameron and Carlo [15] used spreadsheets to perform an assessment for the impact of using modular construction on the equity internal rate of return (IRR), which is the annual rate of return for investment in a specific period of time. This study compared modular and traditional construction by having two scenarios for six buildings. The first scenario assumes these buildings are constructed for rent, while the second is assuming it is constructed for sale. It was concluded that the equity IRR increased from 35.1% to 47.5% for the first scenario when utilizing modular construction compared to traditional construction, while the increase for the second scenario was from 25.75 % to 27.60 %. However, this study did not investigate the effect of changing financial parameters such as the lending interest rate, debt to equity ratio, or the time period of the cash flow. Velamati [16] compared between modular and traditional construction for a case study of a 20-story high rise building in a major east coast project, and it was concluded that the annual IRR for modular construction is 22 % which is higher than the IRR for traditional construction which was 20.44%.

Cazemier [17] performed a comparative financial analysis between a cross laminated timber (CLT) building offsite construction, and traditional steel and concrete building in Australia. This study used the EstateMaster software to input financial parameters such as construction cost, interest expenses, land purchase price, construction contingency, land purchase price, sales revenue, and professional fees to investigate its effects on profit indicators such as IRR, return on equity (ROE) – the total amount of return received for original investment –, development margin, and development profit. It was concluded that CLT developments may result in lower development margin, development profit, and ROE, but it increases equity IRR due to the shorter timeline schedules for offsite construction.

Salama et al. [4] presented a comprehensive literature review for the challenges that any modular developer may encounter to finance MOC projects. These challenges include: 1) the large upfront capital requirements which make manufacturers ask for upfront payment of around 50 percent in advance to purchase materials and if the developer is acquiring this funding from banks, then the current high interest rates would reduce the IRR; 2) perception of ownership: since banks don't issue funding installments until the onsite installation of modules which increase the financial burden for any modular builder; 3) the immature market for MOC since many developers are lacking the experience in this industry while facing uncertainties in scheduling and pricing due to the current supply chain issues and fluctuation in material prices; 4) progress monitoring for manufacturing is considered an increased risk for financing MOC by banks compared to than traditional construction projects; 5) lack of support from authorities and financial sector for MOC industry which makes modular manufacturers pay for materials and then for manufacturing while construction financing is released after modules are delivered to construction site; 6) lending interest rates for MOC are higher than traditional construction since several banks lack the full understanding for it, hence this fact is affecting the IRR for any modular developer; 7) financial impact for using MOC can be different from country to another due to many factors such as the differences in labor costs and the different economical situations between countries; 8) transportation and storage costs might not be considered by lending banks while securing construction financing; 9) Some projects which are publicly funded might not allow some payments certifications in a timely manner for MOC. However, Salama et al. [4] did not conduct a quantitative analysis for assessment of utilizing modular construction on the internal rate of return (IRR) for modular developers.

Assaf et al. [18,19] developed an automated cash flow system that mitigates a number of MOC financial barriers using advanced technologies, such as building information modeling (BIM), blockchain, and smart contracts. However, this study focused on developing an automated payment system MOC projects that consider different procurement approaches more than focusing on financial modeling for MOC to investigate the effect of financial parameters on enhancing the profitability of MOC projects.

3 Research Goal and Methodology

The goal of this study is to develop a novel simulation model to assess the financial dynamics and outcomes of modular offsite building projects. The proposed model would fill the existing research gaps due to the lack of

theoretical understanding and the modeling approaches to comprehend the unique financial implications and requirements of offsite construction projects. System dynamics (SD) modeling [20] was utilized to develop the proposed financial model to account for the time-delayed complex causal loops between the construction work (both offsite and onsite) and the resulting cash flow [21, 22]. System dynamics is a mathematical modeling technique that uses stocks, flows, internal feedback loops, and time delays to understand the nonlinear behavior of complex systems over time. It's a computer-based approach for strategy development and better decision making in complex systems.

To achieve this goal, the study methodology involved three steps. The first step is to develop the prefabrication-construction dynamics (PCD) module that captures the progress of the offsite and onsite work scopes and their resulting expenses. The second step is the development of the cashflow dynamics (CFD) module that accounts for the project financing and revenue cash flows in addition to the construction expenses from the first module. The third step of this study is the validation of the SD model using a real modular project. It should be noted that the two SD modules are interconnected, as explained later. The two modules were implemented in the same SD model using AnyLogic 6.9.0 simulation software.

4 Prefabrication-Construction Dynamics Module

The PCD module was developed around the progress in three main work packages: modular units, site work, and building finishes. The first work package includes the fabrication and installation of the prefabricated modular units. The second package covers the site work, which mainly includes the earthwork, foundation, utilities, pavement, and landscape. The third work scope includes any building finishes that were not done in the prefabricated modules, such as building systems, painting, flooring, facade, windows, doors, etc. As shown in Figure 1, each of these scopes is modeled using stocks (rectangles), flows (arrows with valves), input parameters (small circles with black triangle), and auxiliary variables (circles). All three work scopes start with a work-in-progress (WIP) stock element (*Factory_WIP*, *Site_WIP*, *Finish_WIP*), which holds 100% value of progress at the beginning of the simulation run. Afterwards, the progress in each work scope is modeled using flow elements that transfer the scope from WIP stocks to completed work stocks. The site work scope progress is simulated using the *SiteRate* flow from *Site_WIP* to *FinishedSiteWork* stocks, and its units are % of work completed per time unit (month). The finishes

work scope progress is similarly simulated using the *InteriorRate* flow and the *completedFinishes* stock. However, the work scope of the prefab modules is simulated using more flows (i.e. *ProdRate* and *InstallRate*) and stocks (i.e. *FinishedModules* and *InstalledModules*) to model the fabrication and installation steps. Each work flow is calculated using background functions that consider the following input data: 1) a work rate curve that defines the expected progress rate at a given work completion level; and 2) a table of productivity coefficients for each month to account for holidays and weather impacts. In addition, the value of each flow is controlled by either time or precedence conditions. The *SiteRate* flow will be forced to be null until the simulation time reaches construction planned start time (i.e. *SiteStartTime*) to mark the time of the contractor's access to the site and ability to perform the site work. The *ProdRate* flow will be null until a specific amount of site work progress is achieved (i.e. *FinishedSiteWork* \geq *SiteWorkToProd*) to allow for enough buffer between site and offsite operations. The *InteriorRate* flow will be activated after a specific amount of prefab unit installation is achieved (i.e. *InstalledModules* \geq *InteriorAfterModules*).

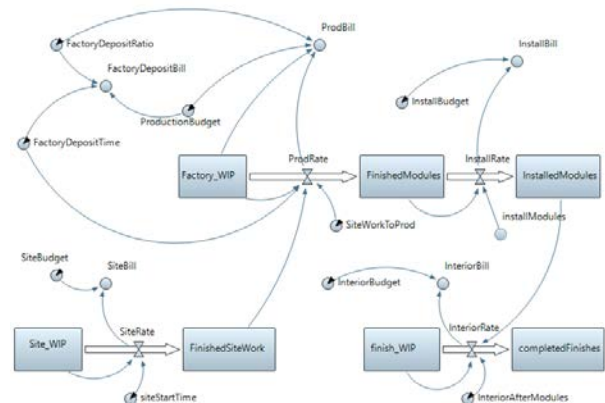


Figure 1. The system dynamics module of the offsite and onsite construction scopes

The model also calculates the construction expenses based on the progress achieved in each work scope. For example, the auxiliary variable *SiteBill* calculates the expenses from the performed work by multiplying the current *SiteRate* value and the inputted *SiteBudget*. A similar logic is used for *InteriorBill*, which is the multiplication of *InteriorRate* and *InteriorBudget*. However, the value and timing of the module production expense billing depends on the deposit payment to the manufacturer to purchase the necessary material (*FactoryDepositBill*). This means that no production billing can be made until achieving a progress that is equivalent to the received down payment. The deposit

payment value accounts for the ratio of this deposit (*FactoryDepositRatio*) to the total module production budget (*ProductionBudget*) and the time of making this deposit payment (*FactoryDepositTime*). These construction expense auxiliary variables will be used in the project cash flow dynamics module, as explained in the next section.

5 Cashflow Dynamics Module

The CFD Module simulates the flow and accumulation of equity, debt, expenses and revenue of the development of the modular building project. As shown in Figure 2, the module utilizes three main stock elements: *Equity*, *Debt*, and *NetCashFlow*.

The *Equity* stock accumulates the *Investment* flow, which can be done in a single installment or multiple installments. The *Equity* invested in the project equals the multiplication of *EquityCapitalRatio* (usually 30 - 70%) and the total capital required for the project *projCapital* (which is the summation of all construction-related budgets).

On the other hand, the *Debt* stock element accumulates two flows (*DebtDraws* and *Interest*) and is reduced by a third flow (*LoanPayoff*). The debt draws *DebtDraws* from the project creditor are calculated as a function of the debt-share of the capital (using *EquityCapitalRatio* and *projCapital*) and the schedule setup of these draws (using *firstDebtDrawTime*, *freqDebtDraw*, and *nDebtDraws*). The Interest flow element calculates the amount of interest in each month based on the current *Debt* level and the *interestRate* value. This project interest rate is usually reduced through refinancing from a higher rate *constrInterestRate* during construction to a lower rate *permInterestRate* once the building is operational. Finally, the *Debt* is serviced using the *LoanPayoff* flow element, which pays off only the due interest on any non-negative values of the *Debt* until the sale time of the constructed property (*saleTime*).

The *NetCashFlow* stock simulates the net accumulation of two positive flows (inflows) and three negative flows (outflows). *NetCashFlow* increases by accumulating: 1) the capital deposits (*CapitalReceipts*) during the construction time; and 2) the generated *NetIncome* from operating the building (when the level of the *CompleteFinishes* reached the targeted *SubstainCompletion*) and selling it at an expected sale time. The value of the *CapitalReceipts* equals the summation of all equity and debit installments made to finance the project, i.e. the values of *Investment* and *DebtDraws* flows. *NetCashFlow* decreases by deducting:

- 1) the construction expenses (*ConstructionBills*) are calculated monthly from the PCD module; 2) the expenses related to the land purchase and soft-cost pre-construction items (design, permitting, marketing); 3) the debt repayment (*DebtPayments*).

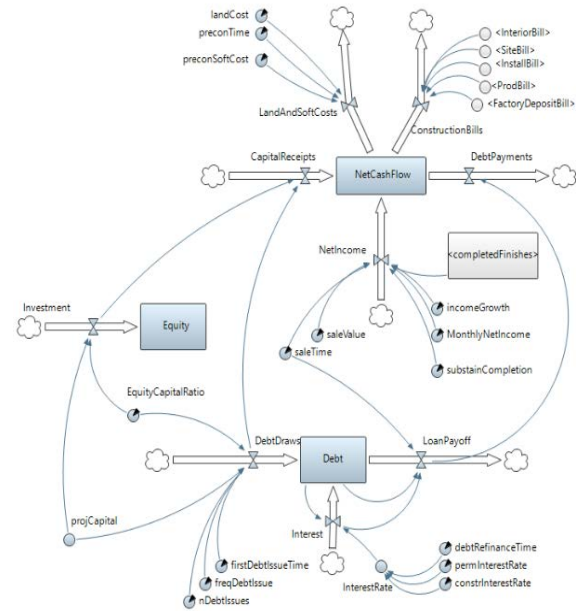


Figure 2. The system dynamics module of the project cashflows

6 Case Study Analysis

The developed SD simulation model is applied to a real case study of a modular residential building project to verify its correctness and validate its relevance. The project is located in California, USA, which entails constructing a 0.4-acre lot into a 74-unit apartment building. Table 1 lists the main time and cost input data of the case study. The debt is structured as follows: a) the first debt draw will happen in the 3rd month, followed by 20 monthly draws; b) the construction debt and permanent debt monthly interest rates are 0.46% and 0.375%, respectively; c) the construction debt will be refinanced in the 30th month using the permanent debt; d) only the monthly interest (both construction and permanent) will be paid without paying for the loan principal amount; and e) the permanent debt will be paid fully when the property is sold (after 84 months).

Table 1. Case study main input data

Input data	Value
Module production budget	\$5,264,570
Module installation budget	\$532,890
Site work budget	\$2,304,954
Finishes budget	\$1,084,940
Land cost	\$6,352,900
Soft (preconstruction) cost	\$537,000
Factory deposit time	0
Factory deposit ratio	0.50
Site work progress rate	8% /month
Module fabrication progress rate	7.5% /month
Module installation progress rate	20% /month
Finishes progress rate	10% /month
Substantial completion	100%
Equity/Capital Ratio	0.45
Preconstruction time	10 months
Time to start site work	5 th month
Property net income	\$94,213/month
Property income growth	0.2% /month
Time to sell the property	84 months
Property sale value	\$23,553,055

The SD simulation output closely matched the spreadsheet-based financial model of the project and provided additional insights and visualization of the construction progress and financial feasibility. Figure 3 depicts the recorded levels of four intermediate and final stock elements in the PCD module: *FinishedSiteWork*, *FinishedModules*, *InstalledModules*, and *CompletedFinishes*. All construction scopes were concluded in the 33rd month. The level of the *FinishedModules* stock stopped its accumulation at the 20th month as it hit a 60% progress milestone, which marked the start of the module installation work. On the financial side, the SD model was run for two cases: the base case as described before, and the base case with no factory deposit. This is due to the fact that the authors had access to the spreadsheet-based financial model that did not account for a factory deposit. The SD model matched the spreadsheet model in the calculated value of the annual internal rate of return (IRR), which was found to be 14.7%. However, a lower IRR value of 13.4% was calculated when accounting for a 50% factory deposit payment (i.e. half of the fabrication budget is paid at time 0). Furthermore, the net cash level from the SD model (i.e. the level of the *NetCashFlow* stock) was visualized as shown in Figure 4 to highlight the timely implications of waiving or requiring factory deposits. Without the factory deposit, the developer will experience cash

shortage between the 30th and 40th months, which need to be covered by additional equity investment or loans. On the other hand, a 50% factory deposit will result in an additional period of negative cash levels with much larger need for capital in the first 10 months.

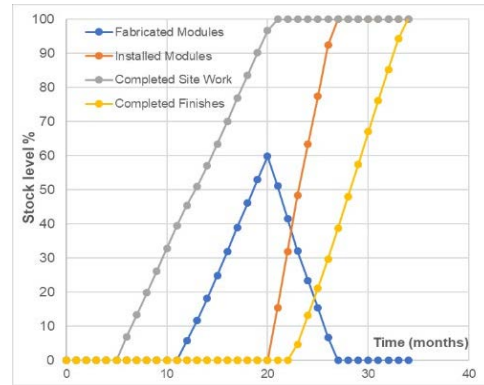
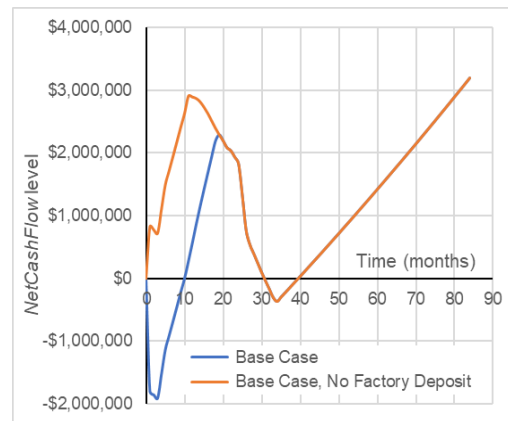


Figure 3. The levels of the major construction work stocks of the analyzed case study

Figure 4. The level of the *NetCashFlow* stock of the analyzed case study under the base case with factory deposit and the same case without the deposit

A sensitivity analysis was performed to further validate the developed SD model and obtain in-depth insights on the influence of different input parameters on the financial outcomes of modular building projects. The previous base case was compared to the +50% and -50% changes in four major financial parameters: factory deposit ratio (*FactoryDepositRatio*), equity/capital ratio (*EquityCapitalRatio*), construction loan interest rate (*constrInterestRate*), and permanent loan interest rate (*permInterestRate*). The other model parameters were not considered in the sensitivity analysis as they are either influenced by the construction plan (fabrication, preconstruction, construction) or the property real estate attributes (property income and sale). Figure 5 shows the sensitivity chart of the IRR value with respect to the change in these four parameters. None of the considered

changes resulted in a negative IRR, but it was found that the profitability of the project is the most sensitive to the change in the equity/capital ratio. Positive and negative 50% changes in the equity/capital ratio value resulted in the IRR value to change to 11.4% and 16.8%, respectively (compared to 13.4% for the base case). It should be noted that low IRR values indicate a profitable project, but the project may not be profitable enough for the developer if IRR is less than their minimal attractive rate of return (MARR). Also, Figure 6 shows the net cash level charts of the sensitivity analysis cases. The equity/capital ratio value had also the biggest impact on the net cash levels of the project. For example, a -50% change of the equity/capital ratio (i.e. 0.225) resulted in a significant negative cash period at the beginning of the project. This negative cash period has to be covered by the developer's own equity, which resulted in the drop in the IRR value to 11.4%.

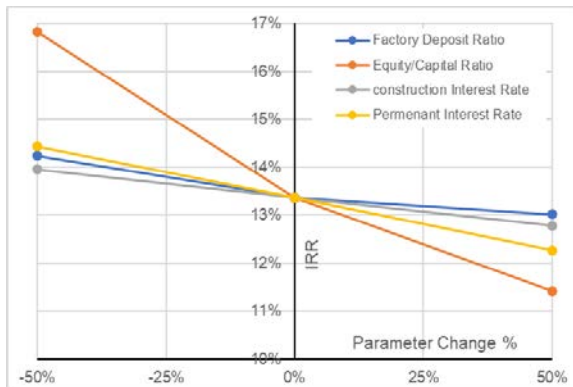


Figure 5. The sensitivity of the internal rate of return (IRR) for the change in four financial parameters

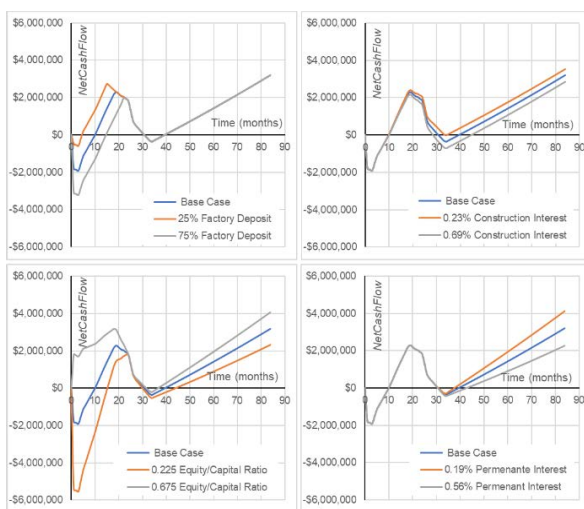


Figure 6. The sensitivity of the *NetCashFlow* level to the changes in the analyzed financial parameters

7 Discussion

The successful application of the SD simulation model to the modular residential building project is evidenced by its close alignment with the spreadsheet-based financial model. The model not only validated the correctness of the financial projections but also offered valuable insights and visualization tools to enhance the understanding of construction progress and financial dynamics. The recorded levels of intermediate and final stock elements provided a visual representation of the project's progression. Notably, the SD model accurately captured the milestones, such as the cessation of *FinishedModules* accumulation at the 20th month, marking the commencement of module installation work. The model's alignment with the spreadsheet model in calculating the annual internal rate of return (IRR) at 14.7% instills confidence in the reliability of the simulation. Introducing a 50% factory deposit payment revealed a nuanced financial landscape. The lower IRR of 13.4% suggests the importance of considering financial strategies related to factory deposits. This observation is further supported by the sensitivity analysis, which also revealed a higher sensitivity to the equity/capital ratio. Increasing the dependency on equity degrades the project's returns, but it is also necessary to cover the negative cash periods of the project due to the hefty deposit required in most modular projects. The combined impact of increasing the factory deposits and the accompanying increase of equity investments results in the deterioration of the financial appeal of modular building projects.

The large upfront capital requirements from manufacturers of modular offsite building projects is one of the main hurdles faced by this industry [4]. Manufacturers ask for an upfront factory deposits of around 50 percent to procure materials in a short period of time for enhancing efficiency of manufacturing [23]. This required large factory deposits from manufacturers may have an effect on bank reserves, hence institutional lenders and banks may require any collateral to reserve some funding on their equity for avoiding regulators' scrutiny [24]. However, Cameron and Carlo [15] conducted an analysis to investigate the impact of utilizing modular construction on the IRR compared to using traditional construction while studying two different scenarios. The first scenario was for a project that includes six buildings which is utilized as rent development, while the second is for the same project but it would be constructed for sale. For first scenario, The IRR was found to increase for the case of using modular construction compared to traditional construction from about 35.1% to 47.5%, and from 25.75 % to 27.60 % for the second scenario. The reason behind the difference between IRR in both scenarios is due to the rental income

for part of the development. Cazemier [17] conducted another financial analysis by comparing a cross laminated timber (CLT) project that was using offsite construction technologies, and traditional steel and concrete building. The CLT project resulted in less development profit, and return on equity (ROE), however the IRR increased due to the reduced investment timeline schedule for offsite building projects compared to traditional construction.

This study contributes to the current body of knowledge by developing a simulation-based model for analyzing the financial performance of modular building considering the complex dynamics and dependencies between the construction and cash flows. The model quantifies the impact of some project financial parameters that have been hypothesized to hinder the growth of the modular building industry due to its disadvantaged setup of higher capital needs early in the project development periods. The model can benefit the current construction practice at different levels, where modular building developers can utilize it to produce more holistic financial projection of their projects and the industry advocating organization can better communicate the need for public policies and incentives to support the rapid delivery of housing through modular building approaches.

8 Conclusion

This study's goal was to develop a simulation-based model for evaluating the financial aspects of modular construction by considering the intricate interactions and dependencies between construction and cash flows. The proposed model quantifies the effect of financial parameters that have been theorized to impede the development of the modular construction sector due to its unfavorable structure of significant early capital requirements. By utilizing a system dynamics (SD) modeling approach, the study examined the impact of variables including factory deposit ratio, equity/capital ratio, and interest rates on the internal rate of return (IRR) for developers in this industry. The model's application to a real-world modular residential building project illustrated its accuracy in aligning with traditional financial models and provided insights into construction progress and cash flow dynamics. Sensitivity analysis further highlighted the significance of specific financial parameters in influencing project profitability. The findings suggest the need for strategies to optimize the equity/capital ratio and mitigate the impact of upfront factory deposits to enhance the financial attractiveness of modular construction projects. Overall, the SD model provides a valuable advanced tool for developers to study different financial scenarios for modular projects and visualize cash flow. Recommendations were drawn for

enhancing developer profit represented in an optimized IRR. It is noteworthy that the broad application of the proposed tool may be limited due to the fact that the construction and real estate industries are not accustomed to using simulation modeling in their practice. As such, the tool is used to generalize observations on the financial modeling of modular building development projects, which can be considered in the industry practice and their existing computer tools. Despite the insights from the sensitivity analysis, advanced stochastic analysis is needed to examine the concurrent interdependent changes in the model parameters. Finally, the study could benefit from considering additional factors such as government incentives, the impact of economies of scale, and the effect of different procurement strategies on financial outcomes.

Acknowledgment

The authors of this paper acknowledge the Research and Creative Activity award (RCA Faculty Awards Program) awarded by Faculty Senate's Research and Creative Activity Subcommittee in California State University, Sacramento for their support to this research.

References

- [1] Sawhney, A., Riley, M., Irizarry, J., & Riley, M. Construction 4.0: An Innovation Platform for the Built Environment. Routledge publishing, Taylor & Francis Group, 526 pages, ISBN-13 : 978-0367027308, 2020.
- [2] Salama, T., Salah, A., Moselhi, O. and Al-Hussein, M. (2017), "Near optimum selection of module configuration for efficient modular construction.", *Automation in Construction*, Vol. 83, pp. 316-329.
- [3] Salama, T., Moselhi, O. and Al-Hussein, M. Overview of the characteristics of the modular industry and barriers to its increased market share. *International Journal of Industrialized Construction*, Vol. 2 No. 1, pp. 30-53, 2021.
- [4] Salama, T., Figgess, G., Elsharawy, M. and El-Sokkary, H. Financial modeling for modular and offsite construction. *Proceedings of the International Symposium on Automation and Robotics in Construction*, Vol. 37, pp. 1082-1089, Kitakyushu, Japan, 2020.
- [5] Salama, T., and Said, H. Agility assessment framework for modular and offsite construction.", *Construction Innovation*, <https://doi.org/10.1108/CI-09-2022-0238>, 2023.

- [6] Wuni, I.Y. and Shen, G.Q. Barriers to the adoption of modular integrated construction: Systematic review and meta-analysis, integrated conceptual framework, and strategies. *Journal of Cleaner Production*, 249, p.119347, 2020.
- [7] Choi, J. O., J. T. O'Connor, and T. W. Kim. Recipes for cost and schedule successes in industrial modular projects: Qualitative comparative analysis. *Journal of Construction Engineering and Management*, 142 (10): 04016055.2016. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001171](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001171).
- [8] The Modular Building Institute (MBI). Permanent Modular Construction Report, 39 pages, 2023. <https://www.modular.org/industry-analysis/>
- [9] Gusmao Brissi, S., Debs, L. and Elwakil, E. A review on the factors affecting the use of offsite construction in multifamily housing in the United States. *Buildings*, 11(1), p.5, 2020.
- [10] Navaratnam, S., Satheeskumar, A., Zhang, G., Nguyen, K., Venkatesan, S. and Poologanathan, K. The challenges confronting the growth of sustainable prefabricated building construction in Australia: Construction industry views. *Journal of Building Engineering*, 48, p.103935, 2022.
- [11] Li, H.X., Al-Hussein, M., Lei, Z. and Ajweh, Z. Risk identification and assessment of modular construction utilizing fuzzy analytic hierarchy process (AHP) and simulation. *Canadian Journal of Civil Engineering*, 40(12), pp.1184-1195, 2013.
- [12] Abdul Nabi, M. and El-adaway, I.H. Risk-based approach to predict the cost performance of modularization in construction projects. *Journal of Construction Engineering and Management*, 147(10), p.04021133, 2021. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0002159](https://doi.org/10.1061/(ASCE)CO.1943-7862.0002159)
- [13] Salama, T. and Said, H. Agility assessment framework for modular and offsite construction. *Construction Innovation Journal*, 2023. <https://doi.org/10.1108/CI-09-2022-0238>.
- [14] Abdul Nabi, M. and El-adaway, I.H. Understanding disputes in modular construction projects: Key common causes and their associations. *Journal of Construction Engineering and Management*, 148(1), p.04021184, 2022. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0002208](https://doi.org/10.1061/(ASCE)CO.1943-7862.0002208)
- [15] Cameron, P.J. and Di Carlo, N.G., Piecing together modular: understanding the benefits and limitations of modular construction methods for multifamily development (Doctoral dissertation, Massachusetts Institute of Technology), 2007.
- [16] Velamati, S. Feasibility, benefits and challenges of modular construction in high rise development in the United States: a developer's perspective, Master of Science, Dept. of Architecture, Massachusetts Institute of Technology, On-line: <http://hdl.handle.net/1721.1/77129>, 2012.
- [17] Cazemier D. Comparing cross laminated timber with concrete and steel: a financial analysis of two buildings in Australia. *Modular and Offsite Construction (MOC) Summit Proceedings*, 2017.
- [18] Assaf, M.K.A., Financial management and procurement approaches of modular construction projects: a blockchain-based bim framework. Master of Science thesis, the Hong Kong polytechnic university, 2022.
- [19] Assaf, M., Hussein, M., Alsulami, B.T. and Zayed, T. A Mixed Review of Cash Flow Modeling: Potential of Blockchain for Modular Construction. *Buildings*, 12(12), p.2054, 2022.
- [20] Stermann J. D. (2000). *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin/McGraw-Hill, New York, NY.
- [21] Dabirian, S., M. Ahmadi, and S. Abbaspour. 2021. "Analyzing the impact of financial policies on construction projects performance using system dynamics." *Engineering, Construction and Architectural Management*, 30 (3): 1201–1221. Emerald Publishing Limited. <https://doi.org/10.1108/ECAM-05-2021-0431>.
- [22] Khaledi, H. (2015). "A Generic System Dynamics Model of Firm Internal Processes." *Proceedings of the 33rd International Conference of the System Dynamics Society, System Dynamics Society, Cambridge, Massachusetts, USA -- July 19-23, 2015*, ISBN: 9781510815056.
- [23] Galante C, Draper-Zivetz S, Stein A. *Building Affordability by Building Affordably: Exploring the Benefits, Barriers, and Breakthroughs Needed to Scale Off-Site Multifamily Construction*. Terner Centre for Housing Innovation, UC Berkeley. Online: http://ternercenter.berkeley.edu/uploads/offsite_construction.pdf, 2017.
- [24] Maher A. *Breaking with tradition*. Special report from AVANA capital. 2018 https://avanacapital.com/wpcontent/uploads/Expert-Speak-Finance-HT_MARCH_ISSUE_-2018.pdf.

Impacts of Multi-skills on Project Productivity and Completion Time: A Building Renovation Case Study

Muhammad Atiq Ur Rehman¹, Sharfuiddin Ahmed Khan², and Amin Chaabane³

^{1,3} Department of Systems Engineering, École de technologie supérieure, Canada

²Department of Industrial Systems Engineering, University of Regina, Canada

muhammad.atiq-ur-rehman.1@ens.etsmtl.ca, Sharfuiddin.Khan@uregina.ca, amin.chaabane@etsmtl.ca

Abstract –

This paper delves into the complexities of integrated supply chain planning decisions and project management in the construction industry, specifically about multi-skill, resource-constrained, and multi-project scheduling problems. To address this issue, we propose a mathematical mixed-integer linear programming (MILP) model that optimizes schedules for multiple projects involving different activities, renewable resources, and skill sets. We test our solution in a real-world case study using data from a construction company in Canada specializing in building renovation. In addition, we conduct a sensitivity analysis to test the robustness of our model and how it can support managerial decision-making. Our results prove that the proposed decision support system (DSS) can help project managers and logistic coordinators achieve optimal integrated project scheduling and supply chain resource planning. Furthermore, it allows for better control over the impact of multi-skill requirements and renewable resources on construction supply chain performance. Finally, our new DSS provides valuable collaborative information to find areas for improved decision-making of project completion and workforce compression times.

Keywords –

Construction project management; Scheduling; Assignment; Optimization; Multi-Skills

1 Introduction

The construction industry has faced many challenges, including rising costs, stagnant productivity, and alarming levels of waste. These obstacles have become even more daunting as the demand for increasingly complex structures, higher sustainability standards, and elevated customer expectations has skyrocketed. Exploring and implementing new analytical models, resolution methods, and decision-support tools from supply chain and logistics perspectives is crucial to

support the industry's ongoing efforts toward industrialization and digital transformation. An integrated approach that accounts for project management and supply chain management (SCM) decisions while allocating multiple resources and workforce skills is essential to promote efficiency, resilience, and a customer-centric approach to construction projects [1].

A real-world scenario involving a construction company based in Canada inspired this research. The company specializes in delivering and repairing modular and traditional construction projects for Canadian clients. They also provide high-quality foundation and structural repair solutions. Their challenge was to allocate multiple resources across different projects while managing different schedules. In search of a new DSS, the company sought to optimize its project scheduling and resource allocation for its current portfolio of projects. Building upon this case, we propose a model to streamline supply chain operations and project management aspects of construction projects, ultimately reducing project completion time and improving resource allocation.

Successful project management requires a combination of knowledge, skills, tools, and technology to execute project activities and achieve predetermined goals effectively. Project scheduling is a crucial aspect of project management, which involves allocating resources and determining timelines for implementing project tasks. The challenge lies in that these tasks follow predetermined precedence relationships and compete for scarce resources, resulting in the optimal schedule for achieving specific goals. The problem is known as the resource-constrained project scheduling problem (RCPSP), first defined by [2], and is widely recognized in project scheduling across various fields, including engineering, construction, manufacturing, software development, and logistics management. It is important to note that the RCPSP is a well-known NP-hard problem, which adds to its complexity.

The primary goal of the Resource-Constrained Multi-Project Scheduling Problem (RCMPSP) is to create a schedule that accounts for the sequence of activities

across multiple projects while also considering resource limitations. This problem is complicated because real-world projects come with various time and cost constraints [3]. These factors are essential to consider, as they significantly affect the overall efficiency of the project construction supply chain (CSC).

So, construction schedulers must carefully manage resources, prioritize activities, and create feasible and effective schedules. However, it is essential to note that some schedules may be infeasible due to limited resource availability. In particular, the availability of renewable resources, such as skilled labor and managers, and non-renewable resources, such as construction materials in building projects, are finite and complicated, and managing several concurrent projects is a sophisticated task that deals with sharing limited resources among multiple projects. Hence, the scheduling process becomes complicated when managing simultaneous projects and resources [4]. Managing several concurrent projects is a sophisticated task that deals with sharing limited resources among multiple projects [5]. After considering these intricacies, the simple scheduling problem becomes a Multi-Skill, Resource-Constrained, Multi-Project Scheduling Problem (MSRCMPSP), which we will denote as a \mathcal{P} for the rest of the paper.

We have selected and compared the latest research papers in reputable journals with our work. We classified them based on multi-project, multi-skill, and supply chain management, considering renewable and non-renewable resources and supplier parameters and whether they applied to real-world projects. Notably, the study conducted by [6] in 2023 comprehensively analyzes skill-based assignments within this domain. They even considered skills familiarity level per renewable resource but should have considered non-renewable resources, supplier selection, and supply chain times in their mathematical model. They used numerical examples to test their model. Another study considered a skill-based assignment and scheduling optimization model [7]. This paper focused on the dynamic nature of the skill level of human resources and its effects on project schedules. They build an optimization model to maximize the value of each renewable resource while optimizing project completion time. They did not consider non-renewable resources. Another paper [8] considered all types of resources. However, they focused on a single project. To encapsulate, no paper in our literature review has simultaneously considered multiple projects, non-renewable resources with supplier selection, and renewable resources based on skill assignment and applied them to a real case study.

Therefore, the main contribution of this paper is to provide a solution method to manage this problem by presenting the MILP model for planning level that optimizes schedules of multiple projects involving

multiple activities and renewable resources with multiple skills assignments applied to actual data. We also show the impact of multi-skills and advanced supply chain planning using a real-life case introduced in the introduction of the building renovation sector. A limited number of papers have explored resource constraint scheduling problems, is shown in Table 1.

Table 1. CSC Optimization Papers Review

Reference	Multi-project	Multi-Skill	SCM	Real Case	Objective (s)
[8]		•	•	•	T, C
[9]			•		T
[6]	•	•			T, C
[10]			•		C, P, R
[11]				•	T, C
[7]	•	•		•	T
[12]				•	P
[13]	•				T
[14]	•			•	T, C
<u>This Work</u>	•	•	•	•	T

SCM: Supply chain management; T: Time; C: Cost; P: Profit; R: Risk

2 Problem Description & Assumptions

This paper describes the problem as a variant of RCMPSP. The problem consists of a set of projects, each of which is made up of a subset of jobs or activities. The construction supply chain's key players are the clients, suppliers, and contractors. Once the contractor wins a project bid from the client, he must complete several jobs or activities. The contractor has a fixed number of shared renewable resources for all concurrent projects as shown in Figure 1.

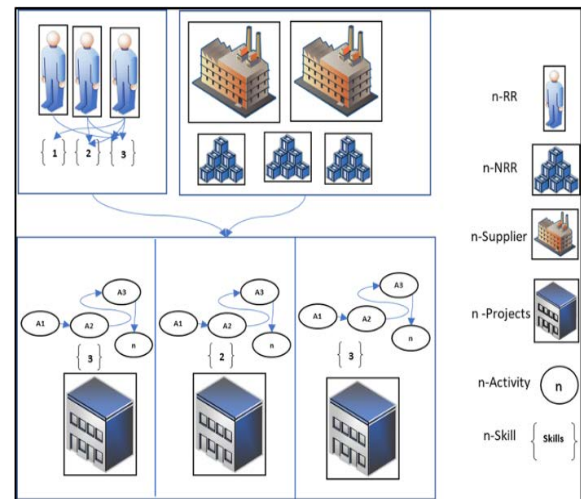


Figure 1. Construction Supply Chain

Each project has a specific release date, the time from which the jobs can start, and a desired date representing the time the project should be completed. Each job within

a project has a processing time and a quantity required for each type of resource. A "Team Leader" and a "Manager" (renewable resources) are required to complete a project. All project activities have established end-start precedence relationships, which means that an activity can only begin after a set of predecessor activities have been completed. A feasible schedule for this problem involves assigning start times to each activity within each project while respecting the precedence relationships between jobs and the availability of resources at each period.

2.1 Model Definitions and Assumptions

To achieve coordination of the construction supply chain, integrated advanced planning and scheduling, including supplier selection and resource assignment optimization, should be solved. To achieve this objective, we model the coordinated operations of the construction supply chain in Figure 1, capturing critical components of the real-life construction project described in Section 1. The supply chain network consists of a set of projects denoted by P . For each project, we consider an independent set of activities across projects. For each project, we must procure K non-renewable resources to each project P by supplier U . Also, we need to assign renewable resources L team leads and M managers to each project based on skill set S requirements. The network diagram is shown in Figure 1.

The model is implemented in the case company, and the model is developed as per their requirements. These requirements are considered assumptions for the model's generality, such as renewable resources (managers and team leads) are assigned to projects, not activities. This work was carried out by the company that assigns these resources to projects. The second assumption is that all non-renewable resources required for the project will arrive before the start of the project's first activity, as in the case of this company. The third assumption is that a team lead can only be assigned one project at a time. The fourth assumption is that once the project is started, it will continue without interruptions from other projects until it is finished. The fifth assumption is that once the team leads and managers are assigned to any project, they will remain assigned until it is completed.

2.2 Mathematical Model Details

We introduce the following notations to formulate the MILP mathematical for problem \mathcal{P} . The details of the model are given below.

2.2.1 Sets

- P : set of projects.
- $A(p), p \in P$: set of activities to be completed in project p .

- S : set of skills.
- U : set of Suppliers for non-renewable resources.
- L : set of team leaders.
- M : set of managers.
- K : set of Non-renewable resources.
- T : set of time periods.

2.2.2 Parameters

The parameters considered for this problem are the data we receive from the case company, which includes scheduling-related data, cost-related data, supplier and non-renewable resources data, team lead and manager skills data, and suppliers' shipping and release time data. The shipping time depends upon the supplier and project location. In contrast, supplier release time depends upon the supplier and type of resource. Different resources require different preparation times once the order is received, which we have described in the parameters subsection.

- $Dur_{pa}, p \in P, a \in A$: Duration of Activity a of Project p
- $Delay_p, p \in P$: Anticipated delay in Project p
- $DS_p, p \in P$: Due starting date of Project p
- $FS_p, p \in P$: Due finishing date of Project p
- $Q_{pk}, p \in P, k \in K$: Non-renewable resource units of k required by Project p
- $Cap_{uk}, u \in U, k \in K$: Capacity of providing Non-renewable resources of k units by Supplier u
- $C_{uk}, u \in U, k \in K$: Cost of Non-renewable resource of k per unit required by Supplier u
- $Req_{ps}, p \in P, s \in S$: Skills required by Project p
- $Avl_{ls}, l \in L, s \in S$: Skills possessed by Team Leader l
- $Mag_{ms}, m \in M, s \in S$: Skills possessed by Manager m
- $Tran_{lp}, l \in L, p \in P$: Distance between Team Leader l and Project p 's location in km
- $MSal_m, m \in M$: Salary of Manager m in CAD (\$)
- $LSal_l, l \in L$: Salary of Team Leader l in CAD (\$)

- $Bud_p, p \in P$: Budget of Project p
- $Rel_{uk}, u \in U, k \in K$: Release time of non-renewable resource k by Supplier u
- $Shp_{up}, u \in U, p \in P$: Shipping time by Supplier u for Project p
- M = The Big M.

2.2.3 Decision Variables

Since the problem we tackle is scheduling and assignment of multiple resources based on multiple skills to multiple projects, the kind of decision variables formed is relevant to the problem; therefore, two decision variables are for non-renewable resource assignment to a project by a supplier, four decision variables are for selection for team lead and manager to project at any period while two decision variables are formed for the scheduling optimization purpose starting time and completion time of project.

- W_{puk} : binary variable, equal one if a supplier $u \in U$ is selected for the project $p \in P$ to provide non-renewable resource $k \in K$
- Z_{puk} : quantity of non-renewable resources $k \in K$ supplied by supplier $u \in U$ to Project $p \in P$
- ST_{pa} : Starting time of Activity, $a \in A$ of Project $p \in P$
- CT_p : Completion time of Project $p \in P$
- Y_{pl} : binary variable, equal 1, if Team Leader $l \in L$ is assigned to Project $p \in P$ (Binary variable)
- X_{plt} : binary variable, equal 1, if Team Leader $l \in L$ is assigned to Project $p \in P$ at Time Period $t \in T$
- R_{pm} : binary variable, equal 1, if Manager $m \in M$ is assigned to Project $p \in P$ (Binary variable)
- E_{pmt} : binary variable, equal 1 if Manager $m \in M$ is assigned to Project $p \in P$ at Time Period $t \in T$

2.2.4 Objective Function and Constraints

The objective of this problem is to minimize the total "completion time" of all the projects simultaneously based on the optimized assignment of resources, as shown in the following equation.

$$\min \sum_{p \in P} ((\sum_{l \in L} \sum_{t \in T} X_{plt} \times t) + C_p) \quad (1)$$

The constraints of the model, which can be divided into various categories, are as follows.

a) Assignment Constraints

$$\sum_{l \in L} Y_{pl} = 1 \forall p \in P \quad (2)$$

$$\sum_{m \in M} R_{pm} = 1 \forall p \in P \quad (3)$$

$$Re q_{ps} - Av l_{ls} \leq 2 \times (1 - Y_{pl}) l \in L, s \in S, p \in P \quad (4)$$

$$Re q_{ps} - Mag_{ms} \leq 2 \times (1 - Y_{pl}) m \in M, s \in S, p \in P \quad (5)$$

$$\sum_{t \in T} X_{plt} = Y_{pl} \forall l \in L, p \in P \quad (6)$$

$$\sum_{t \in T} E_{pmt} = R_{pm} \forall m \in M, p \in P \quad (7)$$

$$\sum_{l \in L} X_{plt} \times t = \sum_{m \in M} E_{pmt} \times t \forall p \in P, t \in T \quad (8)$$

$$\sum_{t \in T} E_{pmt} \leq 5 \forall m \in M, p \in P \quad (9)$$

b) Budget Constraint

$$\sum_{m \in M} \sum_{l \in L} (Y_{pl} \times LSal_l + R_{pm} \times MSal_m + Y_{pl} \times Tran_{lp}) \leq Bud_p \quad (10)$$

$$p \in P$$

c) Scheduling Constraints

$$\sum_{l \in L} \sum_{t \in T} (X_{plt} \times t) \geq DS_p \quad \forall p \in P \quad (11)$$

$$\sum_{l \in L} \sum_{t \in T} (X_{plt} \times t) \geq Delay_p \quad \forall p \in P \quad (12)$$

$$\sum_{l \in L} \sum_{t \in T} (X_{plt} \times t) + C_p \leq FS_p \quad \forall p \in P \quad (13)$$

$$\sum_{l \in L} \sum_{t \in T} (X_{plt} \times t) \leq ST_{pa} \quad \forall p \in P, a \in A, \quad (14)$$

$$ST_{pa} + Dur_{pa} \leq C_p \quad \forall a \in A, p \in P \quad (15)$$

$$ST_{pi} + Dur_{pi} \leq ST_{pj} \quad \forall i \in A, j \in A, p \in P, i \neq j \quad (16)$$

d) Non-renewable Resources Constraint

$$\sum_{p \in P} Z_{puk} \leq Cap_{uk} \quad \forall u \in U, k \in K \quad (17)$$

$$\sum_{u \in U} Z_{puk} = Q_{pk} \quad \forall p \in P, k \in K \quad (18)$$

$$\sum_{k \in K} \sum_{u \in U} (Z_{puk} \times C_{uk}) \leq Bud_p \quad \forall p \in P \quad (19)$$

$$Z_{puk} \leq M \times W_{puk} \quad u \in U, p \in P, k \in K \quad (20)$$

$$\sum_{u \in U} ((Rel_{uk} + Shp_{up}) \times W_{puk}) + 1 \leq \sum_{l \in L} \sum_{t \in T} (X_{plt} \times t) \quad \forall p \in P, k \in K \quad (21)$$

Equation 2 ensures that each project is assigned only one team lead. Equation 3 ensures the same for the manager. Equations 4 and 5 make the model force assign the project to the team lead and the manager with the required skills. Equations 6 and 7 limit the model to select any other team lead and manager except for the ones assigned to it over any time for any project. Equation 8 will ensure that the project's starting time for the assigned team lead and manager should be the same.

While a team lead cannot be assigned to more than one project at any period, a manager can assign a maximum of 5 projects at any time; equation 9 is for this purpose. Constraint 10 in Equation 10 shows each project's budget limit. Since the case company specializes in delivering renovation projects to its customers, it needs to ensure that each project is completed within the allocated budget to maintain profitability and customer satisfaction. Equation 11 will ensure the project does not start before its due date. Equation 12 is for the anticipated project delay, and it will make sure that the project starts after its delay period. Equation 13 is for the projects' finishing due dates. So, it will ensure the project is completed within its due date.

The scheduling decision variable ST should be greater or equal to the starting time of the project; constraint 13 in equation 14 will serve this purpose. Equation 15 is that the completion time of each project should be greater than the starting time of all activities and the duration of its project. Equation 16 is for establishing precedence, and it is only activated on a phased approach. Ultimately, equations from 17 to 19 are assigned to each project and its supplier selection for non-renewable resources. Equation 20 triggers a significant M constraint, which will be activated for suppliers selected to supply non-renewable resources. The last equation, 21, is to ensure that the project will start once all the non-renewable resources necessary are received for that project, incorporating the starting time of the project to be greater than the release time and arrival time of resources on the project location.

3 Case Study

As mentioned, the model is implemented on a real construction company based in Canada. For experimentation purposes, we selected the case of 10 projects, having eight activities in each project with different precedence relationships. The company data is described below. The skills (S) that team leaders (TL) possess are shown in Table 2, while the skills required by the ten projects (P) are shown in Table 3.

Table 2. Team leader's skill availability

Skills / Team Leaders	TL 1	TL 2	TL 3	TL 4	TL 5	TL 6	TL 7	TL 8
S1 Digging	•	•			•			•
S2 Complex piles	•				•			•
S3 Decontamination	•				•			•
S4 Structure	•			•	•			•
S5 Interior drain	•	•	•	•	•			

Table 3. Project skill requirement

Skills / Projets	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
S1	•				•					•
S2							•	•		
S3									•	
S4				•		•				
S5		•	•							

3.1 Initial experimentation

The model is implemented in Pyomo, a Python-based open-source optimization modeling language, and we use the solver Gurobi solver. The solution method used is the branch and bound, which is an exact solution method. For this paper, we tested this model on three instances: 3 projects, 5 projects, and 10 projects. For 10 project instances, we received a solution in around 1 minute, which is a promising result given the fact that it is a type of RCPSP problem. The summary of the first results and their computational time is shown in Table 4.

Table 4. Initial Results

No of Projects	Total Completion Time (Days)	Projects Finishing Time (Days)	Solution Time (Secs)
3	72	37	20,4
5	125	37	35,4
10	313	53	64,8

Upon further analysing the 10 projects, we see the finishing time of the last project, which, in this instance, is 53 days, as can be seen in the Gantt chart of Figure 2. The Gantt chart shows that managers, such as M, and team leaders, such as TL, are assigned to projects. This can be considered aggregate-level planning because we assign projects and track the start, duration, and finish times of these projects.

With this model, we can also check the same for each project's activities. For example, the activities Gantt chart of Project 6th is shown in Figure 3. The sixth project starts on the first day and finishes on the 15th; during this time, all eight activities are executed according to their precedence relationship, as shown in Figure 3. In the next section, we analyze the impact of skill change and the number of team leaders on the total completion time of the 10 projects.

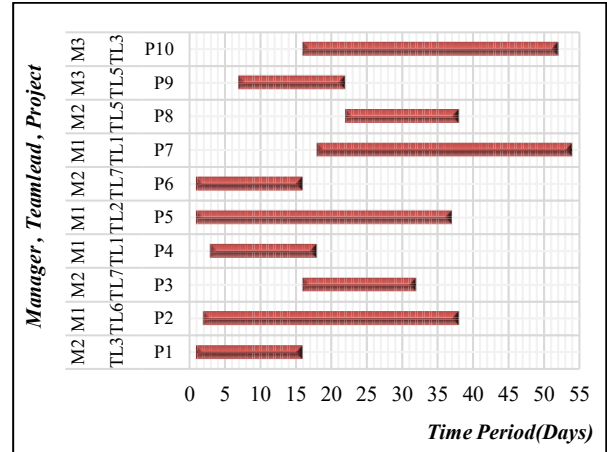


Figure 2. Aggregated Gantt Chart 10 Projects

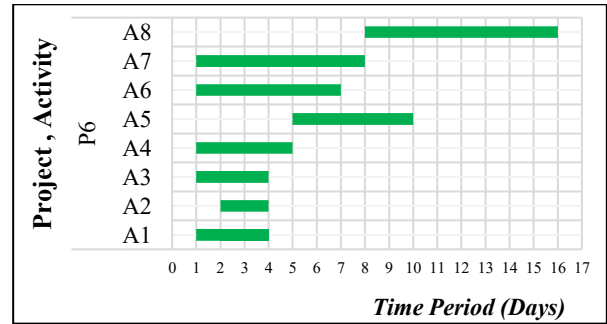


Figure 3. Timeline Gantt Chart of Project 6 with Activities

3.2 Impact of team lead and skills on project scheduling and procurement

For this purpose, an analysis is done concerning the number of team leaders and their skills. How does the availability of a skillful team lead affect project completion time? We test 7 hypothetical scenarios (Sc) on the same 10 project instances described below and compare them with the initial result of Table 3.

- Sc1: Result in Table 3 for 10 projects with 8 TL.
- Sc2: With 4 TL (50% of staff).
- Sc3: With 2 TL (25% of staff).
- Sc4: With 4 TL and all trained-on skill S1.
- Sc5: With 4 TL and all trained-on skill S2.
- Sc6: With 8 TL and all trained-on skill S1.
- Sc7: With 8 TL and all trained-on skill S2.
- Sc8: With 8 TL and all trained-on skills S1 & S2.

In evaluating the resource plan for 10 projects, we found that 8 capable team leaders available. However, it is sometimes difficult to keep this level of staffing (safety reasons, COVID-19 restrictions, disruptions), so we explored the effects of reducing our staff by 50% and 25% on project completion times. Unfortunately, we saw a

significant increase in total completion time for all 10 projects. For example, under Scenario 2, the total project completion time was 406 days, an increase of 93 days from Scenario 1. Under Scenario 3, the completion time rose to 582 days, an increase of 269 days. Ultimately, there were more practical options for improving productivity than downsizing, as it had a negative impact on completion time. This is unacceptable for some customers and would damage the company's reputation. To address this, we tested Scenarios 4 and 5, which focused on providing skill training to staff on the critical skills needed for each project.

To ensure project success, we identified vital skills S1 and S2 and provided training to team leaders lacking

these abilities. Most projects require these skills, making them crucial for the team. After the training, we evaluated the impact of S1 and S2 on project completion time in scenarios S4 and S5. Our findings revealed that training on skill S1 (digging) did not significantly affect project completion time, while training on skill S2 (Complex piles) reduced the time by 36 days, resulting in a completion time of 370 days compared to Sc 2. However, compared to Sc1, solution Sc 5 still had a 57-day increase in completion time. To further test our results, we conducted the same training with 8 team leaders, resulting in scenarios Sc 6, Sc 7, and Sc 8, all of which yielded an optimum completion time of 284 days, a 29-day reduction from Sc1.

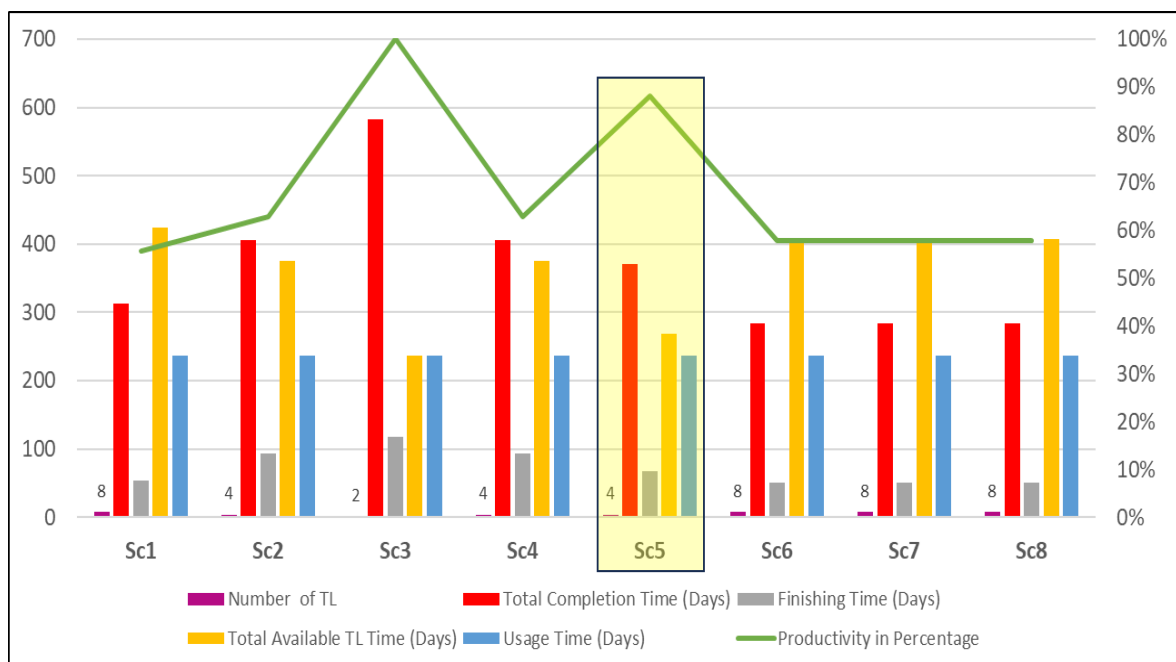


Figure 4. Comparison of scenarios solution

The sensitivity analysis as shown in Figure 4 reveals potential time and cost savings in the current model. Prioritizing customer satisfaction and project completion time suggests keeping the entire staff despite lower productivity (55%) and increased idle time for team leaders. Alternatively, downsizing may be preferable for cost reduction or higher team leader productivity. Scenario 5, with 88% productivity and a total completion time of 370 days, appears as the most optimal solution. Therefore, S2 is identified as the most critical skill requiring training.

Based on this analysis, several factors must be considered when deciding on the “ideal solution” to such challenges. These factors include the cost of keeping and training staff, the duration it takes to complete a project, the availability of staff, and the company's goals. Luckily, our model can manage these complexities to support more thoughtful decision-making.

4 Conclusion

Managing concurrent construction projects is complex and involves both external and internal factors. Externally, it consists of obtaining permits and coordinating with suppliers and clients. Internally, it involves resource management, addressing diverse project needs, skills allocation, and meeting project deadlines.

The proposed mathematical model can be used at the tactical level to estimate multiple project schedules at their operational level. The second novelty is the MILP model, which can optimize and assign multiple project schedules, resource assignments, and supplier selection, thus integrating supplier information into the model. Although the full model is presented in this paper, only the effects of skill and team leaders on project completion time are analyzed. The third contribution is to address the

research gap highlighted by the study of [15], which shows an urgent need for master scheduling at the planning level that integrates resources needed at the operational level. This research emphasizes the significance of effective collaboration between different supply chain stakeholders through information sharing. This approach can lead to more accurate project scheduling, ensuring better resource planning and coordination for a more efficient construction supply chain. Dealing with resource-constrained multi-project reactive scheduling problems amid uncertainties is critical. Disruptive events, such as a new project, changes in resource availability, or transportation delays, can cause uncertainty. The process of rescheduling the detailed plan to obtain an updated schedule is a challenging task and requires further research.

5 References

- [1] Le, P.L., et al., Present focuses and future directions of decision-making in construction supply chain management: a systematic review. *International Journal of Construction Management*, 2020. **20**(5): p. 490-509.
- [2] Dike, S.H., Project scheduling with resource constraints. *IEEE Transactions on Engineering Management*, 1964(4): p. 155-157.
- [3] Gonçalves, J.F., J.J. Mendes, and M.G. Resende, A genetic algorithm for the resource constrained multi-project scheduling problem. *European journal of operational research*, 2008. **189**(3): p. 1171-1190.
- [4] García-Nieves, J., et al., The multimode resource - constrained project scheduling problem for repetitive activities in construction projects. *Computer - Aided Civil and Infrastructure Engineering*, 2018. **33**(8): p. 655-671.
- [5] Kumanan, S., G. Jegan Jose, and K. Raja, Multi-project scheduling using an heuristic and a genetic algorithm. *The International Journal of Advanced Manufacturing Technology*, 2006. **31**: p. 360-366.
- [6] Hosseini, A.H. and V. Baradaran, A two-phase approach for solving the multi-skill resource-constrained multi-project scheduling problem: a case study in construction industry. *Engineering, Construction and Architectural Management*, 2023. **30**(1): p. 321-363.
- [7] Li, Q., et al., The Integrated Problem of Construction Project Scheduling and Multiskilled Staff Assignment with Learning Effect. *Journal of Construction Engineering and Management*, 2023. **149**(8): p. 04023064.
- [8] Mirnezami, S.-A., et al., An integrated chance-constrained stochastic model for a preemptive multi-skilled multi-mode resource-constrained project scheduling problem: A case study of building a sports center. *Engineering Applications of Artificial Intelligence*, 2023. **126**: p. 106726.
- [9] Xie, L.-L., et al., Knowledge extraction for solving resource-constrained project scheduling problem through decision tree. *Engineering, Construction and Architectural Management*, 2023.
- [10] Ghorqi, M., et al., An Integrated Model for Multi-Mode Resource-Constrained Multi-Project Scheduling Problems Considering Supply Management with Sustainable Approach in the Construction Industry under Uncertainty Using Evidence Theory and Optimization Algorithms. *Buildings*, 2023. **13**(8): p. 2023.
- [11] Chauhan, M. and R. Kumar, Integrating the multi-objective particle swarm optimization-based time-cost trade-off model with earned value management. *Asian Journal of Civil Engineering*, 2023: p. 1-11.
- [12] Wang, J., H. Liu, and Z. Wang, Stochastic Project Scheduling Optimization for Multi-stage Prefabricated Building Construction with Reliability Application. *KSCE Journal of Civil Engineering*, 2023: p. 1-16.
- [13] Srimathi, K., A. Padmarekha, and K. Anandh, Automated construction schedule optimisation using genetic algorithm. *Asian Journal of Civil Engineering*, 2023: p. 1-8.
- [14] Ghorqi, M., et al., Integration of resource supply management and scheduling of construction projects using multi-objective whale optimization algorithm and NSGA-II. *Soft Computing*, 2024.
- [15] Zaalouk, A., S. Moon, and S. Han, Operations planning and scheduling in off-site construction supply chain management: Scope definition and future directions. *Automation in Construction*, 2023. **153**: p. 104952.

Integrating the AR-QR Code Approach with Positioning and Orientation Sensors to Facilitate Deploying Drawings Information on Job Sites

Mohsen Foroughi Sabzevar¹, Masoud Gheisari ², James Lo³

¹School of Construction and Design, University of Southern Mississippi, Mississippi, USA

²Rinker School of Construction Management, University of Florida, Florida, USA

³ Department of Civil, Architectural & Environmental Engineering, Drexel University, Pennsylvania, USA
mohsen.foroughisabzevar@usm.edu , masoud@ufl.edu , james.lo@drexel.edu

Abstract –

While general contractors often use tablets to review design and construction information, subcontractors typically rely on paper-based methods. Each of these methods has its own set of advantages and challenges. The AR-QR code approach aims to bridge the gap between paper-based and digital methods, leveraging the power of Augmented Reality (AR) to maintain the benefits of the paper-based approach while circumventing its drawbacks. This approach grants crews direct access to design and construction information by centralizing data and eliminating the need to refer to other drawings and documents. This research study addresses an issue inherent in traditional methods that the AR-QR code approach did not consider. In conventional practices, deploying drawings on-site necessitates that the crews mentally synchronize the drawing's coordination system with the actual job site. Additionally, they must mentally position and orient items to match their actual on-site locations. This process ultimately entails a mental visualization of the items on the job site to ascertain how the designed features correspond with the actual work. Such mentally demanding steps can exacerbate mental strain and increase the risk of errors. This study aims to propose a solution to alleviate this issue by enhancing the capabilities of the AR-QR code approach to facilitate visualization of the drawing's details on the jobsites. The research methodology includes process mapping and workflow development. The outcome is a workflow that offers insights into how the AR-QR code approach could be integrated with a positioning and orientation system to achieve the goal of this study.

Keywords –

AR-QR Code; Sensor; Positioning and Orientation; Augmented Reality; Construction Site; Indoor Positioning

1 Introduction

In this paper, first, we present a condensed overview of a research study on the AR-QR code approach presented in [1] and then focus on a new adaption to enhance the functions of this approach in construction sites. Construction workers use large shop drawings to access design information. Floor plan drawings using symbols and codes refer crews to notes, sections, details, elevations, schedules, and other supplemental information on other sheets or documents [1]. Sometimes, following various symbols and codes for accessing additional information could be complex, leading to potential confusion and errors. Another issue is that these drawings do not include critical construction details like work procedures, materials quantity, equipment and tools, safety, and so on [1]. Specification manuals [2], serving as additional documents to drawings, generally lack detailed instructions for individual building components. These manuals broadly outline the quality requirements for products, materials, and workmanship. When documentation is insufficient, workers rely on verbal instructions, which increases the risk of ambiguity, errors, and the need for costly rework [3]. Avoiding mistakes in interpreting structural drawings is crucial because errors in construction can have severe safety consequences, such as building failures [4,5]. According to statistics [6–9], 70 to 90 percent of construction failures are due to human errors during execution rather than design, with most structural failures arising from construction errors rather than incorrect design [10]. These human errors, which occur during component fabrication or installation, are a significant source of uncertainty and can lead to structural instability during and after construction [11–13]. Poor design and poor design communication are two main factors that lead to errors [14]. Poor design refers to errors originating from the design itself, which are the designer's responsibility [14]. Poor design

communication involves the receiver's improper interpretation and decoding of the design message [15]. Poor design communication can happen due to a lack of supplementary content, inappropriate format, or choosing the wrong channels for conveying design and construction information. Poor design communication can lengthen the process of accessing design and construction information, thereby increasing construction crews' physical and mental workload and raising the risk of human errors [16–21]. Errors can cause safety issues, rework, productivity problems, environmental concerns, and the risk of disputes [3,22–24]. To enhance the accessibility of design and construction information, using Building Information Modeling (BIM) models with a high level of detail (LOD) could be a solution [1]. These models can combine all necessary details, potentially streamlining the information retrieval process [1]. However, significant barriers could be involved such as high cost of developing comprehensive BIM models, screen size limitations of mobile devices (i.e., the complex details of BIM models are intricate to review on the small screens of tablets and smartphones), and construction site workstation challenges (e.g., the use of digital 2D drawings typically requires PCs or workstations, which are cumbersome to move, especially in the early stages of construction where issues such as lack of electrical power, the absence of a sheltered location, and interference with material handling paths can arise) [1,25]. Despite these issues, there are practical reasons for the continued use of 2D paper-based drawings on construction sites: legal approval (e.g., they are often the only legal documents for building construction) [26], large size (i.e., a broader 2D view at once, no need for scrolling up and down on plan views), convenience and resilience (i.e., paper drawings are easier to carry around the typically harsh conditions of a construction site and do not depend on electricity), and they are simple to use, maintain, and can be easily replaced if damaged.

Due to these advantages, 2D drawings remain prevalent, especially among subcontractors' crews, despite the potential efficiencies offered by high-LOD BIM models and digital alternatives. According to site visits [27], general contractor teams, including project engineers, construction managers, and superintendents, typically use tablets for reviewing 2D shop drawings, whereas subcontractor teams, such as foremen and workers, often utilize paper-based shop drawings for their work. To leverage the benefits of both paper-based and digital methods while minimizing their drawbacks, [1] proposed a hybrid approach known as the AR-QR code approach. “The AR-QR code is an approach that uses the advantages of both paper-based and digital-based delivery methods (i.e., a combination approach) but avoids their disadvantages” [1]. “In this way, the

advantages of paper-based 2D plan views (e.g., large size, simply carried, less care needed, etc.) are retained, while the disadvantages (e.g., difficulty in accessing different pieces of information that are not centralized, reliant on information in other sheets and documents, dependency on verbal instructions to transfer design intent and related construction requirements to crews, divided attention between different sheets and documents, facing unrelated information) are avoided” [1].

1.1 Issues Regarding Paper-based Procedures to Access Design and Construction Information

In the traditional approach, drawings frequently require cross-referencing with additional sheets and documents [15,27,28]. For example, if a crew needs to access information regarding a symbol on a plan drawing, they need to read the general notes on the footprint. Sometimes, these notes refer to another sheet, which may reference a detail on a subsequent detail sheet (Figure 1-A, Issue 1). This detail may reference sub-details on other sheets (Figure 1-B, Issue 2). At times, this process can be frustrating. Ultimately, the crew must mentally synthesize all these pieces to understand the complete geometry of the component (Figure 1-C, Issue 3). Concurrently, they must determine the orientation of the final item in relation to the plan drawing (Figure 1-D, Issue 4). Furthermore, construction information is not included on the drawing sheets in the paper-based approach (Figure 1-E, Issue 5). To obtain this information, crews must consult written instructions such as specifications, submittals, and other documents [1]. However, specifications often contain general information not tailored to each building element [1]. Moreover, information is not centralized in these documents and often refers to other documents [1]. Crews must cross through a substantial amount of documentation to find the relevant details, which can lead to scattered attention and a heightened risk of errors [1]. Verbal instructions in construction, usually provided by foremen or superintendents, can result in misunderstandings and inefficiencies (Figure 1-F, Issue 6). Crew members are expected to remember and act on these instructions while working, but this approach can be ineffective, posing potential safety risks, necessitating rework, and ultimately reducing productivity due to ambiguities [3]. To solve these issues, [1] prototyped a new approach (i.e., AR-QR code), and evaluated. The evaluation results showed that the AR-QR Code approach has the potential to improve crew access to design and construction information. The AR-QR code approach allows crews direct access to design and construction information by centralizing information and eliminating the need to refer to other drawings and documents [1].

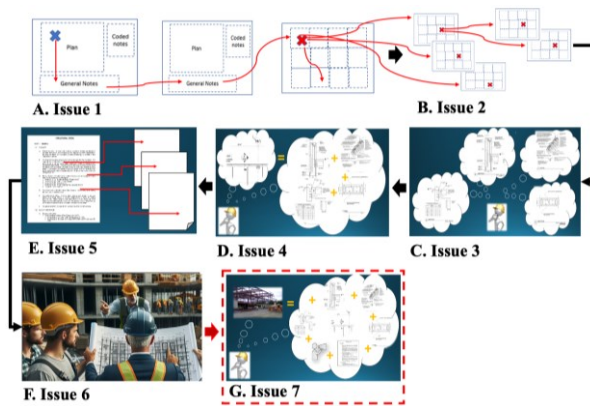


Figure 1. Paper-based procedure to access design and construction information (Adapted from [1])

1.2 Problem Statement

This study addresses another problem associated with the traditional method (Figure 1-G, Issue 7) and suggests that enhancing the AR-QR code approach could offer a potential solution. The problem is that in conventional practices, the deployment of drawings on-site requires the crew to mentally synchronize the drawing's coordination system with the job site. Furthermore, they need to mentally position and orient items to correspond with their actual locations on-site. Ultimately, this procedure involves a mental visualization of the items on the job site to compare the features with the actual work. These mentally intensive steps could increase mental strain and the likelihood of errors. Such challenges might occur prior to the fabrication of components when workers attempt to envision the final work in the field, or after completion when they need to evaluate how the actual work aligns with the design. For both scenarios, merely looking at paper-based drawings to visualize the final work on the job site can be challenging. Even if a 3D model of the element is available, there would still be a kind of non-connectivity between this element on the screen and the actual field environment.

1.3 Aim and Objectives

This study aims to propose a solution to alleviate the cognitive-spatial synchronization issue by enhancing the capabilities of the AR-QR code approach to facilitate visualization of the drawing's details on the job sites. To accomplish this aim, the following objectives were defined: 1) Identify the traditional process for deploying drawing information on the job site; 2) Propose an approach based on tracking systems to enable crews to navigate to the precise position and orientation of the drawing's details on the jobsites at the construction site.

2 Background Information for Method Development

2.1 Communication Elements in Construction

To enhance the AR-QR code approach, it is necessary to identify and incorporate new communication elements into this method. This section reviews the characteristics of various communication elements. The communication elements in construction include messages or information content, senders who encode messages, receivers who decode messages, media or formats of information, channels to transmit messages [29–31] and channel links if more than one channel type is used. Sender and Recipient [28]: In construction, senders may include architects or designers, while recipients could be contractors or craftsmen. Senders encode and transmit messages, and recipients decode them. Media [28]: Media refers to the format of information, which can be verbal or non-verbal, including speech, drawings, and written documents. The choice of media depends on the context and the parties involved. Channels [32,33]: Channels are conduits like air, paper, or electronic devices that transmit content. Both physical and electronic channels are used, with electronic channels being devices like smartphones, tablets, VR/AR headsets, and physical channels being air or paper. Channel Link [1]: A channel link or tracker can be used to bridge a physical channel with a digital one. These links are classified into two types: those that are marker-based and those that are marker-less. Marker-based links can be printed and attached to surfaces [34]. Some examples of marker-based links are dot-based markers [35], QR code markers [36,37], circular markers [38], square markers [39], and alphabetic combination markers [39]. A fiducial marker like a QR code could effectively connect the information available in the electronic channel to the physical channel. Computer vision techniques can easily identify QR codes' distinguishable textures through scanning. A QR code can be created and attached next to the related element for each section, code, or notice related to each element in the plan drawings. This technique can give direct access to the required information in proper formats (e.g., image, 3D model, audio, video, text) for each section, code, or notice related to each element. The marker-less options are further divided into visible and non-visible categories. Visible links are environmental features like edges, corners, and specific points [40], which can be recognized and utilized by algorithms such as SIFT as substitutes for physical markers. In contrast, non-visible links are invisible signals, like sound waves, infrared light, or vibrations, which certain sensors can detect and employ in place of traditional markers. Some features of these sensors and related signals are listed as follows [41]:

GPS sensors are capable of detecting satellite signals to pinpoint a location in three dimensions. However, their functionality is restricted to outdoor settings, and their location accuracy is relatively imprecise [42,43]; Inertial sensors respond to motion and vibrations but suffer from cumulative errors over time, necessitating frequent recalibration [43]. Examples include gyroscopes, which measure orientation angles, and accelerometers, which measure both velocity and changes in motion direction [42]; Compasses consistently indicate the direction of magnetic north [42]; Wi-Fi technology communicates data using radio waves between a device and a network router, offering extensive indoor coverage and location accuracy within a range of 15 to 20 meters indoors [44]; Bluetooth sensors can accurately determine location within a space, achieving 75% accuracy for partial coverage and up to 98% with full coverage, provided the target devices remain stationary [45]; Ultrasonic sensors are affected by temperature variations, obstructions, ambient noise, and require significant infrastructure, offering limited update frequency [46]; Infrared sensors operate over short distances and are constrained by the necessity for a clear line of sight, as utilized in systems like Active Badge [47]; Radio Frequency sensors, operating on standards like IEEE 802.11 or WLAN, have a median location accuracy of 2 to 3 meters but suffer from scalability issues [48,49]; and Ultra-Wideband (UWB) Radio Frequency sensors utilize UWB signals that can penetrate through walls and offer high accuracy, albeit at a higher cost due to the required infrastructure [50,51]. A hybrid channel link combines both marker-based and marker-less systems to compensate for the limitations inherent in each method [52].

2.2 Benefits of the AR-QR Code Approach

In the AR-QR Code approach, crews access information by scanning QR codes associated with each element on the plan sheet. As depicted in Figure 2, this allows the crew to view design details, such as geometries, in augmented reality (AR), virtual reality (VR), and 2D views. AR and VR enable the manipulation of a 3D model from various angles, improving their understanding of the details. AR, in particular, merges the plan view with the 3D model, facilitating a combined review. Construction information is available through a custom user interface, allowing access to specific construction information such as work procedures, safety, material quantities, FAQs, verbal instructions (audio), equipment, tools, and sustainability considerations [1]. This method can significantly streamline the process of accessing design and construction information, thereby reducing time, errors, and complexity [1]. According to a survey conducted by [1], construction professionals believed that adopting the AR-QR code system could greatly mitigate issues arising from the use of paper-

based drawings, especially in terms of reducing queries from the workforce, improving productivity and safety, and addressing sustainability. They found the system to be broadly applicable, user-friendly to a degree, and quite beneficial. They also noted advantages such as better coordination and teamwork across different trades, decreased need for redoing work, more efficient handling of requests for information, less time spent on document review, and improved ease and clarity of information access. These points affirm the perceived merits of the AR-QR code system by those on the construction site.

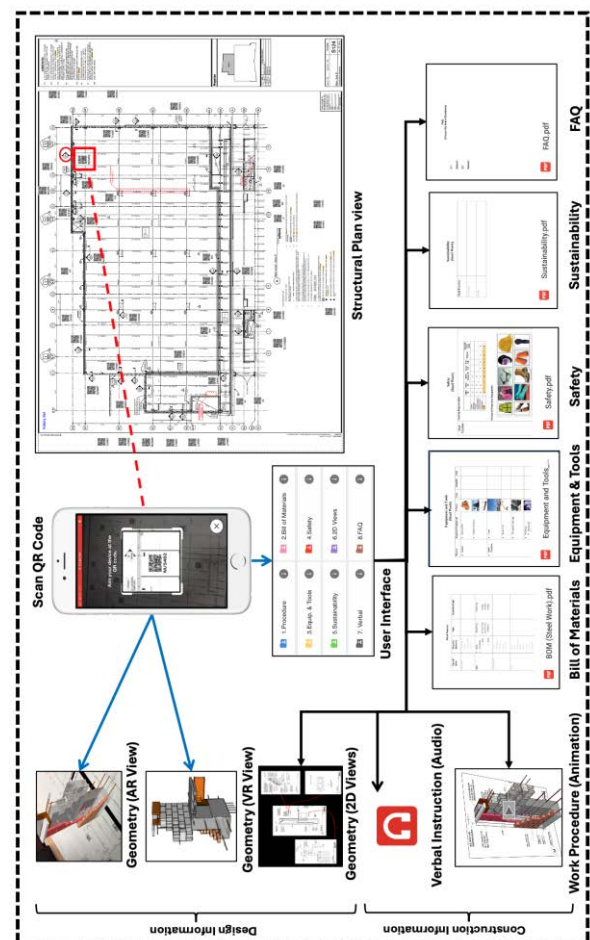


Figure 2. AR-QR Code approach [1]

3 Research Methodology

To fulfill the first objective, a process map was developed through observations and unstructured interviews conducted at five different construction sites in the USA. To meet the second objective, a thorough analysis of positioning and orientation systems was carried out to establish the necessary features. Subsequently, a workflow was developed to demonstrate how these systems could be integrated with AR-QR

codes, serving the specific goals of this study.

4 Adapting the Traditional Procedure of Deploying Drawings Information on Jobsites Using the AR-QR Code Approach with Positioning and Orientation Sensors

Figure 3 illustrates the traditional process for decoding design meaning in contribution with the construction field. In this procedure, the crew needs to match the drawing coordination system with the jobsite coordination system. Then, the crew needs to match the detail's position and orientation with the actual spot on the jobsite. In the end, the crew mentally visualizes the item on the jobsite to be able to compare the features of the designed work with the actual work. Since these steps should be conducted mentally, mental workload and error rates could increase. The suggested adaptation could be eliminating and replacing these steps with a straightforward method.

Suggested Adaptation: An adaptation could be using the AR-QR code approach, scanning the related QR code on the plan drawing, navigating the user to the correct spot on the jobsite, and superimposing the detail element information (e.g., 3D model) with the correct orientation for an accurate position. Figure 3 illustrates the steps that can be replaced with new ones.

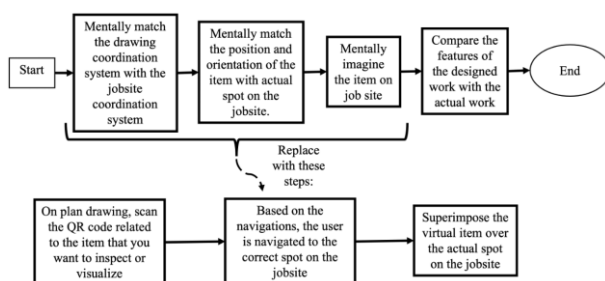


Figure 3. Traditional procedure for deploying drawing information on the job site

To implement this suggestion, the content will be the component geometry properties. The medium will be a 3D model of the items that the crews need to visualize on a jobsite. The physical channel will be a surface or spot on the jobsite where the information needs to be superimposed, the electronic channel's hardware will be a smartphone, and the electronic channel's software will be an AR browser. Among various tracking technologies, ultra-wideband (UWB) systems are considered the most precise for positioning, boasting accuracies up to ± 10 cm [50]. This level of precision significantly surpasses that of conventional systems utilizing Wi-Fi, Bluetooth, RFID, or GPS. According to Pozyx [50], a tag, and

multiple anchors (minimum of four anchors) are needed for the UWB system to function effectively. The tag interacts with the anchors via UWB radio signals capable of penetrating indoor barriers. These anchors serve as reference points, and the system calculates the tag's position by measuring the time-of-flight (TOF) of signals from the tag to each anchor, using the equation: Distance = time of flight \times speed of light, where the speed of light is 299,792,458 meters per second. Positioning is then determined through multilateration [51]. To ascertain three-dimensional orientation, the Pozyx tag is equipped with sensors to measure acceleration, magnetic fields, and angular velocity. Although each sensor has its own limitations, by integrating data from all sensors, the system can accurately determine 3D orientation [50]. Thus, the channel link will be built based on positioning sensors (Ultra-Wide Band wave generators and detectors) in integration with orientation sensors (accelerator, magnetometers, and gyroscope).

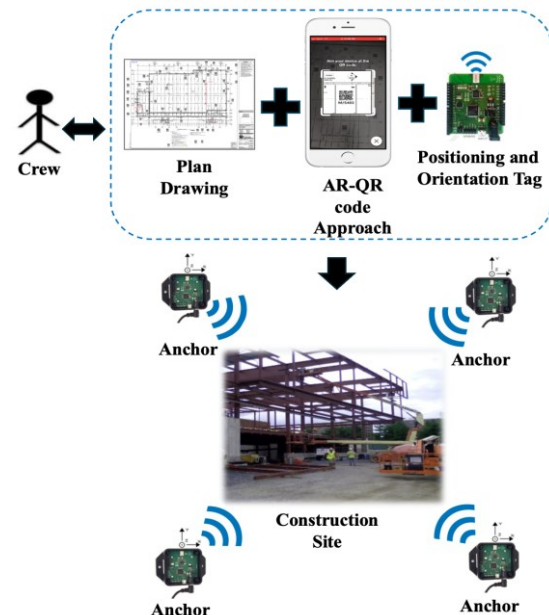


Figure 4. Integrating AR-QR code approach with positioning and orientation sensors

These sensors can help the user adjust the position and orientation of the smartphone where the 3D model and other related information need to be superimposed. To better understand how these 3D positioning and orientation systems can be integrated and implemented for the purpose of this study, a system architecture is shown in Figure 4. As shown, a crew needs to use the AR-QR code approach, scan a QR code related to a section cut on the plan drawing, access the related design and construction information, follow the tag (the tag needs to be attached to the smartphone or tablet by a USB cable to display the navigation signs) navigation signs to

navigate to the correct position and orientation on jobsite, and use the power of the AR to superimpose and review the information on the jobsite. For the positioning estimations, the tag communicates with a minimum of four anchors (i.e., reference points) through ultra-wideband RF signals. These anchors need to be installed on construction sites. For orientation estimations, three sensors are embedded in the tag: accelerometer, magnetometers, and gyroscope, which work together to detect acceleration, magnetic field, and angular velocity to estimate the tag orientation. In future studies, this workflow will be prototyped and tested to evaluate the precision and accuracy of this enhanced AR-QR code approach.

5 Conclusions

This research study focuses on enhancing the capabilities of the AR-QR code approach to facilitate deploying the drawing's information on the job sites. For this purpose, the traditional procedure for deploying drawing information on the job site was depicted, and an adaptation of the channel link of the AR-QR code method was suggested to modify the procedure. The outcome is a development workflow that shows how AR-QR code can be integrated with positioning and orientation sensors to enable crews to navigate to the precise position and orientation of the drawing's details on a jobsite and deploy and visualize drawing information. In future studies, this workflow will be prototyped and tested to evaluate the precision and accuracy of the enhanced AR-QR code system and if this method can reduce the mental workload and errors for workers on construction sites.

Acknowledgments

This research was funded by the National Science Foundation (NSF), award number 1562515, and supported by University of Southern Mississippi.

References

- [1] M. Foroughi Sabzevar, M. Gheisari, L.J. Lo, AR-QR code for improving crew access to design and construction information, *Autom Constr.* 154 (2023) 105017. <https://doi.org/https://doi.org/10.1016/j.autcon.2023.105017>.
- [2] C. on S. of the C. Institute, Preparation of Construction Specifications for Civil Projects, in: American Society of Civil Engineers, 2013. <http://ndl.ethernet.edu.et/bitstream/123456789/64531/1/1348.pdf> (accessed June 13, 2022).
- [3] C. Mourgues, M. Fischer, A work instruction template for cast-in-place concrete construction laborers, 2008. <https://cdn-prod-pdfsimplified-content.azureedge.net/pdfseoforms/pdf-20180219t134432z-001/pdf/construction-working-instruction-template.pdf> (accessed June 12, 2023).
- [4] E. Frühwald, S. Thelandersson, Design of safe timber structures - How can we learn from structural failures in concrete, steel and timber?, 10th World Conference on Timber Engineering 2008. 4 (2008) 1962–1969. <http://www.scopus.com/inward/record.url?eid=2-s2.0-84865730217&partnerID=40&md5=769b4ac9cb95988f5e42a97a96ad8dcf>.
- [5] OSHA, Investigation of the October 10, 2012 Parking Garage Collapse During Construction at Miami Dade College, Doral, FL, (2013) 83. https://www.osha.gov/doc/engineering/2013/2013_r_02.pdf.
- [6] A.S. Nowak, K.R. Collins, Reliability of structures (2nd ed.), CRC press, 2012. <https://doi.org/10.1201/b12913>.
- [7] F. Palmisano, R.T. Ratay, The practice of forensic structural engineering in IABSE member countries: review of a survey, in: IABSE Conference: Structural Engineering: Providing Solutions to Global Challenges, Geneva, Switzerland, September 2015, 2015: pp. 770–773. <https://doi.org/10.2749/222137815818357935>.
- [8] J. Schneider, I.A. for B. and S. Engineering, Introduction to Safety and Reliability of Structures, International Association for Bridge and Structural Engineering, 2006. <https://books.google.com/books?id=sHZaswEACAAJ>.
- [9] T. Vrouwenvelder, M. Holicky, M. Sykora, Modelling of human error, in: Joint Workshop of COST Actions, Ljubljana, Slovenia, 2009: pp. 21–22. https://www.cost-tu0601.ethz.ch/Documents/Meetings/Ljubljana/FACTSHEETS/vrouwenvelder_2.pdf (accessed June 13, 2023).
- [10] G. Alpsten, Causes of Structural Failures with Steel Structures, International Association for Bridge and Structural Engineers. (2017) pp.100–108. <https://doi.org/10.2749/helsinki.2017.100>.
- [11] G. Alpsten, Uncertainties and Human Errors in the Design and Execution of Steel Structures, IABSE Workshop 2017: Ignorance, Uncertainty and Human Errors in Structural Engineering, February 15–16, 2017, Helsinki, Finland. (2017) pp.1–9. <https://doi.org/10.2749/helsinki.2017.077>.
- [12] G. Reichart, How to reduce design and construction errors, Nuclear Engineering and

- Design. 110 (1988) 251–254. [https://doi.org/10.1016/0029-5493\(88\)90030-1](https://doi.org/10.1016/0029-5493(88)90030-1).
- [13] I. Tylek, K. Kuchta, A. Rawska-Skotniczny, Human errors in the design and execution of steel structures—A case study, *Structural Engineering International*. 27 (2017) 370–379. <https://doi.org/10.2749/101686617x14881937385287>.
- [14] C. Eckert, J.-F. Boujut, The role of objects in design co-operation: communication through physical or virtual objects, *Computer Supported Cooperative Work*. 12 (2003) 145–151. <https://doi.org/10.1023/a:1023954726209>.
- [15] G.B. Dadi, P.M. Goodrum, T.R.B. Taylor, W.F. Maloney, Effectiveness of communication of spatial engineering information through 3D CAD and 3D printed models, *Visualization in Engineering*. 2 (2014) 1–12. <https://doi.org/10.1186/s40327-014-0009-8>.
- [16] A. Basahel, M. Young, M. Ajovalasit, The Significance of Verbal and Spatial Attentional Resources on Mental Workload and Performance, in: *CiteSeer*, 2010. <https://citeseerx.ist.psu.edu/pdf/79e9880a3c83cf57e2563ea151c3f34a1cb3c295> (accessed June 13, 2023).
- [17] G. Matthews, D.R. Davies, R.B. Stammers, S.J. Westerman, Human performance: Cognition, stress, and individual differences, *Psychology Press*, 2000. <https://doi.org/10.4324/9781315812809>.
- [18] S.-L. Hwang, Y.-J. Yau, Y.-T. Lin, J.H. Chen, T.-H. Huang, T.-C. Yenn, C.-C. Hsu, A mental workload predictor model for the design of pre alarm systems, in: *International Conference on Engineering Psychology and Cognitive Ergonomics*, Springer, 2007: pp. 316–323. https://doi.org/10.1007/978-3-540-73331-7_34.
- [19] J. Sweany, P. Goodrum, J. Miller, Analysis of empirical data on the effects of the format of engineering deliverables on craft performance, *Autom Constr.* 69 (2016) 59–67. <https://doi.org/10.1016/j.autcon.2016.05.017>.
- [20] J.C. Weeks, L. Hasher, Divided attention reduces resistance to distraction at encoding but not retrieval, *Psychon Bull Rev.* 24 (2017) 1268–1273. <https://doi.org/10.3758/s13423-016-1210-7>.
- [21] S.E. McDowell, H.S. Ferner, R.E. Ferner, The pathophysiology of medication errors: how and where they arise, *Br J Clin Pharmacol.* 67 (2009) 605–613. <https://doi.org/10.1111/j.1365-2125.2009.03416.x>.
- [22] C. Mourgues, M. Fischer, D. Hudgens, Using 3D and 4D models to improve jobsite communication—virtual huddles case study, in: *Proceedings of CIB 24th W78 Conference & 14th EG-ICE Workshop & 5th ITC@ EDU Workshop*, Maribor (Slovenia), 2007: pp. 91–97. <https://citeseerx.ist.psu.edu/pdf/6135a87f49fb763b9ee0e0cdb534e535d9b4536c> (accessed June 13, 2023).
- [23] P.F. Kaming, P.O. Olomolaiye, G.D. Holt, F.C. Harris, Factors influencing craftsmen's productivity in Indonesia, *International Journal of Project Management*. 15 (1997) 21–30. [https://doi.org/10.1016/s0263-7863\(96\)00019-1](https://doi.org/10.1016/s0263-7863(96)00019-1).
- [24] A. Makulsawatudom, M. Emsley, K. Sinthawanarong, Critical factors influencing construction productivity in Thailand, *The Journal of KMITNB*. 14 (2004) 1–6. https://www.thaiscience.info/Journals/Article/TJ/KM/10470220.pdf?fbclid=IwAR1_R51 (accessed June 13, 2023).
- [25] I. Issue, Improving Access to Design and Construction Information Regarding Paper-Based 2D Drawings on Construction Sites, (2016).
- [26] S. Côté, P. Trudel, R. Snyder, R. Gervais, An augmented reality tool for facilitating on-site interpretation of 2D construction drawings, in: *Proceedings of the 13th International Conference on Construction Applications of Virtual Reality (CONVR)*, London, UK, 2013: p. 323.
- [27] M. Foroughi Sabzevar, M. Gheisari, L.J. Lo, Improving access to design information of paper-based floor plans using augmented reality, *Int J Constr Educ Res.* 17 (2020) 178–198. <https://doi.org/10.1080/15578771.2020.1717682>.
- [28] S. Emmitt, C.A. Gorse, *Construction communication*, John Wiley & Sons, 2009. <https://www.wiley.com/en-us/Construction+Communication-p-9781405148726> (accessed June 14, 2023).
- [29] W. Weaver, The mathematics of communication, *Sci Am.* 181 (1949) 11–15. <https://www.jstor.org/stable/24967225> (accessed June 14, 2023).
- [30] M. Feldberg, *Organizational behaviour: text and cases*, 1975. <https://catalogue.nla.gov.au/Record/1197941/Details?> (accessed June 14, 2023).
- [31] J. Fiske, *Introduction to Communication Studies*, Routledge, London, 1990. <https://doi.org/10.4324/9780203134313>.
- [32] A. Dainty, D. Moore, M. Murray, *Communication in construction: Theory and practice*, Routledge, London, 2006. <https://doi.org/10.4324/9780203358641>.
- [33] D. McMurray, J. Arnett, 7. Writing Instructions, (2016). <https://digitalcommons.kennesaw.edu/cg>

- i/viewcontent.cgi?article=1006&context=oertechcomm (accessed June 14, 2023).
- [34] E. Ziegler, Real-time markerless tracking of objects on mobile devices, 2012. http://userpages.uni-koblenz.de/~cg/Bachelorarbeiten/BA_Ziegler.pdf.
- [35] F. Bergamasco, A. Albarelli, E. Rodola, A. Torsello, Rune-tag: A high accuracy fiducial marker with strong occlusion resilience, in: CVPR 2011, IEEE, 2011: pp. 113–120. <https://doi.org/10.1109/CVPR.2011.5995544>.
- [36] T.-W. Kan, C.-H. Teng, W.-S. Chou, Applying QR code in augmented reality applications, in: Proceedings of the 8th International Conference on Virtual Reality Continuum and Its Applications in Industry, 2009: pp. 253–257. <https://doi.org/10.1145/1670252.1670305>.
- [37] K. Ruan, H. Jeong, An augmented reality system using Qr code as marker in android smartphone, in: 2012 Spring Congress on Engineering and Technology, IEEE, 2012: pp. 1–3. <https://doi.org/10.1109/SCET.2012.6342109>.
- [38] L. Naimark, E. Foxlin, Circular data matrix fiducial system and robust image processing for a wearable vision-inertial self-tracker, in: Proceedings. International Symposium on Mixed and Augmented Reality, IEEE, 2002: pp. 27–36. <https://doi.org/10.1109/ISMAR.2002.1115065>.
- [39] S. Han, E.J. Rhee, J. Choi, J.-I. Park, User-created marker based on character recognition for intuitive augmented reality interaction, in: Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry, 2011: pp. 439–440. <https://doi.org/10.1145/2087756.2087839>.
- [40] H. Belghit, N. Zenati-Henda, A. Bellabi, S. Benbelkacem, M. Belhocine, Tracking color marker using projective transformation for augmented reality application, in: 2012 International Conference on Multimedia Computing and Systems, IEEE, 2012: pp. 372–377. <https://doi.org/10.1109/ICMCS.2012.6320245>.
- [41] M. Foroughi Sabzevar, M. Gheisari, J. Lo, Development and Assessment of a Sensor-Based Orientation and Positioning Approach for Decreasing Variation in Camera Viewpoints and Image Transformations at Construction Sites, Applied Sciences. 10 (2020) 2305. <https://doi.org/10.3390/app10072305>.
- [42] H.M. Khoury, V.R. Kamat, Evaluation of position tracking technologies for user localization in indoor construction environments, Autom Constr. 18 (2009) 444–457.
- [43] J.P. Rolland, L.D. Davis, Y. Baillet, A survey of tracking technologies for virtual environments, in: Fundamentals of Wearable Computers and Augmented Reality, CRC Press, 2001: pp. 83–128. <https://doi.org/10.1201/9780585383590-9>.
- [44] M.S. Bargh, R. de Groote, Indoor localization based on response rate of bluetooth inquiries, in: Proceedings of the First ACM International Workshop on Mobile Entity Localization and Tracking in GPS-Less Environments, 2008: pp. 49–54. <https://doi.org/10.1145/1410012.1410024>.
- [45] R. Want, A. Hopper, V. Falcao, J. Gibbons, The active badge location system, ACM Transactions on Information Systems (TOIS). 10 (1992) 91–102.
- [46] P. Bahl, V.N. Padmanabhan, RADAR: An in-building RF-based user location and tracking system, in: Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No. 00CH37064), Ieee, 2000: pp. 775–784. <https://doi.org/10.1109/INFCOM.2000.832252>.
- [47] J. Karlekar, S.Z. Zhou, Y. Nakayama, W. Lu, Z.C. Loh, D. Hii, Model-based localization and drift-free user tracking for outdoor augmented reality, in: 2010 Ieee International Conference on Multimedia and Expo, IEEE, 2010: pp. 1178–1183. <https://doi.org/10.1109/ICME.2010.5583850>.
- [48] G. Deak, K. Curran, J. Condell, A survey of active and passive indoor localisation systems, Comput Commun. 35 (2012) 1939–1954. <https://doi.org/10.1016/j.comcom.2012.06.004>.
- [49] S. Gezici, Z. Tian, G.B. Giannakis, H. Kobayashi, A.F. Molisch, H.V. Poor, Z. Sahinoglu, Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks, IEEE Signal Process Mag. 22 (2005) 70–84. <https://doi.org/10.1109/MSP.2005.1458289>.
- [50] Pozyx, (2017). <https://www.pozyx.io/> (accessed August 11, 2023).
- [51] M. Popa, J. Ansari, J. Riihijarvi, P. Mahonen, Combining cricket system and inertial navigation for indoor human tracking, in: 2008 IEEE Wireless Communications and Networking Conference, IEEE, 2008: pp. 3063–3068.
- [52] F. Dellinger, J. Delon, Y. Gousseau, J. Michel, F. Tupin, Change detection for high resolution satellite images, based on SIFT descriptors and an a contrario approach, in: 2014 IEEE Geoscience and Remote Sensing Symposium, IEEE, 2014: pp. 1281–1284.

Virtual Reality-based Blockchain Application for Optimized Collaborative Decisions of Modular Construction

Mohamed Assaf¹, Rafik Lemouchi¹, Mohamed Al-Hussein¹, Xinming Li^{2*}

¹ Department of Civil & Environmental Engineering, University of Alberta, Edmonton, AB, Canada;

² Department of Mechanical Engineering, University of Alberta, Edmonton, Alberta, Canada.
massaf2@ualberta.ca, lemouchi@ualberta.ca, malhussein@ualberta.ca, xinming.li@ualberta.ca*

Abstract

Game engine technology has been studied extensively lately in offsite construction (OSC) research due to its ability to develop virtual reality (VR) environments that can be used in the decision-making process prior to the actual project implementation. However, accessing these VR scenes typically requires the participant to be in a VR lab or at least possess VR-specialized hardware and software. Also, these models are typically accessed in isolation with no real-time connectivity with other stakeholders, limiting their collaboration efficiency and questioning their applicability in real life. Thus, this study proposes a web-based multi-player framework based on game engines and blockchain technologies to promote collaborative decision-making processes in OSC projects. The developed system allows users to access a cloud simulation-optimization (SO) model to evaluate several decisions based on identified key indicators. The system is validated using an offsite construction (OSC) case study. Two scenarios are defined to ensure efficient connectivity among the OSC stakeholders and the security of the developed network.

Keywords –

Virtual Reality, Blockchain, Operation Management, Modular Construction

1 Introduction

Through the advancement of construction technologies, an expanding range of sustainable and productive construction methods are proposed, including OSC. Modular construction (MC), as a part of OSC methods, has the potential to improve the construction industry's current state, which suffers from huge delays, high injury rates, and cost overruns [6]. MC shifts almost all of the tasks from onsite to offsite factories [4]. This shift provides substantial value to the quality, sustainability, and safety in the construction industry. However, shifting activities from onsite to offsite factories constitutes a complex supply chain that includes

manufacturing, logistics, and construction stages. The fragmented nature of this supply chain leads to a greater demand for collaboration among the MC stakeholders, including manufacturers, logistics companies, and onsite contractors [7]. It also calls for mutual trust and efficient information sharing among the MC stakeholders [2].

Therefore, researchers aimed to use many digital technologies to support a collaborative environment for MC stakeholders [7]. These technologies include building information technology (BIM), game engines, and virtual reality (VR). VR and game engines promote the use of static BIM 3D models by adding interactive scenes. For instance, Zhang, et al. [16] developed a virtual system to support collaborative planning for construction activities of the MC projects. Their study used game engines and BIM models to create virtual scenes. Ezzeddine and García de Soto [4] aimed to integrate different teams in MC projects using game engines and BIM models. Their model targeted design, manufacturing, and logistics stages in MC projects. Wu, et al. [13] developed an information system to share secure data through the construction stage of MC. Their study integrated the blockchain and BIM technologies with the aim of mutual trust in sharing installation data. Besides, IoT applications are also found in the literature and aim to collect near real-time data in order to enhance visualization and traceability. For instance, Wu, et al. [12] developed an information system that facilitated manufacturing management by providing real-time and traceable information flow. In their study, BIM models were integrated with IoT sensors and blockchain technology to prevent the single point of failure in the IoT networks and synchronize BIM changes across different platforms. Hussein, et al. [6] developed a decision support system to optimize the logistics and construction stages of the MC projects. Simulation techniques, namely discrete event simulation (DES) and agent-based modeling (ABM), were adopted in their study to evaluate key performance indicators through MC implementation.

Despite the contributions of these previous studies, several limitations are identified. First, the applications of game engines in developing virtual scenes were only

limited to a single player (user) at a time with no real-time connectivity, questioning the application of these models in real-life practices. Second, when dealing with multi-player functions, the literature lacks a decentralized network (e.g., blockchain networks) that ensures that the virtual assets are secured and participant data are transparent. This aligns with the findings of Bao, et al. [2], who aimed to develop a multi-player safety training model. Their study recommended a decentralized VR model after pointing out that centralized models as a limitation of their study. Third, even with multi-player virtual scenes, the accessibility of the models is limited since the participant needs to be in a VR facility or lab. Hence, remote access to the developed VR models is needed.

In light of these limitations, this study proposes a multi-player collaborative decision-making framework to enhance coordination and connectivity among MC participants. The proposed system combines VR, simulation, optimization, and blockchain (VR-SO-BC) to overcome the mentioned limitations. The objectives of this study are: 1) build a cloud simulation-optimization (SO) model that provides MC participants with appropriate decisions; 2) develop virtual scenes that allow participants to try various selected decisions in VR environments; 3) develop a web-based multi-player model to enhance decision-making processes; 4) validate the proposed model based on a case study.

2 Methodology

Figure 1 illustrates the architecture of the VR-SO-BC-OSC framework. This framework aims for the following: 1) provide critical decisions for the construction, logistics, and planning phases of MC projects; 2) discuss the feasibility of the chosen decisions in virtual environments; 3) provide a fully immersed VR experience for the MC participants to interact with the developed virtual models; 4) develop multi-player virtual scenes governed by blockchain network for collaborative decision making.

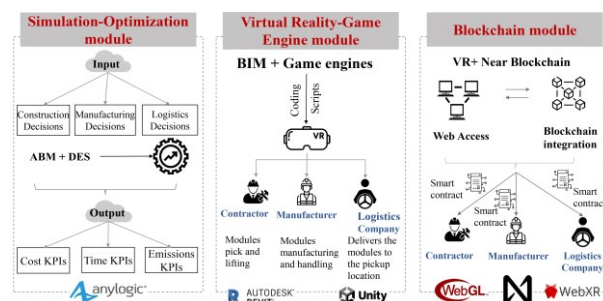


Figure 1: The VR-SO-BC-OSC architecture

The proposed framework integrates simulation and

optimization techniques to estimate analytical values for the critical decisions of the OSC implementation. Even though process simulation can be integrated within the virtual game scenes, it can only represent the flow of the process with no complex interactions and dynamic behavior of the agents [9]. Therefore, the current study's authors chose to develop a hybrid simulation model for this stage that would allow us to consider dynamic behaviors and complex interactions. However, a cloud version of this model is integrated in the following steps with a game engine to assess the physical properties and feasibility of the selected decisions. The proposed hybrid simulation model combines discrete event simulation (DES) and agent-based modeling (ABM) methods to represent the overall MC supply chain activities, considering the dynamic behaviors of the included agents. Furthermore, this module also presents several key performance indicators (KPIs), such as cost, time, and produced CO₂, to measure the performance of the MC supply chain. The hybrid simulation model is then integrated with an optimization model that aims to find near-optimal decisions in each stage of the MC project to enhance the performance of the MC supply chain. The results of the proposed SO model are fed into the virtual reality/game engine module to evaluate the selected decisions in immersive virtual environments.

The virtual reality/game engine module is proposed as a continuation of the developed SO module. Its purpose is to improve decision-making by considering all of the physical properties and feasibility of the chosen decision from the previous module. The developed module integrates the merits of BIM and game engine technologies. Building a 3D BIM model forms the first step in this module, as it determines the design of all of the prefabricated modules and substructure elements. It is worth mentioning that the structural and architectural design of the MC project is beyond the scope of the current study. Game engines are then employed to create virtual environments of the developed BIM model. Virtual environments enable MC teams to visualize each process, allowing them to make proactive and informed decisions [4]. In this module, three main MC teams are integrated: manufacturing, onsite assembly, and transportation teams. Each team has its own virtual environment, and they all share a separate virtual scene. The creation of the virtual environment will be detailed in the following sections. Furthermore, the virtual scenes are developed to enable a fully immersive VR experience. Immersive VR experiences can enable participants to interact highly with virtual environments, facilitating a more efficient decision-making process [15].

The last module in the proposed methodology is the web-based VR/ blockchain module. Web-based VR allows participants to access the developed VR module via remote platforms that do not require them to be in a

VR facility or laboratory [10]. The proposed module allows participants to access the developed virtual environments in screen-based VR and immersive VR (using headsets). Further, a server connection is established between the web screens and the main designed server in the game engine. Despite the capabilities of the web-based VR models in promoting connectivity and facilitating better-informed decisions, they lack data transparency and security of the participants. This issue aligns with the results of Bao, et al. [2], who recommended integrating the web-based VR application with BC technology to ensure the persistency of the virtual scenarios, transparency of the user data, and activity tracking on the web-based VR model. Hence, a BC network is added to the developed Web-based VR models to address these limitations. The developed blockchain comprises several blocks. In which each block includes its own records, timestamps, and hash values of the previous block [11]. In addition, the BC network is supported by a smart contract to facilitate adding participants of any organization to the BC network. Each participant in the web-based network owns a non-fungible token (NFT) that incorporates a unique digital signature, defining the ownership of the virtual character in the virtual environment.

3 VR-SO-BC-OSC Model Development

This section presents the development of the modules mentioned above. Figure 2 shows the sequential steps included in each module. The proposed model is tested on a case study of a modular bridge project. The bridge consists of forty prefabricated steel modules and has a main span of 300 meters and two side spans of 120 meters each. More details on the statistical system of this case study were presented by Yu and Chen [14]. It is worth mentioning that this case study serves as a proof of concept to show the capabilities of the proposed model.

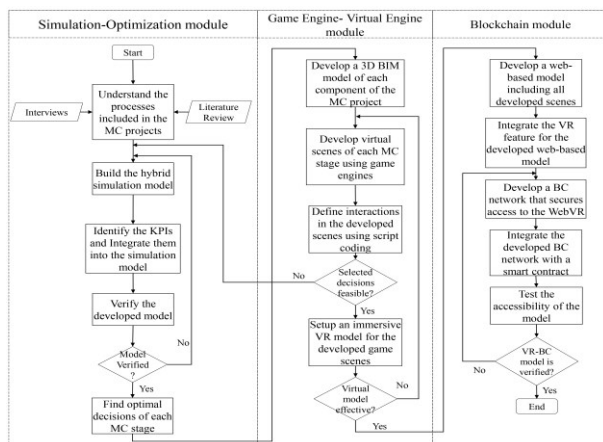


Figure 2: Steps included in the model development

3.1 Simulation-Optimization module

This subsection discusses the development of the SO model. It is worth mentioning that the SO model is an extension of the work done by Assaf, et al. [17], where the optimization of each project KPI was the main focus of that work. However, this study mainly focuses on developing a multi-user decision-making environment for collaborative decisions and participant integration.

As mentioned, a hybrid simulation model of agent-based modeling (APM) and discrete event simulation (DES) techniques is developed to consider all MC activities and possible interactions between the resources. Anylogic tool is used to develop the SO model. Figure 3 shows a sample of SO model development.

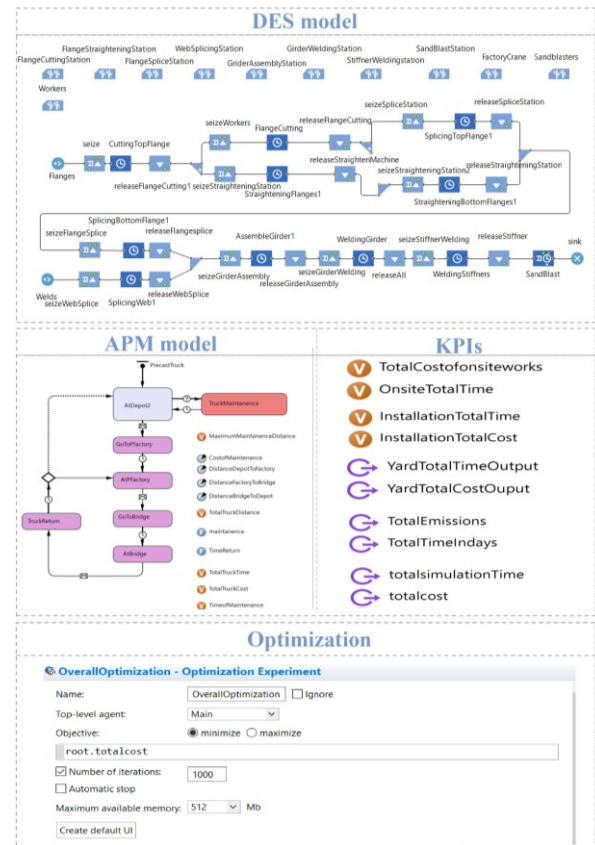


Figure 3: Overview of the developed SO model

The DES model simulates activities related to manufacturing, logistics, and onsite works. The APM supports the interactions between agents (resources) and the DES processes and between one another. Four main APM agents are considered: precast trucks, steel girders trucks, barge, and traffic agents. These agents send and receive messages to and from the DES model and among one another to update specific parameters and trigger certain behaviors. For instance, when a particular stage is reached in the DES model, a “GoToPfactory” message is

sent to the precast agent to travel to the factory based on the distance specified in the model. Similarly, the traffic agent sends messages to the other agents to influence their traveling time based on the traffic conditions.

Several KPIs are calculated within the APM-DES model, including cost, time, and emissions in each MC stage. The change in these KPIs is based on several parameters in all MC stages. For instance, some of the included parameters in the onsite installation stage are the number of mobile cranes and the number of workers. These parameters influence the hybrid simulation model and, subsequently, the resulting KPIs. A sensitivity analysis is carried out to discover which of these parameters are sensitive and have a high impact on the KPIs. An optimization model is integrated into the simulation model to efficiently discover the appropriate values for these parameters. In this regard, the OptQuest optimization tool (supported by Anylogic) is used to run the optimization model. The decision variables of the model are considered based on the sensitivity analysis and include the following: No of steel trucks, No of precast trucks, Inventory capacity, Just-in-Time (JiT) application, No of production lines, No of factory workers, No of onsite mobile cranes. Constraints are defined in the optimization model based on user specifications. For instance, the user can indicate a certain budget that the cost should not exceed. This means that if this budget is exceeded, the generated solution will be infeasible.

3.2 Virtual Reality-Game engine module

This subsection discusses the development of the game scenes that will be used to connect various MC stakeholders together for collaborative decision-making. Further, a cloud version of the developed hybrid SO model by Anylogic is accessible in these scenes to allow users to examine several KPIs based on their input. Two main software tools are used in developing the virtual scenes: Autodesk Revit and Unity 3D. Revit tool (Version 2023) is used to create the 3D BIM models of all of the elements in the MC project. In the BIM models, the needed information (such as module ID and location) is included to be later used in the virtual scenes. The developed BIM models are then exported in FBX format, which is compatible with the Unity3D tool. It is worth mentioning that the developed virtual scenes extend virtual scenes presented by the study authors in Assaf et al. [1]. The Unity3D tool (version 2022.3.11.f1) is used to develop four main scenes, as shown in Figure 4. The first scene, shown in Figure 4(a), is an exploring scene that is accessible to all participants and allows them to do the following tasks: 1) access the Anylogic cloud model to test various decisions, 2) explore the chosen decisions in an interactive virtual environment, 3) review modules details and installation instruction, 4) connect with other

stakeholders in the virtual environment for collaborative decision-making process. The second scene is the onsite scene, as shown in Figure 4(b).

The scene offers the onsite team the following functions: 1) real-time assessment of various KPIs, such as time and CO₂ emissions, 2) real-time assessment of the clearance between the prefabricated component and the surroundings, and 3) connect the logistics team with the onsite team.



Figure 4: Developed virtual scenes in Unity Editor

The third developed scene is the steel manufacturing scene, which is displayed in Figure 4(c). The manufacturing team can perform the following tasks: 1) check the time needed for loading and producing CO₂ emissions in real-time, and 2) evaluate the selected number of modules' storage levels. The last scene is the logistics scene, as shown in Figure 4(d). The

transportation team is able to explore the following features: 1) navigate through city roads leading to the project, 2) assess in real time the clearance of the modules for each chosen road, 3) test the accessibility to the pickup locations, and 4) evaluate the selected safety precautions when maneuvering to the pickup locations. It is worth mentioning that the user is able to navigate the Anylogic cloud model form within Unity Editor to try different scenarios. They can then test the selected decisions in the developed virtual interactive scenes. It is also worth mentioning that users can access virtual scenes in a fully immersive mode. Steam VR plugin is added to Unity3D, and the immersive experience is set using the following hardware tools: Varjo XR headset, HTC Vive controllers, and HTC Vive base stations.

3.3 Multi-Player blockchain module

This subsection discusses the implementation of the web-based virtual model that offers multi-player functions and is supported by blockchain integration. As mentioned before, this module aims to provide a collaborative decision-making environment to test the previously agreed decisions. A multi-player function is added to the Unity editor after developing the virtual scenes using BIM and game engines. Fishnet is a network solution by Unity3D that provides many multi-player functions and supports web-based integration [5]. Fishnet invokes network behavior in the virtual scenes, allowing clients and the server to join the same virtual scenes with real-time connections. Therefore, C# scripts that control all of the desired networking interactions are updated to support multi-player features. Thus, the networking behavior is integrated into Unity Editor using Fishnet functions. The network manager is a vital component in establishing the connection between the server and clients [5]. A player, e.g., a truck, is spawned in the network manager component and is synchronized among the server and clients. Spawning points can be defined to specify the location of networked objects.

To provide authority on who can manipulate the players, a function is added to each network object to only allow clients to control their own players. A canvas is also predefined in the network to support establishing the server and adding clients. Furthermore, the multi-player model also supports broadcasting functions to provide communications among objects without the need for network components. For instance, to connect the logistics team and onsite team, the truck (network object) is added to the network manager, while the crane is configured using broadcast functions. This way, the movement of both crane and truck is synchronized among the server and clients. Furthermore, real-time connection and bi-directional data are configured on the multi-player model. For instance, the onsite team and logistics team can view real-time evaluation of different

KPIs, such as clearance checks and produced emissions. This connectivity and information sharing are tested through chat functions between the server and clients.

Despite the capabilities of the multi-player model to provide collaborative decision-making among project teams, it is centralized in a way that the participants need to be in a VR facility to try the developed scenes. Thus, web functions and decentralized networks are discussed in this subsection to overcome this issue. WebGL is applied in this part to deploy the web-based VR model. WebGL is a JavaScript API that deploys 3D and 2D virtual scenes within any compatible web browser [3]. Unity 3D supports WebGL and can directly build WebGL extensions. However, mobile browsers are not supported by Unity WebGL, which is a limitation since most people prefer touch-based interactions [8]. Unity 3D supports WebXR, allowing users to access the web-based model in immersive environments using headsets.

As discussed in the methodology section, the blockchain network is integrated with the developed model to provide the following functions: access control for the stakeholders, decentralized networks that prevent a single point of failure, and record all of the activities on the web-based model. Among many blockchain platforms, this study focuses on NEAR protocol integration for the following reasons: 1) NEAR is a layer one blockchain network without dependency on other chains [11]; 2) NEAR protocol is developed to have a low transaction cost, making it suitable for networks that have rapid transactions; 3) it is compatible with Unity 3D and allows the users to develop their own smart contract to support the interaction with the blockchain network. Therefore, the NEAR protocol is selected and added to Unity 3D virtual scenes to be later included in the WebGL building process.

A sample of how the smart contract supports the developed blockchain network is shown in Figure 5. One of the prerequisites to establishing the NEAR blockchain is the NEAR wallet. The wallet contains a set of Non-Fungible Tokens (NFTs) that are controlled by a smart contract. The tokens, in this case, are considered the participants in the network. Figure 5(a) shows adding a participant to the blockchain network using the smart contract integration. The information needed when adding a new token (participant) includes the following: token ID, title, description, and image. Figure 5(b) shows how the smart contract secures any activities on the blockchain network with hash value and timestamp to be immutable and permanently saved. As discussed, the NEAR blockchain is integrated with the Unity 3D platform, meaning a smart contract must be assigned to the editor. Figure 5(c) shows how the smart contract is included in the C# scripts to be later on the part of the WebGL build. Figure 5(d) shows the newly added token to the NEAR wallet.

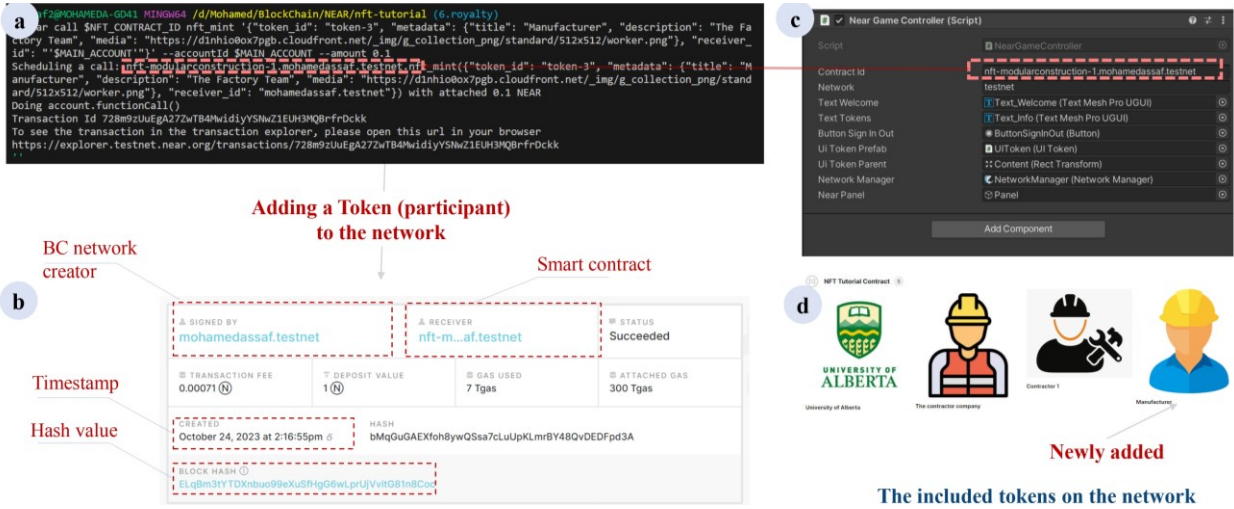


Figure 5: Smart contract integration in the blockchain network and within Unity

Another notable feature of integrating NEAR protocol in Unity Editor is assigning a name header for each participant in the multi-player environment. The information of the tokens (participants) is retrieved from the blockchain network and assigned in the virtual scenes. Table 1 shows a sample of the algorithm used in doing so. As shown in the table, a remote procedure call (RPC) is used to send information across the network. This information is synchronized for both clients and server screens. A NEAR canvas in Unity editor is also added to facilitate the connection to the network and selection of the participants. When a certain token is clicked, the mentioned algorithm is triggered, and the participant information is assigned.

Table 1: synchronizing participant's information

Sample Algorithm: Retrieving Participant information to be added to the virtual scenes
<pre> public class Player: NetworkBehaviour { [SerializeField] private TMPPro.TextMeshProUGUI _ParticipantUsername; public static string localParticipantName; [SyncVar(OnChange = nameof(OnParticipantNameChanged))] public string username; private void OnParticipantNameChanged (string oldValue, string newValue, bool isServer) { _ParticipantUsername.text = newValue; } public override void OnStartClient() { base.OnStartClient(); if(!IsOwner) return; ServerSetName(localParticipantName);} [ServerRpc] private void ServerSetName(string name) {username = name;} </pre>

4 Results and discussions

This section provides a discussion of the results of the proposed framework. The section will focus on two main validation cases: 1) connecting various participants in the exploring scene and 2) connecting the onsite team and logistics team in the onsite scene. Figure 6 shows a screenshot of the accessed model through the editor and web. The right-hand side of the figure shows a participant accessing the multi-player model using the Unity Editor. Establishing the server on the editor is essential to enable other clients to enter the multi-player network. The left-hand side shows a client joining the network from a web window. It is worth mentioning that both participants can do the same functions. For instance, participants can access a cloud version of the Anylogic model to try different scenarios and evaluate the project KPIs. The Anylogic button will only appear when the participant successfully signs into the blockchain network. When the first user enables the network server, they must sign to the blockchain and select the corresponding NFT. The RPC discussed earlier will assign and reflect the participant's name with the potential clients. The client will be asked to join the blockchain network and select the appropriate NFT. The smart contract in the blockchain will record these activities on the network and assign them hash values.

The real-time connectivity between the two participants and the efficiency of the established network is tested as follows. First, the movements of the two characters (with the assigned names) are tracked in the two windows to test the efficiency of the network manager in synchronizing the activities in both scenes.

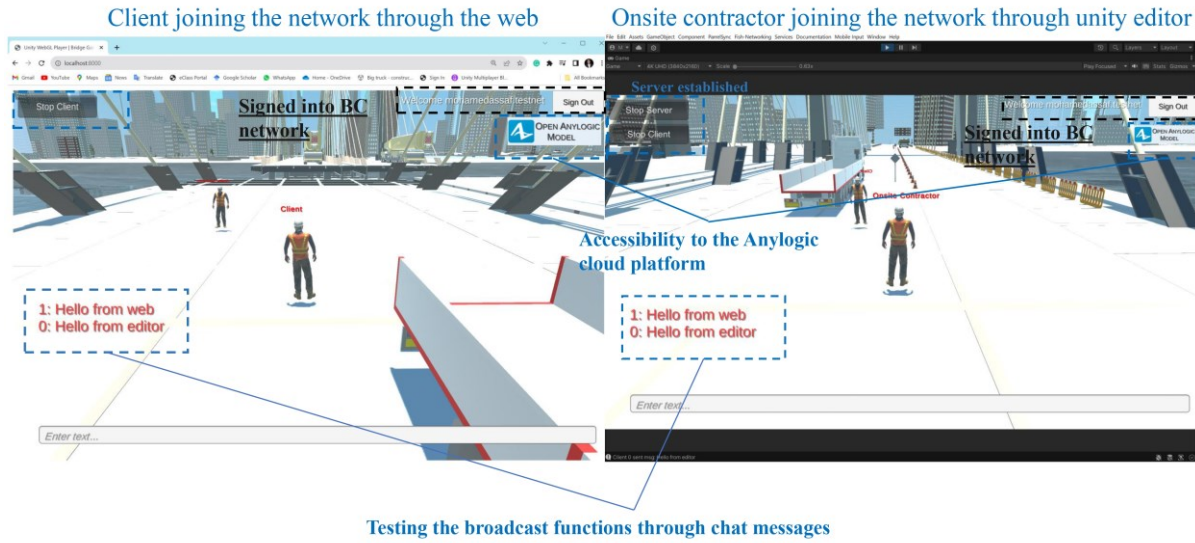


Figure 6: Real-time connectivity between different stakeholders in the web-based model

Second, the broadcast functions are tested through the chat messages between the server and the clients. These tests varied the system's abilities in providing real-time connection among various stakeholders in the developed virtual scenes. When the participant accesses the cloud version of the Anylogic model, they manipulate the sensitive parameters and have an evaluation of the discussed KPIs. Figure 7 shows a sample of the features provided by the cloud SO model. For instance, the figure shows the inventory level and the produced modules changing over time based on the participant's specific set of inputs. In addition, it also shows the modules' installation cost and the onsite work cost over time.



Figure 7: Feature samples provided by the SO model

Another testing scenario is the connectivity between the onsite team and the logistics team. In this case, a truck carrying precast elements is navigating through the city roads to reach the pickup point. The network manager is configured to a new player, which is the truck. Hence, the movement of the truck is synchronized in all of the

screens. When the server is initiated, the client (logistics team) can sign into the blockchain network and join the game scene. The client then can start navigating the truck to pick up points with several metrics evaluated in real time. Some of these metrics are module clearance checks and traveled distance. The onsite team, on the other hand, can view the movement of the truck and control the movement of the onsite mobile crane. The broadcasting function is used to reflect the movement of the crane in the logistics team's view. The connections between the users are also tested using chat messages. The representative of the onsite team and the onsite team can then make collaborative decisions in a virtual replica of the MC project. Further, both teams can also access the SO cloud model to further test several decisions, such as the number of onsite cranes, number of trucks, and number of workers on site, and evaluate these decisions based on several KPIs, including produced emissions, cost, and duration. It is also worth mentioning that there is no limitation on the number of users, which means there can be more than one representative from each team in the multi-player VR model.

5 Conclusion

This study proposes a virtual reality-based blockchain model that enables multi-player networking in the developed virtual scenes. The model also supports access using web browsers to overcome the issue of limited accessibility to the VR scenes. The integration of the blockchain in the multi-player virtual environment enables a decentralized network and provides security and data transparency. It also enables the recording of all of the activities on the VR web-based model to promote trustworthiness among MC stakeholders. The model also

provides participants with access to a cloud SO model that allows them to try different options and have a real-time evaluation based on a set of identified KPIs.

The multi-player networking functions were tested in this study using two main cases: 1) connecting various participants in the exploring scene and 2) connecting the onsite team and logistics team in the onsite scene. Remote procedure calls and broadcasting functions were used to test the connectivity among participants in the multi-player environment. The accessibility of the integrated cloud SO model in the web-based model was also tested and proved the model's efficiency in providing users with an assessment of various decisions. Some limitations of the current work are worth mentioning, including the following: 1) the SO cloud model has no direct connection with the blockchain network. However, both of them are integrated into the VR web-based model; 2) the scheduling data is subjected to the user input. Future work will target integrating IoT sensors to facilitate the flow of near real-time data for better data traceability and visualization.

References

- [1] M. Assaf, R. Lemouchi, M. Al-Hussein, X. Li, A collaborative planning model for offsite construction based on virtual reality and game engines, *Proceedings, Conference on Construction Applications of Virtual Reality*, Nov. 13–16, 2023, Florence, Italy, 2023.
- [2] L. Bao, S.V.-T. Tran, T.L. Nguyen, H.C. Pham, D. Lee, C. Park, Cross-platform virtual reality for real-time construction safety training using immersive web and industry foundation classes, *Automation in Construction* 143 (2022) 104565.
- [3] M. Chessa, C. Bassano, F. Solari, A WebGL Virtual Reality Exergame for Assessing the Cognitive Capabilities of Elderly People: A Study About Digital Autonomy for Web-Based Applications, *ICPR International Workshops and Challenges*, Springer International Publishing, pp. 163-170.
- [4] A. Ezzeddine, B. García de Soto, Connecting teams in modular construction projects using game engine technology, *Automation in Construction* 132 (2021) 103887.
- [5] Fishnet (2019). "FishNet - Unity Networking Evolved." <<https://fishnetworking.gitbook.io/docs>>.
- [6] M. Hussein, A. Darko, A.E.E. Eltoukhy, T. Zayed, Sustainable Logistics Planning in Modular Integrated Construction Using Multimethod Simulation and Taguchi Approach, *Journal of Construction Engineering and Management* 148 (6) (2022) 04022022.
- [7] Y. Jiang, X. Liu, Z. Wang, M. Li, R.Y. Zhong, G.Q. Huang, Blockchain-enabled digital twin collaboration platform for fit-out operations in modular integrated construction, *Automation in Construction* 148 (2023) 104747.
- [8] A.E. Martis, C. Bassano, F. Solari, M. Chessa, Going to a Virtual Supermarket: Comparison of Different Techniques for Interacting in a Serious Game for the Assessment of the Cognitive Status, in: S. Battiato, G.M. Farinella, M. Leo, G. Gallo (Eds.), *ICIAP 2017*, Springer International Publishing, Cham, 2017, pp. 281-289.
- [9] J. Possik, S. Gorecki, A. Asgary, A.O. Solis, G. Zacharewicz, M. Tofighi, M.A. Shafiee, A.A. Merchant, M. Aarabi, A. Guimaraes, N. Nadri, A Distributed Simulation Approach to Integrate AnyLogic and Unity for Virtual Reality Applications: Case of COVID-19 Modelling and Training in a Dialysis Unit, *2021 IEEE*, 2021.
- [10] P.W.S. Putter, J.C. Roberts, P.D. Ritsos, *Immersive Analytics with WebVR and Google Cardboard*, 2016.
- [11] Snehlata, P. Shukla, A.K. Singh, S. Tiwari, Rishabh, V.K. Dwivedi, An Intelligent Blockchain-Oriented Digital Voting System Using NEAR Protocol, *SN Computer Science* 4 (5) (2023) 643.
- [12] L. Wu, W. Lu, R. Zhao, J. Xu, X. Li, F. Xue, Using Blockchain to Improve Information Sharing Accuracy in the Onsite Assembly of Modular Construction, *Journal of Management in Engineering* 38 (3) (2022) 04022014.
- [13] Wu, L., Lu, W., Xue, F., Li, X., Zhao, R., & Tang, M. (2022). Linking permissioned blockchain to Internet of Things (IoT)-BIM platform for off-site production management in modular construction. *Computers in Industry*, 135, 103573.
- [14] Yu, X. and D. Chen, Design and construction of the Tahya Misr cable-stayed bridge in Cairo, Egypt. *Proceedings of the Institution of Civil Engineers - Civil Engineering*, 2020. 174(2): p. 79-88.
- [15] Z. Zhang, W. Pan, Virtual reality supported interactive tower crane layout planning for high-rise modular integrated construction, *Automation in Construction* 130 (2021) 103854.
- [16] Z. Zhang, M.O. Wong, W. Pan, Virtual reality enhanced multi-role collaboration in crane-lift training for modular construction, *Automation in Construction* 150 (2023) 104848.
- [17] M. Assaf, S. Assaf, W. Correa, R. Lemouchi and Y. Mohamed, "A Hybrid Simulation-Based Optimization Framework For Managing Modular Bridge Construction Projects: A Cable-Stayed Bridge Case Study," 2023 Winter Simulation Conference (WSC), San Antonio, TX, USA, 2023, pp. 3094-3105

The Role of Large Language Models for Decision Support in Fire Safety Planning

Dilan Durmus¹, Alberto Giretti¹, Ori Ashkenazi², Alessandro Carbonari¹, Shabtai Isaac²

¹ Polytechnic University of Marche, DICEA Department, Ancona, Italy

² Department of Civil and Environmental Engineering, Ben-Gurion University of the Negev, Beer-Sheva, Israel

d.durmus@pm.univpm.it, a.giretti@staff.univpm.it, oriashke@post.bgu.ac.il,
alessandro.carbonari@staff.univpm.it, isaacsh@bgu.ac.il

Abstract –

This paper explores the integration of Large Language Models (LLMs), specifically GPT-4, in fire safety planning and knowledge-based systems within the Architecture, Engineering, and Construction industry. Focusing on overcoming challenges in expert systems, it presents an AI-in-the-loop model, illustrating how LLMs enhance decision support. The paper introduces a scenario analysis approach, demonstrating the iterative use of LLMs to enrich expert systems' knowledge bases. Two case studies emphasize the practical application of this approach in fire safety planning, showcasing LLM adaptivity, specialized reasoning, and domain knowledge integration. The study addresses the challenges of LLM-induced hallucinations and emphasizes the need for further research to enhance reliability in technical scenarios. Overall, it contributes to advancing fire safety strategies by leveraging the strengths of LLMs in dynamic building environments.

Keywords –

Adaptive Decision Support; Artificial Intelligence Integration; Fire Safety Planning; Knowledge-Based Systems; Large Language Models (LLMs)

1 Introduction

Fire safety planning is essential in the building industry, with the goal of reducing risks, protecting people, and limiting property damage [1]. As architectural trends evolve and cities expand, there is an increasing need for smart and adaptable fire safety approaches. However, traditional methods face challenges in dealing with the dynamic aspects of real-world building environments [2]. New building codes, materials, and occupancy patterns demand a more flexible fire safety approach that can adapt to the ever-changing landscape. Recognizing the mutable nature of buildings as they adapt to multifaceted influences is

fundamental. Underlying factors such as space configurations, material choices and evolving human interactions require not only standard, fixed protocols [1], but also innovative tools that can interactively learn and respond to these changes.

Indeed, Digital twins represent one of the emerging technologies that can deal with such dynamics by means of real-time monitoring, scenario prediction, and proactive measures for optimal functionality. Hence, they cannot rely on predefined, non-adaptive logic-analytic models. [3]. Rather, they must run embedded intelligent tools, capable of processing contextual information and proactively making decisions.

So far, the closest approximation to powerful thinking machines like AI has been expert systems, i.e. knowledge-based systems designed for specific problem-solving tasks. A knowledge-based system (KBS) generally consists of two components: the problem-solving processes (R) and the knowledge base (K). Constructing the reasoning part (R) has been the subject of extensive research and various software tools offer ready-made reasoning engines. Whereas, the almost unexplored challenge lies in creating custom knowledge bases.

Expert systems are inherently knowledge-intensive, representing human expertise in a particular field. Expertise, rooted in extensive domain knowledge, forms the basis of these systems. Constructing each knowledge base is a unique process, starting from scratch and tailored to specific domains or tasks. Acquiring and representing this knowledge is a time-consuming and challenging endeavor that delves deeply into the subject matter, and that must combine a quite high amount of diverse expertise that can hardly be put together on purpose. The underlying philosophy is to create tools that enhance knowledge representation. For this reason, the goal of current research is to develop superior tools for effectively capturing and conveying domain-specific knowledge, acknowledging the importance of a robust foundation for expert systems, which is the point where

AI can provide a valuable contribution to build the knowledge base.

A few studies demonstrate various applications of Artificial Intelligence (AI) in building automation, including expert system design. AI can streamline the design process, improve system operation, and optimize space utilization [4]. Moreover, AI has seen the emergence of powerful tools known as Large Language Models (LLMs), such as OpenAI's GPT-3.5 and GPT-4, that will be investigated in the following of this paper. Trained extensively on text and code, these models understand connections between ideas and offer rich contextual information [5]. This paper explores the potential applications of LLMs in fire safety planning, specifically in "Adaptive Knowledge-Based Systems."

Examining traditional expert systems' performance, the paper highlights the evolving role of LLMs in handling technical complexities and improving fire safety measures in the construction industry. Real-world case studies and National Fire Protection Association (NFPA) guidelines demonstrate how LLMs generate strategic decision-making pathways.

The paper focuses on two NFPA Case Study scenarios, Nightclub Fires (scenario 2 and 3), emphasizing the importance of interacting with AI, using LLMs for strategy generation, and analyzing specific situations.

Highlighting the adaptive nature of artificial intelligence models like ChatGPT-4, this paper underlines their ability to tailor responses to specific conversation contexts. While these AI advancements offer significant potential, caution is advised due to the possibility of hallucinations, where AI models may generate unsubstantiated information, but may even debug the knowledge bases in use and subject to augmentation. The challenges posed by hallucinations and the need for further research to enhance the reliability of AI-assisted fire safety planning will be discussed in the last section of this paper, which aims to contribute to advancing fire safety strategies in the dynamic building industry. Through a comprehensive exploration, it seeks to deepen our understanding of the potential benefits and challenges of adopting AI technologies in fire safety management.

2 Evolution of Decision Support Systems in Technical Domains

Artificial intelligence (AI) systems, including Large Language Models (LLMs) like Generative Pre-trained Transformer (GPT) models, analyze extensive datasets to discern patterns, relationships, draw inferences, provide recommendations, and execute actions. The evolution of conversational AI in the Architecture, Engineering, and Construction (AEC) industry, particularly in Natural

Language Processing (NLP) applications, faces challenges with conventional approaches, limiting user interactions [6].

Large Language Models (LLMs), neural networks with extensive parameters, depart from traditional training methods by utilizing self-supervised and semi-supervised learning on vast datasets. Prominent among LLMs are Generative Pre-trained Transformer (GPT) models, notably the OpenAI-developed decoder blocks. GPT models, evolving from GPT-2 to GPT-4 in 2023, operate on transformer-based architectures, capturing statistical patterns in natural language. Renowned for producing coherent, human-like language, GPT models find practical use in diverse applications, including chatbots, content generation, and machine translation, addressing open-ended queries effectively and serving as valuable tools for natural language communication.

The innate development of communication and inference skills in large language models has been confirmed through experiments demonstrating the efficacy of chain-of-thought (COT) prompting. This technique enhances performance across arithmetic, commonsense, and symbolic reasoning tasks, particularly in addressing complex challenges like multi-step math word problems. Language models, when empowered with the ability to generate coherent sequences of intermediate reasoning steps constituting a chain of thought, efficiently solve problems. The impetus behind COT prompting in knowledge management aligns with construction domain objectives. It allows models to deconstruct multi-step problems, allocating additional computational resources to tasks requiring extensive reasoning. Additionally, a chain of thought provides interpretable insights into the model's behavior, elucidating the reasoning process and offering debugging opportunities. Applicable to various language-based tasks, COT reasoning demonstrates general applicability. The elicitation of chain-of-thought reasoning in large language models is achieved effortlessly, emphasizing its limited effectiveness in smaller models. GPT-3's application of chain-of-thought prompting demonstrates favorable comparisons with prior methods, showcasing its performance in addressing intricate problems without fine-tuning on labeled datasets [7].

The construction industry, reliant on diverse information from various stakeholders, faces challenges in integrating, reusing, and efficiently managing information, impacting industry productivity. To address this, effective methods for representing extensive knowledge are crucial [8]. Expert systems, computer programs guided by specific problem-solving knowledge, are pivotal, emphasizing the need for knowledge directing solutions. The term "expert" signifies narrow specialization and substantial competence, benchmarked against human performance [9]. Presently, expert

systems serve interactive roles, from smart spreadsheets to financial advisors, relying on knowledge bases elicited through effective techniques [10]. Despite their success, expert systems exhibit less flexibility than humans. Large Language Models (LLMs), like GPT models, enhance expert systems by handling vast unstructured information, resembling human-generated text [11].

Table 1 discusses the potential role of LLMs as a tool for improving knowledge management in the building industry. LLMs expand expert systems by extracting information from diverse, unstructured sources, facilitating knowledge base augmentation and updating for advanced decision-making processes.

Table 1. Using LLMs in combination with expert systems to support decision making

	Expert Systems	Large Language Models
Goal	Supporting decision-making with a specific task	Eliciting additional information to broaden out the expert system
Input	Quantitative and structured data	Any type of (even) unstructured data
Approach	Combination of logical rules and a knowledge base to make inferences	Deriving statistical patterns from evidence and performing reinforcement learning
Method	Symbolic AI	Transformers
Interface	Digital	Answers to questions in human-like language / chain-of-thought
Output	Recommendations and advice for specific scenarios	Arguing scenarios representing the dynamics of complex systems

LLMs may provide a solution to some of the existing difficulties in using expert systems to improve knowledge management in the building industry :

- **Data Availability:** The large body of documents on which they are trained may provide an answer to the difficulty in obtaining structured data of high quality, which is often not readily available in the construction industry.
- **Robustness:** The relative resilience of LLMs may allow the drawing of conclusions even when input data is unavailable or inconsistent.
- **Adaptivity:** Unlike expert systems, generative LLMS have shown an ability to adapt to uncoded situations.

At the same time, some challenges are likely to be faced in the use of LLMs for knowledge management. Primary among those is the fact that such models are prone to hallucination, i.e. they give sound and plausible information that is not true, which reduces system performance and users' expectations. Here, their integration with expert systems may provide a solution, given the ability of such systems to provide consistent

inferences based on domain-specific knowledge.

Challenges faced both in the development of expert systems and LLMs are their maintainability, and the costs involved. These stem from the need for large datasets that are complete and accurate and may consequently be difficult and costly to obtain in the first place, and to maintain as changes occur in the construction industry over time.

3 Knowledge Engineering Issue

Knowledge, as employed in knowledge-based systems, encompasses agents' codified experience, serving as the informational foundation for problem-solving. Codification implies the formulation, recording, and preparation of knowledge for practical use. Agents' experience spans well-defined domains, representing fields of interest, and tasks, denoting specific types of work within those domains. Domain knowledge pertains to the general terminology and facts of a domain, while task knowledge involves computational models and facts associated with performing a task.

Two crucial dimensions for describing knowledge are procedural vs. conceptual knowledge and basic, explicit vs. deep, tacit knowledge [12]. Procedural knowledge involves knowing how to perform tasks and encompasses processes, task conditions, task order, required resources, and sub-tasks. Conceptual knowledge relates to knowing that certain relationships and properties exist among concepts, including taxonomies and class membership. Basic, explicit knowledge is consciously processed and focuses on fundamental tasks, relationships, and concepts. In contrast, deep, tacit knowledge resides in the subconscious and is acquired through experiences, leading to automatic activities. Examples include 'gut feel,' 'hunches,' 'intuition,' 'instinct,' and 'inspiration.' Tacit knowledge is demonstrated in activities like driving a car, where experienced drivers navigate effortlessly based on deep, tacit knowledge, contrasting with learners who lack this intuitive understanding [12].

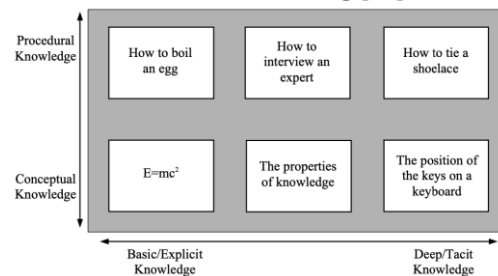


Figure 1. Two important dimensions with which to describe knowledge

3.1 Formulating Expertise

Expertise, typically associated with individuals, goes

beyond mere knowledge sources like books and recordings. Experts inject their unique perspectives into the knowledge they possess due to their individual experiences.

3.1.1 Knowledge Acquisition and Implementation Issues

In artificial intelligence and knowledge engineering, knowledge acquisition involves computer systems obtaining the necessary information to perform specific tasks. This process is a three-role participatory design task, involving users, domain experts, and knowledge engineers [9]. The knowledge acquisition process unfolds through four stages at two levels: *knowledge and symbol levels*.

Knowledge-level analysis precedes symbol-level analysis in the knowledge engineering process. During problem identification, unstructured interviews gather crucial sample problems, while conceptualization involves detailed representations of problem-solving processes using protocol analysis. Symbol-level analysis progresses with formalization, translating semi-formal knowledge into a computationally implementable form through defining an abstract language and computational architecture. The structured approach ensures a systematic transition from conceptualization to implementation, considering computational efficiency and user interface aspects. In addition to these stages, knowledge acquisition faces several issues. Experts often struggle to articulate their problem-solving methods, making knowledge acquisition challenging due to tacit knowledge issues. Building common ground between knowledge systems and end-users is complex, and the brittleness of rule-based expert systems outside their limited expertise can lead to nonsensical answers. The expertise scoping issue arises from oversimplified models in knowledge bases, while maintaining consistency becomes crucial for effective validation and reuse of knowledge and problem-solving methods.

3.2 AI in the Loop

In exploring the integration of AI systems into the development of knowledge-based systems based on logical formalism, this article delves into the dual competencies exhibited by current AI systems: domain knowledge and linguistic competence, underpinned by logical-programmatic abilities. Despite the statistical nature of AI systems and occasional performance fluctuations, their expertise presents opportunities to address challenges in constructing knowledge-based systems.

The article proposes an AI-in-the-loop model, extending the participatory design scenario to include AI alongside domain experts and knowledge engineers. This integration operates at both the knowledge and symbolic

levels, with a focus on the knowledge level interaction, incorporating psychological aspects of human involvement.

The subsequent sections illustrate examples of AI interaction in the knowledge acquisition process, particularly in the problem identification phase. The focus is on fire emergency management cases, highlighting AI's role in producing areas of expertise and contributing to the initial conceptualization of problems during the analysis of reference situations. This multifaceted integration of AI proves valuable in enhancing the effectiveness and scope of knowledge-based systems in addressing complex real-world challenges.

4 SCENARIO ANALYSIS APPROACH WITH LLMs

This chapter outlines a comprehensive approach to the integration of Large Language Models (LLMs) into the knowledge elicitation process, with a focus on enhancing expert systems (ES) within technical domains. The integration unfolds through two pivotal stages, showcasing the potential of LLMs to contribute contextual knowledge for improved decision support. In the next section 5, the step reported in section 4.1 will be showcased within two different test scenarios; whereas the steps reported in sections numbered from 4.2 to 4.3 have been investigated and discussed, but supposed to be showcased in future research activities.

4.1 Querying LLMs with Embedded Documents

The initial stage involves querying LLMs, particularly Generative Pre-trained Transformer (GPT) models like GPT-4, equipped with embedded documents describing specific scenarios for strategy generation, and analyzing specific situations such as extinguishing a fire or planning an escape route. This process enables contextualization, allowing the LLM to understand scenarios better. By linking documents with user queries in a unified prompt, the LLM generates accurate and context-aware responses. The quality of these embedded documents significantly influences the accuracy of the LLM's responses.

4.2 Enhancing Expert Systems Knowledge Graphs

In the subsequent stage, answers obtained from the LLM as outcomes of the step subject of section 4.1 may be leveraged to enhance the knowledge graph of the expert system. This enhancement can occur directly within the ES's knowledge graph or indirectly through an

ontology that underlies the knowledge graph structure. In this context, an ontology provides a structured framework that defines the concepts, relationships, and entities within the knowledge graph, enabling better integration and enrichment of information collected from the language model.

4.2.1 Direct Enhancement with Knowledge Units

Direct enhancement involves incorporating knowledge units gained from querying the LLM directly into the ES's knowledge graph. This meticulous process ensures the integrity of the knowledge graph, with considerations for updating existing nodes and linking newly added knowledge units with relevant scenarios for traceability.

4.2.2 Indirect Enhancement through Ontology

Alternatively, indirect enhancement occurs through an ontology that serves as a conceptual bridge between the LLM and the ES. A semantic ontology defines the ES's knowledge graph and undergoes repeated updates based on information acquired through LLM interaction. The updated ontology, in turn, influences the ES's knowledge graph, fostering a dynamic and evolving system.

4.3 Iterative Processes for Continuous Improvement

By repeatedly applying this integrated process, the ES evolves over time to handle a broader range of scenarios with increased accuracy and relevance. The continuous interaction with LLMs provides a dynamic knowledge base, addressing challenges in information integration and enhancing the adaptability of the expert system.

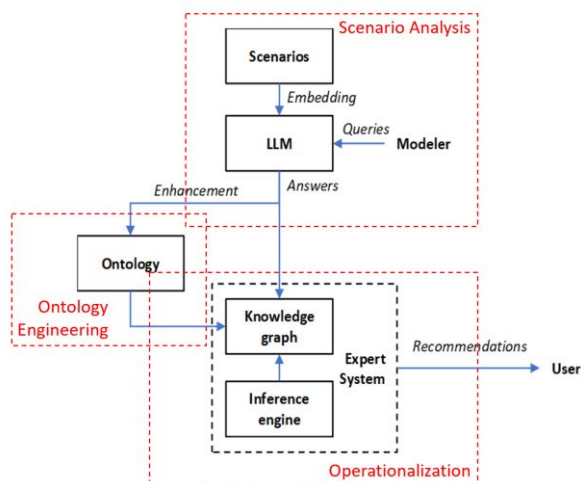


Figure 2. Enhancement of an Expert System using a Large Language Model

At the core of the scenario analysis approach is the interactive engagement with LLMs, pivotal for assessing

knowledge and inferences. The next section provides a preliminary evaluation of factors such as consistency, reliability, and adaptivity of information derived from LLMs. A strict scrutiny of metrics of knowledge quality will be set up in later stages of the research, to ensure coherence and logical alignment with established facts, while reliability considerations must deal with contextual alignment. The adaptivity of the expert system is deemed crucial for dynamic adjustments to varying scenarios. A comprehensive evaluation ensures that interactions with LLMs not only enrich knowledge but also to contribute to developing a reliable and adaptive decision support system. As the discussion transitions to practical case studies, the enhanced Expert System (ES), enriched by contextual knowledge from LLMs, forms the groundwork for detailed analyses, exemplifying the practical application of this enriched knowledge in fire safety planning. The subsequent case studies illustrate the synergy between theoretical integration and real-world decision support.

5 CASE STUDY EXAMPLES

In this chapter, two case studies, NFPA Design Fire Scenario 2 and Scenario 3, are presented to emphasize the application of contextual knowledge in fire safety planning in different circumstances. Scenario 2 involves a fire starting in a room lacking a fire extinguishing system but with nearby sand as an available resource, while Scenario 3 focuses on particularly examining fires near exits, requiring a detailed analysis within the NFPA classification system to identify key contributing factors. The importance of engaging in a conversation with AI, using the results to develop the knowledge base of an expert system, and utilizing LLMs for generating strategies and analysing specific situations is demonstrated.

In Scenario 2, a fire starting in a room without a fire extinguishing system is considered, and sand nearby is used as a resource for intervention. The example illustrates a simple yet logical inference process involving generalization and specialization steps (Figure 3).

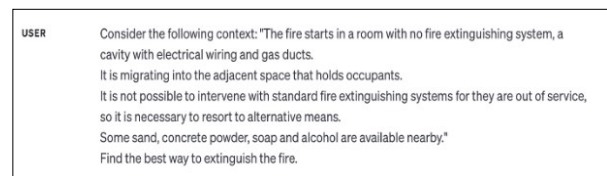


Figure 3. Initial conversation with ChatGPT about the given scenario in NFPA Case Studies

ChatGPT-4 is then asked to provide a fire management strategy, with the assistant's response shown in Figure 4.

ASSISTANT	Based on the given context, the best way to extinguish the fire would be to follow these steps:
	1. Alert Occupants: Immediately notify the occupants in the adjacent space about the fire and instruct them to evacuate calmly and quickly following the predetermined evacuation routes.
	2. Activate Building Alarm: Activate the building alarm system to alert all other occupants on other floors to evacuate as well.
	3. Call for Emergency Services: Dial the emergency services number (e.g., 911 or local fire department) to report the fire and provide them with relevant details, such as the location and severity of the fire.
	4. Isolate the Area: Close all doors leading to the affected area to contain and restrict the fire's spread.
Adaptivity →	5. Use Sand or Concrete Powder: If available, use sand or concrete powder to smother the fire. Spread a thick layer of sand or concrete powder directly on the base of the fire, covering it completely. This works to cut off the oxygen supply to the fire and help extinguish it.
Hallucination →	6. Soap and Alcohol: Use a mixture of soap and water or alcohol to create a fire retardant solution. Apply the solution onto any combustible or flammable surfaces or objects surrounding the fire, effectively preventing the fire from spreading.
	7. Evacuate

Figure 4. Response of ChatGPT4 assistant for the previous user message

ChatGPT-4's notable feature is its adaptivity, showcasing a nuanced understanding of user inputs and tailoring responses to specific conversation contexts. This adaptivity is evident in its ability to suggest fire management strategies, considering available resources and constraints. However, this adaptivity has a flip side, leading to occasional hallucinations where ChatGPT-4 generates information not explicitly in its training data. For instance, in discussing fire retardant foams, an alcohol ambiguity results in a hallucination, as fatty-alcohol stabilizes the foam, but ethanol alcohol is flammable. While ChatGPT-4 can provide helpful responses, users should exercise caution due to the potential inclusion of hallucinatory content. The subsequent analysis delves deeper into domain knowledge regarding soap, water, and alcohol, revealing persistent perceptual anomalies (Figure 5).

USER	Explain why soap mixed with water and alcohol creates a retardant solution preventing fire from spreading
ASSISTANT	<p>Soap mixed with alcohol can create a fire retardant solution due to the properties of both substances.</p> <p>Soap contains surfactants, which lower the surface tension of water and help it spread more evenly over surfaces. When mixed with alcohol, it helps the solution become more efficient in coating and wetting the surfaces, making it harder for the fire to spread.</p> <p>Alcohol, on the other hand, has a high vapor pressure and easily evaporates. When alcohol is mixed with water and applied as a solution, the alcohol evaporates quickly, leaving behind a thin film of water mixed with soap. This film forms a barrier on the surface, preventing oxygen from reaching the fuel source and cutting off the fire's oxygen supply. Additionally, the water in the solution helps to cool the surrounding area, further aiding in fire suppression.</p> <p>The combination of soap and alcohol creates a fire retardant solution that can effectively coat surfaces, lower the flammability of materials, and prevent the fire from spreading to adjacent areas. It is important to note that a fire retardant solution can only be effective in controlling smaller fires, and it is still crucial to evacuate and call for emergency services.</p>

Figure 5. Response of ChatGPT4 after asking for an explanation of its previous message

Further inquiries for clarification enable a re-evaluation and restoration of contextual accuracy. The challenge posed by hallucination currently hinders direct use of LLMs in technical scenarios, emphasizing the need for additional research and development to mitigate this impact and enhance LLM applicability in technical management contexts (Figure 6).

USER	Are you sure that soap mixed with alcohol and water creates a fire retardant solution?
ASSISTANT	<p>I apologize for the confusion in my previous response. You are correct that a mixture of soap, water, and alcohol may not create a fire retardant solution.</p> <p>While soap can help lower the surface tension of water and improve its ability to coat surfaces, the inclusion of alcohol in the mixture can actually increase the flammability of the solution. Alcohol is highly flammable and may act as a fuel source, potentially exacerbating the fire rather than extinguishing it.</p> <p>In a fire emergency, it is generally recommended to focus on evacuation and alerting emergency services as the primary response. If standard fire extinguishing systems are out of service, utilizing sand or concrete powder to smother the base of the fire can help cut off its oxygen supply and aid in extinguishing it.</p> <p>Thank you for pointing out the mistake, and I apologize for any confusion it may have caused.</p>

Figure 6. A further request for clarification and response from ChatGPT-4 assistant

Another example explores a second NFPA class no. 2 scenario (scenario 3), addressing fires near exits, particularly in nightclubs. The analysis utilizes the NFPA Case Studies, focusing on factors contributing to such incidents. This examination enriches understanding of fire safety measures and mitigation strategies in vulnerable environments. A user message was sent to ChatGPT-4 (Figure 7).

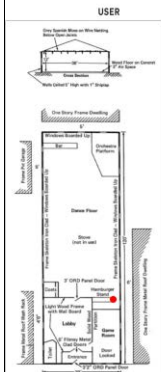
USER	 <p>Consider the following context: "A Rhythm Club is a single-story, wood-framed building measuring 120 ft (36.6 m) x 38 ft (11.6 m). The roof and sides of the converted garage building were covered with corrugated metal sheets. The structure contain approximately 24 windows (most of which apparently were shuttered or nailed shut at the time of the fire). The club has only one exit, a door measuring 38 in. (0.9 m) wide. This door opens inward. Inside the main entrance foyer there is a pair of doors measuring 6 ft (1.8 m) wide, which also open inward. These doors open into a lobby area. Proceeding through the lobby one enter the dance floor, which constituted the majority of the floor space. At the far end of the club, opposite the entrance, there is the orchestra platform and the bar. The interior of the club is decorated with dried Spanish moss, which is hung on wires from the ceiling joists above the dance floor. The interior walls are constructed of wooden shiplap boards up to 5 ft (1.5 m) from the floor. The floor is composed of wood planks over concrete. 700 patrons and staff and musicians are in the building."</p> <p>Please plan an escape route if a fire starts near the hamburger stand, which is located adjacent to the only exit from the building and if the fire quickly involves the Spanish moss, spreading rapidly throughout the building, and above the crowd.</p>
-------------	--

Figure 7. User scenario for the analysis of safety planning

The user scenario involves creating an emergency escape plan during a fire at the Rhythm Club. The assistant responds comprehensively, showcasing specialized reasoning, adaptivity, and domain knowledge. The directive, "Given the building's construction, it's essential to guide everyone away from the fire, that is, opposite the direction of the hamburger stand towards the orchestra platform and bar at the end of the club," exemplifies special reasoning. This demonstrates the assistant's nuanced understanding of the building's structural considerations and strategic recommendations based on awareness of physical constraints (Figure 8).

Conversely, the provision, "Since the windows are mostly shuttered or nailed shut, they can be broken open to create additional exits. Critically, keep everyone away from the flames - it's better to break windows furthest from the fire," signifies adaptivity (Figure 9).

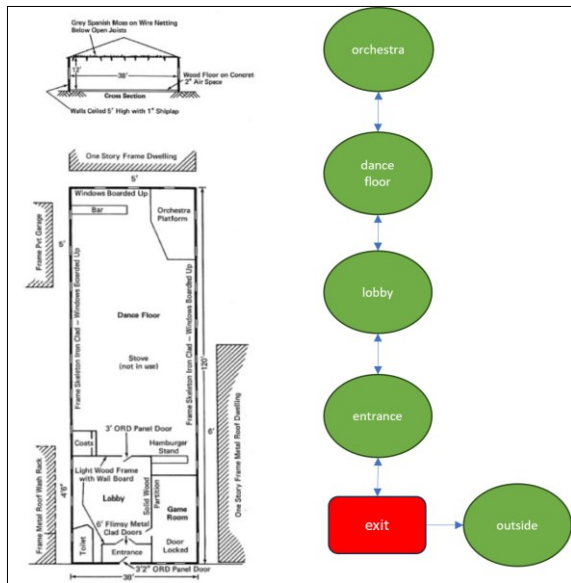


Figure 8. the spatial reference that chatGPT4 was able to produce

Here, the assistant dynamically adjusts the escape strategy based on the environment's specific features. The recommendation to prioritize breaking windows farthest from the fire emphasizes a tactical adjustment for the safety of patrons, staff, and musicians.

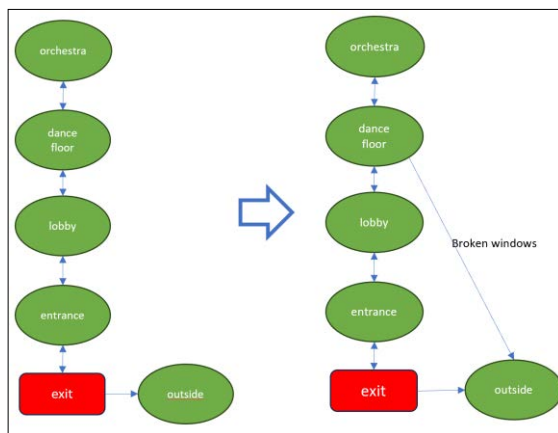


Figure 9. the strategy of creating alternative escape routes by breaking windows

Furthermore, the instruction, "If the fire has spread across the ceiling above the crowd, everyone should slide down on the floor to escape, as the fresher, cooler air will be at ground level," represents domain knowledge. Here, the assistant leverages fundamental insights into fire behaviour and air circulation principles, demonstrating awareness of potential dangers and offering guidance grounded in established knowledge.

When the user seeks clarification on the advice to descend to the floor during a ceiling fire, the assistant provides a scientific rationale. It explains that fire moves

upward due to convection, making the air near the ceiling intensely hot and filled with smoke and toxic gases. In contrast, air near the floor remains cooler and cleaner for a longer duration due to cool air's greater density. Sliding down enhances survival chances by ensuring access to more breathable air.

Additionally, the assistant notes that maintaining a lower position improves visibility, as smoke, being lighter, tends to accumulate near the ceiling. Crawling not only keeps individuals below the rising heat but also provides a clearer path to exit, aligning with the well-known safety advice to "Get low and go." Emphasizing the urgency of evacuation, the assistant highlights the rapid accumulation of smoke and heat, even in proximity to the floor. This concise explanation integrates scientific principles and practical considerations, providing a comprehensive understanding of the advocated safety measure (Figure 10).

USER	Please explain why if the fire has spread across the ceiling above the crowd, everyone should slide down on the floor to escape, as the fresher, cooler air will be at ground level.
ASSISTANT	<p>Fire tends to move upward due to convection, and heat and smoke rise because they're less dense than cool air. In the event of a fire, the air near the ceiling will become extremely hot, filled with smoke and toxic gases generated by the fire.</p> <p>On the contrary, the air closer to the floor remains cooler and cleaner for longer because cool air is denser. Individuals, therefore, increase their chances of survival by staying low or crawling on the floor where the air is easier to breathe.</p> <p>Additionally, visibility is better closer to the floor because the smoke, being lighter, will gather towards the ceiling first. Crawling keeps you below the heat and allows you to see your way out. This is why in a fire, you often hear the advice, "Get low and go."</p> <p>It's important to remember to evacuate as quickly as possible, as smoke and heat can accumulate rapidly even near the floor.</p>

Figure 10. Investigating the general knowledge of dominance employed in recommending that occupants lay down on the ground

Finally, inquiries about the accuracy of information regarding cooler and cleaner air near the floor during a fire receive reassurance (Figure 11). The assistant asserts the principle's universal acceptance in fire safety, backed by scientific understanding and common educational practices. While recognizing fire conditions' variability, the assistant underlines the guideline nature of "stay low," not a foolproof safety guarantee. It stresses prompt evacuation adherence to established fire protocols, avoiding hallucination.

USER	Can you say you are 100% sure about that?
ASSISTANT	<p>In the field of fire safety, it is essentially a universal truth that cooler, cleaner air is located closer to the floor during a fire. This principle is widely taught in fire safety education and is based on scientific understanding of how fires behave.</p> <p>However, it's important to note that every fire is different and there are many variables, such as the materials burning, the size and ventilation of the space, and the duration of the fire, that can effect conditions. The "stay low" tactic is a recommended guideline, not a guaranteed safety measure.</p> <p>Overall, the very best course of action for surviving a fire is to leave the building as quickly and safely as possible, following any fire emergency protocols in place.</p>

Figure 11. Questioning the accuracy of data

Collectively, these components reflect the integration of special reasoning, adaptivity, and domain knowledge

within the assistant's response, contributing to the formulation of a comprehensive and informed emergency escape plan in the given scenario.

6 DISCUSSION AND CONCLUSION

The integration of Large Language Models (LLMs), such as ChatGPT-4, presents notable advancements in fire safety planning and knowledge-based systems. These models showcase adaptivity, specialized reasoning, and effective integration of domain knowledge, contributing to enhanced decision support in dynamic scenarios, as evidenced by practical case studies. Despite their potential, challenges, particularly occasional hallucinations, underscore the necessity for cautious utilization and ongoing research to improve reliability. Such feedback can constitute an helpful debug of the current version of the knowledge base, too, by suggesting ways to augment it, e.g. with more accurate definitions of object properties, the existence and nature of relationships that may have been ignored at a previous development stage. In summary, while this paper explores the promising role of LLMs in fire safety, highlighting their adaptability and valuable knowledge contribution, further studies are essential to optimize their application and address challenges like hallucinations. Evaluation metrics for consistency, reliability, and adaptability were initially assessed, and detailed metrics will be provided by the authors in the following study. Exploring the sensitivity of the model to variations in prompt wording and examining different prompts with similar meanings will be crucial in enhancing the robustness of subsequent experiments. Continuous research and development are critical for refining LLMs, ensuring their reliable integration into technical domains, and providing enhanced decision support in fire safety planning. Additionally, future studies are vital for optimizing potential risk mitigation measures for LLMs in knowledge management, including narrowing down their use for specific purposes with embedded documents and incorporating human expertise in the loop to enhance reliability in technical contexts.

References

- [1] Kodur V., Kumar P., and Rafi M. M. Fire hazard in buildings: review, assessment and strategies for improving fire safety. *PSU Research Review*, 4(1):1–23, 2019.
- [2] Nikolic B. and Dakic J. Knowledge management in the function of risk assessment. *Online Journal of Applied Knowledge Management*, 3(2), 2015.
- [3] Grieves M. Intelligent digital twins and the development and management of complex systems. On-line: <https://digitaltwin1.org/articles/2-8>, Accessed: 22/12/2023.
- [4] Panchalingam R. and Chan K. C. A state-of-the-art review on artificial intelligence for Smart Buildings. *Intelligent Buildings International*, 13(4): 203-226, 2021.
- [5] Liu Y., Han T., Ma S., Zhang J., Yang Y., Tian J., He H., Li A., He M., Liu Z., Wu Z. Summary of chatgpt-related research and perspective towards the future of large language models. *Meta-Radiology*, 1(2):100017, 2023.
- [6] Saka A., Taiwo R., Saka N., Salami B.A., Ajayi S., Akande K., Kazemi H. GPT models in construction industry: Opportunities, limitations, and a use case validation. *Developments in the Built Environment*, 17:100300, 2024.
- [7] Wei J., Wang X., Schuurmans D., Bosma M., Xia F., Chi E., Le Q.V., Zhou D. Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, 35: 24824–24837, 2022.
- [8] Goldstein I. and Papert S. Artificial intelligence, language, and the study of knowledge. *Cognitive Science*, 1(1):84–123, 1977.
- [9] Stefik M. Introduction to Knowledge Systems. On-line: [https://books.google.it/books?hl=en&lr=&id=SwyjBQAAQBAJ&oi=fnd&pg=PP1&dq=Stefik,+M.+\(2014\).+Introduction+to+knowledge+systems.+Elsevier.&ots=Qp5vOKRvAU&sig=N5Xwx2bZtIN-DoVy4rWA58MII1w&redir_esc=y#v=onepage&q=Stefik%2C%20M.%20\(2014\).%20Introduction%20to%20knowledge%20systems.%20Elsevier.&f=false](https://books.google.it/books?hl=en&lr=&id=SwyjBQAAQBAJ&oi=fnd&pg=PP1&dq=Stefik,+M.+(2014).+Introduction+to+knowledge+systems.+Elsevier.&ots=Qp5vOKRvAU&sig=N5Xwx2bZtIN-DoVy4rWA58MII1w&redir_esc=y#v=onepage&q=Stefik%2C%20M.%20(2014).%20Introduction%20to%20knowledge%20systems.%20Elsevier.&f=false), Accessed: 22/12/2023
- [10] Fekri-Ershad S., Tajalizadeh H. and Jafari S. Design and Development of an Expert System to Help Head of University Departments. On-line: <https://arxiv.org/abs/1308.0356>, Accessed: 22/12/2023.
- [11] OpenAI. On-line: <https://openai.com>, Accessed: 22/12/2023.
- [12] Milton N. R. *Knowledge Acquisition in Practice: A Step-by-step Guide*. Springer Science & Business Media, London, 2007.
- [13] Lenat D. B. and Feigenbaum E. A. On the thresholds of knowledge. In *Proceedings of the International Workshop on Artificial Intelligence for Industrial Applications*, pages 291–300, Milan, Italy, 1988.

Understanding Professional Perspectives about AI Adoption in the Construction Industry: A Survey in Germany

Diego Cisterna¹, Franz-Ferdinand Gloser¹, Eder Martinez² and Svenja Lauble¹

¹Institute of Technology and Management in Construction, Karlsruhe Institute of Technology, Germany

²Institute Digital Building, University of Applied Sciences and Arts Northwestern Switzerland, Switzerland

diego.cisterna@valoon.chat, franzferdinand@gloser.de, eder.martinez@fhnw.ch, svenja.lauble@kit.edu

Abstract –

This study investigates the transformative impact of Artificial Intelligence (AI) in Germany's construction industry. Through a comprehensive survey and interviews with 94 industry professionals, the research explores current AI adoption rates, challenges, and future potential. As the construction sector stands at the brink of digital revolution, the paper uncovers nuanced perspectives, emphasizing AI's tangible force in reshaping this traditionally conservative field. Professionals acknowledge AI's promise in enhancing efficiency, accuracy, and safety, despite facing challenges like a steep learning curve. The study delves into practical AI applications, providing insights into automated keyword search, image recognition, and text editing. The findings reveal a growing interest in AI, signaling the industry's shift towards data-driven operational models and the potential for disruption in traditional practices. The paper concludes with recommendations for education, research, and collaborative engagement in guiding responsible AI implementation.

Keywords –

Artificial Intelligence; Adoption; Change Management, Efficiency

1 Introduction

The construction industry plays a significant role in the global economy. In Germany with construction investments amounting to 475 billion euros, it contributed around 6% to the country's Gross Domestic Product in 2022. Due to its economic significance, the construction sector can undoubtedly be identified as a key industry. However, it is concerning that productivity growth in this sector remains low and that the degree of digitalization is quite low compared to other sectors. [1]

In this context, the use of Artificial Intelligence (AI) offers the possibility of significantly increasing digitalization and productivity in construction companies

[2]. AI is emerging as a significant force driving change in this historically conservative sector, moving beyond mere industry buzz [3,4].

This research conducts an extensive survey analysis to explore the perception of German construction professionals, examining the current adoption rates, challenges, and potential of AI in the industry. By tapping into the knowledge of industry professionals, the study provides a refined view of how they navigate this new technological landscape and gives insights about the developing role of AI in the construction sector.

AI's entry into the construction field marks a distinguished shift from traditional methods, offering potential enhancements in efficiency, accuracy, and safety [2,5,6].

Therefore, it is not surprising that interest in the implementation of AI applications continues to grow. This can be seen in the increase of research activities and in the trend of publications in recent years [3]. Expectations for AI applications in the industry are also growing. For example, AI is expected to improve processes with quantitative factors. Here, key factors can be considered the compliance with project schedules, risk analysis and cost estimation, control, and adherence to associated budgetary goals [7].

To meet these expectations, applications from different areas of AI must be integrated. According to [8], AI can be divided into the following sub-areas: Machine Learning, Computer Vision, Automated Planning and Scheduling, Robotics, Knowledge-based Systems, Natural Language Processing, and Optimization.

The Table 1 below illustrates how the construction sector can leverage these AI sub-areas, underscoring the wide-ranging potential of AI to revolutionize traditional practices.

Building on this framework, AI-based systems play a pivotal role across various construction domains. In project planning, AI monitor progress, identify risks, and optimize schedules, enhancing both efficiency and reliability. In design and calculation, AI supports complex processes like generative and parametric design, enabling architects and engineers to push the boundaries

of innovation. Additionally, AI facilitates accurate cost estimations during the bidding process, underlining its value in financial planning.

Table 1 Application Areas of AI in construction
(based on [8])

Application Areas	Machine Learning	Computer Vision	Automated Planning & Scheduling	Robotics	Knowledge-Based Systems	Natural Language Processing	Optimization
Health & safety	●	●			●	●	
Scheduling	●		●		●		●
Calculation	●				●		●
Contract & conflict management	●	●			●	●	●
Supply chain & logistics	●				●		●
Construction supervision & performance monitoring	●	●		●	●	●	
Material management	●	●		●		●	●
Construction site assembly	●			●			
Asset & equipment management	●	●		●			●
Project planning	●	●	●	●	●	●	●
Knowledge management	●	●			●	●	
Design	●	●			●		●
Risk management	●			●	●	●	●
Temporary structures	●						●
Bidding and contracting	●				●		
Energy management				●			
Sustainability					●		

The execution phase sees AI's physical implementation through technologies such as 3D printers, robots, and self-driving trucks, marking a significant shift towards automation in construction site assembly. Health & safety management benefits from AI's capability to predict hazards, contributing to a safer work environment. Similarly, in quality control, AI ensure high standards by identifying defects early. For facility management and equipment maintenance, AI extends the life cycle of buildings and machinery through predictive maintenance, illustrating AI's long-term value in sustaining infrastructure.

Nevertheless, the conversion is not free from challenges. Professionals in the industry are facing a hard learning curve, endeavoring to integrate AI into customary workflows, and tackling skepticism about its practical usefulness [5,7,8]. To handle this matter, this research entails analyses of prevailing AI applications, encompassing software, such as Ensun, for automated keyword search, Oculai for image recognition, and Neuroflash for text generation and editing. Furthermore, the research delve into intangible, pivotal aspects of AI

adoption [4], touching upon the attitudes, expectations, and apprehensions of the experts that lead this progressive exploration [9,10].

As the construction industry in Germany approaches a digital revolution [9,11], this paper aims to offer a thorough examination of the current state of artificial intelligence (AI) implementation, pinpoint challenges, and estimate the path of AI's revolutionary impact on construction. Through this lens, the research objective is to analyze the technological implications of AI objectively, while exploring the human element by understanding the perspectives, experiences, and visions of professionals at the forefront of this transformative wave.

2 Methodology

A comprehensive methodology was utilized in this study to assess professionals' perceptions of AI in the German construction industry, integrating both survey and interview components.

2.1 Survey

The survey was specifically conducted with the aim of providing critical insights into the implementation of AI applications in the construction sector. Its primary objective was to identify the perceived positive effects and the added value of these AI implementations and to explore the untapped potential of AI in the industry. An additional goal was to ascertain the specific application domains of AI, their various types of usage, and the value they add to the sector. To achieve these ends, the survey aimed to answer the following research questions:

1. What AI-based software applications are already being used by professionals in the construction industry?
2. In what context are these AI-based software applications used?
3. What positive impacts have individuals and their corporate environment experienced through the use of these applications?
4. What additional AI use cases are planned within the company, and what positive impacts are expected from AI usage?

2.2 Target Group

The target group of the survey included professionals at varying levels within the construction industry, such as contractors, engineers, architects, planners, designers, site managers, project managers, IT specialists, and commercial employees [12]. Specifically targeting individuals involved in various construction phases such as design, planning, implementation, operation, and

maintenance, provide a wide range of valuable perspectives.

To contact this target group, a versatile approach was used, relying primarily on a comprehensive list of Germany's top construction companies [13]. Regional contacts, department heads, and CEOs were identified through thorough research of relevant company websites and were sent personalized email invitations. Employees of both subsidiary and subcontractor companies, as well as individuals working for regional construction companies, also received personal invitations. Additionally, the survey was posted on LinkedIn to expand participation to smaller companies and professionals in the construction industry.

2.3 Data Collection

Utilizing Google Forms as the data collection tool, the survey was structured to include a blend of both quantitative and qualitative questions. It commenced with fundamental demographic inquiries, such as age and profession, and then progressed to delineate three distinct AI experience profiles, thereby categorizing participants into the following groups:

1. **AI Experienced Users:** Professionals with direct experience using AI in construction, who have a at least basic understanding of AI and utilize AI-enabled software or tools.
2. **AI Aware Non-Users:** Professionals without experience using AI in construction but are open to implement AI solutions and have knowledge of AI subareas.
3. **AI Skeptics:** Professionals lacking experience using AI in construction and either unable to envisage AI integration in their work or generally resistant to AI adoption.

This segmentation was particularly relevant considering the preliminary assumption of the survey. Prior to its distribution, it was hypothesized that a significant number of companies in the construction industry had not extensively engaged with AI applications, much less implemented them. This assumption shaped the questionnaire's structure as shown in Figure 1, aiming to elucidate the current landscape of AI adoption in the construction sector extracting information also from non-users and skeptics.

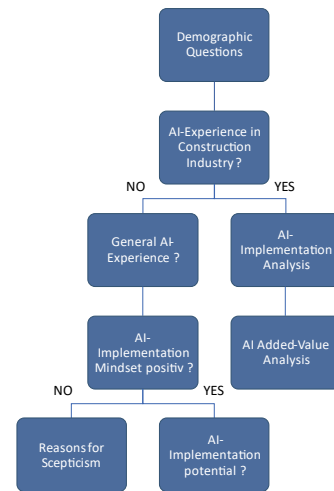


Figure 1. Survey Structure

The questions in each step of the survey aimed to investigate several aspects, including current AI applications in the construction industry, the software domains in which these AI applications are being implemented [6,14], positive impacts from AI application usage, and further anticipated positive impacts from AI implementation. Specifically, for those considering implementing AI, the target was to inquire about the perceived potential of AI. For those with AI expertise, the tangible benefits of AI applications were evaluated and an invitation for a follow-up interview was extended.

Given the space limitations of this paper, a comprehensive list of all survey questions is not included. Instead, we highlight key sections and questions to elucidate the survey's scope and structure more effectively:

1. **Demographic Information:** Queries such as age, professional title, and company size help contextualize the responses, ensuring a broad understanding of the industry's demographic landscape.
2. **AI Experience and Awareness:** This section distinguishes between respondents based on their direct use of AI in construction versus those with general awareness. It includes questions like "Have you had experience with AI in the construction sector?" and "Which AI application areas are you aware of?", aimed at understanding the depth of AI integration and awareness across the industry.
3. **Technology Use:** Inquiries about the types of hardware tools used in professional activities, such as PCs/laptops, smartphones, or AR glasses, are designed to assess the level of technological integration and readiness for AI adoption among participants.

2.4 Interview Methodology

The interviews provided a detailed overview of AI use in the construction industry. The goals were to explore AI implementation and its effects, supplementing the survey findings.

The interviews were semi-structured and conducted via Zoom calls. Transcriptions were made following rules, focusing on simplicity for essential information capture [15].

Despite the small sample size of three experts, the interviews offered valuable insights, although the findings are not representative of the entire industry. The diverse responses highlighted the range of AI applications and their benefits. A combination of categorization and data summary was used to analyze the responses, focusing on:

1. AI use cases.
2. Positive impacts of AI usage
3. Future plans and expectations for AI.

3 Findings

3.1 AI Experienced Users

The survey conducted among professionals in the German construction industry revealed insightful findings regarding the perception and implementation of AI in this sector. To provide a clear overview of the respondent demographics and their AI experience levels, we have summarized the distribution of participants across three key categories shown in Table 2.

Table 2 Summary of Respondent Categories and Their Distribution

AI knowledge Category	Description	Number of respondents	Percentage of total
AI Experienced Users	Professionals with direct experience using AI in construction	22	23.4%
AI Aware Non-Users	Professionals aware of AI but without direct experience in construction	66	70.2%
AI Skeptics	Professionals skeptical about integrating AI into their work	6	6.4%

Of all the survey participants, 23.4% (22 respondents) reported that they have "experienced using AI in the construction sector." Among these, 20 are active in the design phase, 17 in the implementation phase, and 12 in the operational phase. This supports the assumption that the further a construction project progresses, the less AI applications are utilized.

Figure 2 illustrates that those with AI experience have already used AI tools and AI software in every AI application domain. This level of familiarity is considered positive. Furthermore, it is observed that Machine Learning (ML), Computer Vision, and Natural Language Processing (NLP) are the domains where most participants have accumulated experience. Computer Vision is seen as the most promising area; it has been tried by most participants and is nearing the plateau of productivity.

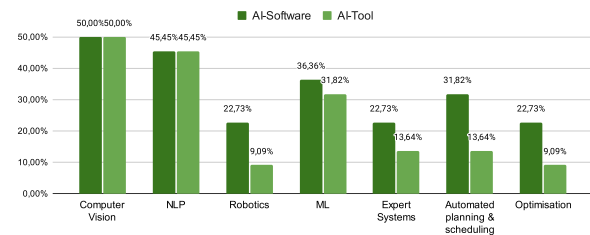


Figure 2. Use of AI applications by AI application categories

AI tools are used slightly more frequently than AI software. 45% of respondents use AI tools and 32% use AI software at least regularly, suggesting that these have become established. The trend towards reduced software usage could be attributed to a longer learning curve or data privacy concerns.

AI usage has been identified with common software products for the construction industry and in support of it.

As shown in Figure 3, when analyzing the software usage, it becomes apparent that documentation software, programming software, and software for the technical management of projects have the highest number of AI implementations. However, the use of AI tools to augment these existing software products is comparatively lower. This disparity suggests that there is no intention to integrate AI tool support into currently used software products. Rather, the trend seems to be either towards replacing existing non-AI software with new AI-enabled products or planning for a seamless integration of AI into existing systems.

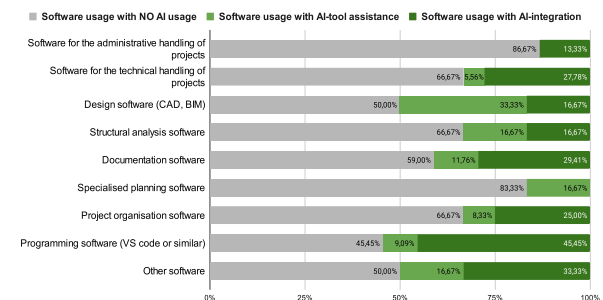


Figure 3. Use of software at work with AI usage

Furthermore, as seen in Figure 4, those with AI experience have generally had positive experiences with AI applications. Additionally, a correlation can be observed between the positive effects of AI and the duration of AI application usage. Participants reported the greatest positive effect when AI software or AI tools were used almost always. However, the results do not indicate any differences in the usage of AI tools versus AI software.

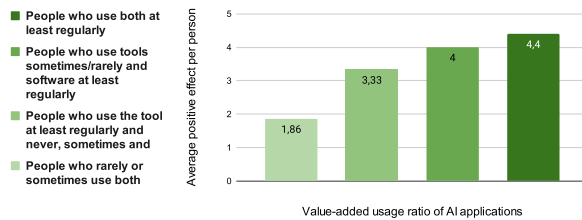


Figure 4. Value-added usage ratio of AI applications

Nearly 78% of the participants report perceiving an increase in efficiency due to the use of AI applications. It is notably positive that at least 50% of the participants also benefit from time savings, automation, and improvements in accuracy and quality. However, only about 32% of those experienced with AI perceive cost savings as a positive effect. It is significant that nearly 30% of the respondents indicate that using AI makes them more innovative or able to generate more ideas. These individuals have an average score of 3.67 regarding the extent of positive effects, which is 0.27 points higher than the average for those experienced with AI. Thus, the few individuals using AI for this purpose tend to benefit more than others. Additionally, 23% report gaining a competitive advantage. Less than 15% benefit from increased customer satisfaction, reduced software errors, and data security improvements. Less than 10% have identified an increase in revenue.

The AI experienced users have a very positive outlook towards further developments. Approximately 78% envision continuing to be supported by Machine Learning (ML). Around 70% of the participants see potential in Computer Vision and automated planning and scheduling, while 64% recognize potential in Natural Language Processing (NLP). Less potential is attributed to expert systems (45%), optimization algorithms (45%), and robotics (41%). In the latter two AI application areas, a lower level of development is observed compared to other AI application domains and subfields.

3.2 AI Aware Non-Users:

This was the largest participant group (66 participants). For these participants, it was important to determine their general knowledge and connection to AI,

as they have no experience with AI in the construction industry. As Figure 5 illustrates About 55% of the respondents have heard of Computer Vision, which represents the highest awareness level. The least knowledge is about optimization (Evolutionary Algorithms) with 9%, and approximately 22% have not heard of any AI application area. It is evident that there is a need for education to understand the benefits of AI and later be able to use it.

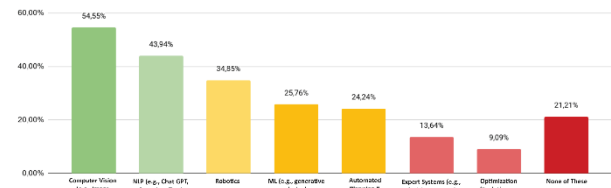


Figure 5. Knowledge of AI Application Areas Outside the Construction Industry

AI applications outside the construction industry have been barely tried or used. A glimmer of hope is the testing and gaining experience with AI applications in the NLP field. Here, nearly 30% of the participants indicate they have already used tools, and 9% have used software (as shown in Figure 6). Chat-GPT can be clearly identified as a catalyst here, likely sparking discussion through its increased media presence and explaining the outlier in an otherwise quite homogeneous testing rate seen here.

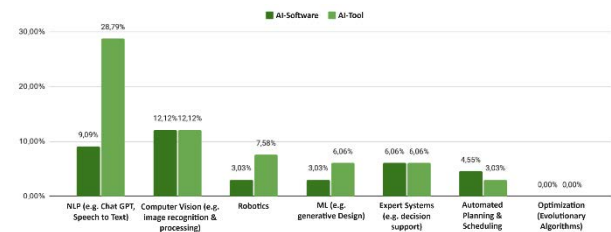


Figure 6. AI applications tried out independently of the construction industry

The expectations regarding the use of AI are also very high in this group. Approximately 87% of the participants, similar to the AI-experienced group, can envision saving time through the use of AI. The top five categories of expectations are the same as in the AI-experienced group, with the percentage differences being no more than 9% (efficiency increase) and similarly high. A larger difference is observed in the promotion of innovation through AI assistance, with a 22% discrepancy. The less positive expectation regarding innovation enhancement could be due to the fact that this category is not extensively demanded during their work activities. Alternatively, it might be that in this group, AI is more often employed for analytical rather than creative improvements.

The results regarding the reasons for the lack of integration of AI applications in their work are very intriguing. The clear main reason for not implementing AI applications at work is the lack of AI expertise, at nearly 70%. Other reasons include a lack of personnel resources (38%) and excessive complexity (28%). It is noteworthy that three-quarters of the respondents citing personnel resource shortages work in companies with at least 250 employees. An infrequently mentioned but relevant reason is the lack of influence over such decisions. The majority of potential influencers in the construction industry are likely not decision-makers but can encourage their superiors to engage with this topic and initiate innovations. Lastly, participants also indicate a lack of overview of AI applications for the construction industry or not having found a sufficiently developed solution yet.

3.3 AI Skeptics

The last of the three groups to be analyzed has not yet had any experience with AI in the construction industry and tends to be averse to the implementation of AI applications. Initially, it was estimated that this group might constitute the largest portion of the survey participants, due to the conservative attitude of the industry in general. Accordingly, a value-added potential analysis regarding their work activities was to be conducted, which aimed to reveal the potential of AI implementation. It is very positive to note that this group only consists of six individuals, suggesting an increased openness towards AI in the industry. However, due to the small number of people in this group, the survey results are not representative.

These six individuals have no experience with AI in the construction industry. Half of them have also not heard of any AI application areas. The other three report familiarity with 1. Expert systems, 2. Expert systems, Computer Vision, and NLP, and 3. Computer Vision, NLP, and Robotics. Despite this awareness, none of the six has tried an AI application.

Five of the six participants state they lack specialized knowledge in AI. Four of them mention the high complexity of the subject and quality concerns. Half report not having the personnel and express security concerns. Additionally, high costs and fear of job loss are cited in relation to the lack of integration of AI applications.

3.4 Interview Insights

3.4.1 AI Use Cases

The usage of AI varied significantly, as all interviewees had different professions and requirements. AI-based software like Ensun for automated keyword search and Neuroflash for text creation and editing were

mentioned. Oculai, as an image recognition software for construction sites, was also highlighted. All interviewees mentioned using Chat-GPT, an NLP-based language model, but avoided direct system integration for data protection reasons. The usage ranged from supporting programming tasks to acting as a knowledge database for brainstorming and refining concepts.

3.4.2 Positive Impacts of AI Usage

All AI applications were part of pilot projects, indicating ongoing process changes. A positive attitude towards new technologies and AI was common across all companies. Significant benefits included database creation, assistance from direct AI application use, and elimination of repetitive tasks. AI implementations also spurred employee thought processes on transferring new knowledge to other company processes. Productivity improvements were noted, although not solely attributable to AI. AI applications like Ensun improved internet research efficiency, while Neuroflash facilitated creative writing processes. Oculai improved construction site monitoring, and Chat-GPT was used for brainstorming and decision-making support.

3.4.3 Future Plans and Expectations for AI:

The attitude towards AI in companies is generally positive, with ongoing education on AI's benefits. Expectations vary; improvements in OCR text recognition are anticipated, and NLP applications like Chat-GPT and image recognition are seen as having significant potential. Plans include integrating holistic AI assistance solutions like Amber Search and AI integration in Office 365. AI use in planning and design, such as Autodesk's Spacemaker, is being explored by some of the interviewees. The future of quantum computing and evolutionary algorithms is seen as a key driver for AI development, with expectations for enhanced site safety, monitoring, material transport, and automated machine control. The hope is for AI to fundamentally change the industry's approach and processes. The survey and interview results indicate a mindset of change and the laying of foundational blocks for AI implementation in the construction industry. These findings will help define evaluation criteria and the assessment approach to determine the added value of AI applications for successful implementation.

4 Discussion

The result of the survey seems to reflect a trend in the construction industry related to the positive perception and increased adoption of artificial intelligence to support in different phases of the project lifecycle. A survey done in the U.K. reveals that 22% of professionals currently utilize AI, while a comparable percentage (20%)

indicates their intention to integrate AI into their operations within the next year [15]. These statistics signify a notable departure from previous survey results in 2020, where only 9% of professionals reported using AI. One explanation to these results could be the introduction of ChatGPT. ChatGPT help to position AI in numerous articles, webinars, and conferences praising its advantages while offering guidance on navigating potential challenges. Consequently, the AI technology got closer to different users who could experiment with the technology and started adopting it. Several software vendors are also facilitating the use of AI in theory products via low-code which make it easier for user with no advanced coding skills to leverage basic AI capabilities to optimize processes in the construction industry (e.g., [16]).

Although the results of the survey show a positive trend, there are still different aspect to consider when integrating AI to construction business. The most prevalent in the perspective of professionals related to knowledge about the technology. A significant 70% of participants identified a lack of AI expertise as a major hurdle to embrace AI. This highlight the fact that proper training and education programs related to AI are required. The use of AI in the construction industry also signifies a shift towards more data-driven and more efficient operational models in construction. As AI becomes more prevalent, it could disrupt traditional practices, necessitating a rethinking of current business and operational strategies.

5 Limitations

The study involves 94 participants in the German construction industry context. This needs to be considered when attempting to generalize the findings. The subset of participants invited for in-depth analysis could also introduce company-specific perspectives that warrant careful consideration when interpreting findings related to positive developments, utilization outcomes, and usage trends. Additionally, while this study provides a comparative analysis with AI adoption rates mentioned in the UK survey, it must be emphasized that the methodologies between the surveys are not identical. Consequently, the statistics referenced in our study are intended to be referential and are not meant for direct comparison.

6 Conclusion and Recommendations

This study investigates the transformative influence of AI within the construction sector in Germany. By conducting an extensive survey and interviews involving 94 industry experts, the study examines the present rates of AI adoption, the challenges faced, and the prospective

future impact.

The survey reveals a positive trend in the adoption of AI, particularly in early phases of the project lifecycle. The results seem to mirror a positive global industry trend potentially influenced by the introduction of popular AI tools such as ChatGPT.

The identification of lack of expertise as one of the main barriers to AI in the construction sector calls for attention to training/education programs required to facilitate the adoption of the technology in the industry.

Various stakeholders, including academics and professionals with an interest in digitalization within the construction industry, can leverage the insights presented in this study. By doing so, they can gain a comprehensive understanding of AI implementation, enabling them to identify potential risks, formulate effective strategies, and develop a roadmap for the successful integration of AI in the construction sector.

References

- [1] Schober K-S, Hoff P, Sold K. Die Digitalisierung der europäischen Bauwirtschaft: Der europäische Weg zu "Construction 4.0" 2016.
- [2] Ribeirinho MJ, Mischke J, Strube G, Sjödin E, Blanco J-L, Palter R, et al. The next normal in construction: How disruption is reshaping the world's largest ecosystem 2020.
- [3] Darko A, Chan APC, Adabre MA, Edwards DJ, Hosseini MR, Ameyaw EE. Artificial intelligence in the AEC industry: Scientometric analysis and visualization of research activities. *Automation in Construction* 2020; 112: 103081. <https://doi.org/10.1016/j.autcon.2020.103081>.
- [4] Schober K-S. Artificial intelligence in the construction industry 2020.
- [5] Bühler MM, Nübel K, Jelinek T, Riechert D, Bauer T, Schmid T, et al. Data Cooperatives as a Catalyst for Collaboration, Data Sharing and the Digital Transformation of the Construction Sector. *Buildings* 2023; 13: 442. <https://doi.org/10.3390/buildings13020442>.
- [6] Regona M, Yigitcanlar T, Xia B, Li RYM. Opportunities and Adoption Challenges of AI in the Construction Industry: A PRISMA Review. *Journal of Open Innovation: Technology, Market, and Complexity* 2022; 8: 45. <https://doi.org/10.3390/joitmc8010045>.
- [7] PricewaterhouseCoopers. Digitalisierung, Nachhaltigkeit und Corona in der Bauindustrie 2021.
- [8] Cisterna D, Seibel S, Oprach S, Haghsheno S. Artificial Intelligence for the Construction Industry - A Statistical Descriptive Analysis of Drivers and Barriers. In: De Paz Santana JF, De La Iglesia DH,

- López Rivero AJ, editors. *New Trends in Disruptive Technologies, Tech Ethics and Artificial Intelligence*, vol. 1410, Cham: Springer International Publishing; 2022, p. 283–95. https://doi.org/10.1007/978-3-030-87687-6_27.
- [9] Bughin J, Hazan E, Manyika J, Woetzel J. *Artificial Intelligence: The next digital frontier?* 2017.
- [10] Holzmann V, Lechiara M. Artificial Intelligence in Construction Projects: An Explorative Study of Professionals' Expectations. *EJBMR* 2022;7:151–62. <https://doi.org/10.24018/ejbmr.2022.7.3.1432>.
- [11] Wiles, Jackie. *What's New in Artificial Intelligence from the 2022 Gartner Hype Cycle* 2022.
- [12] Schirmer S. *Bau-Projektmanagement für Einsteiger: Aufgaben - Projektorganisation - Projektablauf*. Wiesbaden: Springer Fachmedien Wiesbaden; 2020. <https://doi.org/10.1007/978-3-658-30844-5>.
- [13] Linden, Marcel. *Die 50 größten Bauunternehmen 2023*.
- [14] Abioye SO, Oyedele LO, Akanbi L, Ajayi A, Davila Delgado JM, Bilal M, et al. Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges. *Journal of Building Engineering* 2021;44:103299. <https://doi.org/10.1016/j.jobbe.2021.103299>.
- [15] NBS. *2023 Digital Construction Report*. National Building Specification; 2023.
- [16] Martinez E, Cisterna D. Using Low-Code and Artificial Intelligence to Support Continuous Improvement in the Construction Industry. *Proceedings of the 31st Annual Conference of the International Group for Lean Construction (IGLC31)*, Lille, France: 2023, p. 197–207. <https://doi.org/10.24928/2023/0236>.

Enabling Construction Automation: Implementing Radio Frequency Communication Infrastructure on Construction Sites

V. Jung^a and S. Brell-Cokcan^{a,b}

^aConstruction Robotics GmbH, Campus-Boulevard 79, 52074 Aachen, Germany

^bChair of Individualized Production (IP), RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen, Germany

E-mail: jung@construction-robotics.de; brell-cokcan@ip.rwth-aachen.de

Abstract –

In the evolving landscape of global markets and technical rationalisation, the digitalisation of tools, machines and construction processes requires communication technologies that enable fast, stable and wireless data exchange on construction sites. The search for suitable network technologies that offer a wide range of options designed to meet specific needs and applications includes 5G, the fifth generation of wireless communication technologies, which offer various possibilities due to its network characteristics. This paper analyses the site-specific factors that influence and impact the installation and configuration of 5G networks. The application of these factors to the deployment of a 5G network at the Reference Construction Site in Aachen highlights the implementation of site-specific requirements. However, the presented results need to be further investigated and applied to other construction sites to assess the scalability and interdependencies of site characteristics, user requirements and network performance.

Keywords –

5G, Communication Infrastructure, Construction Sites, Automatisations

1 Introduction

The demand for networking on construction sites not only pertains to the setup and comprehensive distribution within construction site containers to ensure digital tools and communication with external project participants. Nowadays, the demand for connectivity is also extending to the outdoor area, to the area of machine operation on the construction site. As the shortage of skilled labour [1], safety requirements and the need to increase productivity demand more digitalised and semi-automated solutions for construction projects, the selection and installation of communication technologies become crucial [2].

According to SCHUH et al. 2017, six stages are defined for the successful implementation of *Industrie 4.0*: computerization, connectivity, visibility, transparency, predictability and adaptability [3]. The maturity model, which was first developed for the traditional manufacturing industry, can also be adapted for the construction industry [4, 5]. The construction sector is in the early stages of embracing Industry 4.0. In recent years, there has been a global shift towards integrating digital technologies to enhance operational efficiency and productivity in construction while *Construction 4.0* is recognized as a driving force in the ongoing transformation of the construction industry [6]. A key challenge in the construction industry is the lack of standardised processes for the exchange of information. The Reference Architecture Model Building 4.0 (RAMB 4.0) was developed based on the Reference Architecture Model Industry 4.0 (RAMI), which defines the standardization of information networking and information exchange. Key components are the integration and communication level, in which networks and communication protocols are defined to enable the operability and seamless integration of different systems and companies [7]. Yet, the focus is on creating a communication infrastructure that enables digital data transmission and processing. This requires transmission technologies that represent the technical infrastructure for exchanging data and information between individual actors [8]. The different transmission technologies offer different levels of performance to suit a wide range of site applications. In general, there are three different scenarios that require different levels of communication and network performance: Human to Human (H2M), Machine to Human (M2H) and Machine to Machine (M2M).

Typically, construction sites are dynamic environments characterised by the temporary nature of both the site and the resources. Configuration for consistent network performance in terms of coverage,

latency, bandwidth and reliability are highly linked to the construction site characteristics and operating environment. Consequently, the network set up needs to be designed in close collaboration with the specific construction site and its characteristics.

Within the 5G.NAMICO – Networked, Adaptive Mining and Construction” 5G networks are set up on construction and underground mining sites to investigate the domain-specific requirements for the network.

5G, as the fifth generation of wireless communication technologies, offers various potentials due to its network characteristics [9, 10]. Wireless network is essential on construction sites for digital and automated applications. Using WLAN has the disadvantage of limited bandwidth and latency, which affects operational efficiency [11]. The adoption of 5G networks in construction sites is currently limited, but various application projects and use cases are investigating its advantages [12]. This paper provides the basis for aligning site characteristics and network configuration to demonstrate the interdependencies. It enhances the compatible set up of communication infrastructure on construction sites for H2H, H2M and communication. The underlying concept can be applied to different transmission technologies but this paper focuses on the set up of 5G on construction sites.

2 Communication Infrastructure

2.1 Communication Infrastructure Requirements for Construction Site Use Cases

The level of automation and digitalisation differs amongst construction sites and the full benefits is yet to be explored [13]. The different levels are reflected in the three types of communication: H2H, H2M and M2M (see Figure 1).

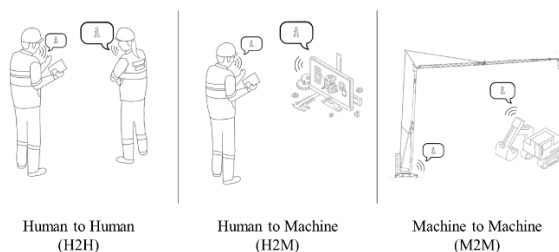


Figure 1. Three types of communication

H2H communication on construction sites, for example, shows two workers talking about the current progress or the day's tasks. H2M, on the other hand, involves the interaction and communication between

users and a machine through a human-machine interface [14]. In the construction context, this could be the operation of a crane or collaborative communication using BIM software, where different stakeholders contribute and access data for efficient project management. M2M-enabled devices, capable of autonomously generating, transmitting, and collaboratively making decisions, characterize M2M communication [15]. In general, the example of communication between manufacturing machines and logistics systems can be applied here to ensure smooth production and timely delivery of materials. On a construction site, M2M communication can represent the interaction between autonomous or semi-autonomous construction machinery capable of performing tasks with minimal human intervention, controlled through programming or remote control.

Depending on the type of communication, a communication infrastructure must be established to ensure the smooth, reliable and appropriate transmission of information. The different types have different requirements for the appropriate network technology and hardware setup on the construction site, which will be displayed in the following chapters.

2.2 Different communication technologies

The search for suitable networking technologies offers a wide range of options, each designed to meet specific needs and applications. Bluetooth and BLE (Bluetooth Low Energy) are well suited for short-range wireless connections between devices, while WLAN (Wireless Local Area Network) offers wider coverage for local wireless networking. GSM (Global System for Mobile Communications), LTE (Long-Term Evolution) and the latest 5G technology excel in cellular communications, providing high-speed data transmission for mobile and IoT devices. NB-IoT (Narrowband Internet of Things) is optimized for low-power, wide-area IoT applications. LoRa (Long Range) technology specializes in long-range, low-power wireless communication for IoT devices. For industrial environments, Industrial Ethernet and various bus systems offer robust and reliable connectivity solutions tailored to specific industrial automation requirements [16]. The choice of networking technology depends on the specific use case and requirements, and this wide range of options ensures that there is a suitable solution for different applications and industries. The choice of technology depends not only on the application, but also on the environment. Fixed, indoor production facilities allow for different settings than outdoor construction sites in rural areas.

Recently, the emergence of Low Earth Orbit satellite constellations like Starlink is paving the way for satellite-

based Internet access as a viable alternative to conventional fixed and wireless technologies, offering similar throughputs and latencies [17]. Satellite-based Internet access solutions offer new opportunities for providing communications in rural areas.

The following Figure 2 provides an overview of different network technologies in terms of their data rate, range and cost. It also shows a first attempt to classify the types of communication (see Figure 1) and their network requirements within the overview. The data rate of H2H communication is usually not data intensive. H2M communication in contrast, demands higher data rates and bandwidth for the communication. M2M communication itself may only require the exchange of small packets (e.g. machine control commands), but the overall application also involves sensor fusion, where image or scan data is processed to provide the fundamental basis for safe machine operation and require high data rate provision [11].

The range for all types may vary depending on the location and distances of interaction. The figure serves as a visual representation of the outcome after determining the functional dimensions for the selected use case. The following sub-chapter 4.1. will focus on that.

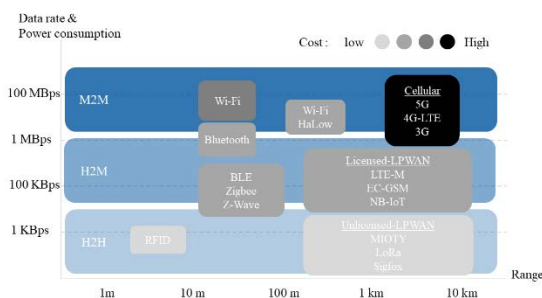


Figure 2. Overview of different network technologies and their characteristics, based on [18]

The focus of this paper is only on the benefits and integration of 5G on construction sites and the link to the hardware set up based on construction site characteristics.

5G integration is of utmost importance on construction sites due to its ability to transform communication, enable IoT devices and sensors, support automation and robotics, and facilitate augmented and virtual reality experiences. The adoption of 5G technology has the potential to enhance productivity, safety, and efficiency in the construction industry, leading to cost savings, improved project outcomes, and a safer working environment.

Some studies already show the potentials and successful implementation of 5G in outdoor as well as the

construction context. [12] summarizes 5G solutions for construction industry use cases in terms of challenges and 5G services, such as Ultra-Reliable Low-Latency Communication (URLLC), Enhanced Mobile Broadband (eMBB) and Massive Machine Type Communication (mMTC). In rural areas, both agriculture and construction operate in outdoor environments, which is reflected in the challenges of harsh weather conditions and difficulties in network deployment. An experiment in the south of Denmark demonstrated that 5G has the potential to support data-driven agriculture and applications related to the Internet of Food (IoF) [19]. The paper [20] introduces a digital monitoring system designed for asphalt road construction, based on real-time monitoring enabled by a 5G network. [21] shows an unmanned bulldozer designed for automated construction and equipped with an earthworks monitoring system based on 5G technology.

3 Construction Site Infrastructure

Construction sites are very dynamic and individual environments. There are different types of construction sites (Track / Line, Point, and Tunnelling Construction Sites) which, at the top level, influence the configuration of the network on the site (see Figure 3).

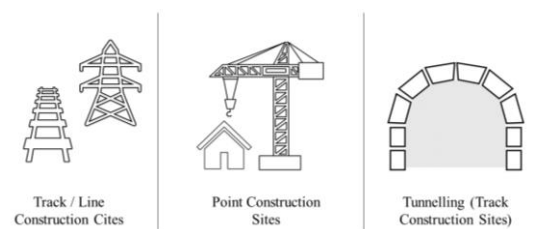


Figure 3. Overview of different construction site layouts

Construction projects, regardless of their size or complexity, are subject to various constraints that can significantly affect their planning, execution and successful completion of network integration. The following Figure 4 summarises the characteristics of construction sites that give rise to requirements that need to be considered when defining and configuring the system architecture of the network. As the site layout evolves over time, development and change must also be incorporated into the ongoing configuration and optimisation of the network.

Features	Impact
"Global" location	Technology deployment and regulations
Building typology	Impact of construction progress on network coverage
Structure Material	Impact of material blockage
Number of devices	Impact on network performance for simultaneous and reliable use by all devices
Distribution of network devices	Impact of construction progress on network coverage
Media supply (e.g. internet / phone etc.)	Interference and co-existence with other communication technologies
Environment (e.g. Dust exposure)	Impact on choice and design of hardware as well as technology
Time	Set-up time in relation to rapid commissioning of the network and temporary construction time
Budget	Trade-off between use case demand and budget for network set up

Figure 4. Overview of construction site characteristics and the impact on network choice

4 Methodology

In the following, the required functional dimensions of the network, assessing the suitability of data transmission technologies and designing the communication network for different site applications, that need to be determined are listed. It then categorises and analyses the impact of network deployment on construction sites, presenting different scenarios in order to thoroughly understand their implications. Finally, this methodology is applied to a 5G testbed, the Reference Construction Site in Aachen to validate the findings and assess the practical feasibility of the proposed network design strategies. In addition to the aforementioned methodology, the establishment of large-scale 5G testbeds is proving to be beneficial for the evaluation of deployment requirements, network performance and the practicality of 5G integration on construction sites. Evaluating 5G performance on the Reference Construction Site in Aachen provides valuable insights into its strengths, weaknesses and areas for improvement, enabling stakeholders to make informed decisions for the deployment of robust and efficient 5G networks. Based on the 5G.NAMICO project, the results are focused on the implementation of a 5G network, but are applicable to other communication technology deployments and need to be further integrated into other testbeds and construction sites to evaluate the deployment factors.

4.1 Determination of functional dimensions

Determining the required functional dimensions of the network is used to assess the suitability of data transmission technologies and the design of the

communications network for different applications on site. Various parameters play a decisive role in optimizing network performance and meeting user requirements [16].

1. Determine data rate needs to size network capacity and avoid bottlenecks.
2. Define transmission reliability, especially for wireless tech, to ensure consistent communication.
3. Decide between wired and wireless links, considering minimum wireless distance.
4. Establish connection range for network coverage within specific geographical areas.
5. Consider data packet size to enhance transmission efficiency and data flow control.
6. Account for latency requirements based on application needs.
7. Specify device density to influence network capacity and performance.
8. Define network size to plan infrastructure and ensure coverage in a specific geographical area.

Overall, these functional dimensions are indispensable parameters to be considered in the design and optimization of communication networks in order to ensure efficient and reliable data transmission and to meet the increasing demands of users. The aforementioned studies by Aktas et al. [2] and Schulze et al. [22] provide important insights and foundations for the development of suitable network concepts.

4.2 Categories and impacts of network set up on construction site

Following the described factors influencing the network configuration on construction sites (use case, network choice and construction site characteristics) Figure 5 demonstrates the categories which have to be taken into account for designing the hardware set up of the network components.

Category	Details
Construction Site Layout	Linear / point construction site
Location of Antenna	Crane, container, other fixed points
Fixed construction site facilities	e.g. containers, crane in relation to construction etc.
Dynamic construction obstacles	e.g. trucks, temporary storage, material transport, etc.
Construction progress	Change of antenna position in relation to construction progress
Construction material	Influencing signal strengths based on material
Requirements of use cases	Different use cases require different network performances

Figure 5. Overview of categories for designing

hardware set up and network configuration on site

In general, the construction site layout has an impact on the location and amount of hardware used on site. In relation to the construction progress, it has to be investigated, if there are positions that are relatively fixed and centrally located, which can serve as a mounting point. At the beginning of the project, during the excavation task, there is rarely a crane, which is centrally located and could serve as a mounting option for the antenna. During that phase, another fixed location, such as a stable tower or a container, would be able to be the carrier for the antenna and the radio unit of the 5G network (see Figure 6). As the position may not be centrally within the excavation pit, it needs to be analysed if the coverage is sufficient for the specific use case, e.g. H2M or M2M communication.

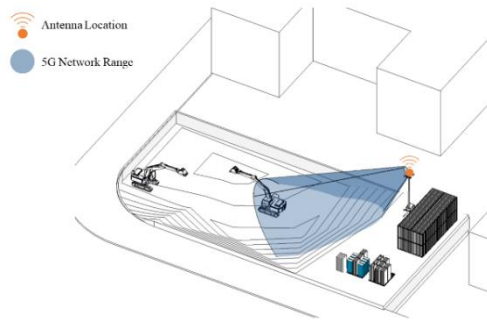
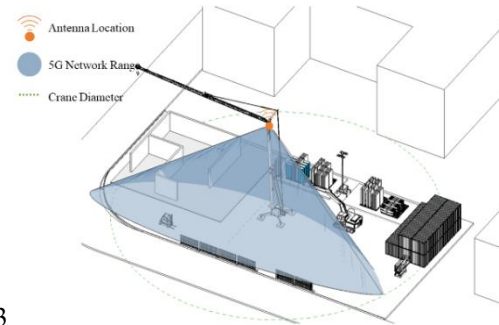


Figure 6. Antenna position during first phase of construction

Once the crane is installed on the construction site, the hardware can be mounted on the crane. As the crane is centrally located, the network's range covers the area of the site where the network is required for the applications (see Figure 7). A distinction can be made between dynamic and static installation points when selecting a crane as a set-up point. Top-Slewing cranes have fixed towers, Down-Slewing cranes have dynamic towers, impacting the choice of hardware in terms of antenna configuration. Omni-directional antennas radiate 360° whether directional antennas radiate mainly in one direction.



73

Figure 7. Antenna position on crane during construction

In addition, there is also the possibility to install radiating cables which also radiate in one direction (see Figure 8). The choice of hardware must therefore be based on the installation point, the location of the structure and the radiating technology.

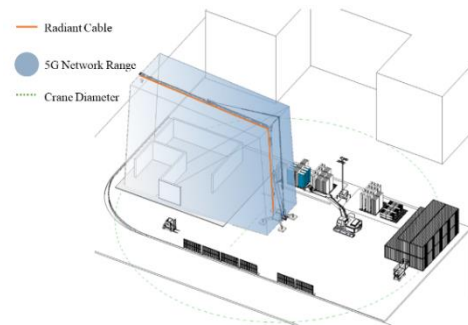


Figure 8. Radiating cable on crane

As discussed previously, there are multiple installation points for the network hardware. Installing the hardware on the container (see Figure 9) would result in range limitations on the site due to the distance between the application location and the hardware installation. This option must be selected with respect to the container and site location if the network requirements for the application can be met.

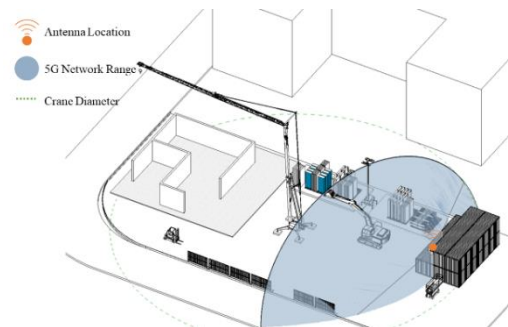


Figure 9. Antenna position on container

Given the multitude of fixed and variable objects such as machines, built structures, and storage at construction sites, it is crucial to recognize their potential impact on network performance, necessitating careful consideration in network configuration and ongoing adaptation. Furthermore, the construction progress and the materials may affect hardware installation location as it may lead to blockage and signal reduction (see Figure 10).

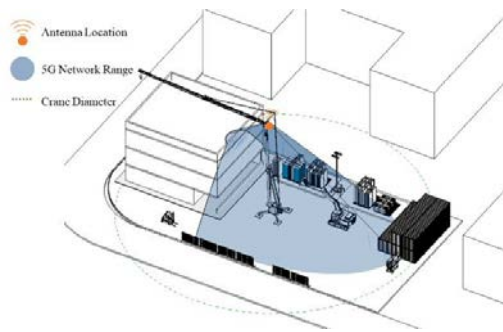


Figure 10. Antenna position on crane during construction progress

Finally, in addition to the characteristics of the site where the network is required, the use case for 5G technology must also be taken into account when configuring and installing the network. 5G offers unparalleled speed, reliability, and low latency, enabling a wide range of applications that can enhance productivity, safety, and communication on construction sites. However, before configuring the network the specific application (H2H, H2M and M2M) need to be analysed to define the network requirements such as latency, jitter, coverage, bandwidth or parallel user in the network.

4.3 Evaluation

A consequently step, the network needs to be evaluated, to test and align the use case requirements with the hardware set up and configuration on site. The evaluation involves two phases: one before implementing the use case to assess the network's performance, and another after implementing the use case to validate its performance. Key network indicators include Coverage, Latency, Packet Loss, and Reliability, which can be tested using methods such as *Ping Test* and *Iperf3 Test* [23]. Upon application of the use cases, a performance assessment can be conducted to analyse how the use case aligns with the network indicators.

5 Implementation

5.1 Hardware Set up on Reference Construction Site

In 2020, Construction Robotics GmbH, Aachen, Germany, opened the first large-scale testbed, the Reference Construction Site in Aachen, aiming to bridge digital gaps across the construction industry, encompassing planning, production, and implementation. Covering 4,000 square meters, this test ground facilitates the transition of research findings to market integration. Equipped with cranes, containers, and demonstrator structures the Reference Construction Site serves as a controlled environment for testing and trials, allowing the identification and resolution of potential risks and issues before arise on real projects.

The 5G.NAMICO project provides the first implementation of a 5G network on the Reference Construction Site. Based on the previous categories, the network was set up as follows (see Figure 11):

- Radiating cable along the entire crane (tower and boom)
- Dual-polarized omni-antenna and radio units on crane tower

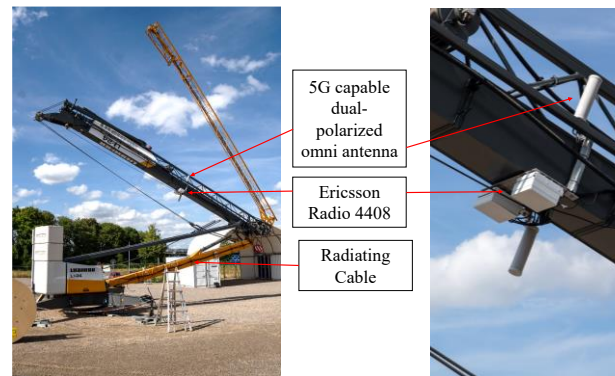


Figure 11. Tower crane with Ericsson radio unit, omnidirectional antenna and radiating cables [24]

The dual-polarized antenna is designed for omnidirectional coverage and is able to send and receive signals in any direction. In addition, radiant cables are installed on the crane. These radiating cables provide an alternative to traditional antennas and can be used in situations where conventional antennas do not provide adequate coverage, such as tunnels or mines. First test with both hardware set up were conducted on the Reference Construction Site. The omni-directional antenna configuration provides extensive coverage with excellent signal quality, with 90% of the readings falling within the -75 to -50 dBm range [24]. Conversely, the

signal quality of the radiating cables is robust in the vicinity of the crane, but degrades significantly as the distance from the crane increases.

5.2 Application

The digitisation of construction site processes often involves the use of multiple sensors or the automation of heavy construction machinery. Data and control commands are then transmitted between these sources and an on-site control centre. In order to assess the performance of the network on site, a set up was established on the Reference Construction Site to model the data communication.

5G and other network technologies for communication may have two different applications on construction sites. Locally, the network is required to enable the automation of construction processes in terms of H2H, H2M and M2M communication. This can be realized with a private and closed 5G network, solely open to participants and users on site.

Based on previous research [12], various use cases and challenges of *Construction 4.0* can be addressed by 5G, such as remotely controlled and autonomous machinery, health and safety at worksites, 3D models, construction processes' management or emissions and waste management.

In addition to local applications, 5G can also be used to connect the construction site to the outside world. Planning and scheduling data can be exchanged seamlessly between different stakeholders. Real-time data transmission enables efficient collaboration, monitoring, and decision-making between local and global resources.

6 Discussion

The integration of new communication technologies in construction offers benefits such as increased efficiency, improved safety and enhanced project management capabilities. However, challenges include initial investment costs, potential resistance to change from workers, and the need for comprehensive training programmes to ensure successful adoption and use of these technologies. Balancing these benefits and challenges is critical to effectively integrating and maximising the potential benefits of new communication technologies on construction projects. Training requirements for site personnel include comprehensive programmes focused on the use of new communication technologies, emphasising practical skills and theoretical understanding. Strategies to promote the adoption of these technologies can include hands-on workshops, educational seminars, and incentives for successful implementation, fostering a culture of innovation and

competence among construction teams.

Although the research focuses primarily on the Reference Site in Aachen, the need for more diverse case studies is evident. Including additional case studies from different geographical locations and project types could improve the generalizability. While the project aimed to deploy a generic 5G network to understand its limitations and challenges, further research should prioritize the exploration of other construction site locations to assess the scalability and adaptability of 5G technology in different contexts.

Furthermore, while the research focused on 5G for construction, it is clear that not every communication use case requires 5G and is also realizable with 5G due to conflicts between regulations and application requirements. An example of this is the TDD patterns that make it difficult to adapt uplink and downlink rates in the outdoor sector in Germany [24]. Future research should focus on identifying the most appropriate communication technologies for different applications in (partially) automated construction. By exploring a wider range of technologies, it is possible to better understand their respective strengths and limitations, and thus make more informed decisions when adopting communication solutions tailored to specific construction needs.

7 Conclusion and Outlook

In the development and implementation of *Construction 4.0*, with the intention of automation and digitalisation, the focus is on creating a communications infrastructure that facilitates digital data transfer and processing. The choice of transmission technologies and the correct hardware configuration in relation to the characteristics of the site are crucial, as they provide the technical basis for the exchange of data and information between the various parties involved. With different technologies providing varying performance levels, the construction site context often necessitates tailored solutions. Broadly categorized into Human to Human (H2H), Machine to Human (M2H), and Machine to Machine (M2M) scenarios, these diverse communication requirements underscore the need for adaptable and efficient network solutions to cater to the distinct demands of each context. To ensure practicality and widespread use on construction sites, it is essential that the network is easy and quick to set up and maintain. Moreover, the systems must embody reliability, incorporating redundant components to ensure that the construction site doesn't come to a halt in the event of a failure in any network system component. Within the TARGET-X project (funded by the Smart Networks and Services Joint Undertaking (SNS JU) under Horizon Europe (funding number 101096614)), the research focus is on using the 5G network for deconstruction use cases

to analyse the impact of site characteristics on network performance, such as material and construction progress or number and type of users. The limitations of this paper, identified in the discussion, will be investigated in further projects. Amongst others, the EConoM project is investigating how 5G campus networks with AI and edge computing applications can provide a robust and nomadic processing and network infrastructure in constantly changing environments, with changing participants.

8 Acknowledgements

The 5G.NAMICO project is funded by the State of North Rhine-Westphalia through the Ministry of Economic Affairs, Innovation, Digitalization and Energy (Grant No. 005-2108-0111). Furthermore, the server infrastructure and other network requirements as well as the use case specification were and will be conducted in the EConoM project, funded by the Federal Ministry for Digital and Transport of Germany within the initiative InnoNT (funding number 19OI22009F).

References

- [1] ifo Institut – Leibniz-Institut für Wirtschaftsforschung an der Universität München e.V., *Fachkräftemangel auf deutschen Baustellen verschärft sich*. [Online]. Available: <https://www.ifo.de/pressemitteilung/2021-10-06/fachkraeftemangel-auf-deutschen-baustellen-verschaerft-sich> (accessed: Dec. 18 2023).
- [2] I. Aktas *et al.*, "Funktechnologien für Industrie 4.0," 2017.
- [3] Schuh, G., Anderl, R., Gausemeier J., ten Hompel, M., Wahlster, W. (Eds.), "Industrie 4.0 Maturity Index. Managing the Digital Transformation of Companies (acatech STUDY)," Herbert Utz Verlag, München, 2017.
- [4] A. A. Tuma Neto and A. Araujo de Souza Junior, "Industry 4.0 Innovations in Construction," *Int J Innov Educ Res*, vol. 10, no. 9, pp. 418–436, 2022, doi: 10.31686/ijier.vol10.iss9.3892.
- [5] M. A. Hossain and A. Nadeem, "TOWARDS DIGITIZING THE CONSTRUCTION INDUSTRY: STATE OF THE ART OF CONSTRUCTION 4.0," *ISEC 10: Interdependence between Structural Engineering and Construction Management*, vol. 6, no. 1, 2019, doi: 10.14455/ISEC.res.2019.184.
- [6] C. J. Turner, J. Oyekan, L. Stergioulas, and D. Griffin, "Utilizing Industry 4.0 on the Construction Site: Challenges and Opportunities," *IEEE Trans. Ind. Inf.*, vol. 17, no. 2, pp. 746–756, 2021, doi: 10.1109/TII.2020.3002197.
- [7] S. Brell-Cokcan and R. H. Schmitt, Eds., *IoC - Internet of Construction: Informationsnetzwerke zur unternehmensübergreifenden Kollaboration in den Fertigungsketten des Bauwesens*, 1st ed. Wiesbaden: Springer Fachmedien Wiesbaden GmbH; Springer Vieweg, 2024.
- [8] F. Jordan, A. Bernardy, M. Stroh, J. Horeis, and V. Stich, "Requirements-Based Matching Approach to Configure Cyber-Physical Systems for SMEs," in *2017 Portland International Conference on Management of Engineering and Technology (PICMET)*, 2017.
- [9] R. Kiesel and R. H. Schmitt, "Requirements for Economic Analysis of 5G Technology Implementation in Smart Factories from End-User Perspective," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, London, United Kingdom, 2020, pp. 1–7.
- [10] Z. M. Temesvári, D. Maros, and P. Kádár, "Review of Mobile Communication and the 5G in Manufacturing," *Procedia Manufacturing*, vol. 32, pp. 600–612, 2019, doi: 10.1016/j.promfg.2019.02.259.
- [11] H. J. Lee *et al.*, "Importance of a 5G Network for Construction Sites: Limitation of WLAN in 3D Sensing Applications," *International Association of Automation and Robotics in Construction, Werkzeugmaschinenlabor WZL der RWTH Aachen RWTH-2022-07117*. [Online]. Available: <https://publications.rwth-aachen.de/record/849931>
- [12] J. Mendoza *et al.*, "5G for Construction: Use Cases and Solutions," *Electronics*, vol. 10, no. 14, p. 1713, 2021, doi: 10.3390/electronics10141713.
- [13] O. Adepoju, "Construction 4.0," in *Springer Tracts in Civil Engineering, Re-skilling Human Resources for Construction 4.0: Implications for Industry, Academia and Government*, O. Adepoju, C. Aigbavboa, N. Nwulu, and M. Onyia, Eds., 1st ed., Cham: Springer International Publishing; Springer, 2022, pp. 17–39.
- [14] G. Johannsen, "Human-machine interaction," *Control Systems, Robotics and Automation*, vol. 21, pp. 132–162, 2009.
- [15] M. Zhao, A. Kumar, T. Ristaniemi, and P. H. J. Chong, "Machine-to-Machine Communication and Research Challenges: A Survey," *Wireless Pers Commun*, vol. 97, no. 3, pp. 3569–3585, 2017, doi: 10.1007/s11277-017-4686-1.
- [16] A. Bernardy, V. Stich, and G. Schuh, "Spezifikation der cyber-physischen Veredelung von Auftragsverfolgungssystemen: Lehrstuhl für Produktionssystematik / Werkzeugmaschinenlabor WZL der RWTH Aachen," Apprimus Verlag; Dissertation, RWTH Aachen University, 2019.

- [Online]. Available: <https://publications.rwth-aachen.de/record/781007>
- [17] F. Michel, M. Trevisan, D. Giordano, and O. Bonaventure, "A first look at starlink performance," in *Proceedings of the 22nd ACM Internet Measurement Conference*, Nice France, 2022, pp. 130–136.
 - [18] A. Wang, *Comparison between LoRa and Other Wireless Technologies*. [Online]. Available: <https://www.mokolora.com/lora-and-wireless-technologies/> (accessed: Dec. 7 2023).
 - [19] S. B. Damsgaard, N. J. Hernández Marcano, M. Nørremark, R. H. Jacobsen, I. Rodriguez, and P. Mogensen, "Wireless Communications for Internet of Farming: An Early 5G Measurement Study," *IEEE Access*, vol. 10, pp. 105263–105277, 2022, doi: 10.1109/ACCESS.2022.3211096.
 - [20] J. Zhang *et al.*, "System Framework for Digital Monitoring of the Construction of Asphalt Concrete Pavement Based on IoT, BeiDou Navigation System, and 5G Technology," *Buildings*, vol. 13, no. 2, p. 503, 2023, doi: 10.3390/buildings13020503.
 - [21] K. You, L. Ding, C. Zhou, Q. Dou, X. Wang, and B. Hu, "5G-based earthwork monitoring system for an unmanned bulldozer," *Automation in Construction*, vol. 131, p. 103891, 2021, doi: 10.1016/j.autcon.2021.103891.
 - [22] D. Schulze, A. Gnad, and M. Krätzig, "Anforderungsprofile im ZDKI Fachgruppe 1: Anwendungen, Anforderungen und Validierung im BMBF-Förderprogramm „IKT 2020 – Zuverlässige drahtlose Kommunikation in der Industrie“ (BZKI),"
 - [23] R. Kiesel *et al.*, "Techno-Economic Evaluation of 5G-NSA-NPN for Networked Control Systems," *Electronics*, vol. 11, no. 11, p. 1736, 2022, doi: 10.3390/electronics11111736.
 - [24] J. Emontsbotz *et al.*, "The Application of 5G Networks on Construction Sites and in Underground Mines: Successful Outcomes from Field Trials," 2024 19th Wireless On-Demand Network Systems and Services Conference (WONS), Avoriaz, France Submitted, 2024.

State of the Art Review of Technological Advancements for Safe Tower Crane Operation

Avi Raj and Jochen Teizer

Department of Civil and Mechanical Engineering, Technical University of Denmark
avira@dtu.dk, teizerj@dtu.dk

Abstract –

This paper focuses on construction accidents, particularly those involving tower cranes. An initial analysis of the modern literature highlights tower cranes as a significant safety concern, with human error identified as a major factor. The paper in the further explores modern technological innovations addressing these human-related risks. Recognizing the potential of technology to improve safety, the review establishes a crucial link between understanding accident root causes, particularly human error, and ongoing efforts to enhance safety through technological interventions. Utilizing a systematic Web of Science keyword searches approach, including a quantitative and qualitative analysis of the respective publications, the review divides tower crane safety into two predominant fields, namely the pre-construction phase that plans, among other objectives, safe worksite layouts, and the construction phase that executes safe operations. The findings provide a comprehensive overview of the most recent technological innovations that involve tower cranes in both project phases. A synthesis informs about future research and industry initiatives for advancing tower crane safety.

Keywords –

Accidents, Construction safety, Tower cranes.

1 Introduction

The construction industry, crucial for global progress, struggles with an enduring challenge that is workplace accidents, particularly involving tower cranes. Between 2011 and 2017, there were 297 reported deaths due to crane accidents in the US alone, averaging 42 fatalities per year [1]. A more detailed analysis of the data from Spain from 1990 to 2000 shows that of the 1.630.452 accidents reviewed, 22.349 or 1,4% of the accidents are related to cranes and lifting equipment of which 165 were fatal while 1.201 are severe and fatal [2].

The construction industry relies on a diverse array of cranes, including tower cranes, mobile cranes,

overhead cranes, and crawler cranes, each playing a unique role in shaping industry dynamics. Tower cranes, particularly, have emerged as a predominant type, garnering widespread use in Europe and gaining popularity in the US [3]. These statistics clearly demonstrate the safety related issues for tower crane operations. This forms the motivation to study what is being done in this field to improve the safety conditions. Therefore, this review paper tries to investigate the following major research questions:

- What is the current state of construction site safety with respect to tower cranes?
- What are the significant factors that cause accidents?
- How can technological solutions improve the safety of tower crane operations?
- What technological solutions are being explored for improving safety of tower crane operations?

This review studies the proportion of accidents related to tower crane operations. It investigates the problem of safety issues. Following that is the methodology of how the publications are found, selected and reviewed are discussed. Then, a quantitative analysis and a qualitative is performed to give the reader an objective idea about the trends being explored in this field, this allows for a better understanding of the trends and their advantages or shortcomings. At last, a conclusion section summarizes the key insights from the review.

2 Background

The construction sites where fixed site lifting mechanisms and lifting structures are used, are hazardous production facilities. The year 2020 saw 35 recorded accidents in Russia [4]. These resulted in 28 fatalities, attributed to the operation of lifting structures at monitored facilities. The majority of accidents related to lifting machinery between 2015 and 2020 were associated with the operation of tower cranes. For instance, out of 25 crane-related incidents, 12 (47%) occurred during the operation of tower cranes, 4 (20%) during the operation of crawler and mobile cranes, 2 (10%) during the use of manipulator cranes, and 2 (7%)

during the use of overhead gantry, and 1 while using the portal cranes as shown in Figure 1. Despite the varied crane types, the focus on tower cranes intensifies when considering occupational accidents.

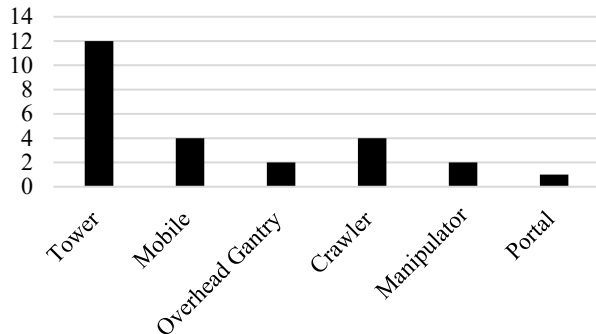


Figure 1: Involved equipment type in fatal crane accident, as an example, Russia in 2020 [4].

In addition to the human cost, crane accidents also lead to fiduciary losses as well. A study among 313 Malaysian registered contractors revealed three key areas that increase the cost of a project when an accident occurs. These being project delays, medical expenses, and damage repairs to completed work and collateral damage. Although the study does not point out an exact number value for the increase in the costs, it demonstrates that out of the contractors who undertook the survey, 97.4% agree that project delays due to accidents increase the project cost, 94.8% agree that the additional cost of repairing completed construction or adjacent structures is significant. 89.4% of the surveys agree that the medical costs for compensating injured workers also increase the cost of the overall project [5]. Moreover, human error accounts for approximately 70% of crane accidents [6].

A preliminary statistical analysis of accidents involving tower cranes, identified contributing factors. Furthermore, a Degree of Influence (DOI) was computed for each factor, providing a more nuanced understanding of their individual levels of significance [7]. According to the findings of this study “Operator Proficiency” has the highest DOI. Whereas the seven factors that can be categorized into human-factor and safety management categories also have high DOIs associated to them.

An further examination of human factor classification revealed key aspects that further refine the understanding of these incidents [8]. These aspects include:

- inattention,
- failure in communication,
- operator error, and
- error of signal person.

Another study arrives at similar conclusions however they do point out that although human error-based accidents have a very high frequency, they rank low in

severity of the accidents [9]. They point out that the most prevalent failures in the dataset are:

- Inattention, a typical human factor that appeared in about 19% of incidents.
- The second most prevalent type (between 10 and 15%) contains three human factors, namely– improper rigging, – signalperson error, and– operator error.
- Environmental factors such as limited visibility and strong wind appear to reside within the least prevalent.

A point-cloud and display based safety feature was tested with five crane operators with work experiences ranging from 8 to 16 years [10]. Although they point out that their solution was still in the preliminary stage and some of the issues would be sorted out in later iterations, the key insights obtained from their tests were:

- The operators generally agreed that the new safety solution was easy to use.
- The operators agreed that a real-time vision-based safety feature would help improve safety. However, the operators were also quick to point out that this does not eliminate the need for a signal person. The main reason was a general insecurity about solely relying on technology as it also put all the responsibility on the operators too.
- The operators could not form a consensus on whether such system necessarily improves efficiency. They pointed out that being able to view the scene in 3D helped them choose an optimum path faster. However, in situations where direct line of sight application was still feasible the system acted as a hindrance.

These cases demonstrate that technological assistance plays a significant role in enhancing the overall safety of workers on a construction site, particularly in the context of crane-related accidents.

3 Methodology

This review utilizes systematic Web of Science keyword searches, similar to [11]. The review is divided in two primary phases (Figures 2 and 3, respectively):

- Pre-construction planning phase, and
- Construction operation phase.

Advanced search options were used that allow for Boolean keyword search operations. Boolean operators like “And” and “OR” were used to identify records that talk about construction safety and technology. For this purpose, following keyword searches were used:

- "ALL=((('crane') AND ('safety') AND ('pre-construction' OR 'planning')))"
- “ALL=(((" tower crane") AND ("safety") and ("improving" or "monitoring"))”

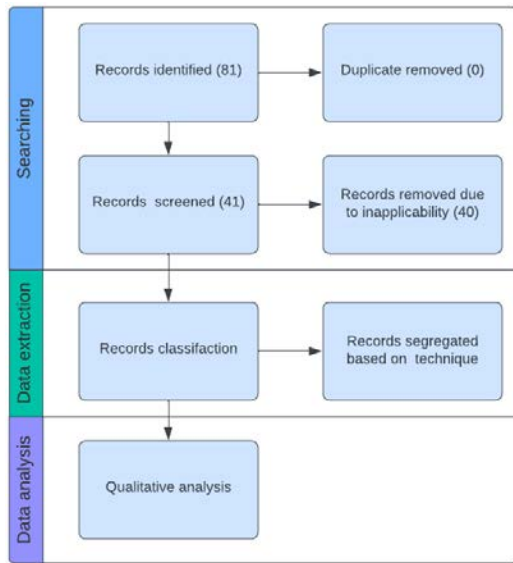


Figure 2: Flowchart for the literature review part of the pre-construction phase.

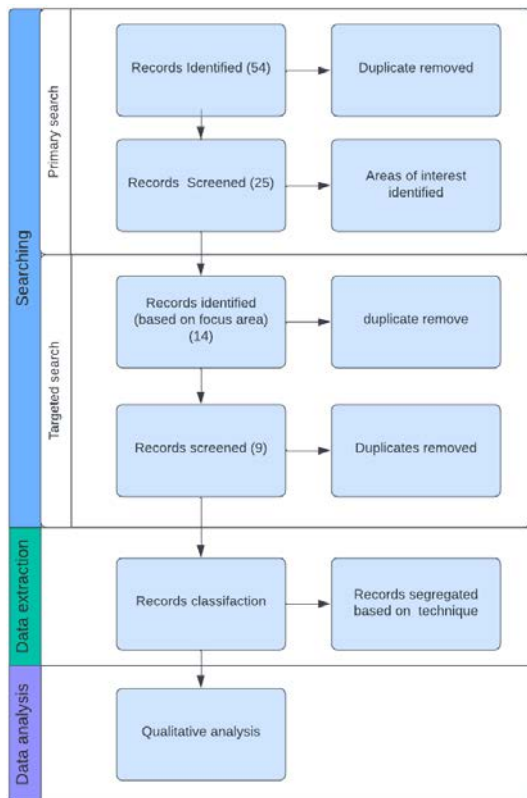


Figure 3: Flowchart for the literature review part of the construction operation phase.

A lot of technological developments exist to make construction sites safer. However, for this review, a strict criterion was chosen to only consider literature that strictly deals with tower cranes and technology that pertains to improving safety on a construction site.

Since the majority of the publications were from or before 2011, 2011 was chosen to be the cut-off date for all records to be included in the review.

The first set of keywords were chosen to target tower crane safety in pre-construction or planning, while the second set focused on tower crane safety during construction. The systematic use of these keywords allowed for a structured exploration of relevant literature, providing insights into technological innovations in both project planning and construction phases. The first planning stage part of the review required only a straightforward approach as shown in Figure 2.

The construction operations part of the review required a more nuanced approach. To address potential human bias and facilitate an unbiased search, technology-specific terms were deliberately omitted for the primary keyword selection. The goal was to prevent assumptions that particular technologies, such as RFID or cameras, exclusively dominated the field. The objective for using this approach was to identify literature that might not be widely known, or avoiding potential gaps in the knowledge base. Subsequently, targeted keyword searches were conducted to comprehensively examine the identified technology fields as shown in Figure 3. Some of the targeted search keywords are mentioned below:

- ALL=((*"crane"*) AND (*"safety"*) and (*"IoT"* or *"Internet of Things"* or *"internet of things"*))
- ALL=((*"crane"*) AND (*"safety"*) and (*"sensor"* or *"sensing"*))
- ALL=((*"crane"*) AND (*"safety"*) and (*"camera"* or *"image"* or *"vision"*))

4 Findings of quantitative analysis

The number of publications reviewed are shown in Table 1. Many of them that contained one relevant keyword were clubbed into a single category of 'others'.

Table 1. Distribution based on journal publications.

Journal and <i>conference</i> name	Publications [No.]
Automation in Construction	18
Computing in Civil Engineering	5
Advanced Engineering Informatics	4
Construction Engineering and Management	4
Advances in Civil Engineering	2
Buildings	2
Sensors	2
ICMSE	2
Construction Research Congress	2
Others	34
Total	75

As seen in Figure 4, between the years 2013 until 2019 fewer publications were found. Of the 75 publications reviewed, 41 focused on the pre-

construction phase while 34 focused on building solutions for real-time applications. Among those 41 papers, 21 looked at crane layout planning, while about one third (14) looked at crane trajectory planning. The remaining 6 papers focused on unique aspects of the planning phase as shown in Table 2. As for the remaining 34 papers, their focus was divided amongst several domains (Table 3).

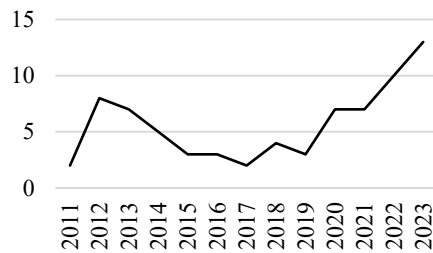


Figure 4: Total publications per year

Table 2. Distribution of publications focusing on the pre-construction planning phase

Focus Area	Number
Single crane layout	6
Multi crane layout	15
Trajectory planning	14
Unique focus areas	6

Table 3. Distribution of publications focusing on the construction operation phase

Technology used	Number
Sensor based	18
Camera based	6
Virtual reality based	4
Digital twin based	1
Model based	2
Simulation based	1
Combined areas	2

5 Discussion

Tables 4 and 5 preview techniques used to implement solutions for each focus area. Traditionally, the task of planning the stages and arrangement of equipment and heavy machinery was solely dependent on the experience and expertise of trained professionals. However, these tasks are also being relegated to automated solutions.

This can help save time, money and most importantly, human lives. The task of planning a construction site w.r.t a tower crane is mainly about identifying two key issues, the optimal position of tower crane and the optimal trajectory for the lift path. Interestingly, trajectory planning is a vital and matured aspect of the robotics field. Several of those algorithms that have been tried and tested. For example, search-based algorithms like L* algorithm [12] and A* algorithms [13] which are were tested in simulations and they show fast execution times

and can successfully generate collision free lift paths. While interpolation-based algorithms like spline curve [14] or waypoint [15] calculations have also shown success, especially in time efficiency as interpolation-based algorithms lead to smoother curves that do not lead to sudden acceleration or jerk.

Moving away from the technologies being laterally transferred from robotics. Some solutions based on fundamental mathematics have also been explored.

Techniques like time-based polynomial functions [16] for accommodating complexities that arise from the double-pendulum created between the jib, hook and the weight attached via cables has shown state of the art performance. Similarly, Lagrange's kinematic equations [17] have been used for swing suppression of the payload when the trolley moves it along the jib. Alternatively, control-theory based state variable models for both double pendulum [21] and anti-swing [22] have also been experimented with. The anti-swing model is tested on an experimental setup while state variable model for double pendulum dynamics only simulated. Both show good results and are successful in meeting their objectives.

Some research has been done into creating process management systems. Prevention through Design (PtD) [18] integrated with a BIM and dependency structure matrix [19] are examples of two techniques used by these type of management system. These typically include a scheduling system, a visualization system, and some form of anti-collision safety. However, they still remain error prone and require further development. A 4D BIM [20] also acting as an overall process management, tries to look at global optimum based on resource optimization over multiple factors, it looks at not only crane movement optimization but also the cost, number of cranes etc.

Simulations and virtual environments have also been used in lift path planning recently. Reinforcement learning based models [23] are being trained in virtual environments, the models can be used to generate a collision free motion for the tower crane and also estimate the lifting time. Simulation models are also being used for near real time planning [24], the model can perform optimization tasks via simulations and provide a day-to-day planning support effectively.

One important factor to understand about most of these type of research related to tower cranes is that, due to the nature of the project real-life experimental validations is often times expensive, infeasible and risky for the construction workers. Hence a lot of research is validated on experimental setups or simulations.

The other main concern of research in the pre-construction phase is the location of the tower crane within the construction site. Mixed integer linear programming [25-28] is shown to be a very effective and popular strategy for the development of tower crane layout algorithms. This approach has shown great results

in terms of collision avoidance, time and resource efficiency and safety. By also paying extra attention to planning the location of source points and destination point along with the tower crane, the amount to be carried per trip can be reduced which means the site requires a smaller crane and saves costs [30].

Table 4. Data analysis techniques for crane safety (pre-construction)

Application	Technical Details	Examples
Crane trajectory planning	Machine Learning-based	[23]
	Mathematical modelling-based	[16], [17]
	Simulation-based	[24]
	Algorithm-based	[12],[13], [14],[15]
	Management system based State variable modelling based	[18], [19], [20]
Crane layout planning for a single crane	AI based	N/A
	Mixed-integer linear programming based	[29]
	Optimization based	[37], [38]
Crane layout planning for multiple cranes	Mixed-integer linear programming based	[25], [26], [27], [28]
	Backpropagated neural network	[39]

Table 5. Data analysis for crane safety (construction operation phase)

Technology	Technical details	Examples
Sensor based	Internet of things platform	[41], [42], [53], [54], [56], [69]
	Embedded systems platform	[43], [48], [49], [50], [44], [45], [46]
	CDMA communication platform	[40]
Camera		[57], [58], [66], [59], [60], [61]
Virtual reality		[62], [63], [64]
Digital twin		[47]
Modelling	Mathematical model	[51]
	Artificial neural network	[55], [52]
Combined		[68], [67]

Artificial Intelligence also provides several alternative solutions to tackle the problem. Firefly algorithm [31] is an optimization algorithm based on the field of AI called swarm intelligence, it is a highly effective method for trying out several alternatives and then ranking them. In this approach a BIM model is integrated with firefly algorithm, and it predicts the best

layout. Generative adversarial network [36] based solutions use image-based solutions where conditions based on all the details of construction site are fed in. The model tries to create an optimized solution in terms of time optimization and the safety of the site by considering the building boundary and load bearing capacity among others.

Genetic Algorithms [35] and agent-based systems [33-34] are also deployed in several solutions. A decision support system utilizing a combination of analytical hierarchy process, a 4D BIM and genetic algorithm [32] was to be capable of dynamic assistance and its capable.

Depending on the scale of the construction project the site may require a single or multiple tower cranes. Mixed integer linear programming can be applied for optimizing the location of a single crane on a small to medium size construction project [29]. Non-linear optimization techniques on models of a tower crane with realistic kinematic dynamics are also used for single crane operations, a solution like this improves safety at the site as it optimizes the location of the crane and supply points such that the trajectory of the lift is smooth with a finite jerk [37]. The firefly algorithm discussed above is also used to create the optimal layout for single cranes with considerations to safety, time and cost efficiency [38]. It uses GIS for accurate localization of the crane and a CAD model of the crane to apply the firefly algorithm to.

While looking at the publications that focus on unique areas, some interesting ideas are seen. The condition of the structure of a tower crane is dependent on not just the load being hoisted but, for instance, the repeated use, corrosion, designed lifecycle. A back-propagated neural network demonstrates that it can predict the lifetime of a tower crane with an error rate as small as 0,001% albeit it has a large training time for every crane [39]. The model can potentially enable better site safety and security as it allows planners to know how dangerous any particular tower crane's chances of collapsing are.

Sensor-based solutions [40] offer users contextual data of the TC, including swing, amplitude, weight, etc. it uses CDMA for communication. Contextual data such as these greatly assist the crane operator to perform their task safely for both the construction worker and crane itself, which indirectly impacts the construction worker, crane operator and everyone around the crane. They can also raise alarms if any safety parameter is being violated during a lifting operation.

Similarly designed systems exist that use IoT (Zigbee) [41, 42]. However, CDMA can be more suited for this type of application as it has extended range, higher data transfer rates, enhanced interference resistance and reliability, as well as capacity and scalability. It can support a large number of users simultaneously, suitable for applications with high device density or scalability requirements.

IoT has also been used for creating anti-collision systems. When operating multiple tower cranes in close proximity with each other, it is potentially a hazardous situation. To mitigate the risks RFID sensors with wireless sensor networks are used to detect incoming collisions and avoid them [53, 54], [56].

Some solutions on the other hand prefer to go for wired embedded system design architectures [43], [48] these allow for extremely high data transfer rates, security and reliability. This can also be easily scaled up with minimal losses to data rates. However, these systems can be more costly as they require more hardware equipment for the setup. Adding a fuzzy logic based PID controller into the system with a human-computer interface allows the operators to adjust the crane movements based on the data being provided to them. Using a PID controller rather than a manual control further assists them to perform safer executions [49]. Similarly, linear interpolation method is used to compute complex torque related calculations for a tower crane which get their input values via sensors and use a PLC based controller to protect the crane from overloading, excessive stress or potential damage to other crane components that can create hazardous situations for workers below or the crane operator [50].

Sensors, when placed in outdoor environments with varying temperature, humidity and other elements, cause the readings to drift. Prolonged load and continued use also cause sensor creep. Both these can introduce significant inaccuracies into the system which can endanger workers' safety. Research to create embedded systems can make such systems' operation safer [44, 45].

Another interesting approach to assist crane operators with situational awareness and contextual data of the crane is to integrate virtual models, kinematic models, and real time data from the site and the crane data. This allows for the visualization and monitoring of the crane and raises an alarm when any safety parameters are violated [46]. This system was shown to have wide acceptance within the workers and operators.

To ensure safe crane operations it is crucial to realize the impact of the environment on the crane as well as the state of the crane. Therefore, mathematical [51] and neural network models [55] can predict the effect of wind loads on tower crane operations. Artificial neural network based models with sensor data are used to monitor the state of the crane and fault diagnosis [52].

Taking that idea a step further, and allowing simulations into the visualization system too allows crane operators to test their hoisting operation plans beforehand. A digital-twin based solution [47] has been designed to perform these testing operations. This idea provides much better insights than a pre-construction trajectory planning as it utilizes a digital twin which is integrated with sensor derived real time data describing the state of

the crane as well as the weather conditions around the site.

Several solutions have been investigated using cameras as well. Some of them focus on the position and displacement of the crane. This information, as mentioned before, provides the operators with crucial information on the real-time situation of the crane and helps him react accordingly. In [57] the crane focuses on detecting the displacement of the jib and the cart, experimental results show good results. While [58], [66] and [59] focus on the hook and provide the operator assistance for precision hoisting.

Apart from using cameras for tracking the crane, it has also been used for other applications. A camera situated in the operator's room is used to analyse their facial images and uses a neural network model to assess the fatigue of the crane operator [60]. This information is very crucial from a safety perspective as a fatigued operator is more likely to make mistakes.

A camera mounted on the tower crane's boom has been used to detect and track construction workers. It assesses the potential danger to the workers around the crane and warns them of it [61]. Ultra-wideband (UWB) sensor-based IoT solutions [67, 69] have also been used to detect and track workers and prevent collision. However, a UWB based solution is a costlier alternative when compared to a camera-vision based detection and tracking system. Both technologies carry high chance of missing a worker because of signal loss or unavailability.

The role of virtual reality has primarily been for training new operators. This is important because it allows the operators to train in a safe environment for longer hours, which otherwise is financially infeasible. Virtual reality based simulations have been used to train operators on how to use the tower crane [62, 63].

Context aware training for a specific construction site is also done by integrating the virtual reality model with the BIM [64]. It has also been used to train operators and workers on how to dismantle a tower crane, which is one of the more challenging aspects of operating a tower crane on a construction site [65].

There are some solutions that have tried to tackle the safety issue by using a combination of multiple technologies, a neural network combined with a simulation model with inputs from the sensor is used to predict the foundation's stability on which a tower crane sits [65]. This creates a reliable solution to predict the safe operations of a tower crane and can warn workers when a crane is too risky to operate. Another such mixed technology solution uses sensors, camera and laser to create accurate navigation paths for a tower crane and assists in blind lifts [68].

Conclusion

This state of the art review, covering mostly academic

publications over the past decade, showed a growing interest in the construction domain to develop specific technologies that can assist in planning and operating tower cranes more safely. While eventually the construction site, its staff and workers, will benefit the most from such availability and use of emerging safety technologies, this review found that it is often leadership in organizations that lacks knowledge on the basic working principles, the benefits, and the shortcomings of applying advanced technology in the field. Furthermore, this study found that scientific work yet has to find effective measures that solve the proclaimed disconnect between technology in the pre-construction tower crane planning and operations phases. This state of the art review found multiple examples that explain some of the reasons for slower than needed technology adaption in the field of safe tower crane planning and operation.

Acknowledgements

This research has been funded by the European Union Horizon 2020 research and innovation program under grant agreement no. 101058548.

References

- [1] U.S. Bureau of Labor Statistics. Injuries, Illness, and Fatalities. Accessed: 10/12/2023. On-Line: <https://www.bls.gov/iif/factsheets/fatal-occupational-injuries-cranes-2011-17.htm>,
- [2] López M., Ritzel D., Fontaneda I., and Alcantara, O Construction industry accidents in Spain. *Journal of Safety Research*, 39(5), 497-507, 2008. <https://doi.org/10.1016/j.jsr.2008.07.006>
- [3] Spirina A., Mironov A., Datkhuzheva R., Maksimov A., and Belova N. Analysis of occupational injuries in construction and offer of a technical solution increasing tower crane stability. *Transportation Research Procedia*, 68, 559-565, 2023. <https://doi.org/10.1016/j.trpro.2023.02.076>
- [4] Shapira A., and Lyachin B. Identification and analysis of factors affecting safety on construction sites with tower cranes. *Construction Engineering and Management*, 135(1), 24-33, 2009. [https://doi.org/10.1061/\(ASCE\)0733-9364\(2009\)135:1\(24\)](https://doi.org/10.1061/(ASCE)0733-9364(2009)135:1(24))
- [5] Dahalan M., and Azmi W. Crane Accidents: Identifying the Impact on Construction Cost among Contractors. In *Intl. J. Academic Research in Progressive Education and Development*, 11(3), 1553–1562, 2022. <http://dx.doi.org/10.6007/IJARPED/v11-i3/15286>
- [6] Haslam R. A., Hide S. A., Gibb A. G., Gyi D. E., Pavitt T., Atkinson S., and Duff A. R. Contributing factors in construction accidents. *Applied Ergonomics*, 36(4), 401-415, 2005. <https://doi.org/10.1016/j.apergo.2004.12.002>
- [7] Shapira A., and Lyachin B. Identification and analysis of factors affecting safety on construction sites with tower cranes. *Construction Engineering and Management*, 135(1), 24-33, 2009. [https://doi.org/10.1061/\(ASCE\)0733-9364\(2009\)135:1\(24\)](https://doi.org/10.1061/(ASCE)0733-9364(2009)135:1(24))
- [8] Raviv G., Shapira A., and Fishbain B. AHP-based analysis of the risk potential of safety incidents: Case study of cranes in the construction industry. *Safety Science*, 91, 298-309, 2017. <https://doi.org/10.1016/j.ssci.2016.08.027>
- [9] Raviv G., Fishbain B., and Shapira A. Analyzing risk factors in crane-related near-miss and accident reports. *Safety Science*, 91, 192-205, 2017. <https://doi.org/10.1016/j.ssci.2016.08.022>
- [10] Fang Y., Cho Y. K., Durso F., and Seo J. Assessment of operator's situation awareness for smart operation of mobile cranes. *Automation in Construction*, 85, 65-75, 2018. <https://doi.org/10.1016/j.autcon.2017.10.007>
- [11] Fang Y., Cho Y. K., and Chen J. A framework for real-time pro-active safety assistance for mobile crane lifting operations. *Automation in Construction*, 72, 367-379, 2016. <https://doi.org/10.1016/j.autcon.2016.08.025>
- [12] Lin X., Han Y., Guo H., Luo Z., and Guo Z. Lift path planning for tower cranes based on environmental point clouds. *Automation in Construction*, 155, 105046, 2023. <https://doi.org/10.1016/j.autcon.2023.105046>
- [13] Burkhardt M. and Sawodny O. A graph-based path planning algorithm for the control of tower cranes. In *American Control Conference*, 1736-1741, 2021. doi: 10.23919/ACC50511.2021.9482797.
- [14] Thomas M., Qiu J. and Sawodny O. Trajectory sequence generation and static obstacle avoidance for automatic positioning tasks with a tower crane. In *Annual Conf. Industrial Electronics Society*, 1-6, 2021. doi: 10.1109/IECON48115.2021.9589398.
- [15] Burkhardt M., Gienger A., and Sawodny O. Optimization-Based Multipoint Trajectory Planning Along Straight Lines for Tower Cranes. *IEEE Transactions on Control Systems Technology*, 2023. doi: 10.1109/TCST.2023.3308762.
- [16] Li G., Ma X., Li Z. and Li Y. Time-Polynomial-Based Optimal Trajectory Planning for Double-Pendulum Tower Crane With Full-State Constraints and Obstacle Avoidance. *IEEE/ASME Transactions on Mechatronics*, 28(2), 919-932, 2022. doi: 10.1109/TMECH.2022.3210536
- [17] Tian Z., Yu L., Ouyang H. and Zhang G. Swing suppression control in tower cranes with time-varying rope length using real-time modified

- trajectory planning. *Automation in Construction*, 132, 103954, 2021. <https://doi.org/10.1016/j.autcon.2021.103954>
- [18] Hu S., Fang Y. and Moehler R. Estimating and visualizing the exposure to tower crane operation hazards on construction sites. *Safety Science*, 160, 106044, 2023. <https://doi.org/10.1016/j.ssci.2022.106044>
- [19] Kim S., Kim S. and Lee D. Sequential dependency structure matrix-based framework for leveling of a tower crane lifting plan. *Canadian Journal of Civil Engineering*, 45(6), 516-525, 2018. <https://doi.org/10.1139/cjce-2017-0177>
- [20] Lin Z., Petzold F. and Hsieh S. H. Automatic Tower Crane Lifting Path Planning Based on 4D Building Information Modeling. In *Construction Research Congress*, 837-845, 2020. <https://doi.org/10.1061/9780784482865.089>
- [21] Ouyang H., Tian Z., Yu L. and Zhang G. Motion planning approach for payload swing reduction in tower cranes with double-pendulum effect. *Journal of the Franklin Institute*, 357(13), 8299-8320, 2020. <https://doi.org/10.1016/j.jfranklin.2020.02.001>
- [22] Liu Z., Yang T., Sun N., and Fang Y. An antiswing trajectory planning method with state constraints for 4-DOF tower cranes: design and experiments. *IEEE Access*, 7, 62142-62151, 2019. doi: 10.1109/ACCESS.2019.2915999
- [23] Cho S. and Han S. Reinforcement learning-based simulation and automation for tower crane 3D lift planning. *Automation in Construction*, 144, 104620, 2022. <https://doi.org/10.1016/j.autcon.2022.104620>
- [24] Al Hattab M., Zankoul E., and Hamzeh F. R. Near-real-time optimization of overlapping tower crane operations: a model and case study. *Computing in Civil Engineering*, 31(4), 05017001, 2017. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000666](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000666)
- [25] Riga K., Jahr K., Thielen C. and Borrmann A. Mixed integer programming for dynamic tower crane and storage area optimization on construction sites. *Automation in Construction*, 120, 103259, 2020. <https://doi.org/10.1016/j.autcon.2020.103259>
- [26] Ji Y. and Leite F. Optimized planning approach for multiple tower cranes and material supply points using mixed-integer programming. *Construction Engineering and Management*, 146(3), 04020007, 2020. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001781](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001781)
- [27] Zhou C., Dai F., Xiao Z. and Liu W. Location Optimization of Tower Cranes on High-Rise Modular Housing Projects. *Buildings*, 13(1), 115, 2023. <https://doi.org/10.3390/buildings13010115>
- [28] Huang C. Wang Z. K., Li B., Wang C., Xu L. S., Jiang, K., ... and Yang H. Discretized Cell Modeling for Optimal Layout of Multiple Tower Cranes. *Construction Engineering and Management*, 149(8), 04023068, 2023. <https://doi.org/10.1061/JCEMD4.COENG-13146>
- [29] Huang C., Wong C. K., and Tam C. M. Optimization of tower crane and material supply locations in a high-rise building site by mixed-integer linear programming. *Automation in Construction*, 20(5), 571-580, 2011. <https://doi.org/10.1016/j.autcon.2010.11.023>
- [30] Moussavi Nadoushani Z. S., Hammad A. W. and Akbarnezhad A. Location optimization of tower crane and allocation of material supply points in a construction site considering operating and rental costs. *Construction Engineering and Management*, 143(1), 04016089, 2017. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001215](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001215)
- [31] Wang J., Zhang X., Shou W., Wang X., Xu B., Kim M. J. and Wu P. A BIM-based approach for automated tower crane layout planning. *Automation in Construction*, 59, 168-178, 2015. <https://doi.org/10.1016/j.autcon.2015.05.006>
- [32] Marzouk M. and Abubakr A. Decision support for tower crane selection with building information models and genetic algorithms. *Automation in Construction*, 61, 1-15, 2016. <https://doi.org/10.1016/j.autcon.2015.09.008>
- [33] Younes A. and Marzouk M. Tower cranes layout planning using agent-based simulation considering activity conflicts. *Automation in construction*, 93, 348-360, 2018. <https://doi.org/10.1016/j.autcon.2018.05.030>
- [34] Khodabandelu A. and Park J. Integrating BIM and ABS for multi-crane operation planning through enabling safe concurrent operations. *Computing in Civil Engineering*, 1128-1135, 2022. <https://doi.org/10.1061/9780784483893.138>
- [35] Yang, Fang, Luo, Liu and Dong. A BIM-based approach to automated prefabricated building construction site layout planning. *KSCE Journal of Civil Engineering*, 26(4), 1535-1552, 2022. <https://doi.org/10.1007/s12205-021-0746-x>
- [36] Yang B., Fang T., Luo X., Liu B., and Dong M. A bim-based approach to automated prefabricated building construction site layout planning. *KSCE Journal of Civil Engineering*, 26(4), 1535-1552, 2022.
- [37] Li R., Chi H. L., Peng Z., Li X. and Chan A. P. Automatic tower crane layout planning system for high-rise building construction using generative adversarial network. *Advanced Engineering Informatics*, 58, 102202, 2023.

- <https://doi.org/10.1016/j.aei.2023.102202>
- [38] Liu C., Zhang F., Han X., Ye, H., Shi Z., Zhang, J., ... and Zhang T. Intelligent Optimization of Tower Crane Location and Layout Based on Firefly Algorithm. *Computational Intelligence and Neuroscience*, 2022. <https://doi.org/10.1155/2022/6810649>
- [39] Tubaileh A. Working time optimal planning of construction site served by a single tower crane. *Mechanical Science and Technology*, 30(6), 2793-2804, 2016. <https://doi.org/10.1007/s12206-016-0346-8>
- [40] Liu C., Zhang F., Han X., Ye H., Shi Z., Zhang J., ... and Zhang T. Intelligent Optimization of Tower Crane Location and Layout Based on Firefly Algorithm. *Computational Intelligence and Neuroscience*, 2022. <https://doi.org/10.1155/2022/6810649>
- [41] Yu Y., Tian Y., Feng N., and Lei M. Research on lifetime prediction method of tower crane based on back propagation neural network. *Advances in Electronic Commerce, Web Application and Communication*, 2, 111-116, 2012. https://doi.org/10.1007/978-3-642-28658-2_17
- [42] Zhang N. Q., Yu M., Zhang X., Fu J., and Wu H. S. Enhancing tower crane safety using condition monitoring system. *Applied Mechanics and Materials*, 418, 80-83, 2013. <https://doi.org/10.4028/www.scientific.net/amm.418.80>
- [43] Zhang Q. C., and Zhang Z. H. Research on wireless sensor network node for tower crane safety monitoring system. *Applied Mechanics and Materials*, 494, 781-784, 2014. <https://doi.org/10.4028/www.scientific.net/amm.494.781>
- [44] Han Z. G., Hao R. Q., and Zheng X. J. The Online Assessment Method of Tower Crane Effective Life Based on Tower Crane Group Monitoring System. *Advanced Materials Research*, 189, 1066-1070, 2011. <https://doi.org/10.4028/www.scientific.net/amr.189.1066>
- [45] Chen G. Z., Zhang B., and Yang Y. Safety monitoring and protection system of tower crane based on ARM. *Applied Mechanics and Materials*, 263, 610-614, 2013. <https://doi.org/10.4028/www.scientific.net/amm.263.610>
- [46] Deng H. L., and Xiao Y. G. Development of General Embedded Intelligent Monitoring System for Tower Crane. *Applied Mechanics and Materials*, 103, 394-398, 2012. <https://doi.org/10.4028/www.scientific.net/amm.103.394>
- [47] Luo X., O'Brien W. J., Leite F., and Goulet J. A. Exploring approaches to improve the performance of autonomous monitoring with imperfect data in location-aware wireless sensor networks. *Advanced Engineering Informatics*, 28(4), 287-296, 2014. <https://doi.org/10.1016/j.aei.2014.08.004>
- [48] Li Y. and Liu C. Integrating field data and 3D simulation for tower crane activity monitoring and alarming. *Automation in Construction*, 27, 111-119, 2012. <https://doi.org/10.1016/j.autcon.2012.05.003>
- [49] Jiang W., Ding L., and Zhou C. Digital twin: Stability analysis for tower crane hoisting safety with a scale model. *Automation in Construction*, 138, 104257, 2022. <https://doi.org/10.1016/j.autcon.2022.104257>
- [50] Zheng M. G. and Zhu X. H. Design of Tower Crane Intelligent Monitoring Management System Based on PLC and WinCC. *Applied Mechanics and Materials*, 184, 1554-1557, 2012. <https://doi.org/10.4028/www.scientific.net/amm.184.1554>
- [51] Zhou Q. H., Li Q. B., and Chen B. J. Study On Intelligent Control System For Tower Cranes Based On ARM. *Advanced Materials Research*, 518, 4449-4454, 2012. <https://doi.org/10.4028/www.scientific.net/amr.518.4449>
- [52] Wang H. PLC-Based Tower Crane Torque Protection System. *Advanced Materials Research*, 694, 2685-2688, 2013. <https://doi.org/10.4028/www.scientific.net/amr.694.2685>
- [53] Jin L., Liu H., Zheng X., and Chen, S. Exploring the impact of wind loads on tower crane operation. *Mathematical Problems in Engineering*, 1-11, 2020. <https://doi.org/10.1155/2020/2807438>
- [54] Yu J. L., Zhou R. F., Miao M. X., and Huang H. Q. An application of artificial neural networks in crane operation status monitoring. In *Proc. Chinese Intelligent Automation Conference: Intelligent Automation*, 223-231, 2015. https://doi.org/10.1007/978-3-662-46463-2_24
- [55] Zhang D., Li S., and Zhao H. Design of Mobile Monitoring System for Tower Crane in Assembly Construction Based on Internet of Things Technology. In *Intl. Conference on Advanced Hybrid Information Processing*, 588-603, 2022. https://doi.org/10.1007/978-3-031-28867-8_43
- [56] Zhong D., Lv H., Han J., and Wei Q. A practical application combining wireless sensor networks and internet of things: Safety management system for tower crane groups. *Sensors*, 14(8), 13794-13814, 2014. <https://doi.org/10.3390/s140813794>
- [57] Li Q., Fan W., Huang M., Jin H., Zhang J., and Ma

- J. Machine Learning-Based Prediction of Dynamic Responses of a Tower Crane under Strong Coastal Winds. *Marine Science and Engineering*, 11(4), 803, 2023. <https://doi.org/10.3390/jmse11040803>
- [58] Sleiman J. P., Zankoul E., Khoury H., and Hamzeh F. Sensor-based planning tool for tower crane anti-collision monitoring on construction sites. In *Construction Research Congress*, 2624-2632, 2016. <https://doi.org/10.1061/9780784479827.261>
- [59] Gutierrez R., Magallon M., and Hernández D. C. Vision-based system for 3D tower crane monitoring. *IEEE Sensors Journal*, 21(10), 11935-11945, 2020. <https://doi.org/10.1109/jsen.2020.3042532>
- [60] Li Y., Wang S., and Li B. Improved visual hook capturing and tracking for precision hoisting of tower crane. *Advances in Mechanical Engineering*, 5, 426810, 2013. <https://doi.org/10.1155/2013/426810>
- [61] Wang J., Zhang Q., Yang B., and Zhang B. Vision-Based Automated Recognition and 3D Localization Framework for Tower Cranes Using Far-Field Cameras. *Sensors*, 23(10), 4851, 2023. <https://doi.org/10.3390/s23104851>
- [62] Liu P., Chi H. L., Li X., and Guo J. Effects of dataset characteristics on the performance of fatigue detection for crane operators using hybrid deep neural networks. *Automation in Construction*, 132, 103901, 2021. <https://doi.org/10.1016/j.autcon.2021.103901>
- [63] Zhang M. and Ge S. Vision and trajectory-Based dynamic collision prewarning mechanism for tower cranes. *Construction Engineering and Management*, 148(7), 04022057, 2022. [https://doi.org/10.1061/\(asce\)co.1943-7862.0002309](https://doi.org/10.1061/(asce)co.1943-7862.0002309)
- [64] Shringi A., Arashpour M., Golafshani E. M., Rajabifard A., Dwyer T., and Li H. Efficiency of VR-Based Safety Training for Construction Equipment: Hazard Recognition in Heavy Machinery Operations. *Buildings*, 12(12), 2084, 2022. <https://doi.org/10.3390/buildings12122084>
- [65] Juang J. R., Hung W. H., and Kang S. C. SimCrane 3D+: A crane simulator with kinesthetic and stereoscopic vision. *Advanced Engineering Informatics*, 27(4), 506-518, 2013. <https://doi.org/10.1016/j.aei.2013.05.002>
- [66] Shringi A., Arashpour M., Dwyer T., Prouzeau A., Li H. Safety in Off-Site Construction: Simulation of Crane-Lifting Operations Using VR and BIM. *Architectural Engineering*, 29(1), 04022035, 2023. [https://doi.org/10.1061/\(asce\)ae.1943-5568.0000570](https://doi.org/10.1061/(asce)ae.1943-5568.0000570)
- [67] Cheng T., and Teizer J. Modeling tower crane operator visibility to minimize the risk of limited situational awareness. *Journal of Computing in Civil Engineering*, 28(3), 04014004, 2014. [https://doi.org/10.1061/\(asce\)cp.1943-5487.0000282](https://doi.org/10.1061/(asce)cp.1943-5487.0000282)
- [68] Lee G., Cho J., Ham S., Lee T., Lee G., Yun S. H., and Yang H. J. A BIM-and sensor-based tower crane navigation system for blind lifts. *Automation in construction*, 26, 1-10, 2012. <https://doi.org/10.1016/j.autcon.2012.05.002>
- [69] Wang Y., Li T., Dong K., Guo Z., and Fu J. The influence of the inclination of lattice columns on the safety of combined tower crane. *Advances in Civil Engineering*, 1-10, 2022. <https://doi.org/10.1155/2022/5072593>

A Delay Liquidated Damages (DLD) Mitigation Model based on Earned Schedule (ES) and Earned Value Management System (EVMS) Concepts

So-Won Choi¹ and Eul-Bum Lee¹

¹Graduate Institute of Ferrous and Eco Materials Technology, Pohang University of Science and Technology (POSTECH), Korea

smilesowon@postech.ac.kr, dreblee@postech.ac.kr

Abstract

After a slowdown in growth following COVID-19, the global engineering, procurement, and construction (EPC) market has been recovering and experiencing sustained growth since 2021. Nevertheless, the profit margins in the global EPC plant industry continue to decline. This study proposes the earned schedule-delay liquidated damages (ES-DLD) model that integrates the ES concept into the earned value management system (EVMS) to manage project schedules and DLD risks. The model was developed and tested using the project data from the Korean 'P' construction company. The result of applying the model to a single critical path of the Hanoi project was to shorten 12 days, saving \$450K compared to the total DLD of \$3.6M. Ten days of acceleration were found to be optimal in assessing multiple critical paths, reducing the potential loss by \$746K among the total DLD of \$3.6M. The model is expected to contribute to lowering DLD risks by accurately predicting project schedule delays through quantitative forecasting during the project execution.

Keywords –

Engineering, procurement, and construction (EPC); ES-DLD Mitigation Model; Delay Liquidated Damages (DLD); Earned Value Management System (EVMS); Earned Schedule (ES)

1 Introduction

The global Engineering, Procurement, and Construction (EPC) market temporarily experienced a decline in growth due to the impact of COVID-19. However, it has been recovering since 2021, and sustained growth is anticipated until 2030 with a projected compound annual growth rate (CAGR) of 5.5% [1].

Nevertheless, despite this overall positive trend, major global EPC players have consistently faced losses

since the oil price collapse in 2014. According to KPMG, a global consulting firm, only 25% of EPC projects met their original deadlines within the past 36 months, and merely 31% of these projects adhered to their budgets during the same period [2]. Despite market growth, EPC companies faced stagnant profits due to delay liquidated damages (DLD) and challenges in managing schedule risks [3].

The key drivers of project losses are attributed to two factors: 1) expensive construction acceleration and 2) significant DLD. These delays negatively impact the owner's critical path (CP) and project success. Critical path means the longest path when each construction activity is connected. Therefore, a delay in this path negatively affects the entire construction period. This schedule management form is called the critical path method (CPM). EPC contractors must apply accurate schedule delay prediction and advanced project management techniques.

When the construction schedule is delayed, the contractor is obliged to refund the client for any damages caused, called DLD [3]. DLD is calculated by multiplying the contract amount, delayed schedule, and DLD Rate, as shown in Equation (1). This value cannot exceed the result of multiplying Contract Amount by Cap.

$$DLD = \text{Contract Amount} \times \text{Delay Schedule} \times \text{DLD Rate} \quad (1)$$

This study aims to predict construction delays despite budget and resource constraints during the construction phase and support project managers in their decision-making process. The authors proposed an DLD mitigation model by integrating the concept of earned schedule (ES) into the earned value management system (EVMS), naming it the ES-DLD mitigation model. This model is aimed at managing project schedule delays and DLD risks.

The ES concept considers DLD, acceleration costs, and construction productivity to output construction delay and EVMS schedule data, which is converted into

time units to improve the accuracy of delay prediction [4].

2 Literature Review

The authors reviewed the prior research in three categories: scheduling models for predicting delays, improving schedule management by optimizing EVMS, and automation tools and techniques.

Kim et al. [5] and Narbaev et al. [6] applied an ES concept to earned value management (EVM) to understand the schedule performance of a project better. Kim et al. [7] developed the DECRIS model for EPC contractors to mitigate risk by examining the adequacy of engineering resources and predicting construction costs and schedule performance. Kim and Lee [8] analyzed the project status regarding cost and schedule by completing the detailed design developed according to the front-end loading theory. Their findings are still relevant as a build-on point towards preventing cost and schedule overrun.

The second category was to improve the construction schedule management by enhancing the function of EVMS. Pascual et al. [9] proposed the enhanced-EVM model, which can detect delayed and advanced projects by converting time into monetary units. Aramali et al. [10] provided a holistic and up-to-date literature review after reviewing 160 publications related to EVM and EVMS over the past decade. In recent literature, "forecasting/prediction" accounted for largest share, followed by "application of EVMS". Avlijaš [11] provided a state-of-science review and critical analysis of the EVM technique by integrating statistical analysis and Monte Carlo simulation.

The third category concerns automation tools and techniques for EVM and EVMS. In recent years, various studies have been conducted using artificial intelligence (AI) technology for schedule management through EVMS. Wauters and Vanhoucke [12] introduced five AI methods for predicting the final duration of a project and then evaluated their performance in comparison to EVM methods. Acebes et al. [13] proposed a stochastic earned duration methodology (SEDM) for monitoring and controlling stochastic projects by applying earned duration management (EDM), and compared SEDM and SEVM through a case study on actual residential house structural work. Nagendra and Rafi [14] analyzed various applications of AI to identify the optimal domains for applying AI in the construction industry.

Despite the above analysis, there were no research cases on an approach that directly links DLD mitigation and EVMS. This study proposes a model integrating EVMS and ES concept to predict DLD delays occurring in EPC projects. Building upon Kim and Lee's research [8], the authors developed a decision model for schedule acceleration for each activity.

3 Research Process

This study consists of three stages: 1) data collection, 2) modelling, and 3) model testing.

Section 4 introduced the data for this study. Section 5 explained a detailed ES-DLD mitigation model developed through this study. The authors explain the theoretical background and development procedure for the ES-DLD mitigation model applying the ES concept to EVMS. In addition, the detailed mechanisms of option model 1 (CP is 1) and option model 2 (CP is n), which are options of the ES-DLD mitigation model, were explained. In Section 5, the authors tested ES-DLD mitigation model using actual project data, and the expected DLD was estimated in case of delay. Figure 1 schematically shows the model development process of this study.

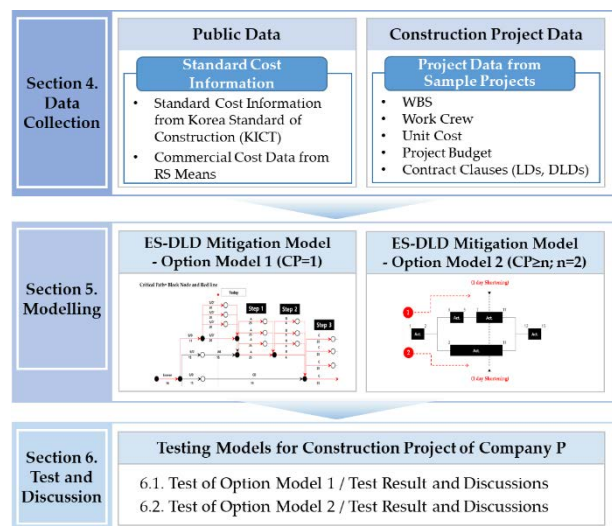


Figure 1. The model development procedure

4 Data Collection

The study used public data such as Korean construction standards and general data from RS Means. The authors also collected the actual EPC project schedule data from P Construction Company in South Korea.

The unit cost information for each assigned work crew was collected through two public datasets. The first was the 2020 Korean Construction Standard [15], organized by the Korea Institute of Construction Technology (KICT) and released yearly. The second was RS Means in the United States [16]. Both datasets were applied to construct the scenario of each working group.

The data required for the construction acceleration assessment was based on the 'A' project executed in Hanoi, Vietnam, undertaken by the 'P' construction

company. The test was implemented with EPC coal-fired power plant project data from Company H, Korea's construction management company. The authors used Work Breakdown Structure (WBS), work crew, budget, the unit price of the work crew, planned value, and earned value information to shorten the project schedule from 'P' company.

5 Modelling

5.1 Integrating ES and EVMS

The authors applied the ES concept to EVMS. The ES is used to overcome the prediction of the time concept in terms of cost, a limitation of EVMS, thus improving the model's accuracy [4].

The ES outputs the planned value (PV) corresponding to the last month (L) and the earned value (EV) corresponding to the current time (AT). Figure 2 shows the EV-L curve in grey and PV-AT in black. At a chosen AT, there is an associated EV and PV. The month delta between these two points is shown in Figure 2 as the horizontal blue dotted line. The month value for each point is shown as a vertical yellow dotted line and a blue dotted line. The month duration between these two points is the I value, also shown as SV_T in Figure 2, a day concept. The derived ES values provide the schedule performance and check for construction delay. For example, if the L value is six and the I value is 0.52, the ES value is 6.52 months. Thus, the earned value duration is 6.52 months, or it is done in 6.52 months' worth of work. If the project has a current L of 7 (ex., started at the beginning of the year and assessed at the end of July), the project has a -0.48 month delay. Figure 2 shows the ES principle and its calculation examples.

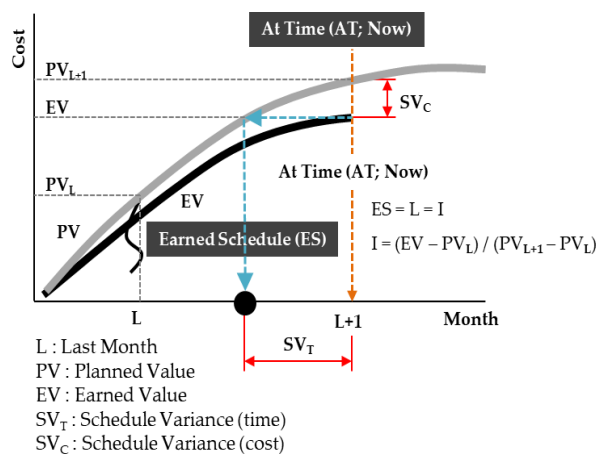


Figure 2. Incorporating ES concept in EVMS

5.2 Developing the ES-DLD Mitigation Model

Through this study, the authors developed an ES-DLD mitigation model, a DLD mitigation model that combines EVMS and ES concept.

First, project managers estimate the schedule delay through EVMS with the ES concept, described in Figure 2. If the construction schedule is ahead of schedule on time, no further analysis is required, and the construction continues according to the schedule. However, if a schedule delay has occurred, the analysis proceeds through the ES-DLD mitigation model. The project manager chooses either Option Model 1 or Option Model 2 based on the number of critical paths. If there is one critical path, select Option Model 1; if not, choose Option Model 2. These models are described below in Section 4.3.1.

The simulation DLD/day and accelerated cost/day (or extra budget/day) are. If the value of DLD per day is less than the average daily shortened costs, the decision to pay DLD is supported. Alternatively, if the DLD per day is larger than the cost to accelerate, schedule acceleration is the supported decision. In case of delayed compensation due to failure to shorten, the provision of credit loss provision supports decision-making.

5.2.1 Option Model 1 (CP=1)

In this study, Option Model 1 is the case where the number of CPs is one. Option Model 1 shortens one CP step by step. Reduction priority is based on the cost reduction per day. The shortening scope of the construction activity is based on the available budget and the maximum number of days. The second shortened activity is shortened based on the remaining budget after the first curtailed activity.

Shown in Table 1 is the input data acquired from P Company Construction of the unit cost of the work crew, execution productivity, and daily work volume of one work crew for each construction activity.

Table 1. Construction activity information used in the scenario

Items	Information on Construction Activity (Ex. Earthwork)
Unit	m3
The total number of work crews available for procurement	40 Groups
1 work crew unit price (daytime)	\$150
1 work crew unit price (3hr nighttime, holiday)	\$84
1 work crew unit price (8hr nighttime, holiday)	\$225
Execution productivity (daytime)	70 %

Execution productivity (nighttime, holiday)	53%
Planned value of 1 work crew (daytime)	\$2,000
Planned value of 1 work crew with execution productivity (daytime)	\$1,400
Planned value of 1 work crew with execution productivity (3hr nighttime)	\$394
Planned value of 1 work crew with execution productivity (daytime+ 3hr nighttime)	\$1,794
Planned value of 1 work crew with execution productivity (daytime+ 11hr nighttime)	\$2,844
Limiting work crew	10 Groups

Based on the input data, the authors calculate the construction period in progress and the remaining construction period and estimate the number of work crews to be put into the planned value in the planned schedule delay scenario. The estimated number of work crews is based on project estimations, shown in Table 2 for this publication. The estimated delayed days and the maximum daily workload based on the work amount are calculated for each working period. The Equation (2) calculates the expected delay date, and the maximum possible reduction is calculated by the Equation (3):

$$\text{Remaining Construction Period} - (\text{Remaining Work} \div \text{1 day Planning Workload}) \quad (2)$$

$$\text{Remaining Construction Period} \times (\text{Maximum Daily Workload} \div \text{1 day Planning Workload}) - 1 \quad (3)$$

The following process calculates the shortened construction period if construction acceleration is chosen. Divide the number of concurrent activities by the total work performed through the changed construction plan. Then divide by the daily work and subtract 1. Subtracting 1 determines the number of days shortened, except for the daytime work. Next, the total result is calculated by listing the possible reduction dates and the cost of the construction schedule, as represented by the following Equation (4).

$$(\sum \text{Construction Workload} \div \text{Number of Concurrent Activities}) \div \text{1 day Planning Workload} - 1 \quad (4)$$

Based on the information of one work crew, it is possible to calculate the construction period delay prediction, the remaining work amount, and the maximum work amount. The cost of the work amount is compared to the DLD to support a decision of acceleration or DLD payment.

5.2.1 Option Model 2 ($CP \geq n$; $n=2$)

In the case of Option Model 2, the number of CPs is two or more, and the model applies to a plurality of instances. If shortening is performed only in one critical path among several critical paths, no shortening is made in the remaining critical paths, as shown in Figure 3. Therefore, no change occurs in the overall construction schedule. All critical paths will need to be reduced to reduce project duration.

In Option Model 2, the project manager (user) enters the schedule data. In Option Model 2, execution productivity is expressed as labor productivity per work area per unit area. (Equation (5))

$$\text{Planned Value} \div \text{Work Area} \quad (5)$$

As shown in Equation (5), if the planned value assigned to the activity is divided by the construction area to which the activity belongs, it can be considered as the labor productivity of the working group per unit area. The premise is not to apply 100% of the planned productivity of the entire activity as in Option Model 1 when establishing Planned Value. Next, the authors find the delay in each activity. Option model 2, unlike 1, looks for deferred activities and derives low priority. The reason for targeting only delayed activities is that there is a risk of being unable to shorten construction due to uncontrollable events such as force majeure. Equation (6) allows Option Model 2 to find the construction work done until the cut-off time.

$$\text{Maximum Workload Per Day} \times \text{Period since construction began} \times \text{SPI} \quad (6)$$

The current work speed is reflected when calculating the construction work because current working speeds indirectly reflect the workforce's capabilities and working environment. SPI is an index defined by the Project Management Body of Knowledge (PMBOK) divided by EV and PV [17]. (Equation (7))

$$\text{Remaining Construction Workload} \div \text{Maximum Workload Per Day} \quad (7)$$

In Equation (7), Option Model 2 predicts the completion time from the remaining construction workload based on the maximum workload per day of the current work crew.

Figure 3 presents a principle for shortening that is applicable in multiple cases.

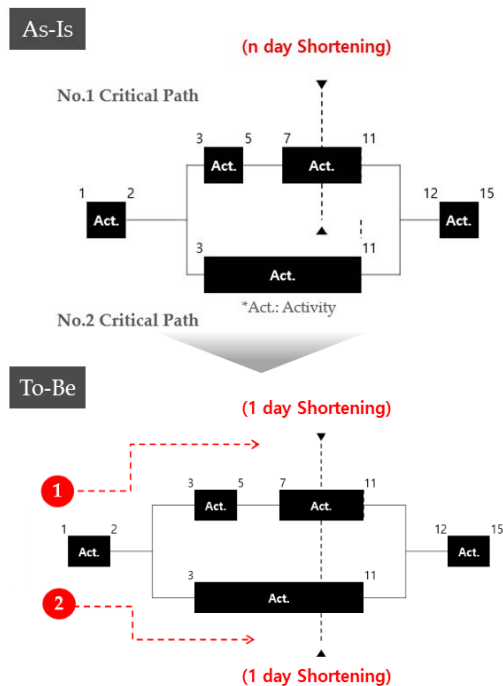


Figure 3. The Developed Principle of Shortening of Multiple Critical Paths

6 Model Test and Discussion

The results of applying the ES-DLD mitigation model to the actual project of P Construction Company are shown below. The test was divided into option models 1 and 2.

6.1 Option Model 1

In the sample project, 72 days are delayed at 75 months into the project, with a cut-off date of April 1, 2021. The developed ES-DLD mitigation model was used to simulate 72 days and derive the results.

In running the ES-DLD mitigation model, the authors assumed there were no problems in procurement and set the total number of work crews to 40. The marginal work crew is introduced by a threshold value that may not reduce the execution productivity of each work crew when additional work crews are added to shorten the construction period. In Equations (2) to (4), the estimated delay days and the maximum daily workload are calculated for each working period of the sample project. The results are shown in Table 2.

Table 2. Applying delay scenario to existing construction plans

Items	Information on construction activity
-------	--------------------------------------

Total construction period	25 Days
Completed construction period	10 Days
Remaining construction period	15 Days
The number of work crews to the planned value (workload)	5 Groups
Planned value	\$175,000
execution productivity	50%
Completed planned value	\$35,000
Remaining planned value	\$140,000
Estimated delayed day	-25 Days
Maximum daily workload with limited work crews	\$7,000
Maximum daily workload with limiting work crews (3hr nighttime)	\$1,969
Maximum daily workload with limiting work crews (8hr nighttime)	\$5,250

The project manager analyses the case study based on delay forecast, remaining work, etc. One of the hallmarks of Option Model 1 is that it tells project managers (users) which activities the ES-DLD mitigation model wants to shorten among those that have not come from now. In other words, it does not only consider activities that have delayed the construction schedule. The next step is to reduce the selected construction activity based on the available budget. In the scenario, the construction activity effectively increased the working time (8 hours at night) and shortened the work without adding a working group. The maximum daily workload is then 9,750. If the remaining construction period is three days, the maximum possible reduction date is 12 days.

In this scenario, 12 of the 72 expected total delays were reduced. In short, it costs \$149,863, and the remaining budget is \$100,137. The late compensation was set at \$50,000 per day. The 60-day delay in the construction period led to a delay compensation, with a total cost of \$3,149,863. If not shortened, this results in a savings of \$450,137 over \$3,600,000. Through this simulation, the validity of the quick decision can be ensured. In addition, it is of great significance to support decision-making by analyzing the DLD settled after the end of the project.

6.2 Option Model 2

This section presents the results of applying the ES-DLD mitigation model to Option Model 2. Table 3 depicts the delay forecasts based on existing construction plan data, using Equations (5) to (7).

Table 3. Applying delay scenario to existing construction plans

Items	Information on
-------	----------------

	construction activity
Work crews	3 Group
1 Work crew unit price	\$1,803/Daytime
Cost per day	\$5,408/Daytime
Labor productivity of 1 work crew per 1m2	40 Unit/Day
Work area	100 m2
Maximum workload per day	12,000 Unit/Day
Remaining construction workload	225,600 Unit
Planned workload	360,000 Unit
Total construction period	30 Day
Completion construction workload	134,400 Unit
The period since construction began	14 Day
Performance index (SPI=EV/PV) against the plan	80 %
Remaining construction period	16 Day
Estimating the period required to complete the construction	19 Day
Delay forecast	-3 Day

The project manager changes the work crew and working hours to develop a new construction plan. At this time, the minimum difference between the target construction work and the construction execution work for 1-day reduction is selected. Here, the shortening success means when the estimated work volume due to the new construction plan exceeds the target work volume for reducing the day. Divide activities with main duplicate processes by the one-day reduction in cost. Through this, it derives the cost of lowering the actual day. Equation (8) finds the activity priority for the catch-up plan in Cost Analysis.

$$\text{Cost for 1-day reduction} \div \text{Number of Concurrent (Overlap) Activities in Critical Path} \quad (8)$$

In Equation (8), activity with duplicate critical paths is divided by the number of identical critical paths from the 1-day reduction cost to find the actual reduction cost. Based on this value, priorities for catch-up are derived. When the calculation is completed, information on the activity to be shortened and the budget required is provided. Finally, with a scenario of \$50,000 DLD, the shortened construction period is ten days, and the cost of shortening the construction period is \$2,853,602. If the construction period is not shortened, \$746,398 can be saved compared to the full DLD.

7 Conclusions

Inadequate management of construction delays in the

early stages of EPC projects leads to DLD issues, excessive resource investments, and cost escalation at the project's completion. This study aims to manage project schedule delays and mitigate DLD risks through the application of the earned value management system (EVMS) integrated model incorporating the concept of earned schedule (ES). As a result, the authors proposed an ES-DLD mitigation model to manage project schedules delays and DLD risks. The ES theory can overcome the limitations of EVMS by converting the cost unit into a time unit when estimating a project's schedule. As a result, it enables highly accurate construction duration predictions depending on the application purpose.

The proposed model is divided into Option Model 1 and Option Model 2 according to the number of critical paths left in the construction stage. The reason for dividing by the number of main processes is because of differences in the shortening model. Also, considering that Option Model 1, which is applied when CP is 1, is near CP when a plurality of main processes are formed during the shortening, the shortening is stopped, and the process proceeds to Option Model 2.

In the result of model test with a case project, the one-day liquidated damage was \$50,000, the budget available to EPC contractors was \$250,000, and the total delay was 72 days. Option Model 1 reduced the simulation by 12 days and \$450,137 compared to the total DLD (\$3.6M). Option Model 2 enabled shorter simulations to save ten days and \$746,398 compared to the total DLD of \$3.6M. These results confirm the developed models' superiority because shortening has not exceeded the total delay compensation.

The ES-DLD mitigation model developed through this study targets the initial execution phases of the EPC projects. The model uses construction information to predict schedule delays and support decision-making. The ES-DLD mitigation model can support project managers in moving beyond experience-based qualitative decision-making to quantitative predictions for project execution. It is expected to contribute to reducing schedule delay risks in EPC projects characterized by complex resource allocation and tight construction schedules.

The limitations and future works of this study are as follows: The validation of the model proposed in this study is based on a single project undertaken by Company P. This raises concerns about the generalizability of the findings. A multi-case study targeting different disciplines and various EPC projects is needed to enhance the broader applicability and robustness of the model. Furthermore, the ES-DLD mitigation model heavily relies on quantitative methods for managing project delays. While such methods offer valuable insights, they have limitations in fully capturing

the complexities of real-world projects. There is a need for research into a more comprehensive approach that integrates qualitative factors such as qualitative elements affecting project delays and unforeseen external events. Finally, it is anticipated that applying machine learning technology to enhance the predictive capabilities of the model could strengthen the DLD mitigation effect and expand the actual application scope in the EPC industry.

References

- [1] Industrial Info Resources. Market Assessment. On-line: <https://www.marketresearch.com/Global-Industry-Analysts-v1039/Oil-Gas-EPC-33869288/>, Accessed: 29/11/2023.
- [2] BrandonGaille. 37 EPC Industry Statistics, Trends & Analysis. On-line: <https://brandongaille.com/37-epc-industry-statistics-trends-analysis/>, Accessed: 29/11/2023.
- [3] Alsharif S. and Karatas A. Framework for Identifying Causal Factors of Delay in Nuclear Power Plant Projects. *Procedia Engineering*, 145: 1-10, 2016.
- [4] Yang J.B. and Lai T.H. Selecting EVM, ESM and EDM (t) for managing construction project schedule. *Eng. Constr. Archit. Manag.*, 2023.
- [5] Kim E., Wells Jr W.G., and Duffey M.R. A model for effective implementation of Earned Value Management methodology. *Int. J. Proj. Manag.*, 21:375-382, 2003.
- [6] Narbaev T. and De Marco A. Combination of growth model and earned schedule to forecast project cost at completion. *J. Constr. Eng. Manag.*, 140, 04013038, 2014.
- [7] Kim M.H., Lee E.B., and Choi H.S. A Forecast and Mitigation Model of Construction Performance by Assessing Detailed Engineering Maturity at Key Milestones for Offshore EPC Mega-Projects. *Sustainability*, 11, 1256, 2019.
- [8] Kim M.H. and Lee E.B. A Forecast Model for the Level of Engineering Maturity Impact on Contractor's Procurement and Construction Costs for Offshore EPC Mega-projects. *Energies*, 12:5-15, 2019.
- [9] Pascual J.L., Rodríguez J.C.M., and Rambaud S.C. The Enhanced-Earned Value Management (E-EVM) Model: A Proposal for the Aerospace Industry. *Symmetry*, 13, 2021.
- [10] Aramali V., Sanboskani H., Gibson G.E., El Asmar M., and Cho N. Forward-Looking State-of-the-Art Review on Earned Value Management Systems: The Disconnect between Academia and Industry. *J. Manag. Eng.*, 38, 2022.
- [11] Avlijaš G. Using Earned Value Management for More Sustainable Project Schedule Control. *Management: J. Sustain. Bus. Manag. Solut. Emerg. Econ.*, 2022.
- [12] Wauters M. and M. Vanhoucke. A comparative study of artificial intelligence methods for project duration forecasting. *Expert Syst. Appl.*, 46 (Mar): 249–261, 2016.
- [13] Acebes, F.; Poza, D.; González-Varona, J. M.; López-Paredes, A. Stochastic earned duration analysis for project schedule management. *Engineering* 2022, 9, 148-161.
- [14] Nagendra S. V. and Rafi N. Application of Artificial Intelligence in Construction Project Management. On-line: <https://www.ijresm.com>: 2018, Accessed: 30/11/2023.
- [15] Korea Institute of civil Engineering and Building Technology (KICT). Standard of Construction Estimate. On-line: <https://www.kict.re.kr/menu.es?mid=a20404000000>, Accessed: 29/11/2023
- [16] RSmeans. *RS Means Cost Data 2020*. RS Means Co, USA, 2020.
- [17] Project Management Institute. *A Guide to the Project Management Body of Knowledge (PMBOK Guide)*, 7th ed. Project Management Institute: PA, USA, 2021.

Impact of Integrated Supply Chain Platforms on Construction Project Management

Giovanni Zenezini¹, Giulio Mangano¹ and Gabriel Castelblanco²

¹Department of Management and Production Engineering, Politecnico di Torino, Italy

² M.E. Rinker, Sr. School of Construction Management, University of Florida, US
giovanni.zenezini@polito.it, giulio.mangano@polito.it, gabriel.castelbl@ufl.edu

Abstract

Supply chain challenges in the construction industry often lead to cost and time overruns, making it imperative to enhance supply chain management. To offer a tool to construction decision-makers, this study investigates the impact of Integrated Supply Chain Platforms (ISCP) on construction project management in the context of Nordic European countries by exploring the effectiveness of integrated software platforms and their influence on project outcomes, focusing on time, cost, and risk management. Utilizing Structural Equation Modeling, the study examines the interplay between ISCP adoption, time management, cost management, and risk management. Results indicate a high satisfaction level among ISCP users, particularly in risk, time, and cost management. The structural equation model analysis demonstrates the positive influence of ISCP adoption on risk and time management, indirectly affecting cost management and overall project performance. The study identifies the significance of prolonged ISCP adoption in improving risk and time management. While ISCP does not directly impact cost management, its indirect influence, mediated through time management, enhances project performance. The findings contribute to understanding the role of digital technologies in optimizing construction supply chains, emphasizing the need for specific sub-frameworks for digital supply chain implementation in the construction industry.

Keywords

Integrated Supply Chain Platforms; Construction; Risk Management; Cost Management; Time Management

1 Introduction

Information constitutes a key instrument in decision-making for the supply chain. Gaining power from information would likely allow industries dealing with

supply chain challenges to exert more control over their suppliers and improve their supply chain management capabilities [1].

In the supply chain, information is frequently used to achieve two objectives [2]: first, organizing everyday operations related to manufacturing, storage, positioning, and transportation; and second, using forecasting and planning to predict future demand and determine the necessary steps to requirements. Effective information dissemination and transmission can enhance all supply chain management components. Information Technology (IT) thus plays a major role in the digitalization of SC information management processes [3].

Planning and operational decision-making can benefit greatly from accurate and timely information managed through information technologies under the umbrella of Enterprise Resource Planning. This is an integrated information system that unifies internal data processing operations, merges operational data and integrates enterprise internal function working processes [4].

Data produced by various Enterprise Resources Planning systems are widely used to create integration with customers, suppliers, or both since they deliver accurate and timely information that will result in supply chain integrity. An Enterprise Resource Planning system offers high levels of cross-functional integration across sales, marketing, and other departments, employs tried-and-true business processes for decision-making, and has the capacity to connect clients and suppliers into a full supply chain [5, 6].

The construction industry traditionally has conceived the supply chain as troublesome and a complicated process often resulting in cost and time overruns. The construction sector is known for its poor performance and limited margins of profit [7]. Waiting, handling materials, and other indirect labor accounted for more than 80% of the typical Swedish craftsmen's working day [7].

Construction sector supply chains may be quite intricate due to the complex interrelations between a vast array of suppliers and subcontractors with heterogeneous objectives, priorities, and organizational values. Supply

chain disruptions trigger multiple underperformance on construction projects, as reflected during the recent global crises. Increasing construction project productivity requires strengthening the supply chain throughout the project's life cycle [8-10].

Digital technologies can improve supply chain processes to guarantee a prompt response to construction project needs. Digital supply chains are conceptualized as a set of inter-organizational systems that companies adopt to digitize transactional and collaborative processes with their supply chain partners, including upstream suppliers and downstream customers [11].

The primary technological advancements with the most significant impact on the digital supply chain are big data and cloud computing [12].

The integration of technology into supply chain management can result in various benefits, such as reducing operational costs and improving stakeholders' satisfaction allowing an efficient, adaptable, and responsive supply chain, which leads to shorter lead times and greater product availability [13]. There is a significant research gap when it comes to evaluating the effect of digital technology applications on construction supply chains in a holistic and objective manner [14]. There are heterogeneous policies, approaches, and practices for implementing digital supply chains within the construction industry. Further research is necessary to develop specific sub-frameworks for digital supply chains in the construction industry. Hence, having a proper guideline and framework for implementing a digital supply chain is essential to facilitating the digital transformation of the construction industry [15].

The primary objective of current research is to measure the effectiveness of integrated software platforms in the management of the supply chain in the construction industry and examine their effectiveness in project outcomes. The research problem for this topic is to understand the influence and implications of integrated supply chain platforms on construction project management criteria in the context of construction projects in Nordic European countries (Denmark and Norway) using Structural Equation Modeling. These countries were chosen because of their advanced digital infrastructure and high levels of technology adoption, their similar economic structures, regulatory environments, and construction standards.

Among the criteria for project management, this study focused on time, cost, and risk management as the parts that could be influenced by the implementation of the integrated software platforms. By using a survey, the validity of the proposed model will be examined, and the degree of influence of the software utilization on different aspects of the project management within the construction industry.

2 Research Methods

The conceptual causal model that illustrates the relationship between the Integrated Supply Chain Platform Adoption and Project Performance was first designed based on theoretical background.

Based on the literature review conducted, key parameters were considered; each represents a critical aspect related to the integrated supply chain platform adoption (ISCPA) in the context of time management (TM), cost management (CM), and risk management (RM) within construction projects, as shown in Table 1. Based on these parameters and their interactions and relations, the latent variables and indicators were built.

Table 1. Acronyms Indicators Analyzed

Symbol	Indicator
Q1	Performance (timelines, scope, budget)
Q2	Stakeholder engagement and satisfaction
Q3	Construction time management
Q4	Lead times for materials
Q5	Project delays
Q6	Order errors
Q7	Unnecessary supply chain processes
Q8	Cost Management
Q9	Unnecessary inventory
Q10	Scheduled budget
Q11	Bringing back the cost to the scheduled budget
Q12	Risk Management
Q13	Real-time project tracking
Q14	Risk identification and mitigation

The Integrated Supply Chain Platform Adoption represents the extent to which construction firms adopt and integrate digital supply chain platforms into their project management processes. This latent variable is composed of indicators that consider factors like technology adoption, information sharing, collaboration, visibility, and transparency.

Time Management assesses the efficiency and effectiveness of project scheduling, project delays, and overall project time performance within construction projects. This latent variable is evaluated based on various time-related metrics and indicators, including project duration, critical path analysis, schedule adherence, and time-related deviations.

Cost Management assesses financial control and expenditure monitoring within construction projects derived from the management of total project costs, avoiding cost overruns, controlling expenses, and adhering to budgetary constraints. This construct encompasses factors such as cost estimation accuracy, cost tracking, cost overruns, and financial control mechanisms.

Risk Management focuses on the identification,

assessment, mitigation, and response to risks within construction projects. This latent variable can be measured through Risk Exposure (a measure of potential project risks) and Risk Mitigation Effectiveness (the effectiveness of risk management strategies implemented during the project).

The model tests eight hypothesis regarding relationships among our latent variables: Integrated Supply Chain Platform Adoption, Time Management, Cost Management, and Risk Management.

The first hypothesis is related to the influence of Integrated Supply Chain Platform Adoption on Cost Management [16]. The second hypothesis asserts that there exists a positive relationship between the Integrated Supply Chain Platform Adoption and Risk Management [17]. The third hypothesis supports that a higher level of Integrated Supply Chain Platform Adoption will increase the level of coordination among stakeholders and thus positively influence Time Management within construction projects [18]. The fourth hypothesis establishes that Cost Management will result in better overall performance in construction projects, through affecting time, cost, and risk management. The fifth hypothesis establishes that Risk Management can enhance Cost Management. Proficient Risk Management is anticipated to have a positive impact on Cost Management by preventing incurred costs due to risk occurrence. The sixth hypothesis presumes that Risk Management can enhance Time Management. The seventh hypothesis asserts that Effective Time Management is expected to have a positive influence on construction Cost Management. The last hypothesis establishes that Time Management will result in a better overall performance in construction projects, through affecting time, cost, and risk management. Time management, facilitated by the adoption and integration of Integrated Supply Chain Platforms, directly influences project timelines, ensuring projects are completed within scheduled durations. Indirectly, by enhancing efficiency and coordination, it reduces costs associated with delays and mismanagement. Furthermore, it aids in risk identification and mitigation by allowing more proactive project management.

The current study used a survey research design. A questionnaire was developed for data collection and testing hypotheses. The questions were formulated based on relevant extant literature supporting the measure of each construct [1,4,6,11,14,15]. The questionnaire included different sections representing items of each construct. A five-point Likert scale is used to measure each item. In this study, the unit of analysis stands for the professionals having working experience in organizations dealing with construction projects in two Nordic European countries.

3 Findings

In order to test our hypothesis on the effects of the ISCP on project management in the construction study, we used a questionnaire as the basis for building the Structural Equation Model.

3.1 Characterization of Respondents

Out of the 594 questionnaires sent out, 197 answers were recollected, representing a response rate of 33% percent, which is larger than similar studies that achieved rates of 16%, 27%, and 28%, respectively [19,20,21]. The initial cleaning process, eliminated 25 of the responses, mainly because either the job position was not related, or they did not use the Integrated Supply Chain Platform in their company or the countries were outside the desired ones.

In the first stage, it was investigated the qualitative characteristics of the Integrated Supply Chain Platforms users concerning their job position, country, level of work experience, project size, software platforms utilized, and time since such a system has been used in the construction project organization. For each variable, statistical values are calculated, accompanied by graphical analyses for a comprehensive understanding.

Analysis of this variable indicates that the job position of users falls into 18 different categories, with the highest frequency belonging to the Project Manager, job category. The relative frequency of this job category is 61.6%, which is in line with the main objectives of the research. Following that, "project director" has the next highest frequency at 17.4%, and "PM assistant" constitutes the majority of users' job frequencies with 8.1%.

To deal with this variable more conveniently, the users' experience level was categorized into 3 main categories. In this research, users are classified into three skill levels: Junior (between 1 to 4 years of experience), Senior (between 5 to 9 years of experience), and Experienced (10 years and beyond). Senior respondents represent the majority of the study's population, accounting for 51.2% of the users. This high percentage suggests that over half of the participants have relevant experience. Experienced respondents make up 35.5% of the participants, indicating that more than a third of the users have extensive experience. Conversely, Junior respondents represent a smaller portion, constituting only 13.3% of the participants, indicating that a lesser fraction of the respondents are at an incipient stage of their professional journey in the field.

The size of the projects undertaken by the Integrated Supply Chain Platforms users is another qualitative variable considered in the present study. Users' projects fall into three categories: Less than 1 M Euros, Between 1 to 10 M Euros, and More than 10 M Euros. According

to users, the highest frequency of economic project scale is, in order, in the category Between 1 to 10 M Euros at 44.8%, followed by Less than 1 M Euros at 34.9%, and More than 10 M Euros at 20.3%. Overall, 79.9% of the respondents are managing projects with values less than 10 M Euros.

The next variable under qualitative examination is the Software platforms employed by users to manage the supply chain in their projects. 16 different software platforms were identified. This variety indicates that the data encompasses responses from different user groups employing a broad spectrum of software solutions for their supply chain management needs. Among these systems, Primavera has the highest frequency percentage (12.8%), surpassing all other systems used. Following that, Oracle Fusion Cloud and NetSuite, with frequencies of 12.2% and 11% respectively, have garnered the most significant representativeness within the construction industry.

The duration of the Integrated Supply Chain Adoption is another crucial variable and has been considered in the present study for its significance. This focus is underpinned by the assumption that an extended period of software adoption correlates with improved organizational proficiency in utilizing and integrating the software. Prolonged engagement deepens users' comprehension and operational proficiency, which enhances understanding not merely about mastering the software but about integrating it effectively within the complex ecosystem of organizational processes and supply chain dynamics, thereby contributing to improved operational outcomes and strategic objectives.. For this purpose, the duration of adoption has been categorized into four groups: less than 1 year, between 1 to 3 years, between 3 to 5 years, and more than 5 years. The results from the frequency table indicate that the majority of users (89.8%) have utilized these platforms for more than 1 year, with the highest adoption period being between 1 to 3 years at 35.5%, followed by between 3 to 5 years at 30.8%.

3.2 Impact of ISCPA on time, cost, and risk management

The impact of the Integrated Supply Chains Platform Adoption on time, cost, and risk management variables, the impact of management variables on each other, and their collective influence on the overall performance of ISCP users.

According to the survey, the users of Integrated Supply Chain Platforms have relatively high satisfaction with the implementation of these systems, as most users have assigned responses of 4 (within a 5-Likert rate).

The users of Integrated Supply Chain Management also demonstrated relatively high satisfaction with the improvement of the system's performance in risk

management, as most users have assigned responses of 4 to the risk management-related questions.

Along the same line, respondents demonstrated high satisfaction with the improvement of the system's performance in time management and cost management, as most users have assigned responses of 4 to these questions.

3.3 Structural Equation Model

The causality of the conceptual model was assessed. The null hypothesis of noncausality was assessed at the usual 5% significance level, and the results are presented in Table 2. Eight hypotheses, each of which indicates a relationship among two variables, were tested. For example, it was found that Integrated Supply Chain Platform Adoption influences Risk Management at a significance level of 5% by rejecting the null hypothesis H2. On the other hand, the null hypothesis H1 was not rejected, resulting in no causal relationship between the two variables (Integrated Supply Chain Platform Adoption and Cost Management).

Table 2 Causality test results

Null hypothesis	T statistic	p-value	Identified causality
ISCPA does not influence CM	1.186	0.236	
ISCPA does not influence RM	2.241	0.025	ISCPA influence CM
ISCPA does not influence TM	3.029	0.003	ISCPA influence TM
CM does not influence performance	2.128	0.034	CM influence performance
RM does not influence CM	3.348	0.001	RM influence CM
RM does not influence TM	4.335	0.000	RM influence TM
TM does not influence CM	3.977	0.000	TM influence CM
TM does not influence performance	1.177	0.240	

Based on the confirmed causal relationship, a Structural Equation Model including all the associated components (variables and indicators in Fig. 1) was developed, and analysis was performed to explore complex relationships, investigate multiple interactions, and evaluate the connectedness of the entire network (model). Within the context of Structural Equation Modeling terminologies, five variables (Integrated Supply Chain Platform Adoption, Risk Management, Time Management, Cost Management, and performance)

correspond to latent variables, and they are graphically enclosed in a circle. Indicators (Q2 - stakeholder engagement and satisfaction, Q4 - lead times for materials) correspond to observed variables, which are defined as a square. Fig. 1 illustrates the overall Structural Equation Model. For estimating the path coefficients, the algorithm iteratively adjusts them to make the model's predicted covariance matrix as close as possible to the observed covariance matrix derived from your data. This involves calculating the derivative of the likelihood function with respect to each parameter and solving these equations to find the parameter values that maximize the likelihood and minimize discrepancy. The widely-used program IBM SPSS version 26 was used for the implementation.

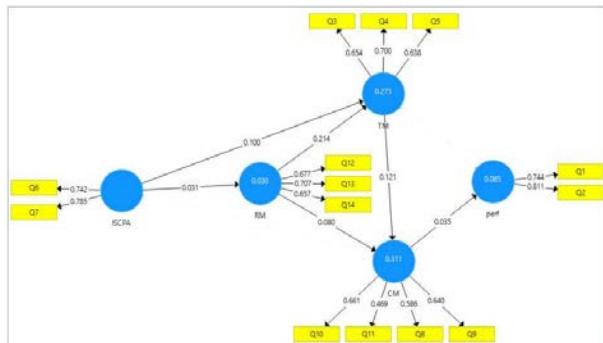


Figure 1. Structural Equation Model Analysis

Analyzing the results of the Structural Equation Model based on the standardized path coefficient (0.742 between Q6 - Reduction in Order Errors and Integrated Supply Chain Platform Adoption in Fig. 1) led to the following two observations. First, the trends of the five variables (i.e., Integrated Supply Chain Platform Adoption, Time Management, Risk Management, Cost Management, and Performance) presented in Fig. 1 align with the results of the Structural Equation Modelling. For example, a standardized path coefficient of 0.87 implies that Risk Management is strengthened as the Integrated Supply Chain Platform Adoption increases. Second, three Cost Management metrics (Q8, Q9, and Q11) and one Time Management metric (Q5) were found to have insignificant relationships, as represented by low standardized path coefficients (0.586, 0.640, and 0.469, respectively) and (0.638). In our study, the cutoff criterion was set to 0.65 (standardized path coefficients lower than 0.65 represent an insignificant relationship). Consequently, Cost Management is driven by a scheduled budget, which is critical for the financing of the project (line of credit) while time management is driven by the lead times for materials.

The study model investigates the impact of implementing the Integrated Supply Chain Platform

Adoption system on users' projects and its role in the intermediary variables of management (cost, and risk management). Subsequently, the role of management variables on the overall project performance was evaluated. To analyze the final structural equation model, two stages must be considered.

Firstly, measurement models need to be examined, and the quality and appropriateness of indicators in fitting variables should be assessed. In the next stage, the fit of structural equations in the final model was assessed. By examining fit indices, validity, and reliability of the final model, the appropriateness and accuracy of the model fit will be assessed.

After drawing the structural equation model and estimating regression coefficients, the results of fitting the proposed structural equation model, separated by variables, are as follows.

The factor loadings related to the studied constructs were examined. If these values exceed the critical value of 0.4, it indicates the appropriateness of the item (the question in question) in measuring the latent variable. As observed, appropriate values for standardized factor loadings have been obtained for all constructs, and all considered items in the structural equation model.

In the Integrated Supply Chain Platform Adoption, the factor loadings for order errors and unnecessary supply chain processes (0.742 and 0.785) have been included in the model considering that both demonstrated p-values < 0.05.

There were also three controlling variables for the Integrated Supply Chain Platform Adoption, such as the time of adaptation, the size of the projects, and the experience level of the project team that performed factor loadings (0.654, 0.7, and 0.638) with p-values < 0.05.

4 Discussion and Conclusions

The previous studies on Enterprise Resource Planning systems in general were mainly focused on customizations and selecting the proper modules for each firm. To advance construction supply chain management, this study focused on the impact of integrated supply chain platform adoption on the performance of construction projects. The main findings of this study can be summarized as follows:

- The study found no significant differences in software system perception among different groups from two northern European countries, regardless of the country or project firm size.
- Information systems are considered commodities, leading to similar benefits across different project sizes and firm scales.
- Longer use of software systems in organizations leads to greater positive impacts on construction risk and time management.

- Introducing new methods and changing routines in firms require time to be effectively implemented, especially with integrated systems that require collaboration across various company functions.
- Adopting supply chain software positively influences risk management in the construction industry by integrating functions such as engineering, procurement, logistics, and finance.
- Integrated software platforms help construction project managers control supply chains, mitigate risks, and improve time and cost management.
- The Integrated Supply Chain Platform Adoption improves time management through better procurement and delivery control, but it does not directly influence cost management.
- Time management, influenced by the integrated platform adoption, indirectly affects overall project performance through cost management.
- The construction industry's reliance on commodified goods means the integrated platform helps manage failed deliveries and source alternatives, though this does not directly reduce costs or construction times.
- Overall, the Integrated Supply Chain Platform Adoption indirectly leads to better project performance through improved time and risk management, subsequently enhancing cost management.

Based on the findings from the current study, several directions for future research can be outlined to expand understanding and application in the field. First, further research should aim to include a more diverse geographic sample to explore whether the observed consensus extends globally. Second, as the study suggests that longer adaptation periods for software systems result in more significant benefits, future research could track organizations over time to document the progression of impacts on risk, time, and cost management, offering a more detailed understanding of the adaptation process. Third, as technology evolves rapidly, future studies should consider the impact of new software solutions, such as artificial intelligence and blockchain, on supply chain management, risk mitigation, and project performance in the construction industry. Lastly, exploring the comparative effectiveness of different supply chain management software solutions could help identify key features and practices that drive success in risk, time, and cost management.

References

- [1] Alimohamadian S. and Abdi F. Analyzing the effects of information technology on supply chain integration: The role of ERP success mediator. *Management Science Letters*, 4 (4): 799–806, 2014, doi: 10.5267/j.msl.2014.2.003.
- [2] Mashreghi M. and Nahavandi N. The Role of Information Technology on Supply Chain Performance with an Emphasis on the Integration and Flexibility Case Study of Automobile Parts Company. In *First International Conference on Management, Innovation and Entrepreneurship*, 2010.
- [3] Calatayud, A., Mangan, J., Christopher, M. The self-thinking supply chain. *Supply Chain Management: An International Journal*, 24: 22–38, 2019.
- [4] Adaileh M. and Abu-alganam K. The Role of ERP in Supply Chain Integration. *Journal of Computer Science and Network Security*, 2010.
- [5] Habib M. M. Supply Chain Management: Theory and its Future Perspectives, 2010. On-line: <https://www.researchgate.net/publication/323166975>
- [6] Kashyap A. Impact of ERP implementation on Supply Chain Management. *Journal of Computer Applications in Engineering Sciences*, 2011.
- [7] Cox A. and Ireland P. Managing Construction Supply Chains: The Common Sense Approach. *Engineering Construction & Architectural Management*, 2009.
- [8] Lambert D. M. and Cooper M. C. Issues in supply chain management. *Industrial Marketing Management*. 65–83, 2000.
- [9] Wisner J. D. and Tan K. C. Supply chain management and its impact on purchasing. *Journal of Supply Chain Management*, 36 (3): 33-42, 2000.
- [10] Zhao X., Huo B., Flynn B., and Yeung J. The impact of power and relationship commitment on the integration between manufacturers and customers in a supply chain. *Journal of Operations Management*, 26(3): 368-388, 2008.
- [11] Farahani P., Meier C., and Wilke J. Digital Supply Chain Management Agenda for the Automotive Supplier Industry. *Shaping the Digital Enterprise Cham*: 157-172. Springer International Publishing, 2017. doi: 10.1007/978-3-319-40967-2_8.
- [12] Rai A., Patnayakuni R., and Seth N. Firm performance impacts of digitally enabled supply chain integration capabilities. *MIS Quarterly*, 2006.
- [13] Deloitte. Industry 4.0 Challenges and solutions for the digital transformation, 2015.
- [14] Yevu, S.K., Ann, T., Darko, A. Digitalization of construction supply chain and procurement in the built environment: Emerging technologies and opportunities for sustainable processes. *Journal of Cleaner Production* 322, 2021.
- [15] Büyüközkan G. and Göçer F. Digital supply chain: Literature review and a proposed framework for future research. *Computers in Industry*, 97: 157-177,

- 2018.
- [16] Kelle P. and Akbulut A. The role of ERP tools in supply chain information sharing, cooperation, and cost optimization. *International Journal of Production Economics*, 93: 41-52, 2005. doi: 10.1016/j.ijpe.2004.06.004.
 - [17] Mandičák T., Mésároš P., Kanáliková A., and Špak M. Supply chain management and big data concept effects on the economic sustainability of building design and project planning. *Applied Sciences (Switzerland)*, 11 (23), 2021. doi: 10.3390/app112311512.
 - [18] Chen, Q., Hall, D.M., Adey, B.T., Haas, C.T. Identifying enablers for coordination across construction supply chain processes: a systematic literature review. *Engineering, construction and architectural management*, 28:1083–1113, 2020.
 - [19] Chen, Y., Dib, H., Cox, R.F., Shaurette, M., Vorvoreanu, M. Structural Equation Model of Building Information Modeling Maturity. *Journal of Construction Engineering and Management*, 142(9), 2016.
[https://doi.org/10.1061/\(asce\)co.1943-7862.0001147](https://doi.org/10.1061/(asce)co.1943-7862.0001147)
 - [20] Almarri, K., Alzahrani, S., Boussabaine, H. An evaluation of the impact of risk cost on risk allocation in public private partnership projects. *Engineering, Construction and Architectural Management*, 26(8):1696–1711, 2019.
<https://doi.org/10.1108/ECAM-04-2018-0177>
 - [21] Yang, R. J., Shen, G. Q. P. Framework for Stakeholder Management in Construction Projects. *Journal of Management in Engineering*, 31(4), 04014064, 2015.
[https://doi.org/10.1061/\(asce\)me.1943-5479.0000285](https://doi.org/10.1061/(asce)me.1943-5479.0000285)

Automated Construction Site Safety Monitoring Using Preidentified Static and Dynamic Hazard Zones

Kepeng Hong and Jochen Teizer

Department of Civil and Mechanical Engineering, Technical University of Denmark, Denmark
keho@dtu.dk , teizerj@dtu.dk

Abstract

Ensuring workers' safety at construction sites is complicated as protective measures often involve the tasks of planning, monitoring, and mitigation at the same time. Despite traditional methods during the pre-construction and construction phases that require time-consuming and manual efforts, poor risk assessment and situational awareness can easily lead to unplanned mishaps in detecting and eliminating risk. Semi-automated rule-based risk assessment approaches as they predominantly exist in research (ref. SafeConAI) are capable of designing out known hazards before they appear in the workplace. These, however, tend not to be interoperable with other emerging technology tasked to monitor how well safety is practiced on the construction site. This paper presents a method for enhanced safety incident analysis by fusing preidentified hazard zones that remain in construction schedules (after SafeConAI has been applied to a 4D BIM) with high-precision trajectory data (using RTK-GNSS) of pedestrian workers and heavy construction equipment. A real-life case study validates the method's feasibility yielding, aside from basic statistical spatiotemporal counting of incident numbers and precise locations between the pedestrian workforce and construction equipment, also new insights into the right size of the so-defined protective safety envelopes that should surround the construction machinery. These promising results still require further investigation into the practical applicability, for example, testing the effectiveness of sharing the detailed personalized feedback that becomes now available.

Keywords –

4D Building Information Modeling (BIM), Construction site safety planning, Construction hazards, Pedestrian workers, Proactive personalized feedback, Hazard control measures, Protective equipment envelopes, Safety risk assessment, RTK-GNSS location tracking, visualization.

1 Introduction

Of all workplaces, construction sites have the highest accident and fatality rates. As many as 20 percent of accidents and deaths are from construction sites in the European Union [1]. Construction sites are confined environments where workers are exposed to different types and sources of hazards, which can be categorized into static and dynamic hazards [2]. For instance, static hazards can come from built-in design, where pedestrian walkways transverse equipment driveways, while dynamic hazards come from workers and construction equipment interaction. One effective approach to prevent accidents is to pre-identify these hazards at the planning stage, and implement preemptive measures, such as safe path planning and guardrail installation. Nevertheless, workers are often inevitably under exposure to the hazards, e.g., when workers have to interact with construction equipment. Monitoring the movement of workers and construction equipment and detecting safety incidents, when workers are in proximity to the preidentified hazards, can prevent accidents to a further extent. Ensuring workers' safety at construction sites is complex and challenging, which involves a set of intrinsically coherent and holistic tasks across the planning and construction phases. Despite the fact that a myriad of research focuses on the digitalization and automation of individual tasks, studies on the interconnections between tasks are limited in quantity, practical implementation and consideration of hazard types on construction sites.

Therefore, the study proposes a framework for automated safety incident detection by combining safeBIM models and real-time location tracking systems. It first extracts the location data of static and dynamic hazard zones to the tracking platform. Meanwhile, the real-time locations of workers are fed into the tracking platform for safety incident detection. As a case study, this framework is tested on a real-life construction project. Aiming to bridge the research gap in the interconnection between safety planning and monitoring, this study further investigates refining safety planning with analysis results from safety monitoring.

2 Background

Traditional safety work at the construction site is manual and laborious, such as drawing hazard zones at the printed site layout [3] and conducting daily job hazard analysis by safety managers [4]. With the digitization of construction planning, extensive research has arisen in an effort to automate traditional manual methods. Hazard zone identification algorithms have enabled a more efficient identification and more accurate representation of hazard zones in BIM models. Many algorithms have been developed to identify static hazard of different categories (e.g., fall and struck-by hazards) in the design. Zhang et al. developed a rule-checking algorithm in safety planning that identifies fall hazard zones [5]. Johansen et al. designed a 4D BIM-based tool (ref. SafeConAI) to prevent two severe and frequent accidents: falls from heights and falling objects by prevention through design and planning (PtD/P) [6].

Safety monitoring tracks the status and movement of workers, materials, and equipment at the construction stage. One of the goals is to detect safety incidents in static and dynamic environments before the incidents evolve into accidents [7]. Real-time location tracking is an important safety monitoring method, allowing for checking whether safety rules have been violated or conflicts exist between resources spatiotemporally. Researchers have adopted active and passive location-tracking systems for different scenarios on construction sites. For example, Costin and Teizer used passive Radio Frequency Identification (RFID) to locate workers within the indoor environment [8]. Ultra-wideband (UWB) is used to track workers in both indoor and outdoor environments [9]. Global Positioning Systems (GPS) have also been widely adopted for outdoor location tracking at construction sites [10, 11]. While GPS tracking often faces the issues of low accuracy and location drifting, Real-Time Kinematics Global Navigation Satellite System (RTK-GNSS) is developed from GPS. Via receiving signals from more satellites and corrections from static receiver, RTK-GNSS can provide an up-to-cm-level accuracy [12]. RTK-GNSS has been widely adopted on the applications requiring high accuracy, such as survey. So far, there have not been applications of RTK-GNSS in construction safety monitoring to the best of our knowledge.

Although Preemptive hazard identification and mitigation in safety planning can effectively mitigate many static hazards, it does not resolve recurring hazards. These hazards can only be mitigated by close safety monitoring and analysis. One important indicator is the observation of safety incidents, where workers enter unauthorized zones (static hazards) or the safety envelope of construction equipment (dynamic hazards). Previously, safety incident detection is inaccurate because of lacking precise static hazard zone and

dynamic hazard geometry. For one thing, the static hazard zones identified in BIM are not explicitly reflected at construction sites, resulting in mitigation equipment being improperly installed and workers being unaware of the exact authorized working areas. One approach to tackle the problem is to import zone information into safety monitoring and integrate it with location tracking. Pfitzner et al. derived floorplans from BIM models and tracked the location of workers from vision data [13]. Costin and Teizer fused RFID data with BIM models for improved accuracy and trajectory visualization [8]. A challenge is to integrate as-planned models and location tracking data efficiently and accurately, which requires not only precise sensors but also accurate system alignment and integration.

For another, the tracked object is, in practice, oversimplified into a point or a circle without considering the actual geometry of the tracked object. Consequently, proximity-based safety incidents are overlooked if the proximity radius is too small or overrated if the proximity radius is too large. Teizer and Cheng define a polygonal equipment representation and warning zone around the equipment for a dynamic hazard zone [2]. Golovina et al. develop a similar polygonal resource boundary and protective envelopes [15]. In [16], we implemented a circular protective envelope to detect safety accidents, and it only provides rough results with incidents. The dynamic hazard zones have to be defined for safety monitoring for further analysis of incidents, which requires careful design of dynamic hazard zone geometries and knowledge of the location-tracking sensor installations.

3 Methodology

The section describes the methods we adopted for automated detection of safety incidents that are induced by static and dynamic hazards. Figure 1 shows the steps connecting BIM-based safety planning and RTK-GNSS location tracking. As a prerequisite, SafeBIM, where assigned and hazard zones are identified at the as-planned safety model, is generated with SafeConAI, a rule-based safety planning algorithm developed in [6].

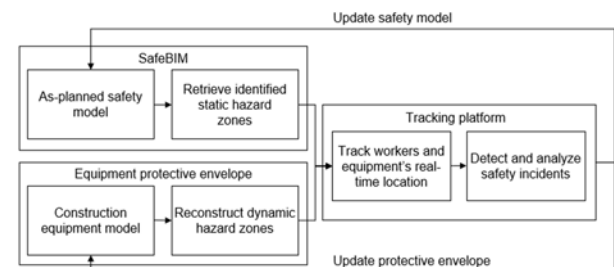


Figure 1. Workflow for connecting safety planning and monitoring.

3.1 Hazard zone retrieval from safeBIM

SafeBIM is generated from the as-planned model. Workers of different trades and equipment are only allowed to move in assigned zones. Static hazard zones, such as fall and struck-by hazard zones, are identified based on safety rules, such as leading edges and height differences in active working space. Mitigative protective equipment (e.g., guardrails) and measures are advised to be installed at hazard zones in safeBIM.

As a starting point, safeBIM is generated in the safety planning stage. A 2D tracking platform is created by retrieving and mapping zones identified in safeBIM. Industry foundation class (IFC), which is an open standard used in BIM, is applied in this study for its interoperability and extensive open-source support. In particular, we use *Ifcopenshell* for reading, writing, and modifying the IFC model [17]. *Ifcopenshell.geom* allows us to efficiently query geometric elements regarding their id, information, and coordinates. We extracted the coordinates of assigned zones and hazard zones in safeBIM. The coordinates of three dimensions in the BIM local coordinate system are flattened to two dimensions at the level where workers and construction equipment move in the tracking platform. To later integrate the trajectory on the tracking platform, zones must be correctly georeferenced. The global coordinate of the origin and true north vector can be retrieved in the correctly georeferenced IFC models. For models that are not georeferenced, one approach is to survey two reference points (P_{1G} and P_{2G}) in the global coordinate system at the construction site and then map the corresponding points (P_{1L} and P_{2L}) in the local coordinate system of the model. The locations contained in the trajectory are translated and rotated using a transformation matrix (M).

$$M = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & \Delta x \\ \sin(\theta) & \cos(\theta) & \Delta y \\ 0 & 0 & 1 \end{bmatrix}$$

The rotation angle (θ) is calculated from the angle between the true north vector and $(0,1)$ in the georeferenced model or the angle between $\overrightarrow{P_{1G}P_{2G}}$ and $\overrightarrow{P_{1L}P_{2L}}$. The translation consists of two-directional movement, the distance between the origins in two systems Δx in the x-direction and Δy in the y-direction.

3.2 Dynamic hazard zone reconstruction

One source of dynamic hazards is the movement of equipment at the construction site. Workers often have to work close to moving equipment, which poses struck-by hazards to pedestrian workers. One common approach is to monitor the proximity between workers and construction equipment. According to OSHA, workers should always maintain a distance of six feet (1.8 meters)

from the equipment and not enter the protective envelope. The protective envelope is created at a predefined safety distance from construction equipment, represented with a polygon in this study. On the tracking platform, the protective envelope creates dynamic hazard zones.

Due to the limited number of sensors, the construction location is only represented with one point on the tracking platform. Therefore, the reconstruction of dynamic hazard zones requires not only the location of the sensors but also the position of the sensors in the construction equipment as well as the heading direction of the equipment. As shown in Figure 2, the centroid of construction equipment is set as its origin point and the heading direction as the x-axis so as to determine the positions of the sensor (x_s, y_s) and the polygon's vertices (x_{vi}, y_{vi}). For the forklift in the following case study, the centroid of the particular equipment equals the centroid of the driver's seat, where the GNSS antenna is installed. If the heading direction (\vec{v}_h) is not available from the inertia measurement unit, the heading direction is calculated from the difference in consecutive coordinates in the trajectory.

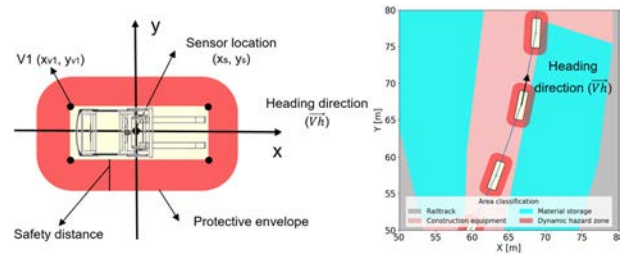


Figure 2. Simplified protective safety envelope for forklift (left image) applied to selected trajectory data (right image).

3.3 Tracking platform setup

After the RTK-GNSS system is set up at the construction site, the location data of workers and construction equipment is streamed and recorded in the global coordinate system using the World Geodetic System (WGS) format, i.e., longitude, latitude, and altitude. The data is first converted to the format of the Universal Transverse Mercator (UTM) coordinate system, a cartesian coordinate system in alignment with the BIM model coordinate system. The global coordinates are then transformed into the local coordinate system in the BIM model. The transformation matrix includes translation and rotation in the two-dimension space. The real-time location and safeBIM are combined to create a safety monitoring tracking platform. For models that are not georeferenced, reference points are required from the construction site. The transformation matrix can be retrieved.

Two types of safety incidents are detected and analyzed on the tracking platform, as illustrated in Figure

3. One is unauthorized entry to static hazard zones, i.e., when workers leave assigned zones or enter hazardous zones. The other is proximity to construction equipment, i.e., when a worker is within dynamic hazard zones. The worker is reduced to a point on the tracking platform so that the safety incident detection algorithm checks whether the point is within the polygonal area. The frequency of safety incidents by worker and location is analyzed to understand the causes of incidents regarding personalized safety performance and incomprehensive safety planning. The measures of utilizing analysis results in enhancing safety planning and training are briefly discussed as well.

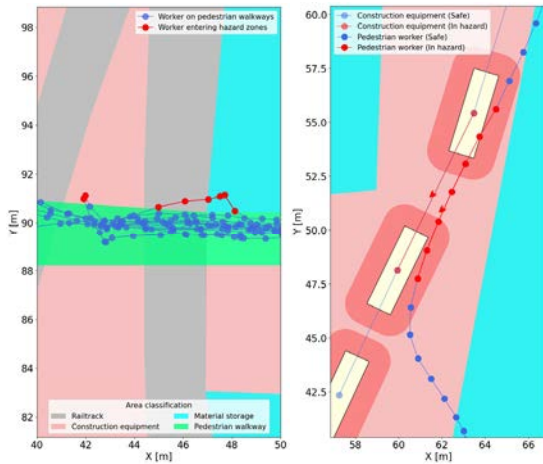


Figure 3. Two examples of safety incidents: unauthorized entry of pedestrian worker in potentially hazardous work site areas (left image), and too close proximity to construction equipment passing by (right image).

4 Case study

A case study was conducted to test the applicability of the framework and methods. The chosen site is a staging area for the rail track replacement project, where workers help equipment to load, unload, and recycle construction materials (e.g., track, sleeper, and ballast). In addition to the construction equipment, railcars also pass through the site to transport the materials to the track replacement site, increasing the complexity of the environment. A safety as-planned model is created to assign zones for different resources and activities (e.g., workers and equipment movement and material storage), based on which a safeBIM is generated to identify hazard zones. On this site, hazards mainly come from the proximity between workers and equipment.

We installed the RTK-GNSS system on the site and tracked the movement of workers and equipment. Real construction activities are monitored, where workers assist in loading and unloading materials. In addition, we

also test the method with simulated scenarios in a controlled environment when there are no construction activities. In the simulated scenarios, two pedestrian workers walk inside and outside pedestrian walkways while cars drive across the pedestrian walkways.

4.1 SafeBIM and RTK-GNSS setup

The contractor provides a BIM model for site layout, where zones are assigned for different resources and activities, such as material storage, pedestrian walkways, and equipment moving zones. A safe BIM model is created on the basis of the as-planned BIM model. It identifies the junctions of the different zones for pedestrian walkways, equipment, and rail tracks. In the safeBIM, the protected areas are identified to indicate where equipment is not allowed to traverse, and railings should be installed. Zones with inevitable built-in hazards are also marked in safeBIM, for example, where construction equipment has to move across pedestrian walkways. The geometry and coordinates of identified hazard zones are retrieved on the 2D tracking platform, where location data is later integrated.

We conducted a precision test for the RTK-GNSS system in comparison with stand-alone GNSS receiver. Both receivers were placed statically at the same spot with clear view to the sky, for 3 hours continuously. In total, 10800 location data points were retrieved from each receiver and a random selection of 3600 data points were plotted and analyzed, as shown in Figure 4. The test results showed that the 50 percentile of the location data falls within a radius of 0.01422 m and 0.5244 m, respectively from the RTK-GNSS receiver and the stand-alone GNSS receiver. And, for the 95 percentiles of the location data, the results were 0.03351 m and 1.2096 m. In addition, the data acquired from the standalone GNSS receiver exhibits a pronounced drift, which tends to skew disproportionately towards certain directions. The results demonstrate that the precision of RTK-GNSS is within the centimeter range, as opposed to the meter-level precision observed in the standalone GNSS solution.

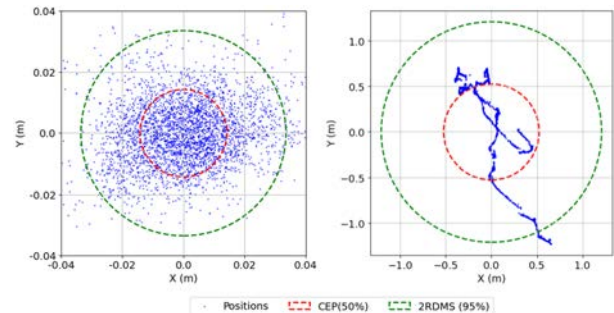


Figure 4. Circular error probable test results of rovers from 3-hour long recordings: RTK-GNSS receiver (right image), (b) Stand-alone GNSS receiver (left image).

Figure 5 shows that the setup of the RTK-GNSS system consists of a static base station and moving rovers. The base station and rovers both contain GNSS antenna and LoRa, whereas the former is used to receive GNSS signal, and the latter is used for the communication between the base station and rovers to correct the rover location. The base station is placed statically at the top of the container at the staging area, and rovers are carried by workers and installed on the construction equipment. The location data is streamed to the digital twin platform for prospective real-time hazard interference. The data is stored locally in the case study for later safety incident detection and analysis.



Figure 5. Setup of RTK-GNSS system on pedestrian workers and equipment.

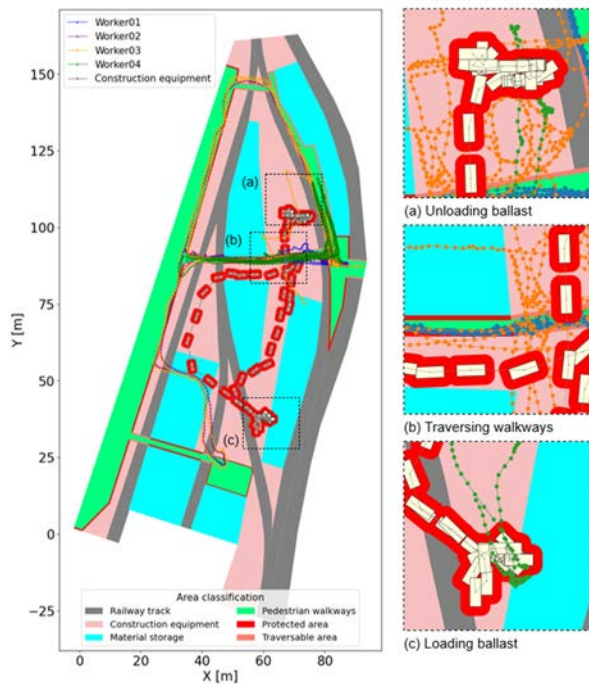


Figure 6. Monitored worker's trajectories (left image) and observed activity of forklift transporting ballast from material storage zone to railcars. (right image) on the 2D tracking platform.

4.2 Tracking platform integration

The 2D tracking platform is created with the input of the safeBIM model, from which we extracted the zone coordinates and created the polygons. Figure 6 (left) shows that four assigned zones and two hazard zones are created on the tracking platform. The origin (0, 0) of the safeBIM model of the site is located at (694697, 534392, 32, U) in the UTM system, and the true north of the site model is (0, 1). The RTK-GNSS system receives location data in WGS. These location data are converted to local coordinates and then mapped on the tracking platform.

To reconstruct the protective envelope of the forklift, we took a measurement of its length (4 m including fork length), width (1.2 m), and the installed location of the GNSS antenna. The protective envelope of the forklift is set at 1 m offset from the polygonal boundary of the construction equipment when it is not driving, and the driving orientation is calculated from the trajectory. Figure 6 (right) shows the observed activities of the forklift transporting ballast from the material storage zone to railcars. The red zone surrounding the forklift indicates the dynamic hazard zones.

4.3 Safety incident detection and analysis

4.3.1 Unauthorized entry

In the simulated scenarios, pedestrian workers were required to walk within and outside the area of pedestrian walkways while one car carrying the RTK-GNSS sensor traveled across the transversal area. Figure 7 shows the result of unauthorized entry incidents from four workers. The causes of the incidents are related to safety planning and workers' safety behavior. From Figure 7(a), it can be seen that worker 03 violated predefined pedestrian walkways more frequently than others. The causes of worker 03 need to be investigated on the actual construction site, which can be due to a lack of sufficient safety training or missing authorization of entry to specific zones. Personalized feedback and training are envisioned to be further provided in the training environment devised.

Other than incidents due to individual-related reasons, some incidents can result from safety planning that fails to reflect the construction site in time. As shown in Figure 7(c), locations with lighter colors have higher safety incident occurrences. Besides, some routes chosen by workers deviate substantially from the safety planning in the BIM model, such as the route connecting the pedestrian walkway and container office, as shown in Figure 7(b). The coordinates of the spots with high incident occurrence are extracted and included in the safety planning model, as the 1x1 meter cells where the occurrence of safety incidents exceeds the average 8 times in this study are marked out in Figure 7(d).

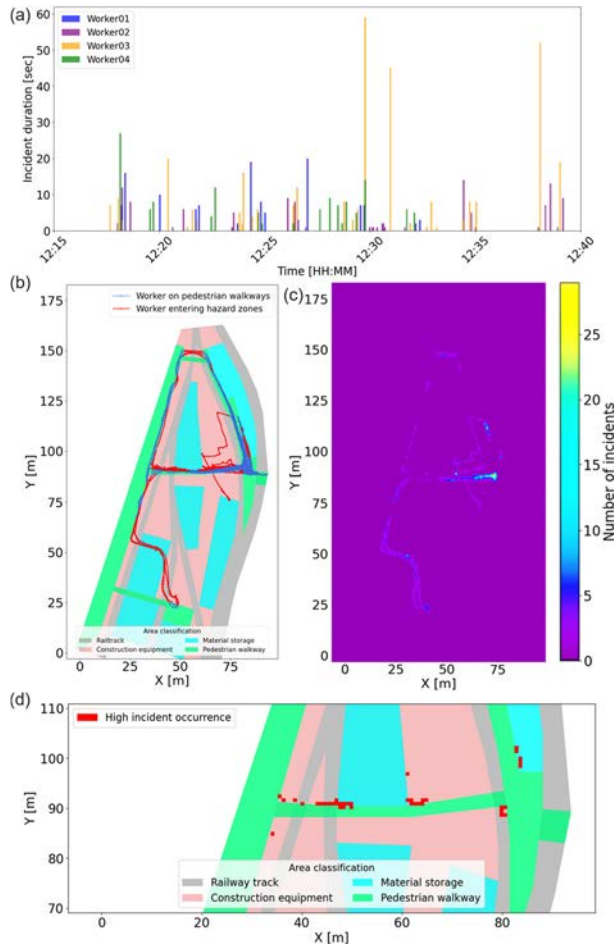


Figure 7. Safety incidents of workers entering unauthorized zones in the simulated scenario: (a) safety incidents count and duration by workers, (b) locations where pedestrian workers leave pedestrian walkways, (c) incident frequency over the construction site, and (d) marked locations with high incident occurrence.

4.3.2 Proximity to construction equipment

For workers working in proximity to construction equipment, we detected the safety incidents from the observed activities of one worker assisting the forklift in transporting ballasts. We compared the detection results using two protective envelopes, circular and polygonal safety envelopes. The protective envelope's safety distance (the radius to the sensors) is set as 3.2 meters long (1m offset from a 2.2 m wide envelope). Figure 8 shows that 11 incidents are detected with the polygonal envelope and 20 incidents with the circular envelope. Figure 9 (left) shows that the circular envelope overlays the polygonal envelope. Hence, the incidents detected with the polygonal envelope are expectably all included in the incidents with the circular envelope.

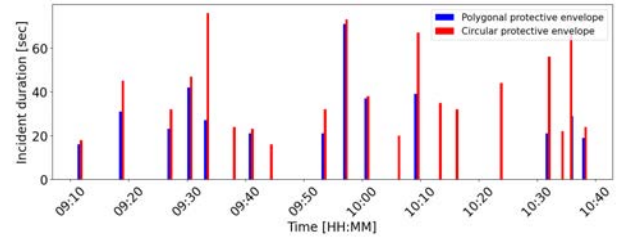


Figure 8. Frequency and duration of a worker in proximity to equipment using the polygonal and circular protective envelope of the forklift.

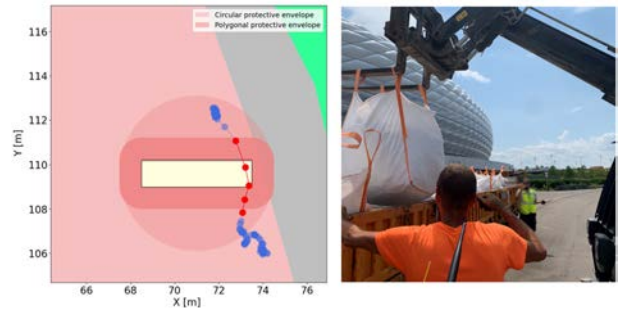


Figure 9. Construction worker walking underneath the forks of the telehandler, as detected on the tracking platform (left image) and as about to be observed on the site (right image).

On the other hand, a circular shape overlooks the geometry of the equipment and identifies incidents even when workers work one meter away from the equipment. A polygonal shape more precisely represents the contour of the equipment and makes the investigation of the incidents more informative. Figure 9 shows an example of a safety incident investigation. It can be seen in Figure 9 (left) that the spotter worker guided the forklift to load the ballast onto the railcar and walked underneath the telehandlers, which is also observed in Figure 9 (right). In comparison, such detailed safety incident investigation is difficult to retrieve with oversimplified equipment geometry and protective envelopes.

The safety distance of the protective envelope is relevant to workers' working distance to the equipment and site compactness. We set different safety distances for the protective envelope and observed the occurrence of proximity safety incidents. The results are displayed in Figure 10, which shows that with the increase in safety distance, the total duration that workers are involved in safety incidents also increases. However, with the safety distance to the forklift increasing from 0.5 meters to 5 meters, the occurrence of safety incidents first increases and then decreases, reaching the highest at 2 meters. The result indicates that the worker mostly works within a 2 m protective envelope to assist the forklift in loading and unloading materials. While at the 1m protective envelope, the occurrence increases drastically, showing that workers mostly keep a 1-meter distance from the forklift.

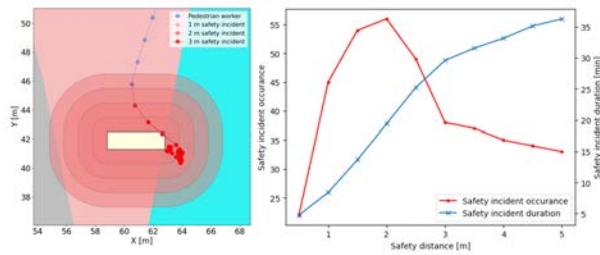


Figure 10. Worker within the protective envelope of different safety distances (left image), frequency and total duration of safety incidents (right image).

5 Discussion

The results validate the applicability of the framework and demonstrate the prospective applications in safety monitoring. The output from safety planning can be efficiently extracted for safety monitoring, thus improving safety incident detection accuracy.

The precision of the framework is dependent on several aspects, including the precision of georeferenced BIM model, location-tracking technology and geometrical measurement of construction equipment. In the case study, we investigated the precision of RTK-GNSS, which can reach cm-level accuracy if there is no obstruction to the receivers to clear sky view. While the precision of the BIM model is not at the scope of this study, there has been research on increasing the georeferencing precision of BIM model. Zhu and Wu devised an approach to geo-reference BIM models with reference points and transformation matrices in the geographic information system (GIS) [17]. Reversely, locations tracked in a global coordinate system can also be converted back to BIM models as described in our method. However, further tests still need to be conducted on the precision of the framework and the methods it adopted, such as validation with other monitoring methods. In addition, the framework is dependent on the availability of as-planned models and the precision is subject to the quality and granularity of the model. Regarding the methods that we adopted in the case study, there are also limitations on the applicable scenarios while alternatives exist for different methods. For instance, although RTK-GNSS is limited to outdoor location tracking, other tracking systems (e.g., Ultra Wideband (UWB), Bluetooth-Low Energy (BLE)) can substitute it when it comes to indoor tracking scenarios.

As a starting point, the framework shows the positive impact of imposing safety planning output on safety monitoring. Compared to other attempts in using BIM-based safety plan for safety monitoring on construction sites, the framework also investigates how safety monitoring results can provide feedback to safety

planning and training in return as a further step. Locations with recurring safety incidents are marked as hazard zones in the safeBIM, while workers with high safety incident rates may be advised to receive safety training. Further investigation can be conducted to understand the causes of safety incidents. Despite better performance in safety incident detection and investigation, the forklifts' geometry and protective envelope can further be enhanced with consideration of driving velocity and the movement of telehandlers. In addition, the protective envelope can further consider blind spots, where hazard risks are higher than in other places. We only studied forklifts in the case study, while other construction equipment can have higher degrees of movement freedom. For instance, excavators can also rotate which creates a more sophisticated protective envelope. When reconstructing the dynamic hazard zone for such equipment, polygons with higher degrees of movement freedom are needed to reflect all components' motion, which also requires consideration of where the sensors should be installed on the machine. As an outlook, the framework should provide a platform where additional (and even robotic) human-machine interaction can be further studied [18-19].

6 Conclusion

This paper presented a framework that automates safety incident detection and monitoring based on technologies that were originally purposed for tasks in detailed model-based safety planning and resource location tracking. The first technology (ref. SafeConAI) generated a model of a safe construction site layout plan based on the construction project's BIM model at a given time in the construction schedule (aka. 4D BIM). Here, the scope was limited to finding and modeling the geometry of simplified objects of so-defined 'static hazard zones' where pedestrian workers' entry is – by following an existing safety rule – restricted. These in reality three-dimensional spaces, for simplicity in this paper reduced to two-dimensional zones, were later in field trials fused with real-time location tracking data of both multiple construction workers and equipment. Small wearable RTK-GNSS tags were designed to function as part of smart safety protective equipment and additional tags were also deployed on the heavy construction equipment that were present in the highly congested work environment. The combination permitted a basic but live safety status monitoring in partially simulated experimentation on a real construction site. Further computational data analysis focused on successfully detecting two particular types of incidents: workforce entering restricted work areas and their proximity to construction equipment. The achieved results give good reasons to conclude that detecting such incident types,

including a timely report of their safety information (i.e., respective locations and frequency) on a first of a kind BIM-based safe construction site layout model, is technically feasible. Furthermore, this work created much-desired information on the parameters of the so-defined (virtual) protective safety envelopes (e.g., size and shape) of construction equipment. These, in future work, might be used to surround the construction equipment and allow generating autonomous and automatic warnings or alerts before mishaps occur and the equipment can seriously harm a pedestrian worker.

7 Acknowledgment

The research presented in this paper has been funded by the European Union Horizon 2020 research and innovation program under grant agreement no. 958310.

References

- [1] Eurostat. Accidents at work - statistics by economic activity. <https://ec.europa.eu/eurostat/>, 04/12/2023.
- [2] Teizer J., Cheng T. Proximity hazard indicator for workers-on-foot near miss interactions with construction equipment and geo-referenced hazard areas. *Automation in construction*, 60: 58-73, 2015, <https://doi.org/10.1016/j.autcon.2015.09.003>
- [3] Kiviniemi M. et al. BIM-based safety management and communication for building construction. *VTT Tiedotteita*. 1-123, 2011.
- [4] U.S. Dept. Health and Human Services. Job Hazard Analysis. Online: <https://www.hhs.gov/about/agencies/asa/foh/ehss/job-hazard-analysis/>, 04/12/23.
- [5] Zhang S., Sulankivi K., Kiviniemi M., Romo I., Eastman C.M., Teizer J. BIM-based fall hazard identification and prevention in construction safety planning. *Safety Science*, 72: 31-45, 2015, <https://doi.org/10.1016/j.ssci.2014.08.001>
- [6] Johansen K.W., Schultz C., Teizer J. Hazard ontology and 4D benchmark model for facilitation of automated construction safety requirement analysis. *Computer-Aided Civil and Infrastructure Engineering*, 38: 2128-2144, 2023, <https://doi.org/10.1111/mice.12988>
- [7] Rao A.S., Radanovic M., Liu Y., Hu S., Fang Y., Khoshelham K., Ngo T. Real-time monitoring of construction sites: Sensors, methods, and applications. *Autom. Constr.*, 136, 104099, 2022, <https://doi.org/10.1016/j.autcon.2021.104099>
- [8] Costin A.M., Teizer J. Fusing passive RFID and BIM for increased accuracy in indoor localization. *Visualization in Engineering*, 3(1):1-20, 2015, <https://doi.org/10.1186/s40327-015-0030-6>
- [9] Cheng T., Teizer J., Migliaccio G.C., Gatti U.C.. Automated Task-level Productivity Analysis through Fusion of Real Time Location Sensors and Worker's Thoracic Posture Data. *Automation in Construction*, 29:24-39, 2013, <https://doi.org/10.1016/j.autcon.2012.08.003>
- [10] Pradhananga N., Teizer J. Cell-based construction site simulation model for earthmoving operations using real-time equipment location data. *Visualization in Engineering*, 3(1):1-16, 2015, <https://doi.org/10.1186/s40327-015-0025-3>
- [11] Vasenev A., Pradhananga N., Bijleveld F.R., Ionita D., Hartmann T., Teizer J., Doree A.G. An information fusion approach for filtering GNSS data sets collected during construction operations. *Adv. Eng. Inf.*, 28(4):297-310, 2014, <https://doi.org/10.1016/j.aei.2014.07.001>
- [12] Wielgocka N., Hadas T., Kaczmarek A., Marut G. Feasibility of using low-cost dual-frequency GNSS receivers for land surveying. *Sensors*, 21(6):1956, 2021, <https://doi.org/10.3390/s21061956>
- [13] Pfitzner F., Braun A., Borrmann A. Towards data mining on construction sites: Heterogeneous data acquisition and fusion. In *European Conference on Product and Process Modelling (ECPM)*, 516-524. 2023, <https://doi.org/10.1201/9781003354222>
- [14] Golovina O., Perschewski M., Teizer J., König M. Algorithm for quantitative analysis of close call events and personalized feedback in construction safety. *Automation in Construction*, 99:206-222, 2019, <https://doi.org/10.1016/j.autcon.2018.11.014>
- [15] Speiser K., Hong K., Teizer J. Enhancing the realism of virtual construction safety training: Integration of real-time location systems for real-world hazard simulations. In *Construction Applications of Virtual Reality (CONVR)*, Florence, Italy, 156-167, 2023, <https://doi.org/10.36253/979-12-215-0289-3.15>
- [16] IfcOpenShell. The open-source IFC toolkit and geometry engine. On-line: <https://ifcopenshell.org/>, Accessed: 08/12/2023.
- [17] Zhu J., Wu P. A common approach to geo-referencing building models in industry foundation classes for BIM/GIS integration. *International Journal of Geo-Information*, 10(6): 362, 2021, <https://doi.org/10.3390/ijgi10060362>
- [18] Hong K., Teizer J. A data-driven method for hazard zone identification in construction sites with wearable sensors. In *Proc. CIBW099W123*, 41-48, <https://doi.org/10.24840/978-972-752-309-2>
- [19] Teizer J., Hong K., Larsen A., Nilsen M.B.. Robotic assembly and reuse of modular elements in the supply chain of a learning factory for construction and in the context of circular economy. In *Construction Applications of Virtual Reality (CONVR)*, Florence, Italy, 568-577, 2023, <https://doi.org/10.36253/979-12-215-0289-3.55>

Innovative standardized cost data structure: application on price list document for estimating public tendering

Jacopo Cassandro, Chiara Gatto, Antonio Farina, Claudio Mirarchi and Alberto Pavan

Department of Architecture, Built Environment and Construction Engineering, Polytechnic of Milan, Italy
jacopo.cassandro@polimi.it, chiara.gatto@polimi.it, antonio.farina@polimi.it, claudio.mirarchi@polimi.it,
alberto.pavan@polimi.it

Abstract –

In the construction industry, managing cost data represents a major challenge. One of the main issues is the inaccurate cost estimation of a project. This study, based on an innovative approach, standardizes and restructures unit cost items defining a new cost ontology starting from fundamental resources (materials, labor, and equipment) and then moving to construction work. A new methodology has been proposed to reduce errors, verify the uniqueness of the cost items, and ensure the correctness of cost estimates. This approach will allow, both humans and machines, to read and use information more easily and accurately. In conclusion, this new structure will provide a standardized and structured cost domain.

Keywords –

Cost Item, Cost Management, Price List, Price Analysis

1 Introduction

Building construction is a complex and time-consuming process involving various participants in the Architecture, Engineering, and Construction (AEC) industry. Specifically, cost estimation is a crucial task in construction projects, involving the calculation of quantities, project costs, and product classification, where all participants must interact with each other to share these various information [1]. In order to minimize the risk of human error when choosing the cost items useful for cost estimation processes, the study proposes creating a structured cost ontology that can be relied upon and applied consistently. This will enable the identification of a standardized procedure for creating, storing, and selecting cost information.

In Italy, a price list is utilized to manage cost data during public tendering. Each region has its price list, which contains construction materials, equipment, labor, services, and construction work. These price lists are used to process economic offers and regulate payments in public contracts. However, the current price lists have

a hierarchical and inflexible structure designed for printing purposes, which makes it challenging to link and cross-reference the various components of the price list. Furthermore, an analysis of the ontology and semantics of the cost items in the AEC/FM sector revealed several shortcomings and critical aspects.

This study aims to establish the basis for developing a standardized cost data structure that will form the foundation of the new digital platform for public works prices in the Lombardy Region to ensure better cost management. The cost data have been broken down and transformed into a new ontology, resulting in more detailed and granular information.

This new standardized structure will contain both elementary resource information and construction work items obtained from elementary resources and based on price analysis. Data structures will be based on specific fields/attributes, consider cost items no longer only as text strings but as replicable computer classes, and ensure future validation and comparison between attributes of a certain cost class and the attributes of the object to which it is associated. This structure will be interrogatable, standardized, and understandable by humans and machines. This will allow to structure not only the data within the proprietary context, as is already the case thanks to the tools currently available (Oracle, SAP, etc.); in the construction sector, in fact, the association of these elements outside their context "owners" remains an unresolved problem.

Therefore, a cost ontology can be used to model and represent knowledge about the cost domain specific to construction projects. It is a system that describes the relevant entities of a domain and the relationships between them in a shared, explicit, and formal way.

Currently, data on cost items are described in natural language without specific structure, which makes them unclear to digital tools and leaves a subjective interpretation of the data open by users such as professionals, architects, and quantity surveyors. This causes errors and a waste of time in understanding, choosing, and evaluating cost items to be associated with geometric objects.

2 Background

According to the Project Management Body of Knowledge (PMBoK) as outlined by the Project Management Institute (PMI), project cost management consists of all those procedures related to planning, estimating, budgeting, financing, funding, overseeing, and regulating expenses to ensure projects are accomplished within their allocated budget [2].

Before the start of the construction stage, cost management is centered on cost estimation and cost planning. In the implementation of this phase, work breakdown structure (WBS), bill of material (BOM), and bill of quantity (BOQ) are carried out.

The inaccuracy of the quantity take-off and bills of quantities was demonstrated to be one important factor affecting the cost performance. The management of elementary resources is therefore fundamental in the planning activity of a production process[3], [4]

Planning and optimizing resource utilization can lead to a significant reduction in the duration and cost of construction projects [5]. Within this section an investigation is performed related to how elementary resources are managed in the state of the art during the production phases, not only focusing on the AECO context but also extending the analysis to other industrial sectors.

Currently, the Bill of Materials (BOM) tool is one of the most widespread and used for the quantification and identification of elementary resources in the production processes of various sectors. It involves the classification and quantification of elementary resources concerning the examined project.

Han et al. [6] research activity propose a framework that analyzes design changes' impact on project cost and emphasizes controlling changes during the project. This framework ensures that no information is missed or lost among 2D drawings, 3D models, and the Bill of Materials (BOM).).

Qiao et al. [7] simplified the project planning process by using a genetic algorithm Petri net cell rules to create a Bill of Materials (BOM) view. By adjusting the optimal start time of non-critical activities, the resource can be balanced when multiple projects are running in parallel. The study makes it possible to decompose the plan process into a BOM, maintaining the leveling of the elementary resources.

In a further theoretical study, the focus is placed on the importance of correctly choosing labor, product, financial, and informative resources. The resource that most significantly impacts the success of a project is the human resource. A competent professional team ensures 75% of the project's success. Many find it wrong that financial resources are the most important. The mistake lies in the way financial resources are invested, no matter how "rich" they are, if spent inefficiently they are as if

non-existent [8].

Wang et al. describe BOM, in the field of mechanical product manufacturing, as a central datum in the complete life cycle management of complex products. In this study, a framework for building the BOM of a complex product is developed, to correctly integrate design information that often varies during the various steps preceding the assembly stage. The developed methodology integrates all BOM information using digital twin models [9].

Regarding the construction sector in the cost estimation and cost control phase, greater interest is placed especially on digitizing the process to increase the accuracy of the cost estimation process using BIM methodologies. The state of the art is affected by a general deficiency in the state of the art regarding the absence of frameworks or guidelines to support practitioners in decision-making, especially when integrating time-related aspects into BIM models for cost management in projects [10]. Hillebrandt et al. study collects some examples of the consequences of failure to plan and indications of instances where resource planning was ineffective[11].

Li et al. [12] found that project planners in the AECO sector overlook details activities while focusing on the master program. They lack a mathematical method to analyze resource allocation. This paper proposes using virtual prototyping to optimize construction planning schedules. Building detailed 3D models requires significant resources. Selecting problem models to be built as detailed models is crucial for saving resources. Building a construction model and resource model database is also crucial for future projects.

An interesting research activity describes an Intelligent Scheduling System (ISS) that uses simulation techniques to allocate resources and prioritize activities to create an optimal project plan that aligns with goals and constraints.[13].

Poshdar et al. [14] developed the Multi-objective Probabilistic-Based Buffer Allocation (MPBAL) method for complex projects. It uses a visual representation of mathematical optimization results to help decision-makers achieve the best outcome. Tested on a bridge project, MPBAL outperformed numerical analysis.

Said et al. [15] developed the Automated Multi-objective Construction Logistics Optimization System (AMCLOS) to help contractors plan material supply and storage efficiently. By automatically gathering project data from scheduling and BIM files and optimizing material supply and site decisions, AMCLOS reduces overall logistics costs.

In the AECO sector, the primary focus is on scheduling on-site activities to complete projects efficiently. However, there seems to be a general lack of standardization when it comes to the knowledge

delivered by cost estimation tools such as BOM and BOW. This creates a missed opportunity to use uniformity of information as a cost control methodology, which has not been explored in any study so far.

3 Methodology

The followed methodology is based on an in-depth analysis of the current state of the art related to the resource analysis theme, and subsequent price analysis, to structure the cost item data which currently is conveyed in the form of unstructured natural language. These themes, as noted in the previous section, do not only concern the construction sector but are found in different fields (automotive, etc.) having a common theoretical basis, albeit with different complexities due to the variability of the working environment.

The research aims to define a new cost item data structure (materials, equipment, labour, and construction work), to ensure the standardisation of information, the definition of a new cost domain, and the possibility in the future to validate the correctness of the cost item information and the objects or documents to which it is associated (geometric model, time schedule, informational specifications). In addition, this new structure will make it possible to obtain cost items that are also understandable and queryable by machines because they are characterised by structured information and no longer in natural language.

The definition of a standard cost data sheet, based on predefined attributes, that can be used for the various types/gender of cost items (material resources, equipment, labour and construction work) and the relations between this information are crucial, in the logic of price analysis.

According to the objectives, the research structure develops in these different stages:

- Study of the state of the art and identification of the problem statement;
- Analysis of current cost items related to construction works and resources in the building sector linked to the civil and infrastructure part;
- Definition of a new data structure for cost items in natural language through meetings and dialogues with industry experts (practitioners, manufacturers, national associations, etc.) and a detailed analysis of cost items;
- Identification of a new price analysis structure based on the link between the new identified data;
- Validation of the structure through application in a practical and real case (the project of digitization of the price of the Lombardy Region);
- Discussion of the results obtained, limitations and possible future developments.

4 Framework

This study is a crucial part of a larger project that aims to create a new digital pricing platform for the Lombardy Region. The research focuses on the structure and standardization of the cost items related to Material, Equipment, and Labor Resources, which together will form the structure of the Construction Works. It is noteworthy that the project will establish several categories of resources and construction works (Figure 1 and Figure 2). To estimate the cost of a project, it is important to have accurate data on the resources needed to complete the project, as well as the associated prices. This information is obtained through a process of analyzing the prices of resources and quantifying the amount of resources needed to complete a unit of work. Therefore, it is crucial to structure the data starting from the Resources, as they are the fundamental building blocks of any project.

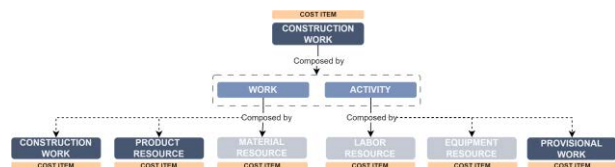


Figure 1. Structure of construction work and the resources/construction work that compose it

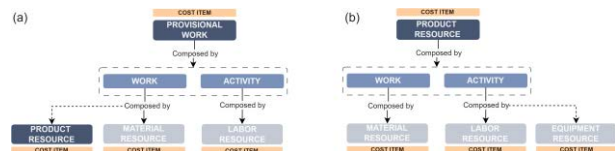


Figure 2. (a) Structure of provisional work and the resources/ product resources that compose it; (b) Structure of the product resources and the resources that compose it

Cost items are currently identifiable in materials, labor, equipment, and construction. Two main groups have been defined to structure the data (Figure 3): one relating to elementary resources (material resources, labor resources, equipment resources) and one to construction works (product resource, provisional work, and construction work).

Elementary resources are the basic elements/components for defining construction works; the latter, on the other hand, are composed of two abstract entities, activity, and work, which together constitute exactly the cost of the construction work. The material resources will allow defining the group "work" while the human and instrumental resources will allow defining the group "labor".

Construction work is the result of the process of

combining work and activity (Slab formation).

The work corresponds to the physical part of the construction work, what will be realized and created (for example, the "slab"); the work consists of elementary resources products/materials (hollow clay blocks, prefabricated joist) or other construction works (Stiffening layer formation).

The activity corresponds to the action that must be performed to carry out the construction work (for example in a new construction activity the "formation/creation" of the slab or in a maintenance activity the "restoration or demolition"); the activity will consist of elementary resources such as equipment (tower crane), labor (construction worker) and other provisional work (Formwork formation, Props formation).

The peculiarity of the structure allows for each identified work ("Slab") to be identified and associated with different activities ("formation, restoration or demolition").

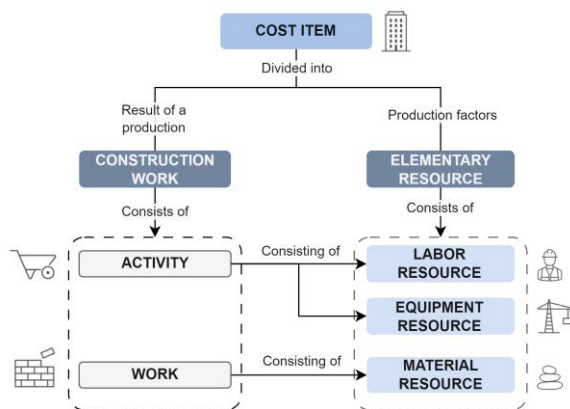


Figure 3. Two main groups of Cost Items

Cost items are structured using a cost data sheet with specific attributes for Elementary Resources (Materials, Labor, Equipment), Product Resources, Provisional Work, and Construction Work.

Cost data sheets are important documents that provide key information about cost items and their descriptions.

To such purpose, a first system of classification of the construction works is implemented to be able to structure them according to a common logic. There are three macro-categories in which to classify a construction work (Figure 4). The first allows to identify which family, category, and object of construction work is referred to (for example family=constructions, category=buildings, object=residence); the second allows to identify to which category, family, and object of activity refers (for example family=production, category=execution, object=formation/creation); finally, the third allows to identify to which family (discipline), category (system), object (technological unit) and matter of work are referred (for example discipline=architecture,

system=slab systems, technological unit=slab, material=concrete and clay-brick). This example is visible in Figure 4.

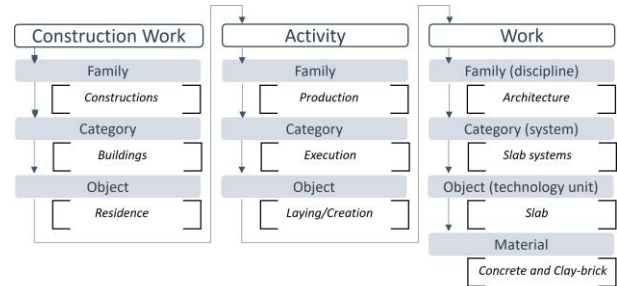


Figure 4. System of classification of the construction works

When linked to a cost value and a prefix consisting of a regional abbreviation, year of publication, and the current version, new price items can be created based on structured data (Figure 5).

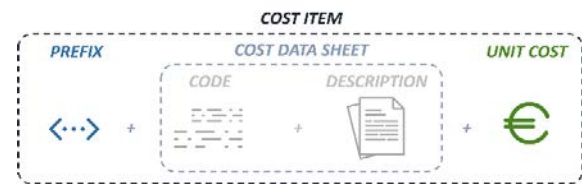


Figure 5. Structure of new price items

About 60 experts/practitioners and trade associations in the construction industry were interviewed at least 3 times in a year each. The goal was to identify the essential information for each type of resource and construction work. This was part of a previous research activity where the research group analyzed and defined attributes through an iterative interactive process with practitioners [16]. Through this qualitative and quantitative analysis of the data currently available (about 20,000 lists of resources and construction work in the price list of the Lombardy Region), the various attributes in which to deconstruct the current information and to define the useful data to define the new structured cost items have been identified. Each of these attributes allows to storage of specific information useful to define a cost item.

The attributes were initially classified as mandatory. Subsequently, a validation was envisaged based on:

- type of choice: single/multiple;
- type of compilation: default values/free (with prevalent use of standardized language to limit the adoption of an unstructured natural language).

The mandatory attributes allow you to identify the coding associated with the description. These are gender,

family, category, object, and material.

Other fundamental attributes are the "characterizing component", and the "auxiliary component" which is fundamental for defining the relations between the various cost data sheets.

Therefore, resource records (which will be associated with a price found following a market analysis) can be instantiated within other resources if it is necessary to define composite records or for the definition of relationships in the context of price analysis for the structuring of the work.

5 Validation and Results

This study is a crucial part of a larger project that aims to create a new digital pricing platform for the Lombardy Region. The validation of the structure has been performed on a dataset of 20 thousand cost items including elementary resources (materials, equipment, and labor) and construction work that are currently contained in the price list of the Lombardy Region.

5.1 Price List

The process of estimating costs for public tendering varies between different States. In Italy, the estimation of prices for public works is determined using a price list. Each region has its price list, which is a catalog containing price items. These price items are the basis for the preparation of the economic offer and regulate payments in public contracts. The determination of the unit prices of the works is based on former Article 32 of the D.P.R. 5 October 2010.

The price items consist of the following:

- A unique code for each processing;
- A description in natural language of the work;
- The unit of measurement of the work;
- The price of the construction work;
- The impact of labor equipment and materials on the unit price.

The protocols mentioned here are specific to the Italian context, but the methodology proposed can be implemented with any cost estimation approach. The factors that determine the cost of construction work, remain essentially the same across any approach, including labor, equipment and materials, overheads, and profits. These factors form the basic components of the final cost and help determine the total processing amount.

5.2 Break down price items and structuring of cost items

It is important to follow this process of breakdown because, from the initial analysis, it was found that the descriptions were unclear.

PRICE LIST ITEM IN NATURAL LANGUAGE	
	Plain slab in reinforced concrete and brick blocks meeting the Minimum Environmental Criteria set out in the Decree 23 June 2022 of the Ministry of Ecological Transition, parallel ribs, cast in place. Including the brick monoblocks, concrete with C20/25 resistance cast in place and vibrated, joists and screeds at the joint, the stiffening layer of a thickness of not less than 5 cm, the formwork and the temporary support frame up to 4,50 m high from the support surface, the relative disassembly. Excluding reinforcing bars.
	- total height 21 cm (16 brick blocks + 5 stiffening layer)
	Plain slab total height 21 cm
	brick blocks meeting the Minimum Environmental Criteria set out in the Decree 23 June 2022 of the Ministry of Ecological Transition, brick monoblocks, height 16
	in reinforced concrete, concrete with C20/25 resistance, stiffening layer of a thickness of not less than 5 cm, height of stiffening layer 5
	parallel ribs, cast in place; cast in place and vibrated
	Including the joists and screeds at the joint, the formwork and the temporary support frame up to 4,50 m high from the support surface, the relative disassembly. Excluding reinforcing bars.

Figure 6. Nonlinear description of a slab cost item

This lack of clarity can lead to misunderstandings, inaccuracies, and errors when selecting and associating price items with their respective geometric objects during cost estimation.

Based on Figure 6, the description provided is non-linear, and information related to a particular processing object can be scattered throughout the description.

Through breakdown, to achieve this, a set of attributes has been defined to compile the information already expressed in the item descriptions.

Various cost data sheets have been identified that differ from each other by type and number of attributes according to the attribute "gender" of the cost item analyzed (construction work, material, etc.).

Once the phase of the breakdown of the data inside of the data sheet of cost, these are reorganized to define the new descriptions and the code of the cost items.

The process of structuring and standardizing cost items began with the analysis of elementary resources (materials, equipment, and labor) and then the construction work.

5.3 Elementary Resources

Elementary resources include material resources, equipment, and labor. Elementary resources are the basic elements/components for defining construction works.

The structuring of the data starts from the attributes contained in the cost sheet. As regards the description of elementary resources, two levels of information structuring have been identified: general and detailed.

The general level provides for the general information of the main resource (for example the data associated with the total ready-mix concrete).

The level of detail describes the constituent elements of the main resource and further specifies the data associated with the cost item (for example, the data associated with binders and ready-mix aggregates); the latter is not necessarily present in all cost items but only where necessary for a greater specification of the cost item. Each of these two levels, if within the same "kind" of resource, is based on an identical standard structure, facilitating the subsequent reading, and understanding of

the data also by the machine. From these data structures, it is possible to define both the description of the cost item and the relative coding. Both are obtained through rules defined from the attributes of the cost data sheet.

Figure 7 shows an example of the cost item of ready-mixed concrete.

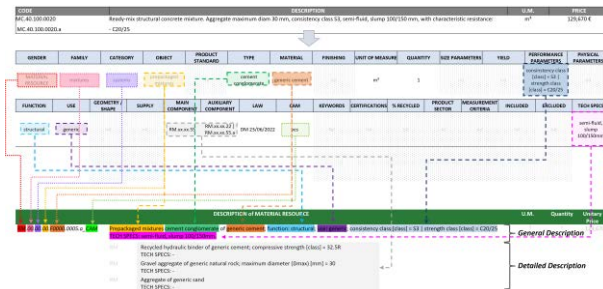


Figure 7. Cost item structure of Resource

5.4 Construction Work

Among the works, there are construction works, provisional works, and product resources.

Unlike the Elementary Resources, structuring the cost data in the workings is more complex.

The architecture of the construction works cost items is based on the logic of price analysis. From the single elementary resources, it is possible to quantify the number of resources (material, equipment, and labor) necessary to obtain the unit working and the relative associated unit price.

However, the research is not limited to the application of the traditional method of price analysis but aims to further structure the cost data to facilitate the understanding of the item and ensure a possible link of the new cost domain with other domains such as tasks, geometries, etc.

The new structuring of construction works involves the addition of a new intermediate level between the elementary resource and the construction work (section “4 Framework”); these intermediate levels breakdown the construction work into work (physical entity - floor stiffening layer) and activity (the activity to be performed on the physical entity - formation, concrete casting). Both have a standard but different cost data sheet.

This subdivision is useful and allows to structure of the information of the cost item that describes the geometric object contained in the 3D model (work) and the activity contained within the list of activities for the management of time 4D (activity). This will then allow to carry out checks on the correctness of the cost item concerning the geometric object to which it has been associated and concerning the timing established by schedule for the performance of the given task.

The two data sheets of work (Figure 8) and activity

(Figure 9) costs are related to each other within the datasheet of construction work cost through specific attributes. This creates a more complex architecture for the cost item, structured, interrogatable, and standardized.

Finally, from these data structures, it is possible to define both the description of the cost item and its coding. Both are obtained through rules defined from the attributes of the cost data sheet.

In Figure is shown an example of a cost item relating to a stiffening layer of a slab.

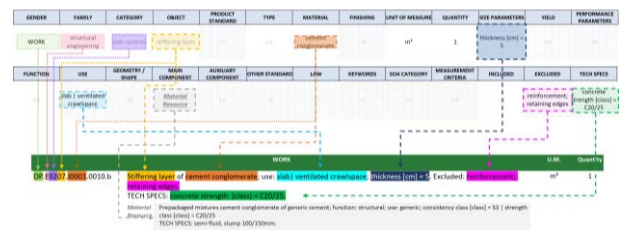


Figure 8. Cost data sheet structure of a Work (stiffening layer)

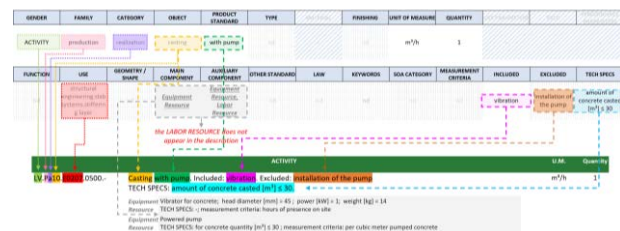


Figure 9. Cost data sheet structure of an Activity (formation)

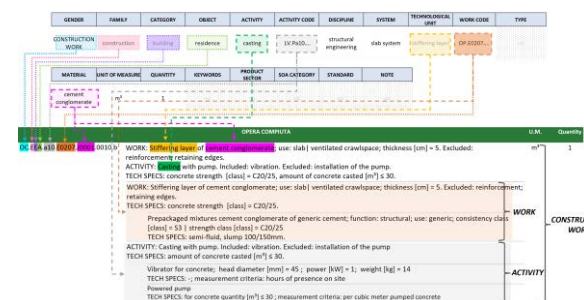


Figure 10. Cost data sheet structure of a Construction Work = Work + Activity (stiffening layer formation)

5.5 Relationships between cost data sheets and price analysis

From the relationships defined within the individual cost data sheets and the construction work data sheet, it is possible to obtain the structure of the associated price analysis. The structure will be created automatically

based on the relationships and information defined and inserted in the different cost data sheets; these are all related to each other. Here is an example of price analysis for the realization of a slab (Figure 11).

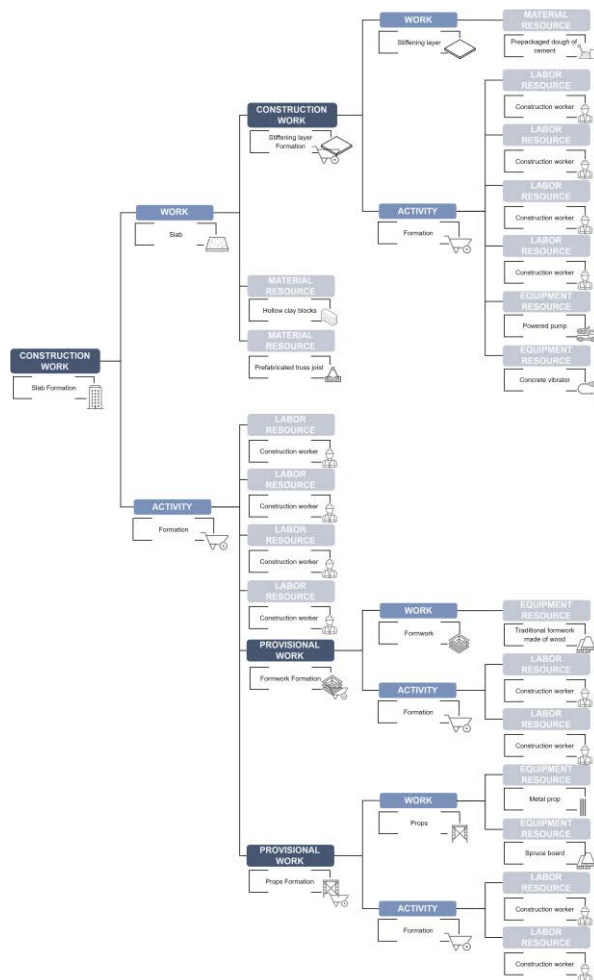


Figure 11. Relationships between cost data sheets and price analysis of “Slab Formation”

6 Discussion

The research has led to the definition of a new data structure that allows to build in a semi-automatic way the price items, relating them to the resources that characterize them. This new structure allows to standardize the information of the cost items that are currently described in natural language and not understandable by machines.

By analysing the state of the art and developing this study it was clear that the lack of uniformity of the language used in the different price items represents a main issue; this makes it difficult to understand and correctly choose cost items during the phases of cost

estimation. A new structure based on cost data sheets with standard attributes is proposed to overcome these problems. This will facilitate the search of data both for users, as it will allow to query the system by general topics or by specific attribute searched, either by any future tools that will want to verify the correctness of the data associated between the cost domain and any other domains (geometric-3D, timing-4D, etc.).

In addition, to overcome these problems, it was also necessary to standardize the terms and define a clear vocabulary to use, avoiding this way duplication and misunderstandings.

Around 180 meetings and interviews were held with practitioners, manufacturers, and trade associations to develop a data structure that could be easily understood and used by end users. Those meetings proved to be necessary to refine the initial structure set and to integrate data on current cost items, leading to modifications and as a result a clearer structure more comprehensible for everyday use. This stage also demonstrates a deep understanding by the end user of the hypothesized structure, which was strengthened through active collaboration and a thorough understanding of the underlying logic of the new cost data sheets.

7 Conclusion

The complexity of the construction projects and the many activities that characterize them, entail a series of difficulties in managing the entire process that can therefore lead to important errors.

To reduce the conflicts that may exist between the design phase, planning, and project management, within this article, a new methodology of structuring elementary resources data (materials, labors, and equipments) has been developed, tested, and demonstrated to be useful in the realization of the construction work through price analysis. In this study, we propose a new architecture that defines the cost domain using a new ontology and semantics of costs. To create this new cost architecture, it is necessary to break down the current cost item into attributes, representing the fundamental parameter that substantiates the unit price value. This helps to increase their generality and abstraction, resulting in elementary constituent cost elements. These elements help to ensure order and standardize the prices, making them replicable in a variety of contexts. The study defines new descriptions of each cost item to make them more understandable, by humans and machines, and reduce the possibility of information misunderstanding. It also established an intuitive and standardized structure and defined a detailed and comprehensive price analysis for each construction work.

During the study, some limitations were identified in the implementation of a standardized methodology for

costs. The first limitation is the market barrier; even if the structure proved to be understood by practitioners, the breaking down process has been performed to assist them. However, the development of a platform for creating, managing, and storing new cost items is currently underway, allowing an easier end-to-use by users. The second limitation is related to the use of terms that are commonly used to define an object but are often misused. To address this issue, a synonymous attribute has been added to facilitate the search. Additionally, the meaning of certain words can vary depending on the region, making it difficult to standardize without imposing specific definitions. The third limitation is due to the industry culture and resistance to change. This can make it challenging to implement new methodologies and practices.

For future work, it is essential to analyze the actual results of the new cost data structure. This analysis will help in estimating costs and verifying their correctness against other crucial project documents, such as 3D models and 4D time management. The research team has already carried out some tests in their previous work [16], [17]. However, these tests need to be extended to more complex case studies.

8 References

- [1] H. Adeli, A. Karim, and N. York, *Construction Scheduling, Cost Optimization and Management*. 2001.
- [2] W. Lu, C. C. Lai, and T. Tse, "BIM and Big Data for Construction Cost Management," *BIM and Big Data for Construction Cost Management*, Oct. 2018.
- [3] Y. M. Cheng, "An exploration into cost-influencing factors on construction projects," *International Journal of Project Management*, vol. 32, no. 5, pp. 850–860, Jul. 2014.
- [4] H. Doloi, "Cost Overruns and Failure in Project Management: Understanding the Roles of Key Stakeholders in Construction Projects," *J Constr Eng Manag*, vol. 139, pp. 267–279, Mar. 2013.
- [5] K. El-Rayes and O. Moselhi, "Optimizing Resource Utilization for Repetitive Construction Projects," *J Constr Eng Manag*, Feb. 2001.
- [6] J. Han, S. H. Lee, and P. Nyamsuren, "An integrated engineering change management process model for a project-based manufacturing," *Int J Comput Integr Manuf*, vol. 28, pp. 745–752, Jul. 2015.
- [7] L. Qiao, Z. Zhang, and Z. Huang, "Construction of Multi-Project Network Planning Based on BOM and Its Resource Leveling," *IEEE Access*, vol. 9, pp. 71887–71899, 2021.
- [8] A. Ioana, M. Costoiu, D. Tufeanu, A. Semenescu, and D. Marcu, "Management elements of conception and development of scientific research projects," *IOP Conf Ser Mater Sci Eng*, vol. 591, no. 1, p. 012092, Aug. 2019.
- [9] Y. Wang, W. Ren, C. Zhang, and X. Zhao, "Bill of material consistency reconstruction method for complex products driven by digital twin," *International Journal of Advanced Manufacturing Technology*, May 2022.
- [10] M. A. Vigneault, C. Botton, H. Y. Chong, and B. Cooper-Cooke, "An Innovative Framework of 5D BIM Solutions for Construction Cost Management: A Systematic Review," *Archives of Computational Methods in Engineering*, vol. 27, no. 4, pp. 1013–1030, Sep. 2020.
- [11] P. M. Hillebrandt and J. L. Meikle, "Resource planning for construction," *Construction Management and Economics*, pp. 249–263, 1985.
- [12] H. Li, N. Chan, T. Huang, H. L. Guo, W. Lu, and M. Skitmore, "Optimizing construction planning schedules by virtual prototyping enabled resource analysis," *Autom Constr*, vol. 18, no. 7, pp. 912–918, Nov. 2009.
- [13] S. M. Chen, F. H. Griffis, P. H. Chen, and L. M. Chang, "Simulation and analytical techniques for construction resource planning and scheduling," *Autom Constr*, vol. 21, Jan. 2012.
- [14] M. Poshdar, V. A. González, G. M. Raftery, F. Orozco, and G. G. Cabrera-Guerrero, "A multi-objective probabilistic-based method to determine optimum allocation of time buffer in construction schedules," *Autom Constr*, vol. 92, pp. 46–58, Aug. 2018.
- [15] H. Said and K. El-Rayes, "Automated multi-objective construction logistics optimization system," *Autom Constr*, pp. 110–122, Jul. 2014.
- [16] C. Gatto, A. Farina, C. Mirarchi, and A. Pavan, "Development of a framework for processing unstructured text dataset through NLP in cost estimation AEC sector," *Proceedings of the 2023 European Conference on Computing in Construction and the 40th International CIB W78 Conference*, Jul. 2023.
- [17] J. Cassandro, M. G. Donatiello, C. Mirarchi, C. Zanchetta, and A. Pavan, "Reliability of IFC classes in ontology definition and cost estimation of public procurement," *Proceedings of the 2023 European Conference on Computing in Construction and the 40th International CIB W78 Conference*, Jul. 2023.

A Proposed Framework to Implement Advanced Work Packaging (AWP) with the Support of Blockchain

Slim Rebai ^{1*}, Olfa Hamdi ², Zoubeir Lafhaj ¹, Wassim AlBalkhy ¹, Rateb Jabbar ³, Hamdi Ayeche ¹, Ahmed Moukhafi ¹, Mohammed Chadli ¹, Pascal Yim ¹

¹ Centrale Lille, CNRS, UMR 9013-LaMcube, Lille, France

² Concord Project Technologies, 1900 S. Norfolk St., Suite 350, San Mateo, CA, 94403, USA

³ KINDI Center for Computing Research, College of Engineering, Qatar University, Doha, Qatar

*Corresponding Author: slim.rebai@centralelille.fr

Abstract

The construction industry faces persistent challenges around efficiency, transparency, and project performance. Advanced Work Packaging (AWP) offers a structured methodology to improve project outcomes through planning alignment and work packaging. However, AWP implementation itself encounters obstacles related to information security, access controls, and supply chain optimization. As an emerging technology with innate capabilities around decentralization, security, and transparency, Blockchain holds promise in addressing these barriers. This paper puts forth a framework for an integrated Blockchain-supported AWP system to enhance construction project delivery. The proposed three-stage framework maps key AWP activities in conceptual planning, detailed engineering, and site execution to Blockchain tools for strengthened information flows, analytics, and coordination. A case study of a firm's digitalization efforts provides real-world context. By bridging AWP and Blockchain, this research aims to spur innovation at the intersection of project management philosophies and next-generation technologies.

Keywords

Blockchain, AWP, Advanced Work Packaging, Construction Project Management, Framework

1 Introduction

1.1 Challenges in the Construction Industry

The global construction industry, pivotal to economic expansion, is experiencing unprecedented growth fueled by rapid urbanization and the increasing demand for

infrastructure. However, this growth is hampered by persistent challenges such as cost overruns and pervasive corruption, which compromise the efficiency and integrity of large-scale construction projects[1][2][3][4]. These systemic issues highlight the critical need for innovative project management methodologies beyond traditional barriers to efficiency and transparency. These challenges indicate the necessity of integrating AWP with technologies capable of addressing these impediments.

1.2 Blockchain Technology: A Solution for AWP Challenges

Blockchain technology, characterized by decentralization, immutability, and enhanced transparency, presents a promising avenue to overcome these hurdles[7][8]. Integrating Blockchain with AWP can revolutionize construction project management, fostering greater efficiency, accountability, and trust among all stakeholders [9][10].

In addition to the benefits above, timestamps' immutability within the Blockchain is pivotal. This ensures that once a transaction or data entry is made on the Blockchain, the timestamp associated with this action cannot be altered, providing an unassailable record of every transaction. This permanence is crucial for construction projects where the timing of decisions, updates, and completions is vital for project scheduling, legal compliance, and dispute resolution. The Blockchain's ability to offer immutable timestamps enhances the reliability and integrity of project data, further solidifying the case for its adoption in managing and executing AWP workflows.

1.3 Synergistic Potential of AWP and Blockchain

This manuscript explores the synergistic potential of AWP and Blockchain within the construction industry, proposing a theoretical framework for their integration. It seeks to provide valuable insights for AWP practitioners, researchers, and Blockchain developers by examining existing applications, pinpointing challenges, and highlighting the transformative benefits of this integration. Ultimately, the paper advocates for adopting integrated AWP-Blockchain frameworks to significantly elevate industry standards, enhance project delivery outcomes, and usher in a new era of efficiency and transparency in construction project management.

The paper's structure articulates the fundamental principles of the AWP methodology, analyzes the advantages of Blockchain for construction project management, and provides a comprehensive overview of Blockchain's applications in the AEC (Architecture, Engineering, and Construction) field. It identifies specific AWP challenges that Blockchain can mitigate and proposes a framework for implementing Blockchain-supported AWP, underscoring the benefits of leveraging emerging technologies to refine and advance construction project management practices. The broader implementation of integrated AWP-Blockchain frameworks is a crucial strategy to enhance project efficiency, reduce costs, and build trust across the construction industry.

2 LITERATURE REVIEW

2.1 AWP

The AWP methodology originated from Workface Planning (WFP), one of the best practices developed to address challenges in oil and gas construction projects in Alberta, Canada[11]. The Construction Owners Association of Alberta (COAA) defines WFP as "the process of organizing and delivering all necessary elements, before starting work, to enable craft persons to perform quality work safely, effectively, and efficiently. This is accomplished by breaking down construction work into discrete work packages that fully describe the project scope to utilize resources and track progress efficiently" [11]. In 2011, a joint research initiative between COAA and the Construction Industry Institute (CII) aimed to review various methodologies including WFP to develop an integrated planning and execution model[12]. This research incorporated case studies, literature reviews, expert interviews, and experiential learning to create the AWP methodology[12][13].

The key principles of AWP include: collaboration between construction and engineering teams during

planning to optimize project sequencing; breaking down overall project scope into construction work packages (CWPs) and engineering work packages (EWPs); early identification and removal of constraints; and integrated planning and work packaging over the entire project lifecycle[11][14]. The central AWP deliverable encompassing these tools is the Path of Construction (POC), prepared through multiple iterations of constraint analysis and sequencing reviews to determine optimal workflow. The POC guides the progression of construction activities from site preparation through mechanical completion and system turnover.

The AWP implementation process contains three main sequential stages:

1. **Conceptual planning** : to define work package deliverables at a high level and divide the project into CWPs aligned with overall execution strategy along with corresponding EWPs. This also involves charter alignment between all stakeholders.
2. **Detailed engineering** : to refine stage 1 work packages into detailed specifications, discipline-specific schedules, and multi-system integration alignment using three week lookahead processes.
3. **Construction planning and execution of installation work packages (IWP)**: issued 3 weeks before start and approved by frontline leads, with quality assurance by owner representatives throughout the build process.

While AWP significantly enhances project delivery outcomes, its broader application in the construction industry encounters several challenges. These include traceability issues, where tracking project progress and changes become cumbersome; security concerns related to safeguarding sensitive project data; and coordination difficulties among a wide array of stakeholders, leading to project inefficiencies [12][15]. These challenges not only hinder the optimal execution of AWP but also signal the need for innovative solutions to address these gaps.

2.2 Use of Blockchain in Construction

In response to these challenges, Blockchain technology emerges as a promising solution. Its features of decentralization, immutability, and enhanced transparency offer novel ways to address the critical issues identified in AWP implementations.

Recent research has examined the potential applications of Blockchain technology in the construction industry. Shojaei and San et al. initiated comprehensive inquiries into leveraging Blockchain for construction[16][17]. Nanayakkara et al.'s study provided insights into using Blockchain for smart contracts, supply chain management, and Building Information Modeling (BIM) systems, emphasizing Blockchain's ability to address persistent issues in

construction like trust and transparency deficits and cumbersome administrative procedures[18]. Additional research by Kim et al. and Dakhli et al. examined Blockchain's profound influence on real estate dynamics and construction project management paradigms[19][20].

There has been growing advocacy for integrating Blockchain to improve engineering supervision information dissemination and data authenticity[21]. Zhao et al. emphasized the importance of offline functionalities and rapid synchronization in innovative frameworks like ChainPM for Construction Project Management Digital Twins[22]. Lu et al. proposed a pioneering trust-building Blockchain framework to facilitate technological and organizational advances in construction[23].

Through collective effort, research has revealed the symbiotic relationship between technological developments and supporting organizational frameworks. Table 1 proposes an overall overview of the potential application of the Blockchain in the construction industry.

Recent scholarship endorses Blockchain integration to enhance transparency, trust, data security, administrative efficiency, and information dissemination in the construction industry. Frameworks facilitating organizational change have been proposed to leverage technological capabilities fully. Continued cross-disciplinary collaboration is illuminating the immense potential of Blockchain applications in construction. Further research is needed to explore its applications in supply chain management, building information modeling, and contract management[24].

Blockchain's potential in construction extends beyond administrative efficiency to tackle AWP's challenges directly. For instance, its immutability and transparency enhance traceability, allowing for a secure and verifiable record of all project changes and updates. This directly addresses the traceability issue in AWP implementations. Furthermore, Blockchain's secure nature can significantly bolster information security, protecting against unauthorized access and data breaches. Lastly, smart contracts on Blockchain platforms facilitate improved coordination among stakeholders by automating and enforcing project agreements and tasks, thus addressing the coordination challenges within AWP frameworks.

Table 1. Use of Blockchain in Construction Industry

Year	Authors	Topic in Construction
2019	San et al.	Potentials, and Impacts for the Construction
2019	Nanayakkara et al.	Smart contracts, supply chain management, and BIM
2019	Dakhli et al.	Real Estate
2020	Kim et al.	Project Management

Systems		
2021	Akinradewo et al.	Applications for the construction
2023	Chen	engineering supervision information dissemination and data authenticity
2023	Zhao et al.	rapid synchronization in Construction Project Management Digital Twins
2023	Lu et al.	framework to technological and organizational advances in construction

Despite the theoretical benefits of integrating Blockchain within AWP practices, there is a notable gap in empirical research demonstrating this integration's real-world applications and benefits. The literature predominantly focuses on Blockchain's potential without providing detailed case studies or evidence of its practical application in overcoming AWP's specific challenges.

Therefore, future research should aim to bridge this gap by providing empirical evidence and detailed case studies on integrating Blockchain technology with AWP. Such research would validate the theoretical advantages discussed and offer practical insights into leveraging Blockchain technology to enhance AWP practices in the construction industry.

3 Proposed Solution

Blockchain technology has shown promise in addressing the myriad challenges faced by construction projects despite its early stages of development. This study proposes a flexible framework incorporating Blockchain technology into monitoring construction projects, utilizing the AWP methodology. A mixed-methods approach was employed to develop this framework, starting with a comprehensive literature review to identify the gaps in AWP practices. Successful implementations of Blockchain technology in similar industries were then analyzed, from which relevant strategies and technologies were extracted. Theoretical modeling was also utilized to tailor these findings to the specific needs of AWP in construction, ensuring that the proposed solution is both innovative and practical. This approach ensures that the framework is grounded in empirical evidence and is tailored to meet the unique challenges of the construction industry.

The Ethereum Blockchain was selected to develop the proposed framework due to its robust smart contract capabilities, widespread adoption, and active developer community. Ethereum's smart contract platform enables the automation of complex project management workflows, which is critical for implementing AWP effectively. The need for transparency and accountability

in construction project management guided the decision to use a public versus private Blockchain architecture. Public Blockchains offer unparalleled transparency, allowing an immutable record of all transactions and interactions. However, considering the sensitivity of some project data, a hybrid approach was proposed, where Ethereum serves as the public layer for transparency, complemented by private, permissioned layers for sensitive information. This hybrid model ensures the security and privacy of critical project data while maintaining the benefits of transparency and immutability for auditability and trust.

Furthermore, the choice between permissioned versus permissionless systems was made to balance accessibility with control. A permissioned Blockchain layer, built on top of Ethereum's public network, allows project stakeholders to have controlled access and permissions tailored to their role in the project, enhancing security and efficiency in managing the supply chain and project workflows.

By integrating a hybrid Blockchain model combining Ethereum's public network with private, permissioned layers, this framework addresses the unique challenges of AWP implementation in construction projects. This approach leverages the strengths of both public and private Blockchain architectures, ensuring that the proposed framework is adaptable, secure, and capable of fostering trust among all project stakeholders.

As the AWP implementation encompasses three primary stages, this framework focuses on the initial and subsequent stages. It begins with receiving project confirmation from the client and initiating the project, as illustrated in Figure 1. The integration of AWP follows, involving the establishment of high-level work package deliverables through conceptual planning and dividing the project into Construction Work Packages (CWPs).

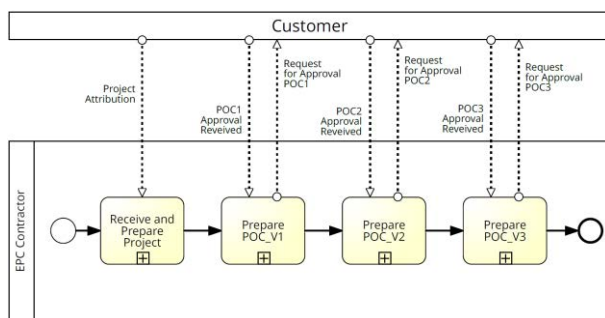


Figure 1: The principle of framework developed of the integration of the AWP by a support of Blockchain

In the conceptual planning stage, Blockchain technology can be applied to securely document and verify the establishment of high-level work package deliverables. Blockchain allows all stakeholders to access a tamper-proof and transparent record of the

planning process, enhancing trust and accountability. For the division of the project into CWPs, Blockchain can facilitate real-time tracking and updates of work packages. Smart contracts automate the approval process of CWPs, ensuring that all relevant parties immediately reflect and approve any changes.

Detailed engineering then provides the specifications and discipline-specific schedules and aligns multi-system integration. The Blockchain is a decentralized repository for specifications and schedules during the detailed engineering stage. This secures the data against unauthorized changes and enables seamless multi-system integration, ensuring that all engineering work is accurately aligned with project requirements.

Termed collectively as the 'Path of Construction' (POC), this three-step process facilitates early contractor monitoring. The project manager manages multiple tasks, including data verification, risk assessment, internal departmental approval, regulatory/legal approval, and POC preparation over three versions. The 'POC' benefits from Blockchain through enhanced traceability and accountability. Each step of the construction process is recorded on the Blockchain, offering an immutable history of project progress. This facilitates early contractor monitoring and provides a reliable performance evaluation and dispute resolution basis.

This structured approach to detailing the implementation process will make the application of Blockchain technology within AWP more explicit, directly addressing the need for refinement and providing a comprehensive understanding of the proposed framework's benefits.

3.1 Current Scenario

The internal organizational procedures of the Project Design and Engineering department focus primarily on project management and oversight. Additionally, this department provides support services to other departments (Figure 1). When a project is transferred from a customer to the Project department, the customer also provides relevant inputs to facilitate the project workflow. The current workflow process relies heavily on manual email and physical document procedures. Each document possesses distinct properties and characteristics.

To critically evaluate this complex workflow and establish an improved framework, the Business Process Model and Notation (BPMN) 2.0 standard was utilized for modeling and visualization [25]. The analysis indicates substantial interconnections and integration challenges in terms of data management within the existing workflow process. Consequently, implementing a Blockchain-based platform could effectively digitize and streamline this process.

After outlining the existing project management

challenges, including reliance on manual processes and the inefficiency of current digital tools, we propose specific Blockchain applications. For instance, Blockchain's decentralized nature allows for a unified platform that streamlines document management, risk assessment, and regulatory approvals, reducing reliance on manual email and document exchanges.

3.1.1 Receive and Prepare Project

Figure 2 illustrates that the study scenario commences with the Project Handover Phase. Upon project attribution, customers are notified and a Project Handover meeting is conducted, overseen by the Projects Sales Department (PSD). A well-defined handover format and a designated signatory for acknowledgment are established. The Project Manager (PM) emails the Project Handover meeting minutes in PSD format.

Upon receiving all project documents, the PM initiates the process of quantitative risk assessment. This evaluation considers technical, contractual, commercial, and environmental hazards. The risk assessment document is updated throughout the entire project, from execution to commissioning. The PM and the engineering team subsequently compile Pre-Qualification Submittals, Company Profile, Project Organization Chart, Professional License and ISO Certificates, QHSE documentation, Initial Design Drawings, and an updated Project Reference List.

The PM seeks internal verification from the Program Manager following the submission of Pre-Qualification Submittals. The Program Manager validates, reviews, and approves the Project Documents. Each revision incorporates additional Project Risk Assessments. Upon obtaining the customer's approval, the project manager proceeds to submit the Pre-Qualification documents to the third-party consultant responsible for issuing the legal Safety Authority Certificate required for Project Permit Approval.

All submitted documents undergo a thorough review process upon request for additional reviews. The Prime Minister collaborates with the consultant to make revisions until the Letter of Acceptance is received.

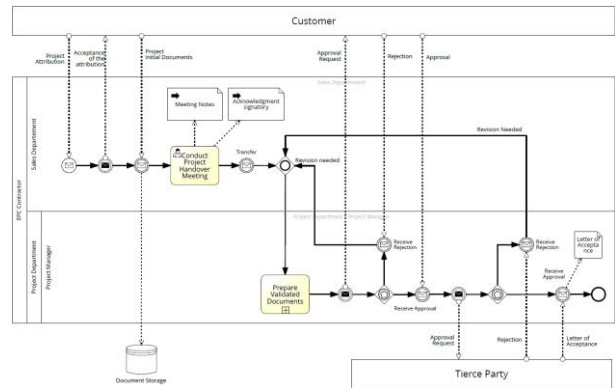


Figure 2: Receive and Prepare Project Process Modeling

3.1.2 Integrating AWP through POC

At this point, the process of integrating the AWP begins, focusing on its primary output, which is the Path Of Construction (POC). The POC will be developed in three iterations, each iteration providing an overview of the integration process and the necessary deliverables to achieve the desired outcome. The AWP Champion plays a crucial role in overseeing and coordinating all the processes at this stage of the AWP methodology implementation.

3.1.3 POC 1 : First iteration of POC

As depicted in Figure 3, the AWP Champion initiates the implementation of POC1 by requesting the Engineering Lead to create the conceptual drawings. Simultaneously, the AWP Champion commences three tasks:

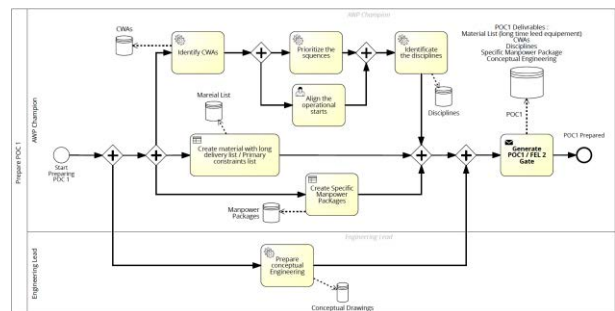


Figure 3: First iteration Prepare POC 1

1. Firstly, the identification of the Construction work areas (CWAs) allows for the prioritization of sequences and the simultaneous alignment of operational activities, facilitating the identification of different disciplines.
2. Compiling a list of materials with extended delivery periods, while considering the primary constraints list.
3. Subsequently, generating the necessary specialized workforce packages.

Upon completing this step, the Path of Construction 1 (POC 1) will be generated and assembled, incorporating the various deliverables produced in each task, and the customer's validation will be requested.

3.1.4 POC 2 : Second iteration of POC

The process to follow for preparing POC2 is illustrated in Figure 4. The AWP Champion will request the Engineering lead to commence work on the Basic Engineering. Simultaneously, the AWP Champion will prepare the Construction Work Packages (CWPs) while also identifying the Engineering Work Packages and the Procurement Work Packages (PWPs). This will enable them to work on preparing the turnover packages and identifying the operating systems concurrently. After completing these tasks, the AWP Champion will identify the testing packages and provide the commissioning packages.

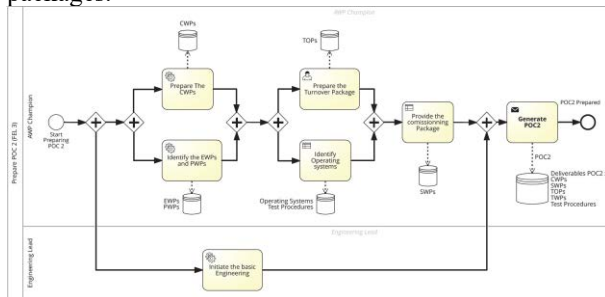


Figure 4: Second iteration to Prepare POC 2

After this step, POC 2 will be created and put together, integrating the different deliverables generated in each task, and the customer's validation will be sought.

3.1.5 POC 3 : Third iteration of POC

The current version of POC preparation incorporates the final set of tasks, as illustrated in figure 5, to fully integrate the AWP into the project monitoring process.

The Engineering Lead will now finalize the fundamental engineering tasks. Simultaneously, the AWP Champion will ensure a one-to-one relationship between the EWPs and PWPs, with regards to the CWPs, in order to guarantee their alignment. Subsequently, the AWP Champion was able to establish a connection between the System Work Packages and the CWPs.

These steps will enable the AWP Champion to develop a comprehensive schedule, assess its feasibility with the subcontractors responsible for different disciplines, and allocate resources accordingly to ensure the active participation of the subcontractors. These steps will result in the successful completion of preparing the deliverables for POC3 and obtaining customer confirmation.

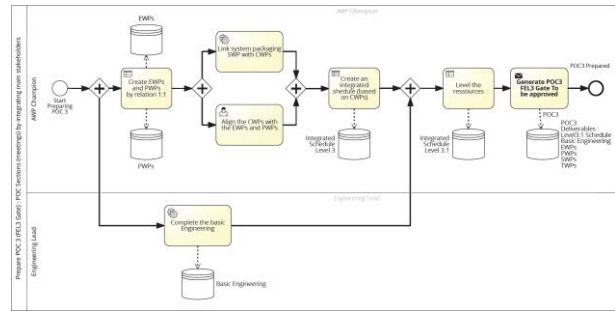


Figure 5: Third iteration Prepare POC 3

After outlining the application of Blockchain within the AWP framework, it becomes clear that there are potential implementation challenges, including technological readiness, stakeholder resistance, and regulatory compliance. Targeted training programs are essential to ensure all participants become well-versed in the functionalities and benefits of Blockchain, fostering a culture of innovation and acceptance.

Conducting empirical validation through pilot projects will be instrumental in demonstrating the practical viability and effectiveness of our proposed Blockchain-AWP framework. These pilots will enable adjustments based on real-world applications, ensuring the framework meets the diverse needs.

3.2 Case Study

3.2.1 Actual digitalization Actions

Various measures were taken to advance and improve the digitalization of this process. An Enterprise Resource Planning (ERP) System was implemented as part of a digitalization project aimed at encompassing all aspects of the company's EPC contractor process. The project is scheduled to be implemented and tested over a period of 4 years. Therefore, the project was divided into three main phases: firstly, the implementation and digitalization of the workflow process, which covers every stage from project attribution to project preparation, including instant notifications and an email system. Secondly, the system will be expanded to integrate the AWP across the three versions of POC. Finally, the success of this phase will lead to the inclusion of the integration of various departments throughout the company.

The study primarily focuses on the initial stage of the digitalization process, which commenced within the past 12 months. Additionally, tests of the implemented solutions have been conducted for a period of 3 months. The successful execution of this phase will serve as a catalyst for the company to proceed with the implementation of the remaining two phases. An analysis was conducted to assess the significance of the action taken and its financial and temporal impact on the

organization.

3.2.2 Phase Analysis

According to table 2, the project cost for this phase was set at 80,000 USD, and the planned duration for the project's cost depreciation is 5 years. The organizations opted to procure its own dedicated server to store all data and essential information, along with implementing the necessary infrastructure system. The company oversaw a total of 40 projects annually, and the estimation for conducting and analyzing the digitalization processes was derived from the assumption that each managed project required approximately 150 interactions.

Table 2. Indicators about the digitized actions

Estimation	Amount
Phase 1 Digitalization Project Cost	80 000 USD
Depreciation Duration	5 Years
Projects managed per year	40 Projects
Interactions per project	150

The analysis primarily examines the duration of interaction to facilitate and improve the process of decision making, as well as the transaction cost required for furthering the interaction. As a central server, it efficiently handles data sharing, email communication, notifications, and real-time updates for new documents and steps in the process. It seamlessly involves various stakeholders. The transaction cost will be determined by dividing the total project cost by the depreciation period, and then dividing that by the number of projects managed annually, and finally dividing it by the number of interactions per project. After performing all the necessary calculations, the transaction cost for the digitalization solution amounted to 2.66 USD.

3.2.3 Proposed Approach

In order to enhance the company's digitalization efforts across its three process phases, a proposal is being put forth to adopt a comprehensive and cutting-edge solution utilizing Blockchain technology, as illustrated in Figure 6. This solution comprises four levels, spanning from level 0, denoting the starting point with no progress, to level 4, indicating a fully developed solution with multiple phases and the capability to assess transaction cost and time.

To advance from Level 0 to Level 1, a Blockchain framework was utilized in a comprehensive event-logging solution. The utilized Blockchain protocol is based on Ethereum technology. To advance from Level 1 to Level 2, integrating Smart Contracts enables the establishment of mappings and the specification of events in alignment with different BPMN Diagrams.

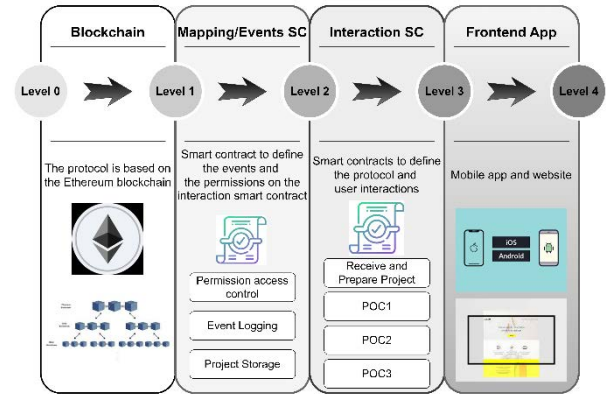


Figure 6: Second iteration to Prepare POC

Within this particular framework, Smart Contracts are responsible for overseeing various functions, including managing user access control, recording events, storing project documents and dates, and enforcing permissions for interactions between Smart Contracts. Consequently, advancement was made to the next level, precisely level 3. After the creation of all the smart contracts, the interaction between these contracts and users is governed by a specialized protocol that aligns with the different stages of the modeled process, starting from receiving and preparing the project up to POC 3.

Table 3. Cost Decomposition of the developed solution

Estimation	Amount
Ether Price	2047,07 USD
System total time response	9 Sec
Cumulative cost of transactions	145 USD
Deployment expenses	360 USD

The preceding work is commonly known as the backend of the developed solution. Hence, the task is to develop the user interface application, encompassing both the mobile application and website, to employ the solution effectively and attain Level 4. Upon examination of the implemented solution's transaction cost and response time, the transaction cost was quantified at 2.9 USD, accompanied by a total response time of 9 seconds. Consequently, this resulted in a cumulative cost of 145 USD. In terms of development, the associated cost stands at an unspecified value. Deployment expenses are explicitly stated at 360 USD. Furthermore, the operational cost is articulated through the formula: $\text{cost} = 0.00000002 * \text{eth_price} * \text{Number of gas}$. The cost includes all expenses related to transactions, development, and deployments, as outlined in Table 3.

3.3 Comparison and Future Directions

In our own analysis, we find Blockchain systems have higher infrastructure costs compared to traditional

databases due to their hybrid architecture requiring both a private Blockchain network and web interface.

However, traditional systems also bear difficult to quantify costs related to potential cyber-attacks, data breaches, and conflicts over data responsibility. When considering the enhanced security and transparency benefits of Blockchain alongside raw infrastructure expenses, we find the cost-benefit tradeoff depends on the application and associated security needs.

Blockchain provides data provenance through encrypted ledgers and consensus protocols, reducing risks of unauthorized alterations and disputes over accountability. While we cannot precisely estimate cost savings from prevented attacks, Blockchain offers clear security advantages. Additionally, the immutable and transparent nature of Blockchain transactions facilitates use cases like supply chain tracking requiring detailed histories. Determining the most suitable system still depends on the specific use case and security priorities. For applications where data accuracy and responsibility tracing are critical, the extra infrastructure costs of Blockchain may be warranted over traditional databases.

4 Conclusion

This paper has laid the foundation for integrating blockchain technology with AWP to enhance construction project management significantly. Through the proposed framework, the introduction of blockchain capabilities into managing and executing AWP workflows for construction firms and project managers is facilitated and optimized. Key features such as security, transparency, and coordination are systematically integrated into AWP's lifecycle, leading to notable improvements in traceability, accountability, and efficiency. The collaboration of blockchain and AWP has been shown to elevate project outcomes and set new standards for managing construction workflows.

Moving forward, the validation of this framework's effectiveness through comprehensive case studies and the development of specialized software systems tailored for the construction industry are identified as pivotal next steps. Furthermore, a concerted effort to quantify the anticipated economic and productivity benefits will substantiate the framework's value proposition, encouraging broader adoption across the sector.

The need for a deeper exploration and empirical validation of the proposed solutions has been recognized, reflecting the feedback received from industry practitioners. Our commitment to expanding research to include a broader array of case studies and deeper engagement with field experts is unwavering. This future research aims to validate the theoretical model and refine it based on practical insights and the complexities encountered in real-world applications of construction

project management.

By striving to bridge the gap between theoretical innovation and practical application, we aim to contribute a robust, empirically grounded framework to academic and professional communities. This endeavor seeks to catalyze a shift towards more efficient, transparent, and accountable practices within the construction industry. In doing so, we chart a course towards an innovative, data-driven future that fully leverages the potential of blockchain technology in harmony with AWP methodologies, thus making a substantial contribution to the field.

References

- [1] M. J. Ribeiro *et al.*, "The next normal in construction," 2020. [Online]. Available: [https://www.mckinsey.com/~media/McKinsey/Industries/Capital Projects and Infrastructure/Our Insights/The next normal in construction/The-next-normal-in-construction.pdf](https://www.mckinsey.com/~media/McKinsey/Industries/Capital%20Projects%20and%20Infrastructure/Our%20Insights/The%20next%20normal%20in%20construction/The-next-normal-in-construction.pdf) (accessed 10 February 2024).
- [2] X. Q. Zhang, "The trends, promises and challenges of urbanisation in the world," *Habitat Int.*, vol. 54, no. November 2015, pp. 241–252, 2016, doi: 10.1016/j.habitatint.2015.11.018.
- [3] B. Flyvbjerg, "What You Should Know About Megaprojects and Why: An Overview," *Proj. Manag. J.*, vol. 39, pp. 28–42, 2008, doi: 10.1002/pmj.
- [4] M. Sohail and S. Cavill, "Accountability to Prevent Corruption in Construction Projects," *J. Constr. Eng. Manag.*, vol. 134, no. 9, pp. 729–738, 2008, doi: 10.1061/(asce)0733-9364(2008)134:9(729).
- [5] B. Copeland, "Digital advanced work packaging – a case study in construction-led delivery," *APPEA J.*, vol. 63, no. 2, pp. S169–S172, May 2023, doi: 10.1071/AJ22046.
- [6] Y. A. Al Balushi, I. M. B. Ishaq, A. A. Al Naamani, and B. N. Seth, "Advance Work Packaging and Workface Planning ADAA," *Day 2 Tue, Novemb. 01, 2022*, 2022, doi: 10.2118/210850-MS.
- [7] R. Jabbar, M. Krichen, M. Shinoy, M. Kharbeche, N. Fetais, and K. Barkaoui, "A Model-Based and Resource-Aware Testing Framework for Parking System Payment using Blockchain," *2020 Int. Wirel. Commun. Mob. Comput. IWCMC 2020*, pp. 1252–1259, Jun. 2020, doi: 10.1109/IWCMC48107.2020.9148212.
- [8] V. Ciotta, G. Mariniello, D. Asprone, A. Botta,

- and G. Manfredi, "Integration of Blockchains and smart contracts into construction information flows: Proof-of-concept," *Autom. Constr.*, vol. 132, Dec. 2021, doi: 10.1016/J.AUTCON.2021.103925.
- [9] N. Adeeb, A. Abdu, and Z. Wang, "Blockchain based Security Solutions with IoT Application in Construction Industry," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 614, no. 1, p. 012052, Dec. 2020, doi: 10.1088/1755-1315/614/1/012052.
- [10] Ž. Turk and R. Klinc, "Potentials of Blockchain Technology for Construction Management," *Procedia Eng.*, vol. 196, pp. 638–645, Jan. 2017, doi: 10.1016/J.PROENG.2017.08.052.
- [11] Hamdi Olfa, F. Leite, and W. J. O'Brien, "Advanced Work Packaging from Project Definition through Site Execution: Driving Successful Implementation of WorkFace Planning," 2013.
- [12] Y. S. Halala and A. R. Fayek, "A Framework to Assess the Costs and Benefits of Advanced Work Packaging in Industrial Construction," 2018.
- [13] CII - RR272-11, "Construction Industry Institute ® Enhanced Work Packaging: Design through Workface Execution," 2013.
- [14] S. Ponticelli, W. J. O'Brien, and F. Leite, "Advanced work packaging as emerging planning approach to improve project performance: case studies from the industrial construction sector," 2015, doi: 10.14288/1.0076409.
- [15] C. Wu *et al.*, "Hybrid deep learning model for automating constraint modelling in advanced working packaging," *Autom. Constr.*, vol. 127, p. 103733, Jul. 2021, doi: 10.1016/J.AUTCON.2021.103733.
- [16] K. M. San, C. F. Choy, and W. P. Fung, "The Potentials and Impacts of Blockchain Technology in Construction Industry: A Literature Review," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 495, no. 1, p. 012005, Apr. 2019, doi: 10.1088/1757-899X/495/1/012005.
- [17] O. Akinradewo, C. Aigbavboa, A. Oke, and I. Mthimunye, "Applications of Blockchain Technology in the Construction Industry," *Lect. Notes Networks Syst.*, vol. 276, pp. 275–282, 2021, doi: 10.1007/978-3-030-80094-9_33.
- [18] S. Nanayakkara, S. Perera, D. Bandara, T. C. Scientific, and T. Weerasuriya, "Blockchain technology and its potential for the construction industry," *AUBEA Conf. 2019*, no. November, 2019, doi: 10.6084/M9.FIGSHARE.10315331.V3.
- [19] Z. Dakhli, Z. Lafhaj, and A. Mossman, "The Potential of Blockchain in Building Construction," *Build. 2019, Vol. 9, Page 77*, vol. 9, no. 4, p. 77, Apr. 2019, doi: 10.3390/BUILDINGS9040077.
- [20] K. Kim, G. Lee, and S. Kim, "A Study on the Application of Blockchain Technology in the Construction Industry," *KSCE J. Civ. Eng.*, vol. 24, no. 9, pp. 2561–2571, Sep. 2020, doi: 10.1007/S12205-020-0188-X/METRICS.
- [21] Z. Chen, "Enhancing the engineering supervision process in China: A solution enabled by integrating hybrid Blockchain system," *Innov. Green Dev.*, vol. 2, no. 4, p. 100091, Dec. 2023, doi: 10.1016/J.IGD.2023.100091.
- [22] R. Zhao, Z. Chen, and F. Xue, "A Blockchain 3.0 paradigm for digital twins in construction project management," *Autom. Constr.*, vol. 145, p. 104645, Jan. 2023, doi: 10.1016/J.AUTCON.2022.104645.
- [23] W. Lu, L. Wu, and R. Zhao, "Rebuilding trust in the construction industry: a Blockchain-based deployment framework," *Int. J. Constr. Manag.*, vol. 23, no. 8, pp. 1405–1416, 2023, doi: 10.1080/15623599.2021.1974683.
- [24] M. S. Kiu, F. C. Chia, and P. F. Wong, "Exploring the potentials of Blockchain application in construction industry: a systematic review," *Int. J. Constr. Manag.*, vol. 22, no. 15, pp. 2931–2940, Nov. 2022, doi: 10.1080/15623599.2020.1833436.
- [25] Thomas Allweyer, "BPMN 2.0 Modeling," 2016, [Online]. Available: <https://books.google.nl/books?hl=nl&lr=&id=sowaDAAAQBAJ&oi=fnd&pg=PA11&dq=bpmn+an+introduction+to+the+standard&ots=5w51Hp7hyL&sig=h7wxFXs-gS07ZI8RuTsr49HXrrw%0Ahttps://www.itp-commerce.com/bpmn-modeling/bpmn-2-0-modeling/>

Construction worker fatigue load management using IoT heart rate sensing system

Lun-Wang Wu¹, Hui-Ping Tserng¹ and Xiu-Zhen Huang¹

¹Department of Civil Engineering, National Taiwan University, Taipei, Taiwan.

d10521040@ntu.edu.tw, hptserng@ntu.edu.tw, r10521710@ntu.edu.tw

Abstract

Many studies consider excessive fatigue as one of the reasons for accidents among construction workers, especially in special hazardous work environments such as working at heights, heavy physical labor, and confined spaces. Unfortunately, due to factors such as complex environments, unstable equipment, and frequent movement of workers, there is little research on safety management of confined spaces and fatigue loading of construction workers in the construction industry. Therefore, this study developed an Internet of Things (IoT) heart rate sensing system, which has been verified in the real field and can be applied to the physical and mental health management of tunnel workers in the construction industry. In addition, a fatigue interval fitting model for construction workers was established by applying the percentage of heart rate reserve (%HRR). The maximum value of the anticline point %HRR in the model was defined as the fatigue alert value for construction workers, which facilitates project managers to monitor the abnormal conditions of workers' physiological load in a timely manner.

Keywords –

Worker Fatigue Load; Confined Space; Internet of Things Heart Rate Sensing System; %HRR

1 Introduction

The construction industry employs about 7% of the global employment, and 100,000 workers die on construction sites every year, which is about 35% of the global occupational fatalities [1]. Many studies have shown that occupational accidents in the construction industry are associated with overwork and fatigue among workers, as inattention may affect their awareness of environmental hazards or cause accidents during the operation of construction machinery [2-4]. In previous studies, heart rate (HR) was commonly used to measure

worker fatigue in the construction field [5]. Some studies have also used relative standards to measure the workload or training load of an individual, such as the percentage of heart rate reserve (%HRR), %VO₂max, etc. [6-8]. The relative standard focuses on the management of the output of the percentage of body energy relative to the load, which is more conducive to the precise management of an individual's physical and mental conditions.

With the development of wearable devices, more and more physiological parameters can be easily collected [9], including electromyography (EMG) [10,11]. Electroencephalography (EEG) [12,13] and wrist-worn photoplethysmography (PPG) devices [14-16]. In these techniques, EMG and EEG are weak bioelectrical signals that are susceptible to interference from a variety of noises. In addition, they are invasive in nature and lack convenience [17]. The wristband PPG heart rate sensor not only can monitor the heart rate of workers in real-time but also does not cause discomfort to the staff, so it has more potential [18].

Despite these advances, the collection of physiological parameters and personalized fatigue management is still a challenge due to complex environments, unstable equipment, and frequent worker movement, and there is a lack of a basis for judging the safety of fatigue load management for special hazardous operations such as confined spaces in the construction field. Therefore, this study develops an Internet of Things (IoT) heart rate sensing system for the limited space in the construction industry. In addition, based on the physiological values of construction workers, this study establishes the coordinate axes with the horizontal axis as the cumulative percentage of heart rate reserve (HRR) and the vertical axis as the cumulative percentage of working time [CPWT] (%) to establish a fatigue interval fitting model for construction workers. It defines the maximum value of the hyperbolic point in the model as the fatigue alert value of the construction workers, which is convenient for project managers to monitor the

abnormalities of the workers' physiological loads in a timely manner.

Unlike previous studies, the heart rate sensing system of the construction industry's Internet of Things (IoT) can be used for continuous sensing during the whole working day, which overcomes the complex environmental interference in the submerged shield tunnel and achieves stable data transmission. In addition, the sampling frequency, calculation time window, and measurement verification method of this study are also different. In the data analysis stage, instead of adopting the medical concept of Field resting heart rate (FRHR), this study calculated the FRHR at the construction site and then calculated the %HRR. Python was used to fit the relationship between cumulative heart rate reserve consumption percentage (% HRR) and working time percentage. The coefficient of determination R^2 (R-squared), root mean squared error (RMSE), and mean absolute error (MAE) of the fitted curve were used to evaluate the results of the regression curve equation, and then to find out the anticline of the cumulative curve of each day, and the largest anticline of %HRR was used as the warning value for this worker. The maximum anticline %HRR is used as the warning value for this worker.

2 Literature Review

2.1 Physiological indicators of workload in the construction industry

The causes of excessive workload and fatigue are very complex. Some studies have used relative physiological indicators to measure an individual's workload, focusing on the management of the output of the percentage of the individual's physical capacity close to the load, which is more conducive to the measurement of individualized fatigue in laborers and the need for safety management in the workplace, for example, emotional heart zone, the percentage of heart rate reserve, and %VO₂max [5-8]. Among them, %HRR is considered suitable for assessing the physical demand of labor tasks and applied to measure the workload of workers [19,20]. However, few studies have defined the specific safety interval of %HRR for construction workers. Chen & Tserng (2022) considered that the %HRR of workers and the percentage of cumulative working hours show an S-curve relationship, and its distribution location and anticurve point have the significance of the safety management of workers' load [21]. In this study, the concept is continued, and a fatigue interval simulation model is established to find out the specific value of the anticurve point of %HRR as the fatigue warning value of construction workers, which can

also get the safety interval of the workload of construction workers.

2.2 Development of IoT real-time heart rate sensing systems

The Internet of Things (IoT) heart rate sensing system has only begun to be practically applied to actual construction sites in recent years, including intelligent sensing, cloud-based IoT technology, and real-time heart rate monitoring and management using big data. The sensing data collection and processing has also expanded from hours and minutes to continuous sensing technology [20,22]. However, Anwer et al. (2021) argued that the challenge of using real-time physiological measures to assess physical fatigue in construction workers is the limitation of the validity of the physiological values used to determine fatigue, and there is a severe lack of information on fatigue due to data omission. [23].

Currently, there is still a lack of physiological sensing systems for the monitoring and management of tunnel workers, such as access control, localization in tunnels, and display of vital signs. Therefore, this paper evaluates the development of a physiological sensing system suitable for long-term monitoring functions and conducts system validation in the confined space of a submerged shield tunnel site.

3 Methodology

This research mainly consists of four main phases as depicted in Fig. 1: (1) Data Acquisition Overview (2) Design of IoT heart rate sensing system (3) Reserve heart rate percentage (%HRR) calculation (4) Personal Safety Interval Fitting Model.

3.1 Data Acquisition Overview

Tunnel construction is a high-risk construction operation with more stringent labor safety requirements. This study was conducted with the consent of the construction company under the condition of not interfering with the construction workers' work. The test subjects were all the construction workers in the construction area, working from 07:00 to 19:00, and their heart rate was monitored and recorded throughout the whole process—recruitment of test subjects: 23 people, all male. Due to the demand of project tasks and personnel mobility, not all 23 workers were on the job or recorded continuous heart rate data, so in this study, only P1, P2, P3, P4, P5, P6, P8, P9, P11, P13 will be used, a total of 10 workers' data, as shown in Table 1.

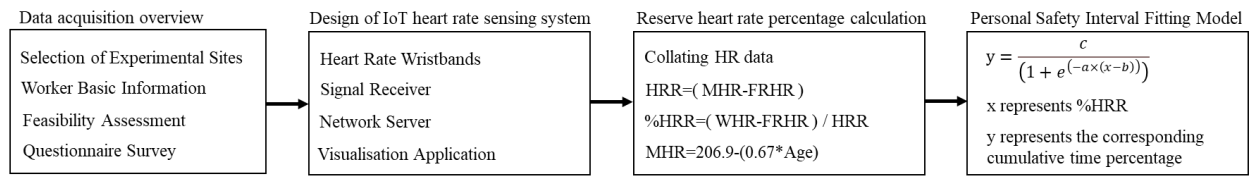


Figure 1. The methodological flow chart of this study

Table 1. Basic data table of construction site testers

Subject	Age	Height cm	Weight kg	Experience Years	Cardiovascular Disease	Smoking Years	Drinking Habit	Conscious Fatigue
P1	36	179	86	10	N	0	N	N
P2	28	164	85	5	N	14	Y	Y
P3	41	172	83	2	N	20	Y	N
P4	45	178	85	19	N	31	N	NS
P5	45	161	74	5	N	20	Y	NS
P6	26	172	73	1	N	12	N	Y
P8	35	168	90	2	N	25	Y	NS
P9	23	165	43	1	N	10	Y	Y
P11	42	168	60	10	Y	10	Y	N
P13	31	186	71	6	N	17	Y	NS

3.2 Design of IoT heart rate sensing system

3.2.1 Sensing system equipment and specifications

The environment in the confined space of the construction industry is complex, with poor quality network signals, darkness, humidity, loud noise, high mobility of workers, frequent job rotation, and high difficulty in mechanical operation. In the view of these characteristics, this study develops sensing devices and constructs an Internet system suitable IoT-PPG for this field in order to increase the feasibility and acceptance of physiological sensing device applications. The equipment and specifications of the heart rate sensing network system in this study are shown in Table 2.

Table 2. Heart Rate Sensing System Equipment and Specifications

Device Name	Functional Specifications
Heart Rate Wristbands	<ul style="list-style-type: none"> Detection frequency > 500 Signal Frequency >10 times/sec. Power Durability >12 hours Waterproof rating IP 67
Signal receiving and transmission equipment	<ul style="list-style-type: none"> BLE Router: Raspberry Pi 3B+ Low Power Bluetooth Transmission Simultaneously receiving 20 Wristbands 100Mbps Ethernet interface
Server Hosting Application Software	<ul style="list-style-type: none"> Ubuntu 16 Java Script HTML

3.2.2 Planning of sensing systems in actual field

The experimental field of this study is the construction site of the tunneling project, in order to make the sensing bracelet broadcasting signal can be received smoothly, the system adopts the Bluetooth BLE V4.2 transmission standard and configures the signal receiving equipment in the construction location where the personnel is permanently stationed, and aims at the system data coverage rate of more than 80%. According to the characteristics of the construction operation of the shield project, this study divides it into three sensing areas, A, B, and C. It allocates eight signal-receiving devices, as shown in Fig. 2, which are as follows: A. Ground working area, including four receiving sensors for a workshop, rest standby, machine maintenance, and material lifting; B. Working shaft area, with one receiving sensor at the bottom of the shaft; and C. Tunneling area, including three receiving sensors for the front, middle, and tail sections of the shield machine. Tunnel dull area, including three receiving sensors for the front, middle, and tail sections of the potential shield machine. As shown in Fig. 3, it is a worker wearing a heart rate wristband and a Bluetooth receiver transmitter in a shield tunnel.

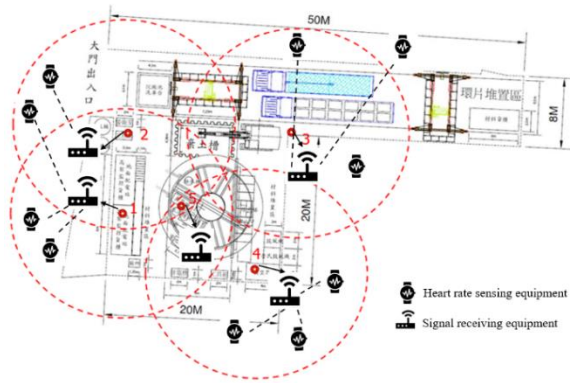


Figure 2. System equipment configuration plan of ground operation area A and working well area B

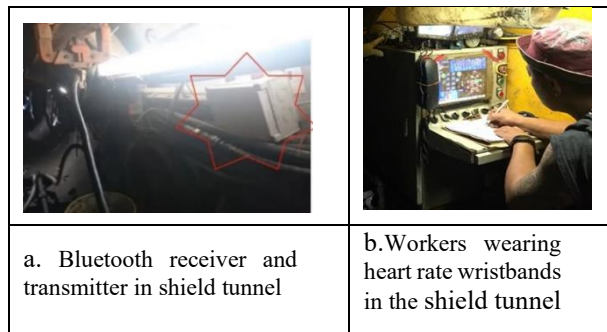


Figure 3. Subjects wearing wristwatches work in a small shield tunnel environment in the experimental field

3.3 Reserve heart rate percentage (%HRR) calculation

In this study, based on the theoretical development of cardiovascular loads (CVL) [24] and heart rate reserve (HRR) [7] [8] [25], the PPG heart rate was used as the basis for calculating the real-time %HRR index value, which reflects the degree of personal physical output of construction workers. The %HRR index value reflects the degree of individual energy output of construction workers and serves as an indicator for the real-time management of the workload of construction laborers. The significance of the %HRR index refers to the rate of depletion of personal heart rate reserve, and the higher rate of depletion indicates a higher workload, and the related calculation equations are shown in (1) and (2):

$$HRR = (MHR - FRHR) \quad (1)$$

$$\%HRR = (WHR - FRHR) / HRR * 100\% \quad (2)$$

MHR=206.9-(0.67*Age), Maximum heart rate value,bpm. FRHR= Rest heart rate in the construction

site area, bpm. WHR= Rest heart rate in the construction site area, bpm. The estimated maximum heart rate was calculated using an age-corrected estimation formula proposed by Jackson et al. (2007) [26]. The resting heart rate was measured using the field resting heart rate (FRHR) [21].

3.4 Personal Safety Interval Fitting Model

A sigmoid function, also known as the Logistic function fitting curve, is a function that maps the independent variable to a function between 0 and 1. It is often used to categorize a problem or to represent the growth trend, with formulas such as (3), [27] [28].

$$y = \frac{c}{(1+e^{(-a \times (x-b))})} \quad (3)$$

Where x represents %HRR, y represents the corresponding cumulative time percentage, a represents the control of the growth rate of function, b represents the center point of the function and the slope of the sigmoid function, and c represents the upper limit value of y when x approaches infinity.

This study used Python to fit the relationship between cumulative heart rate reserve consumption percentage (% HRR) and working time percentage. The same worker was statistically counted for more than seven days, and the data were organized into a 5-minute average %HRR time series. The coefficient of determination R^2 (R-squared), root mean squared error (RMSE), and mean absolute error (MAE) of the fitted curve were used to evaluate the results of the regression curve equation, and then to find out the anticline of the cumulative curve of each day, and the largest anticline of %HRR was used as the warning value for this worker. The maximum anticline %HRR is used as the warning value for this worker.

4 Results and Discussion

The results of the fitted regression curves for the safety interval of fatigue loads for construction workers in this study are discussed. Taking the fitting results of P1, P4, and P9 workers as an example, they have continuous heart rate data for more than seven days, and their job types are construction site managers, crane operators, and trolley drivers.

The main task of P1 is site management, and the %HRR data are simulated for nine days. The coefficient of determination R^2 is above 0.97, which is highly interpretable, and the root mean square error (RMSE) and the mean absolute error (MAE) are less than

5%, which makes the regression curves highly reliable. The maximum %HRR reversal point occurred on November 9, P1_1109, with a %HRR of 36.4%. The overlapping curve area is bounded by P1_1101, P1_1103, and P1_1109, as shown in Fig. 4, which can be regarded as the safety zone of fatigue load for P1's past work.

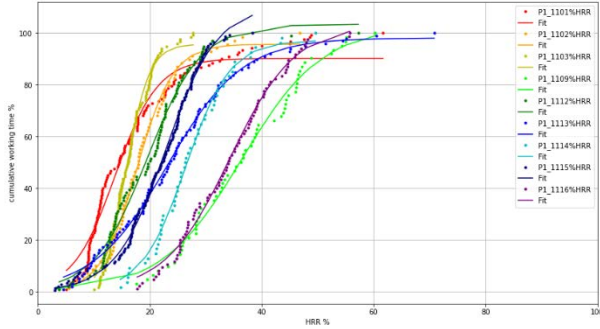


Figure 4. Fatigue Load Data Point Plots and Fitting Curves for Worker P1

The main work item day of P4 is the overhead crane operation, and 13 days of monitoring %HRR data are simulated. The coefficient of determination R^2 all falls above 0.98, with high interpretation power, the root mean square error RMSE and the mean absolute error MAE of %HRR obtained from the simulated curves are less than 4%, and the regression curves have a high degree of feasibility. The maximum %HRR anticlip occurs in P4_1102 on November 2, with a %HRR of 36.4%. The overlapping curve area of each day is bounded by P4_1109, P4_1112, P4_1102 and P4_1107 as shown in Fig.5, and this area can be regarded as the fatigue load safety zone of P4 in the past work.

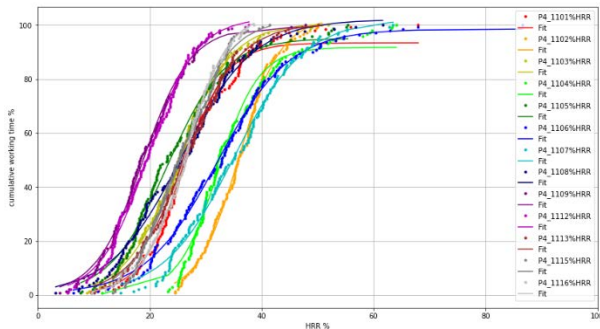


Figure 5. Fatigue Load Data Point Plots and Fitting Curves for Worker P4

The primary working day of P9 is cart driving, which belongs to the category of mechanically operated work with lower loads. Eleven days of %HRR data were simulated, and except for one day, P9_1106, which has a relatively small amount of data, with a coefficient of determination of R^2 of 0.94, an RMSE of 6.5%, and an MAE of 5.4%, and the regression result is a little bit

lower, most of the other days' R^2 falls above 0.97, which is of high interpretative power. The root mean square error (RMSE) and the mean absolute error (MAE) of %HRR obtained from the fitted curves are mostly less than 4%, which means that the regression curves have a high degree of reliability. The most significant %HRR inversion point occurs on November 12, P9_1112, with a %HRR of 40.4%. The overlapping curve area of each day is bounded by P9_1106 and P9_1112, as shown in Fig.6, and this area can be regarded as the fatigue load safety area of P9 in the past work.

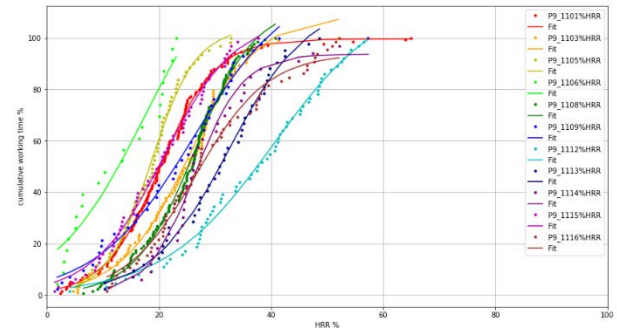


Figure 6. Fatigue Load Data Point Plots and Fitting Curves for Worker P9

5 Conclusion

This study develops a customized wearable IoT-PPG heart rate bracelet and wireless sensing system, which can be used for continuous sensing during the whole working day. It is field-proven to be applied to the physical and mental health and safety management of tunnel workers in a submerged tunnel, which is better than traditional tunnel personnel management. In addition, in this study, the accumulated physiological values of workers were used to synthesize the equation $y = \frac{c}{(1 + e^{(-a \times (x-b))})}$, which is composed of three unknowns, using Python. This equation has a high explanatory power and a low error, and it can be used to superimpose the cumulative curvilinear inversion point of %HRR over a few days as a warning value to establish a model of the safety zone for workers' fatigue load in the confined space of the construction industry. When the %HRR of a worker's future work exceeds the warning value, it is possible to pay appropriate attention to the worker's workload situation and adjust the working hours to reduce the occurrence of fatigue, which can be used as a reference for future related studies.

In the future, the data processing algorithm of this study can be extended and applied to other workplaces with harsh environments such as high temperature, high altitude, etc., to develop a logistic regression model for

continuous %HRR workload prediction so as to establish the prediction of workers' residual physical capacity and the management of precise working hours.

References

- [1] Chiang, Y. H., Wong, F. K. W., & Liang, S. (2018). Fatal construction accidents in Hong Kong. *Journal of Construction Engineering and Management*, 144(3), 04017121.
- [2] Kirk, P. M., & Sullman, M. J. (2001). Heart rate strain in cable hauler choker setters in New Zealand logging operations. *Applied Ergonomics*, 32(4), 389-398.
- [3] Ricci, J. A., Chee, E., Lorandeanu, A. L., & Berger, J. (2007). Fatigue in the US workforce: prevalence and implications for lost productive work time. *Journal of occupational and environmental medicine*, 1-10.
- [4] Cavuoto, L., & Megahed, F. (2018). Advancing safety surveillance using individualized sensor technology (ASSIST): Final progress report. American Society of Safety Professionals.
- [5] Lee, Wonil & Migliaccio, Giovanni C. (2014). Field use of physiological status monitoring (PSM) to identify construction workers' physiologically acceptable bounds and heart rate zones. *Computing in Civil and Building Engineering*, ASCE, pp.1037-1044.
- [6] Wu, H.C. & Wang, J. M. J. J. (2002). Relationship between maximum acceptable work time and physical workload. *Ergonomics*, 45 (4), pp.280-289.
- [7] Daniels, Jack (2013). *Daniels' Running Formula*, third ed. 48. United States of America: Human Kinetics. 2013. ISBN-13: 978-1450431835.
- [8] Hwang, Sungjoo & Lee, SangHyun (2017). Wristband-type wearable health devices to measure construction workers' physical demands. *Automation in Construction*, 83 (2017), pp.330-340.
- [9] Sun, F., Yi, C., Li, W., & Li, Y. (2017). A wearable H-shirt for exercise ECG monitoring and individual lactate threshold computing. *Computers in industry*, 92, 1-11.
- [10] Dong, H., Ugaldey, I., & El Saddik, A. (2014, May). Development of a fatigue-tracking system for monitoring human body movement. In 2014 IEEE international instrumentation and measurement technology conference (I2MTC) proceedings (pp. 786-791). IEEE.
- [11] Yousif, H. A., Rahim, N. A., Salleh, A. F. B., Zakaria, A., Alfarhan, K. A., & Mahmood, M. (2019, July). A study of lower limb muscles fatigue during running based on EMG signals. In 2019 IEEE International Conference on Sensors and Nanotechnology (pp. 1-4). IEEE.
- [12] Ke, J., Zhang, M., Luo, X., & Chen, J. (2021). Monitoring distraction of construction workers caused by noise using a wearable Electroencephalography (EEG) device. *Automation in Construction*, 125, 103598.
- [13] Xing, X., Zhong, B., Luo, H., Rose, T., Li, J., & Antwi-Afari, M. F. (2020). Effects of physical fatigue on the induction of mental fatigue of construction workers: A pilot study based on a neurophysiological approach. *Automation in Construction*, 120, 103381.
- [14] Jankovský, M., Merganič, J., Allman, M., Ferencík, M., & Messingerová, V. (2018). The cumulative effects of work-related factors increase the heart rate of cabin field machine operators. *International Journal of Industrial Ergonomics*, 65, 173-178.
- [15] Kim, J. H., Jo, B. W., Jo, J. H., & Kim, D. K. (2020). Development of an IoT-based construction worker physiological data monitoring platform at high temperatures. *Sensors*, 20(19), 5682.
- [16] Russell, A., Heneghan, C., & Venkatraman, S. (2019). Investigation of an estimate of daily resting heart rate using a consumer wearable device. *medrxiv*, 19008771.
- [17] Ni, Z., Sun, F., & Li, Y. (2022). Heart Rate Variability-Based Subjective Physical Fatigue Assessment. *Sensors*, 22(9), 3199.
- [18] Hwang, S., Seo, J., Jebelli, H., & Lee, S. (2016). Feasibility analysis of heart rate monitoring of construction workers using a photoplethysmography (PPG) sensor embedded in a wristband-type activity tracker. *Automation in construction*, 71, 372-381.
- [19] Maman, Zahra Sedighi; Yazdi, Mohammad Ali Alamdar; Cavuoto, Lora A. & Megahed, Fadel M. (2017). A data-driven approach to modeling physical fatigue in the workplace using wearable sensors. *Applied Ergonomics*, 65, pp.515-529.
- [20] Hashiguchi, Nobuki; Kodama, Kota; Lim, Yeongjoo; Che, Chang; Kuroishi, Shinichi; Miyazaki, Yasuhiro; Kobayashi, Taizo; Kitahara, Shigeo & Tateyama, Kazuyoshi (2020). Practical judgment of workload based on physical activity, work conditions, and worker's age in construction site. *Sensors*, 20, 3786.
- [21] Chen, Wei-Cheng & Tserng, H. Ping (2022). Real-time individual workload management at tunnel worksite using wearable heart rate measurement devices. *Automation in Construction*, 134(2022) 104051.
- [22] Coolbaugh, Crystal L.; Jr, Stephen C. Raymond & Hawkins, David A. (2015). Feasibility of a dynamic web guidance approach for personalized physical activity prescription based on daily information from wearable technology. *JMIR Res.Protoc.*, 2015

- Apr-Jun, 4(2), pp.e67.
- [23] Anwer, S., Li, H., Antwi-Afari, M. F., Umer, W., & Wong, A.Y. L. (2021), Evaluation of physiological metrics as real-time measurement of physical fatigue in construction workers: state-of-the-art review, *Journal of Construction Engineering and Management*, 147(5): 03121001-19. DOI: 10.1061/(ASCE)CO.1943-7862.0002038.
 - [24] Lunde, Lars-Kristian; Koch, Markus, Veiersted, Kaj Bo; Moen, Gunn-Helen; Wæ rsted, Morten & Knardahl, Stein (2016). Heavy physical work: cardiovascular load in male construction workers. *International Journal of Environment Research and Public Health*, 13(2016), pp.356.
 - [25] Karvonen, J. & Vuorimaa, T. (1988). Heart rate and exercise intensity during sports activities. *Sports Medicine*. 5 (1988) 303–312.
 - [26] Jackson, A. (2007). Estimating maximum heart rate from age: Is it a linear relationship? *Medicine & Science in Sports & Exercise*, 39(5), pp.821.
 - [27] Dreiseitl, S., & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: a methodology review. *Journal of biomedical informatics*, 35(5-6), 352-359.
 - [28] Menard, S. (2002). *Applied logistic regression analysis* (No. 106). Sage.

Evaluation of Effect of Different Safety Training Styles on Mental Workloads Using Electroencephalography

Peng Liu¹ and Chengyi Zhang²

^{1,2}Department of Civil and Architectural and Construction Management, University of Wyoming, USA

pliu3@uwyo.edu, chengyi.zhang@uwyo.edu

Abstract –

In the construction industry, diverse safety training styles, such as verbal instructions, pictorial representations, radio-based, video, Virtual Reality (VR), and Augmented Reality (AR) training, have been employed to improve workers' hazard recognition. However, the impact of these varied training styles on mental workload has been understudied. This research aims to investigate two primary aspects: (1) the influence of different safety training styles on participants' mental workloads, and (2) the variation in mental workload across different training phases. For this purpose, an experimental study was conducted with 20 participants (10 males and 10 females), who were randomly assigned to receive safety instructions either through reading or listening. The training was divided into four phases, each focusing on a specific hazard. Participants' brain activity was monitored using electroencephalography (EEG), and their mental workloads were assessed based on engagement and arousal indices. The results indicated no significant differences in engagement and arousal indexes between the reading and listening groups. However, variations in mental states were observed across different training phases, with increasing arousal levels noted as the training progressed. This study offers insights into the efficiency of different safety training styles and provides valuable recommendations for designing more effective safety training programs to enhance hazard recognition in the construction industry.

Keywords –

Safety training styles; EEG; Engagement index; Arousal index

1 Introduction

The construction industry is a vital component of the global economy, characterized by its dynamic and often high-risk work environments [1]. The construction

industry continues to report a significant number of workplace accidents and fatalities each year [2]. This situation underscores the critical need for effective safety measures and training programs tailored to address the unique hazards of construction sites.

Traditionally, construction safety training has relied on conventional methods such as the dissemination of printed manuals and verbal instruction [3]. However, the effectiveness of these methods has often been questioned, particularly in terms of the engagement. Recently, the industry have seen a shift towards more interactive and multimedia-based training approaches, including the use of auditory training tools [4]. However, there remains a lack of comprehensive data to assess the overall impact of these diverse training styles.

Effective safety training in the construction industry is not merely a regulatory requirement but also a fundamental aspect of ensuring workers' safety and preventing accidents [5]. The complexity of construction work and the diversity of risks present on construction sites, make comprehensive safety training indispensable. In addition, effective safety training programs are essential to enhance construction safety, which can contribute to reducing the incidence of workplace accidents.

While there is a consensus on the importance of safety training, there is a noticeable gap in research regarding the differences of workloads when people receive different safety training methods. Traditional evaluation methods have primarily focused on subjective assessments or post-training performance metrics. The application of neurophysiological tools, such as Electroencephalogram (EEG), remains largely unexplored. This gap hinders the development of evidence-based training programs that can enhance safety in the construction sector.

EEG presents a novel approach to understanding and evaluating training effectiveness by providing objective measures of cognitive engagement [6]. Although EEG has been successfully employed in cognitive research to assess learning processes, its application in evaluating

construction safety training is pioneering. By analyzing EEG data, this research aims to uncover the workloads of effective learning in safety training, offering a new method to training evaluation.

This study seeks to evaluate and compare the workloads among people undergoing different forms of safety training, specifically traditional manual reading versus auditory training methods. The objective is to determine the differences of workload when people receive different training styles. The research questions were formulated as following:

Research question (RQ 1): Do people show different level of workloads when receiving different safety training styles?

Research question (RQ 2): How do people's workloads change at different phases?

2 Literature Review

2.1 Construction safety training styles

The construction industry, characterized by its inherently hazardous environment, necessitates the implementation of robust safety training methods to mitigate workplace accidents and fatalities [1]. There has been a concerted effort to develop and refine effective safety training techniques to address this need. A variety of safety training programs have been integrated into the construction sector, each with its unique approach and methodology.

Traditional safety training within the construction domain typically relies on verbal instructions and pictorial representations [3]. This approach often involves job hazard identification conducted by professionals utilizing two-dimensional drawings, photographs, or static representations, supplemented by verbal explanations [7]. Such methods are foundational but may have limitations in conveying complex scenarios.

Radio-based training, while less prevalent, provides an auditory learning experience that can be particularly effective in reinforcing verbal instructions. Video training, conversely, has seen a surge in popularity due to its capacity to visually demonstrate procedures and scenarios. A notable example of this is the OSHA-10hr video training program, widely recognized in the construction industry [8].

Recent technological advancements in VR and AR have led to significant innovations in construction safety training. Xu and Zheng [9] have developed an immersive, interactive multiplayer-based training platform utilizing VR technology, aimed at enhancing workers' safety awareness. This platform represents a shift towards more engaging and realistic training environments.

Furthering this trend, Li et al. [10] introduced an optimization-centric methodology for creating

personalized training scenarios within a VR framework, specifically designed for construction safety education. Their research demonstrated that participants receiving personalized guidance in the VR setting showed more significant improvements in hazard identification skills.

Additionally, the study by Pereira et al. [11] explored the development of a virtual reality environment that augments real-world training resources, such as those provided by OSHA's Susan Harwood Grant. This approach underscores the potential of VR in replicating and enhancing traditional training methods.

In summary, the construction industry has witnessed the application of a diverse array of safety training methods, ranging from conventional techniques to cutting-edge VR and AR-based platforms. These methods collectively contribute to the overarching goal of enhancing construction safety, underscoring the industry's commitment to reducing accidents and fatalities through innovative and effective training solutions.

2.2 Application of EEG on construction issues

The advent of wearable sensing technologies has significantly broadened the scope for enhancing safety and health measures in the construction industry [12]. These technologies comprise a wide range of devices, such as Inertial Measurement Units (IMU), Electrocardiography (ECG), Photoplethysmography (PPG) sensors, Electrodermal Activity (EDA), Electromyography (EMG) sensors, eye-tracking technology, and Electroencephalography (EEG). Among these, EEG stands out for its unique benefits in addressing safety issues in construction: (1) Real-time cognitive state monitoring; (2) Non-intrusive nature; and (3) Training effectiveness assessment.

The significance of EEG in studying construction safety issues is increasingly recognized in research. Jebelli et al. [13] developed a method for automatically identifying stress levels in construction workers by analyzing EEG signals, paving the way for early stress detection. Wang et al. [14] introduced a novel approach combining kinematic data with EEG to derive vigilance measurement indices, enhancing understanding of workers' risk perception and safety management on construction sites. Additionally, Wang et al. [15] created a technique for monitoring mental fatigue in workers using EEG signals and machine learning, facilitating real-time and continuous fatigue assessment, thereby improving safety management practices.

This section underscores the state-of-the-art state in construction safety training methods and the application of EEG in addressing construction-related issues. However, there remains a gap in research regarding the impact of different safety training methods on workers' workload. Addressing this gap, the current study aims to

employ EEG technology to measure the workload experienced by individuals subjected to various training methods, thereby contributing to the optimization of safety training in the construction industry.

3 Methodology

This section presents the methodology of this study, including participant, experimental procedure, EEG signal pre-processing, and feature selection. Figure 1 presented the process of this study.

3.1 Participants

In this study, 21 people were invited from the university to participate in the experiment. Among them, one was excluded because of the missing data. Finally, the eligible sample size is 20, including 10 (female) and 10 (male). All participants range from 18 to 40 years, including 18-25(3), 26-30 (11), and 31-40 (6). In addition, all participants not only have normal or corrected-to-normal vision and hearing but also normal mental health.

While inviting 20 participants to an EEG study on the effects of safety training on mental workload is a good starting point, especially for pilot studies or exploratory research, this number may be on the lower end for ensuring robust statistical power, detecting small to moderate effect sizes, and achieving generalizable results.

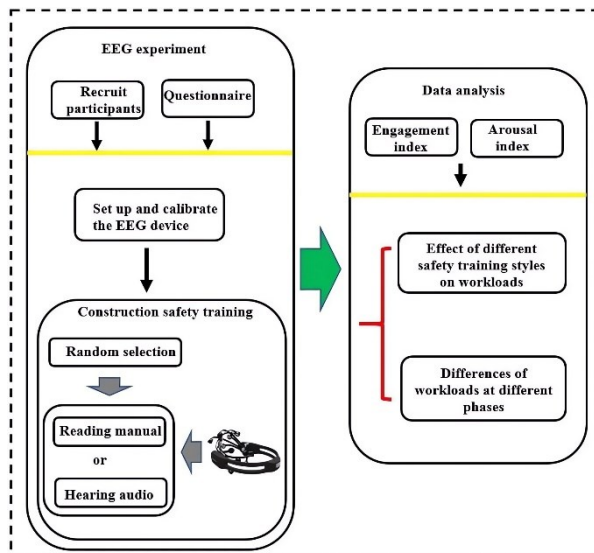


Figure 1. Procedure of this study

3.2 Experiment procedure and apparatus

Before the formal experiment, the participants were required to complete the questionnaire regarding their basic information, including gender, age, and their

mental health. Then, the staff then prepared the device and attached the electrodes on the participant's scalp. During the process of calibration, participants were required to relax. The formal experiment started when the quality of EEG signal reaches 100%.

During the formal experiment, the participant was randomly assigned a task, which required the participant to read a safety training manual or hear a safety training audio to enhance the safety awareness. Figure 2 shows an example of the test set up. Both of the two training styles have the same content and the only difference between them is the presentation style. The training concentrates on different hazards at construction sites to enhance people's safety awareness. This training was divided into four phases each phrase corresponds to one common hazard: phase 1- fall hazard, phase 2- tripping hazard, phase 3-electrical hazard, and phase 4- caught-in/between hazard. For example, phase 1 in the safety training manual describes the fall hazard and presents an example of this hazard through a figure. The selection of fall hazards, tripping hazards, electrical hazards, and caught-in/between hazards for inclusion in safety training programs is based on their ubiquity and potential severity within workplace environments, reflecting a strategic approach to mitigating prevalent risks that contribute significantly to workplace accidents and injuries. Time window for each phase depends on the length of the description. For example, phases 1 about the fall hazard lasts 30 seconds because it has 56 words and the audio play speed is 120 words/ minute. There is a five-second relaxation period between the two phases. The detailed procedure of the experiment was presented in Figure 3.

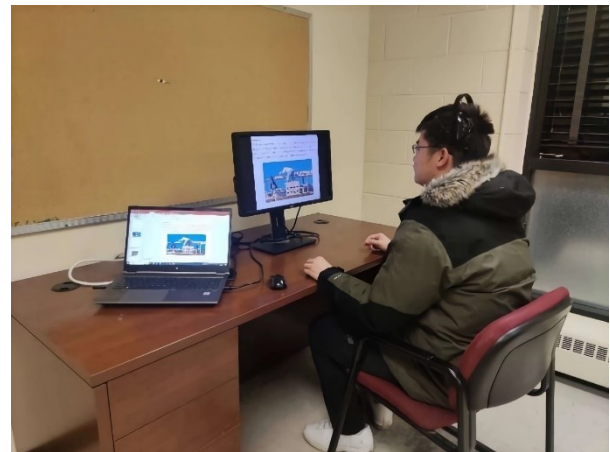


Figure 2. An example of the test setup

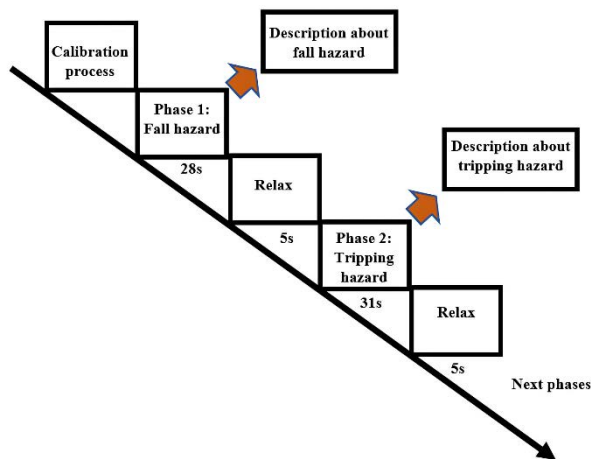


Figure 3. Experiment procedure

In this study, brain activity of the participants was meticulously monitored and recorded using the EMOTIV EPOC+ 14-Channel Headset, a wireless EEG signal amplifier. This device operates at a sampling frequency of 256 Hz, ensuring high-resolution data capture. The continuous EEG signals were acquired through 14 strategically placed electrodes (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4). Two additional electrodes function as referential anchors, specifically aligned with the P3 and P4 regions. These are designated as the Driven Right Leg (DRL) and Common Mode Sense (CMS), respectively, playing critical roles in enhancing signal fidelity and mitigating interference. Figure 4 shows the position of 14 electrodes of Emotive EEG device [16].

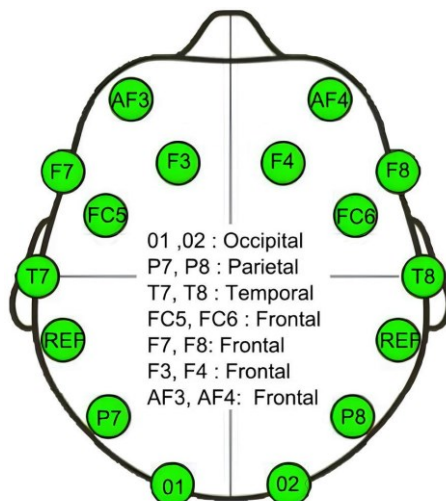


Figure 4. Location of the 14 electrodes of Emotive EEG device

The device is worn compliance with the International 10-20 system. This system is globally recognized for its standardized approach in electrode placement on the scalp, primarily aimed at facilitating accurate mapping of brain regions, with a specific emphasis on the cerebral cortex. The EEG signal was referenced using the P3 and P4 electrodes, a method that ensures reliable and stable signal quality by reducing noise and enhancing the accuracy of cerebral activity measurement.

3.3 EEG signal pre-processing

The raw EEG data often contains various types of noise and artifacts, making pre-processing a critical step to ensure the accuracy and reliability of EEG analysis [17]. When the EEG starts recording the brain activities, the user is required to record baseline, including relax with your eyes open and relax with your eyes closed. This step stems from the following reasons: (1) The EMOTIV EEG devices, like many other EEG systems, require calibration to the individual user's brain wave patterns. Opening and closing the eyes can help the device in calibrating and recognizing the user's unique brainwave patterns. This process ensures that the device accurately captures and interprets the EEG data; (2) Artifacts Detection: Eye movements, including blinks, can cause artifacts in EEG recordings. By asking the user to open and close their eyes, the system can identify and potentially filter out these artifacts from the actual brainwave data. This helps in improving the accuracy of the EEG readings; (3) Baseline Establishment: The difference in brain activity when the eyes are open versus when they are closed (such as changes in alpha waves) can be significant. This contrast helps in setting a baseline to better understand and interpret subsequent brain activity during the EEG recording.

The initial processing of EEG data involves the application of a bandpass filter, specifically within the frequency range of 1–30 Hz. This filtering stage is a critical component in EEG data pre-processing, as the raw EEG signals typically encompass a variety of fluctuating time-dependent curves, incorporating multiple channels and frequency bands that are not pertinent to the intended analysis. The primary function of filtering is to eliminate extraneous noise and to isolate the signal of interest from a specified frequency band for subsequent examination. A bandpass filter is particularly instrumental in this context, as it allows the passage of signals within a designated frequency range—here, 1–30 Hz—while attenuating signals that fall outside this spectrum. Employing a bandpass filter in the 1–30 Hz range is a common practice in the initial stages of EEG signal processing, given that the data of relevance are predominantly contained within these frequencies.

3.4 Feature selection

Cognitive or mental workload, integral to understanding human performance and learning, reflects the mental effort required to perform tasks, influenced by task complexity, individual capabilities, and environmental factors [18]. It is assessed through subjective self-reports, behavioral metrics, and physiological indicators, offering insights into the demands placed on cognitive resources [19]. Training outcomes significantly affect cognitive workload, with effective training reducing workload over time through skill acquisition and efficiency gains [20]. Adaptive training, tailored to individual differences and optimized through strategic feedback, can enhance learning efficiency and effectiveness, demonstrating the critical interplay between cognitive workload management and optimized training designs for improved task performance and learning.

Engagement index and arousal index are often considered key predictors of successful learning outcomes [21]. Engaged learners are typically more attentive, motivated, and invested in the learning process, which can lead to better understanding, retention, and application of the learned material. Higher engagement indices are usually associated with more focused attention and active cognitive processing, which are crucial for effective learning [22, 23]. High level of arousal index represents that the phase is more stimulating but low arousal index indicates that process is arousal [24]. Therefore, this study adopted engagement index to measure the participants' brain activities. It is calculated as follows:

$$\text{Engagement index} = \beta / (\alpha + \theta) \quad (1)$$

$$\text{Arousal index} = (\beta(F4) + \beta(F3)) / ((\alpha(F4) + \alpha(F3))) \quad (2)$$

Where, θ , α , β , and represents the EEG frequency bands, namely 4-8 Hz, 8-12 Hz, and 12-30 Hz.

The process for extracting the three frequency bands from the EEG signal involves several steps. Initially, a one-second segment of the EEG signal is processed using a Fast Fourier Transform. This method facilitates the separation of the signal into distinct frequency bands: theta (θ), alpha (α), and beta (β). Subsequently, a composite value representing these three bands is calculated. This is achieved by aggregating their respective values across the measured regions of the brain.

In the final step, the EEG engagement index and arousal index at a given instant time is determined. This is accomplished by calculating the average of each ratio within a sliding window preceding the instant T. This approach provides a dynamic measure of workloads, reflecting changes over the observed period.

4 Result and Analysis

This section presents result about the effect of different training styles on people, which responds to two RQs.

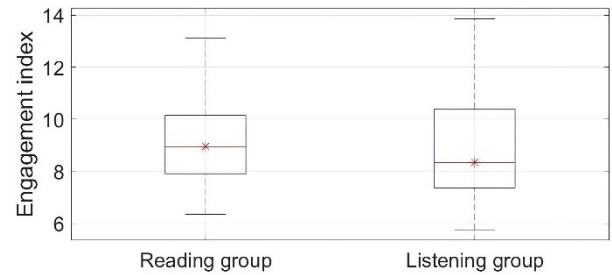
4.1 Learning outcome

Table 1 shows the EEG descriptive for two metrics, including engagement index and arousal index.

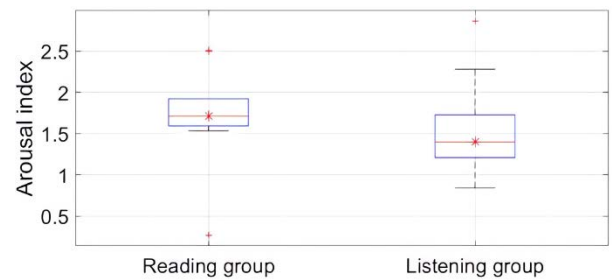
Figure 5 demonstrates the engagement index and arousal index of each group on different safety training styles. A Mann-Whitney U test was conducted to test the difference of the index between two groups. As the p-value = 0.33 > 0.05, engagement index in reading group ($mdn=8.95$) and listening group ($mdn=8.34$) do not show significant differences between the reading and listening groups. Arousal index performs the similar result.

Table 1 EEG descriptive for each index

Indices	Training styles	Mean	SD	SEM
Engagement index	Reading	9.10	1.88	0.56
	Listening	8.95	2.54	0.76
Arousal Index	Reading	1.72	0.62	0.19
	Listening	1.57	0.59	0.18



(a) Engagement index between two groups

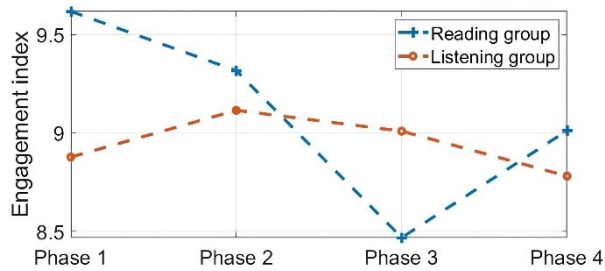


(b) Arousal index between two groups

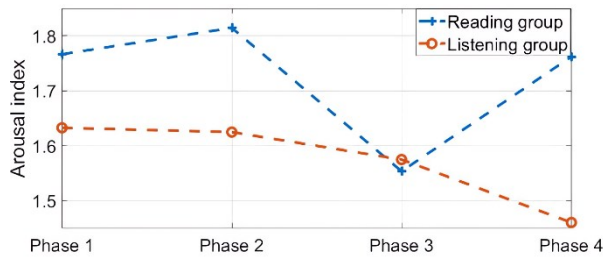
Figure 5. Engagement index and arousal index between reading and listening groups

4.2 Workloads at different phases

Figure 6 shows the workloads at different phases. According to the figure, people have lowest engagement index at phase 3 when reading the manual and at phase 4 when listening the safety manual. For the arousal index, people have lowest arousal value at phase 3 when reading. When listening to the manual, arousal index goes down as the learning process proceeds.



(a) Engagement index at different phases



(b) Arousal index at different phases

Figure 6. Workloads at different phases

Engagement index and arousal index can measure the immersion level while learning. High level of engagement is an indicator to reflect the level of stressful state. In addition, high level of arousal index represents that the phase is more stimulating but low arousal index indicates that process is arousal. According to the result, this study found that 1) people have the most relaxed mental state at phase 3 while reading but at phase 4 while listening; 2) In general, as the learning process proceeds, individuals experience a progressive enhancement in their mental tranquillity; 3) the hazard at phase 3 is more stimulating while reading but at phase 4 when listening; 4) As the listening learning process continues, the hazard becomes more and more stimulating, which means more and more arousal.

The study on the differential impact of reading and listening on engagement and arousal indices during construction safety training reveals nuanced insights into learner responses. It indicates that the mode of content

delivery—whether visual or auditory—significantly influences learners' mental states and their perceived stimulation from the material. Specifically, learners exhibit a more relaxed state during phase 3 while reading and during phase 4 while listening, suggesting that content complexity and delivery mode interact uniquely with individual psychological responses. Furthermore, the progressive enhancement in mental tranquility across the learning process highlights the importance of acclimatization to the training material, while the increased arousal in listening scenarios underscores the cumulative engagement effect of auditory delivery. These findings point to the critical role of individual learning preferences, content delivery methods, and the inherent characteristics of the training material in designing effective safety training programs, emphasizing the need for tailored approaches that consider the diverse responses of learners to optimize training outcomes.

The study underscores the multifaceted nature of learning engagement and arousal, influenced by factors that include individual learning preferences, prior knowledge and experience, and environmental conditions. Individual learning preferences play a pivotal role, as they dictate how learners process and engage with information, whether through auditory or visual means. Participants' prior knowledge and experience with the subject matter can also significantly affect their response to the training, potentially reducing stress and increasing comprehension. Lastly, environmental factors, such as the presence of distractions and the physical or digital nature of the learning environment, contribute to the overall effectiveness of safety training by influencing learners' ability to focus and engage with the content. In conclusion, these factors highlight the necessity of adopting a nuanced and learner-centered approach in the design and implementation of safety training programs to enhance efficacy and learner outcomes.

5 Conclusion

This research aimed to evaluate the cognitive workloads associated with various safety training methodologies, utilizing electroencephalography (EEG) technology. To achieve this, an experimental framework was established to monitor and record cerebral activity in participants during the training sessions. The primary metrics for assessing cognitive workload were the engagement index and the arousal index. The findings of this study revealed two key insights: firstly, there was no statistically significant variation in the engagement index and arousal index between participants engaged in reading-based and listening-based training methods; secondly, the study observed differential cognitive workloads among phases, particularly noting that as the

learning process proceeds, there was a corresponding rise in mental tranquillity, suggesting a more profound cognitive processing. These results have significant implications for the design of more effective safety training programs, potentially enhancing workers' ability to recognize and respond to hazards in their work environment.

This study illuminates the impact of different safety training styles on learner engagement and arousal within the construction industry, highlighting the potential for tailored training approaches. By demonstrating that reading and listening variably affect mental states and stimulation levels across different phases of safety training, the research suggests a strategic blending of content delivery methods to optimize learning outcomes. This approach not only accommodates diverse learning preferences but also utilizes the instructional strengths of each training style to enhance learner engagement, reduce stress, and improve safety protocol. Such insights advocate for a more informed and adaptive design of safety training programs in the construction industry, aiming to foster safer work environments.

This study, while yielding insightful findings, is subject to certain limitations that must be acknowledged. Firstly, the participant pool was a little small, comprising only 20 individuals. This limited sample size may not adequately represent the broader population, potentially affecting the generalizability of the results. Additionally, the participants were not frontline construction workers, a factor that could influence the applicability of the findings to the intended target group. Moreover, the study's reliance on just two metrics – the engagement index and the arousal index – to assess cognitive workloads might not capture the full spectrum of cognitive responses to safety training.

For future research, several avenues are suggested to enhance the understanding of cognitive workloads during safety training. These include: (1) investigating the detailed trends in cognitive workload throughout the duration of safety training sessions; (2) employing a more comprehensive set of indices to measure cognitive workloads, thereby providing a more nuanced understanding of the cognitive processes involved; (3) examining the effectiveness of various safety training modalities, such as video, VR and AR in order to determine the most effective methods for imparting safety knowledge and skills. These future studies could significantly contribute to the optimization of safety training strategies, particularly in high-risk industries like construction.

References

- [1] H. Lingard, "Occupational health and safety in the construction industry," *Construction management and economics*, vol. 31, no. 6, pp. 505-514, 2013.
- [2] C. Nnaji and A. A. Karakhan, "Technologies for safety and health management in construction: Current use, implementation benefits and limitations, and adoption barriers," *Journal of Building Engineering*, vol. 29, p. 101212, 2020.
- [3] M. Afzal, M. T. Shafiq, and H. Al Jassmi, "Improving construction safety with virtual-design construction technologies-a review," *Journal of Information Technology in Construction*, vol. 26, 2021.
- [4] S. Thallapureddy, S. Bhandari, and A. Albert, "Are Multimedia Training Interventions Effective within Traditional Safety Training Frameworks?," in *Construction Research Congress 2022*, pp. 264-273.
- [5] L. G. Mollo, F. Emuze, and J. Smallwood, "Improving occupational health and safety (OHS) in construction using Training-Within-Industry method," *Journal of Financial Management of Property and Construction*, vol. 24, no. 3, pp. 655-671, 2019.
- [6] Y. Zhang, M. Zhang, and Q. Fang, "Scoping review of EEG studies in construction safety," *International journal of environmental research and public health*, vol. 16, no. 21, p. 4146, 2019.
- [7] A. D. u. Rafindadi, N. Shafiq, I. Othman, and M. Mikić, "Mechanism models of the conventional and advanced methods of construction safety training, is the traditional method of safety training sufficient?," *International journal of environmental research and public health*, vol. 20, no. 2, p. 1466, 2023.
- [8] A. J. Caban-Martinez *et al.*, "Physical exposures, work tasks, and OSHA-10 training among temporary and payroll construction workers," *Journal of occupational and environmental medicine*, vol. 60, no. 4, pp. e159-e165, 2018.
- [9] Z. Xu and N. Zheng, "Incorporating virtual reality technology in safety training solution for construction site of urban cities," *Sustainability*, vol. 13, no. 1, p. 243, 2020.
- [10] W. Li, H. Huang, T. Solomon, B. Esmaeili, and L.-F. Yu, "Synthesizing personalized construction safety training scenarios for VR training," *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 5, pp. 1993-2002, 2022.
- [11] R. E. Pereira, M. Gheisari, and B. Esmaeili, "Using panoramic augmented reality to develop a virtual safety training environment," in *Construction Research Congress 2018*, 2018, pp. 29-39.
- [12] C. R. Ahn, S. Lee, C. Sun, H. Jebelli, K. Yang, and B. Choi, "Wearable sensing technology

- applications in construction safety and health," *Journal of Construction Engineering and Management*, vol. 145, no. 11, p. 03119007, 2019.
- [13] H. Jebelli, S. Hwang, and S. Lee, "EEG-based workers' stress recognition at construction sites," *Automation in Construction*, vol. 93, pp. 315-324, 2018.
- [14] D. Wang, H. Li, and J. Chen, "Detecting and measuring construction workers' vigilance through hybrid kinematic-EEG signals," *Automation in Construction*, vol. 100, pp. 11-23, 2019.
- [15] Y. Wang, Y. Huang, B. Gu, S. Cao, and D. Fang, "Identifying mental fatigue of construction workers using EEG and deep learning," *Automation in Construction*, vol. 151, p. 104887, 2023.
- [16] X. Hou, Y. Liu, O. Sourina, and W. Mueller-Wittig, "CogniMeter: EEG-based emotion, mental workload and stress visual monitoring," in *2015 International Conference on Cyberworlds (CW)*, 2015: IEEE, pp. 153-160.
- [17] M. Wang, Y. Zhao, and P.-C. Liao, "EEG-based work experience prediction using hazard recognition," *Automation in Construction*, vol. 136, p. 104151, 2022.
- [18] F. Dehais, A. Lafont, R. Roy, and S. Fairclough, "A neuroergonomics approach to mental workload, engagement and human performance," *Frontiers in neuroscience*, vol. 14, p. 519228, 2020.
- [19] P. Vanneste *et al.*, "Towards measuring cognitive load through multimodal physiological data," *Cognition, Technology & Work*, vol. 23, pp. 567-585, 2021.
- [20] D. Stefanidis, M. W. Scerbo, P. N. Montero, C. E. Acker, and W. D. Smith, "Simulator training to automaticity leads to improved skill transfer compared with traditional proficiency-based training: a randomized controlled trial," *Annals of surgery*, vol. 255, no. 1, pp. 30-37, 2012.
- [21] M. Alimardani, S. van den Braak, A.-L. Jouen, R. Matsunaka, and K. Hiraki, "Assessment of engagement and learning during child-robot interaction using EEG signals," in *Social Robotics: 13th International Conference, ICSR 2021, Singapore, Singapore, November 10-13, 2021, Proceedings 13*, 2021: Springer, pp. 671-682.
- [22] C. Berka *et al.*, "EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks," *Aviation, space, and environmental medicine*, vol. 78, no. 5, pp. B231-B244, 2007.
- [23] A. Kamzanova, G. Matthews, A. Kustubayeva, and S. Jakupov, "EEG indices to time-on-task effects and to a workload manipulation (cueing)," *International Journal of Psychological and Behavioral Sciences*, vol. 5, no. 8, pp. 928-931, 2011.
- [24] S. Saedi, A. A. F. Fini, M. Khanzadi, J. Wong, M. Sheikhhoshkar, and M. Banaei, "Applications of electroencephalography in construction," *Automation in Construction*, vol. 133, p. 103985, 2022.

SafeSense: Real-time Safety Alerts for Construction workers

Sivakumar S K¹, Vishnu Sivan¹ and Grigary C Antony¹

¹Tata Consultancy Services, India

sivakumar.sk@tcs.com, vishnu.sivan@tcs.com, grigary.a@tcs.com

Abstract

Our research presents a cutting-edge solution for enhancing safety in construction sites: the Smart Helmet-based Wearable Personnel Proximity Warning System, specifically designed to prevent collisions & zone-based alerts in hazardous areas. Through the use of Bluetooth beacons strategically placed on heavy machinery and hazardous areas, the system communicates with a smart helmet worn by workers. Featuring an ESP32 module and Radio Frequency Identification (RFID) technology, the helmet provides real-time alerts with minimal effort from the user. When unsafe zone is notified by the safety manager in the safety command center, smart helmet receives this signal and turns the Light Emitting Diode (LED) indicators flash red and are accompanied by both auditory alerts and Heads Up Display (HUD) based directions. The inclusion of Bluetooth Low Energy (BLE) technology also increases the accuracy of proximity detection, ensuring reliability even in dynamic work settings. To ensure ease of use, the system incorporates a lithium polymer battery for portability. Laboratory based controlled environment is chosen for study of the Proof of Concept (PoC) to validate the functionality with 2 mock-up workers. Dynamic alerts, Communication to workers in real time, Low cost, High utility & Easy ergonomics are the five parameters experimented. The proposed model has the potential to be widely adopted in the construction industry and reduce the accidents using dynamic alerts to the workers in the real time. Present study also highlights the associated limitations of the smart helmet where future study can happen to overcome these limitations.

Keywords –

Smart Helmet, Personnel Proximity Warning System, Safety sign indications, Collision Prevention, Bluetooth Beacons, HUD

1 Introduction

The construction work force constitutes only 29 % to total industry workforce, but the contribution of accidents is 40% to the total industry. This clearly indicates accident exposure is so high in construction industry and it tops the list of all industries [1]. This is due to the dynamic nature of construction activities and exposure to open areas. Big equipment like cranes, dozers, moving vehicles & the need to work on higher elevations make construction site highly accident prone. Fall from height, struck by, caught between, electrical shock are the fatal four categories of construction accidents. In such a hazardous work environment, safety helmets, part of Personal Protective Equipment (PPE), are crucial in the construction industry. Other popular safety precautions are use of sign boards to display specific field conditions like Safe zone/ unsafe zone / Use ear protectors etc., Designated work areas are marked using signboards, to indicate area which are not safe to work. This approach of making an area based on geographical co-ordinates is known as Geo-fencing. This geo-fencing will not remain the same, it will change with time and location in the construction industry. For example, activities like electrical cabling or plumbing or excavation demands different kinds of geo fencing and depend on the construction stages.

Though PPE monitoring and sign boards work effectively in other industries, they are not so effective in the construction industry because of dynamic changes happening in the field. One approach to address both the issues is to embed the warning messages into the safety helmet itself [2][3][4].

Another critical difference of construction industry is difficulty involved in making a uniform risk level. Consider the case where multiple activities masonry and carpentry works on going, they are exposed to different risk levels, though they operate in same area. For example, mason and electrician may be having different risk exposure. While making a generalized risk exposure and associated safety rules is possible in other industries, this is not possible in the construction industry. In short, workers are dynamic, conditions are dynamic,

geographical co-ordinates are dynamic. This forces big challenge to build any safety alert system for construction industry.

Considering the advancements in mobile communication, use of mobile phones could be an approach, however construction work nature and ergonomics requirements of this industry prohibits use of mobile phones which is widely accepted in other industries.

Lastly any proposed safety approach has to be cost effective and able to use current infrastructure than green field approach of making new PPE

All the above-mentioned safety challenges posed the following research question:

“Is it possible to build safety helmet which can take care of highly dynamic and unsafe conditions of construction field maintaining the cost, utility and ergonomics demands using advancements in the electronics and computer science filed to reduce the accident counts in the construction?”

2 Literature Survey

To address the research question, literature review is carried out to critically examine the studies, comparing methodologies, outcomes, and implications. A comprehensive understanding of the evolving landscape of safety helmets that use Personal Proximity Warning Systems (PWS) in other industrial settings is studied.

One study[2] proposed use of smart glasses and BLE signals from beacons in a mining site. The study demonstrates that the smart glasses based PWS effectively provided visual proximity alerts with accompanying images of approaching equipment. The average BLE signal recognition distances were reported for different scenarios, and the workload analysis using NASA-TLX criteria indicated lower mental efforts and improved work efficiency when using smart glasses compared to a smartphone-based PWS.

Another study [3] explores indoor localization solutions for safety applications, with a focus on meeting system requirements and adhering to existing standards and regulations. This study discusses various technologies for worker localization and emphasizes the potential of RFID, particularly passive UHF-RFID technology, in ensuring worker safety. The study considers the properties required for an indoor localization system, considering safety directives and machinery assemblies.

Another study[4] used personal PWS with visual and tactile proximity warnings using an AI smart helmet equipped with a camera module. The system analyzes image data on a cloud server to evaluate recognition distance and hazard warning accuracy. The study

demonstrates that the AI smart helmet achieved recognition distances varying with the input minimum diagonal length and maintained high hazard warning accuracy at different face angles between the camera module and mining equipment. The proposed smart helmet based PWS focuses on preventing collision accidents with moving equipment and suggests future extensions, including the integration of additional sensors for worker health monitoring and environment sensing.

AI Based safety helmet [7], Long Range Radio (LoRA) based communication[8], beacon based alarm[9][10], Situation aware PPE [11], smart PPE [12], proximity warning [13] are the other few state of the art possibilities in this area. In below table 1, present study compares the features available towards construction safety and established new solution to address all the gaps.

Present study extends this exploration by proposing a new safety helmet module integrating an ESP32 microcontroller, BLE receiver, LEDs and buzzers offering unique perspective on hazard detection and communication.

The study investigated three proposed solutions [2], [3], [4] for providing visual alerts in various environments, including construction sites, mining sites, and machinery assemblies and summarized the features and gaps in the Table 1.

Table 1: Comparison of Technologies

Communication techniques	BLE	RFID	WiFi	Present Study
Dynamic alert	✓	X	✓	Yes
Realtime notification	✓	X	✓	Yes
Ergonomic comfort	✓	✓	X	Yes
Less Work obstruction	X	✓	X	Yes
Low Cost	X	X	✓	Yes
Profile based notification	X	X	X	Yes
Large coverage area	X	X	✓	Yes
Brown Field approach	✓	✓	✓	Yes

3 Methodology & Experimentation setup

The proposed system utilizes BLE technology for data transmission and communication between a network of helmets worn by workers and central control unit. Each helmet is equipped with a receiver module comprising an ESP32 microcontroller for data collection and analysis, as well as a wave-shaped acrylic sheet housing LEDs for visual alert. The central control unit serves as safety command center. This will be operated by the safety manager, who manages the network of helmets and transmits hazard information through color changes and blink patterns of the LEDs. Front end application- a web interface - enables the safety manager to monitor and update the status of different areas in real-time, facilitating him to send prompt alerts to the workers in the detected hazard zones.

3.1 System Architecture

The proposed smart helmet system is depicted in Figure 1. It has two components: One is hardware helmet worn by all workers and other one is centralized control unit. Helmet is ergonomically designed so that worker will still feel comfortable. Alerts are sent via wave shaped LEDs using color codes so that his work is not impacted, and alerts are seen clearly. In addition to the visual alerts, tiny buzzers are added in the hardware which will serve as speakers to alert the workers. The components of the hardware are a) Blue tooth module b) TP4056 Battery management system c) Wave shaped acrylic LED Holder d) LED e) ESP Module f) Battery

The safety command center will be operated by the safety manager, who manages the network of helmets and communicates hazard information to the helmets via BLE signals.

Each helmet will have designated unique identifier - Media Access Control Identifier (MAC ID). The MAC ID is registered on the ESP32 module when it is created in the factory, it is permanent for each helmet. Authorization is carried out based on MAC ID. Only registered MAC IDs will be authorized to get the alerts from the safety command center. This MAC ID is attached with the worker's employee identification number. There by analytics data about the worker can be collected using the MAC ID of the helmet. This MAC ID also plays crucial role in sending the profile based alerts. Present study used Electronic Design Automation (EDA) tool - a software application to create design, simulation, and testing of electronic circuits and systems. Detailed circuit designing is carried out using EDA to show schematic capture, layout, routing, and 3D visualization.

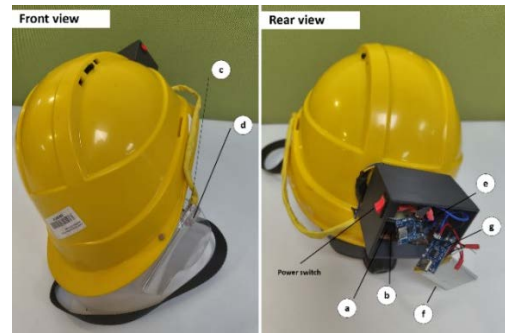


Figure 1: Safety helmet front view and rear view

3.2 Data Transmission and Communication

The BLE technology is utilized for establishing a communication network between the helmets and the central control unit. This enables efficient and low-power data transmission, ensuring minimal impact on the battery life of the helmets. The central control unit is responsible for transmitting hazard information to the helmets, which is then displayed to the worker wearing the helmet through color changes and blink patterns of the LEDs.

3.3 Real-time Hazard Monitoring

Real-time hazard monitoring is addressed using a web interface which safety command centre (Figure 2.1 & 2.2). The interface allows the safety manager to map different areas within the environment with BLE transmitter information. When a hazard is detected in a specific area, workers will notify the safety manager, then he can update the status of that area through the web interface easily as the system allows use of easy drawing and marking tools. This information is then transmitted to the corresponding helmets, prompting immediate visual alert to the workers.

An experimental setup was constructed within the laboratory environment to assess the performance of the PoC. The software aspect of the system was intricately designed to simulate three distinct zones: Room_1, Room_2, Room_3. Safety managers were granted the functionality to dynamically modify the status of each room in real-time. This feature allowed for the designation of specific areas as "unsafe zone" and all unmarked area is considered as "safe zone"



Figure 2.1: Safety command center software safe zone marking

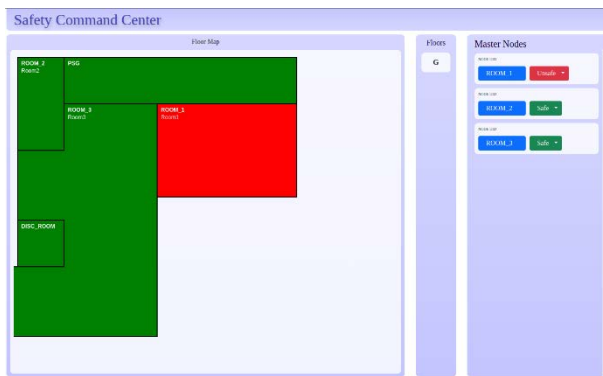


Figure 2.2: Safety command center software unsafe zone marking

When workers enter into such designated zones, workers would observe an immediate change in the LED color within their wave shaped shield of their helmets, transitioning to red to signify the hazardous nature of the environment.

In Figure 2.1, All the three rooms, Room_1, Room_2 and Room_3 are marked as safe and indicated with green color in the UI. In Figure 2.2, the Room_1 marked as unsafe using the controls. That specific room (Room_1) is changed to red color in the UI. Whenever the worker enters to that room the color of the LED will change to red for indicating the unsafe area. Figure 3(Left part) shows the worker wearing the helmet and located in a safe area and Figure 3(Right part) shows the worker wearing the helmet located in unsafe area. Helmet showing different LED color signals – green and red respectively for safe and unsafe zones.



a. Safe Zone

b. Unsafe Zone

Figure 3: Worker wearing helmet in safe zone (a) and unsafe zone (b)

4 Solution

The safety helmet designed for hazardous environments operates seamlessly and efficiently to detect and communicate potential hazards in a timely manner.

The following steps outline the key functionalities of the system (Figure 4):

- Workers wear the safety helmet, which serves as the primary interface for hazard detection and communication.
- The safety helmet is outfitted with BLE receiver (B), responsible for collecting the information related to the environment.
- BLE transmitters (A) are strategically positioned throughout the workplace to establish a comprehensive network, particularly in areas marked as potential hazard zones.
- A central control unit (D) continuously monitors the workplace for potential hazards. This task is done by safety manager who gets inputs from all the workers and management about the safe zones of construction. Based on this input, safety manager sets the alert in the system, the server promptly sends this real-time hazard information to active BLE receivers within the identified hazard area remotely through internet/intranet
- Upon receiving hazard information, the safety helmet communicates with the user through red LED lights / blinking lights to visually alert the user about the presence of a hazard.

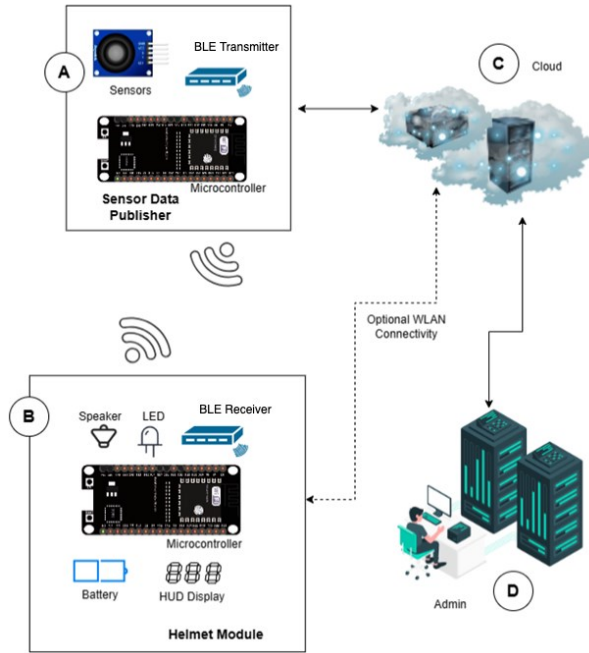


Figure 4: Solution architecture

The system architecture comprises two primary components: hardware and software. The hardware segment primarily focuses on data transmission and visualization utilizing BLE technology and LED indicators, respectively. The software part is dedicated to controlling the system overall and marking hazardous environments by the safety manager.

The entire flow of the working principle of smart helmet is depicted in the Figure 5. The steps involved are

- Deploy BLE transmitters across site
- Smart helmet receives the signal from transmitter
- MAC ID is sent to server for verification
- Server responds validity
- Server sends alerts – Location, Battery, Profile Risk
- Alert signal received
- Alert signal notifies using LED and Buzzer

3.7V 500mA Lithium Polymer Battery is the critical component of entire system. The LiPo battery is a rechargeable battery makes the helmet a lightweight and demands only compact power source. The compact size of the battery makes the system design thin, light weight and easy to deployable.

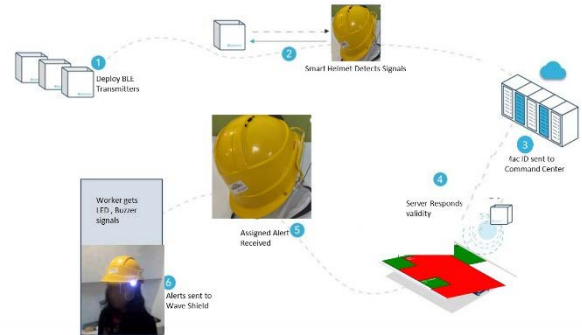


Figure 5: Solution workflow

4.1 Hardware Design

The hardware system comprises of two components: the transmitter and the receiver. The transmitter component is affixed in each hazardous environment, establishing a connection with Wi-Fi or Ethernet to gather information from the safety command center. Subsequently, it transmits the data via BLE signals to the smart helmets. The smart helmet is equipped with a receiver module, responsible for capturing the BLE signals and translating them into LED indications, which are then presented to the laborers.

4.1.1 Transmitter

The transmitter circuit comprises an ESP32, Bluetooth HC-05 module, and a Battery Management System (BMS) circuit. In the schematic (Figure 6), the ESP32 is connected to the Bluetooth module via the RX and TX pins. The BMS circuit includes a TP4056, which interfaces with the ESP32 through the GND and VIN pins, establishing connections with the positive and negative terminals of the BMS circuit. The battery is linked to the BMS circuit through the BAT+ and BAT- pins on the TP4056, completing the entire transmitter circuit.

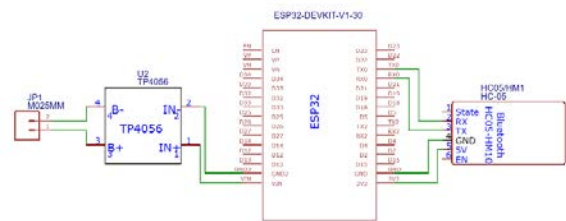


Figure 6: Transmitter circuit diagram

4.1.2 Receiver

The receiver circuit includes an ESP32, Bluetooth HC-05 module, BMS circuit, OLED display, WS2812B addressable LED, and a buzzer. In the schematic (Figure 7), the ESP32 is linked to the Bluetooth module via the RX and TX pins. The TP4056 serves as the battery management system, connecting to the ESP32 through the GND and VIN pins, establishing connections with the positive and negative terminals of the BMS circuit. The battery is interfaced with the BMS circuit through the BAT+ and BAT- pins on the TP4056.

For functional components, the OLED display (optional) is connected to the ESP32 using the SCL and SDA pins, while the WS2812B addressable LED and buzzer are linked to the ESP32's I/O pins for actuation.

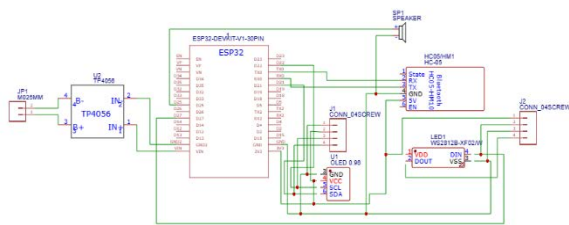


Figure 7: Received circuit diagram

4.2 Software Design

The software aspect of the system is designed using React.js which offers a seamless experience for users to visualize the entirety of the designated area, facilitating the clear demarcation of safe and unsafe zones. Moreover, an administrative dashboard has been implemented to provide real-time monitoring of the status of connected helmets.

In addition to its user interface capabilities, the system boasts a robust control panel (Refer Figure 2.1 in section 3.3) positioned conveniently on the right-hand side. This feature empowers safety managers with the functionality to swiftly designate areas as either safe or unsafe, thereby enhancing overall operational efficiency. Subsequently, this critical information is seamlessly relayed over the internet to the BLE transmitter.

Furthermore, the system is enriched with an array of analytical tools (Figure 8) designed to provide comprehensive insights into helmet utilization patterns. These analytics not only afford administrators the ability to gauge the frequency and duration of helmet usage but also enable them to identify potential areas for optimization and improvement.

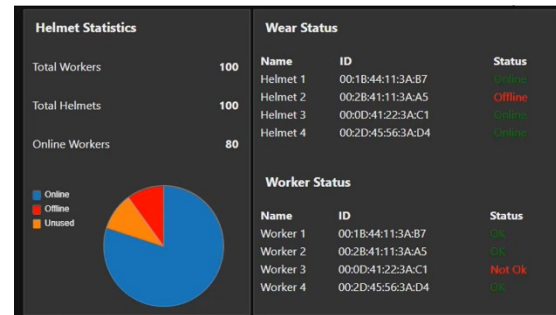


Figure 8: Software Dashboard

When an unauthorized MAC ID is registered in the system, verification fails and that MAC ID will not get any alerts. Further details about such MAC ID will be logged in the system to find the trespasser and security hacker

5 Results

When the PoC was carried out in our lab setup, below results were observed.

When the worker was loaded with other tasks, suddenly alert was sent by setting that zone as unsafe. Immediately red LED was powered and red light was shown and since it was easily noticeable, worker was able to understand the alert and moved out of the area till the LED warning sign turns back into normal color.

Secondly when the battery was used for more than 5 hours and then checked the safety command center to observe the battery condition of this worker. As expected in the UI, alert was triggered to safety manager and helmet condition changed to Red

Third condition, two different workers are observed in two regions. Worker 1 – Room_1, Worker 2 – Room_2 is the initial setup. When Room_1 was set as unsafe zone, checked whether worker 1's helmet is giving red signal, and when he moved to Room_2 (safe zone) his helmet signboard turned into green. Again, from the safety command center, Room_2 is marked as unsafe and turned Room_1 to safe zone, expected behavior of both signboards of helmets are turning into red is observed correctly. While asking only Worker 1 to move into Room_1 (safe zone now), Keeping the worker 2 in unsafe zone Room_2. It is observed that additional triggers are being sent to the safety manager after 3 minutes as worker 2 has not acted upon the signal warning.

Profile based alert or person-based alert was also tested. Though a worker is in safe zone, due to his higher

risk exposure, sent an alert to worker 2 from the system, then helmet of the worker 2 received the alert and warning red LED was on.

Continuous monitoring over long duration is carried out to know the battery capability. It is observed that helmet battery can work for 5 hours in continuous manner.

We asked workers about ergonomic comfort like weight and disturbance caused by electronic signal. Both the workers responded that additional weight is negligible which is around 150 grams and electronic signal and LED signboard are not causing any trouble to their regular activities. However, there is a concern over charging the helmet on regular basis which is a new additional task to them.

6 Discussions & Future scope

Five major challenges that affect building any safety system in construction work force- Dynamic nature, communication, cost, utility & ergonomics, ability to collect data. The present study proposes BLE powered smart helmet which can overcome all the specified challenges to some extent. Because of connectivity established, any area in the construction site can be marked as safe or unsafe zones easily within minutes of time. By using BLE beacons and BLE receivers inside the helmet communication gap related challenges can be addressed. The present study is carried out on experimental study on lab setup, so detailed cost comparison is not yet made. However considering the huge number of accidents in this sector, it justifies the investment on new technology-based safety helmet. Still our calculations show that this helmet will be comparable enough costing 1500 Rupees (INR-Indian currency). Present study carefully designed the safety warning using color-based alert using tricolor LEDs. Transparent acrylic board with LED is ergonomically positioned in such a way that it will not affect the works being carried out, at the same time able to alert the worker quickly to get full attention. A major advantage of the proposed study is its ability to collect the data from all the construction workers and helps in forming safety analytics dashboard like how many workers entered into unsafe zone and how long they were inside unsafe zone. Data collection helps to provide reasonable Rewards and Recognition to the safety compliant workers. Experimental results showed that all the critical challenges can be addressed by this proposed study, however the same need to be validated with field implementation in the future work.

Though there are lot of benefits in the proposed approach, there are also few limitations. Those issues are privacy, battery charging and dependency of zone data entry. Present study assumes safety manager will be knowing all the data about safe and unsafe zones and able to enter them into safety command center instantly. This assumption may not be practically possible. Hence auto detection of safety zones will be studied in the future study. Location data about workers' movement is continuously recorded in the system, so privacy consent must be collected from workers to adhere to Global data privacy and regulation (GDPR). In the future study, data purging, consent collection and other requirements of GDPR will be built. Battery powered helmets are having inherent issue of battery charging regularly. Though there is an alert sent to workers whenever battery charge goes below 20% and there is data collection about the workers' devices which are having drained batteries, charging still has to be done. Since construction work is of open nature, in the future study possibility of using solar Photo Voltaic (PV) will be studied to overcome this critical challenge. Installing beacons and Wi-Fi routers across the construction site could be a tedious task and may not even be possible in some extreme scenarios. Profile based alerts opens up the possibility of handling various risk exposure which also changes with time. Constant wear of BLE powered helmet may expose Specific Absorption Rate (SAR) which is a measure of the rate at which RF energy is absorbed by the human body when exposed to a wireless signal from devices. IoT devices are exposed to high software security threats and attacks, the present study follows just MAC address-based approach. In the future, detailed white listing approach can be studied.

7 Conclusions

Considering the research question, the study's objective was to develop a proof-of-concept smart helmet and measure its ability to address the critical challenges faced by construction industry. A BLE beacon-based communication was adopted, and an ergonomically designed warning board with tri-color LED was adopted. Results demonstrated that dynamic zone can be indicated to workers instantly using both software and hardware. Using front end GUI, whenever safety manager marks any zone as safe or unsafe and helmet hardware instantly gets this alert. However, at this stage, the contraption is just a proof-of-concept to demonstrate the capability of IoT based smart helmet to understand the possibilities and limitations associated with it. Nevertheless, this study is a step closer to realizing the potential of such smart helmets in making highly unsafe construction field into a safe area to work.

8 References

- [1] M. Ibrahim Shire, G. T. Jun, and S. Robinson, "The application of system dynamics modelling to system safety improvement: Present use and future potential," *Safety Science*, vol. 106. Elsevier B.V., pp. 104–120, Jul. 01, 2018. doi: 10.1016/j.ssci.2018.03.010.
- [2] J. Baek and Y. Choi, "Smart glasses-based personnel proximity warning system for improving pedestrian safety in construction and mining sites," *Int J Environ Res Public Health*, vol. 17, no. 4, Feb. 2020, doi: 10.3390/ijerph17041422.
- [3] Landi, L., Buffi, A., Marracci, M., Stecconi, A., Leone, P.D., Bernardini, F. and Didonato, L., 2021. Localization systems for safety applications in industrial scenarios. In *Proceedings of the 31th European Safety and Reliability Conference* (pp. 2413-2419). Research Publishing.
- [4] Kim Y, Choi Y. Smart Helmet-Based Proximity Warning System to Improve Occupational Safety on the Road Using Image Sensor and Artificial Intelligence. *International Journal of Environmental Research and Public Health*. 2022; 19(23):16312. <https://doi.org/10.3390/ijerph192316312>.
- [5] "ESP32", Accessed: Dec. 25, 2023. [Online]. <https://www.espressif.com/en/products/socs/esp32>
- [6] Institute of Electrical and Electronics Engineers, 2019 IEEE Workshop on Metrology for Industry 4.0 and Internet of Things : proceedings : Naples, Italy, June 4-6, 2019.
- [7] Lalitha K, Ramya G, Shunmugathammal M. AI-Based Safety Helmet for Mining Workers Using IoT Technology and ARM Cortex-M. *IEEE Sensors Journal*. Volume: 23, Issue: 18, September 2023.
- [8] Sandi Kumar Reddy and Anil S Naik. An Enhanced IoT and LoRa-Based Communication System for Underground Mines. *Signals, Machines and Automation* (pp.513-521)
- [9] Ying Huang, Kong Fansheng, He Yifei. Applying beacon sensor alarm system for construction worker safety in workplace. *IOP Conference Series Earth and Environmental Science*. December 2020.
- [10] Kong, F.; Ahn, S.; Seo, J.; Kim, T.W.; Huang, Y. Beacon-Based Individualized Hazard Alarm System for Construction Sites: An Experimental Study on Sensor Deployment. *Appl. Sci*. 2021,11,11654.
- [11] Zhe Zhanga, Brian H.W. Guoa, Alice Chang-Richardsb, Ruoyu Jinc, Yu Han. Digitalisation-based Situation Awareness for Construction Safety Management – A Critical Review. 38th International Symposium on Automation and Robotics in Construction (ISARC 2021).
- [12] Sina Rasouli, Yaghoub Alipouri, Shahin Chamanzad. Smart Personal Protective Equipment (PPE) for construction safety: A literature review. *Safety Science* Volume 170, February 2024.
- [13] J. Baek & Y. Choi. Development of a low-cost Proximity Warning System for mine equipment using smartphone and bluetooth beacons. *Mining goes Digital* (pp.715-719).

Towards Automation in 5s – Classification and Detection of Images from Construction Sites

Viral Makwana¹ and Ganesh Devkar²

¹Faculty of Technology, CEPT University, India

²Faculty of Technology, CEPT University, India
ganesh.devkar@cept.ac.in, viral.pcm21404@cept.ac.in

Abstract –

Traditional construction management practices often suffer from inefficiencies and waste, leading to increased costs and project delays. Lean Construction (LC), by emphasizing waste reduction and customer value prioritization, has emerged as an effective construction management system. In recent years, efficiency enhancement approaches such as Lean are being further augmented by automation and AI tools in the Architecture, Engineering, and Construction (AEC) industry.

This work develops an ML-based image processing algorithm that can be used within the Lean Construction framework. This algorithm classifies construction images based on the 5S Lean tool's sort and set in order principles, utilizing image analysis techniques. The algorithm that leverages convolutional neural networks and TensorFlow/Keras libraries, shows promise in effective image classification under the 5S framework. This study also proves that YOLO models can integrate lean concepts with AI Tools to foster improvements in productivity, safety, and sustainability in the construction industry.

Keywords –

5s; Image Classification; YOLO Lean Construction; Automation; Object Detection; Image Classification;

1 Introduction

The construction industry is facing increasing pressure to improve efficiency, productivity, and safety, and the avenues offered by automation, and Artificial Intelligence tools such as machine learning, object detection, and image analysis are seen as key to achieving these goals. Construction projects around the world are confronted with numerous challenges. Automation and AI tools can help streamline and optimize construction processes through data-driven decisions [1]. Object

detection and image analysis can be used to improve the accuracy and efficiency of tasks such as classification and site inspection.

For the past two decades, Lean Construction is being used to overcome the inefficiencies and waste associated with traditional construction practices[2]. The efficacy of Lean construction can be greatly enhanced with the implementation of AI tools and there is increasing academic and industrial interest in this synergy[3]. The integration of automation, machine learning, and advanced analytics aligns with the lean principles of waste minimization by facilitating real-time monitoring, continuous improvement, and data-driven decision-making [2].

Cisterna et al., (2022) explored the benefits of integrating of Lean Construction (LC) methodologies and Artificial Intelligence (AI) approaches within the realm of production processes in projects and corporate settings. They used a theoretical framework and provided illustrative examples of application scenarios specifically within the Architecture, Engineering, and Construction (AEC) industry. The study's findings showed that the incorporation of Lean principles in construction initiatives can enable the assimilation of AI technologies. This was attributed to the capacity of AI-driven automation and support solutions to alleviate individuals from tasks characterized by tedium or complexity. Furthermore, the research showed that the interplay between AI and Lean methodologies can generate mutual growth and value addition[3].

One Lean construction methodology that can benefit from AI integration is 5S. 5S is a cyclical methodology based on a visual management approach, that emphasizes on organization and cleanliness in the workplace. To ensure the site is following 5S principles, personnel trained in the 5S tool periodically inspect the site and provide suggestions for improvement. Site photographs are frequently taken and reviewed in meetings to identify areas for improvement[4]. These improvements are categorized as sort, set in order, shine, standardize, and sustain. The identification of areas of improvements (which is also called as Kaizens) requires considerable

involvement of human beings for analysis of site photographs from the perspective of 5S principles. Often, this process of analysis would be time consuming and tiresome. Utilizing advanced AI-driven image processing and machine learning tools can significantly enhance efficiency, and can be used in the sort and set in order categories[3].

The objective of this study was to develop an image processing algorithm that can classify construction images according to the sort and set in order principles of 5S Lean tool. The algorithm was designed to use image analysis techniques to identify specific features in the construction images and categorize them based on the 5S methodology. The goal was to provide a useful tool for construction professionals to improve their processes and increase productivity.

2 Literature Review

The Lean philosophy emphasizes the importance of continuous improvement, and 5S lean tool is considered a crucial prerequisite for implementing lean principles. According to Obulam & Rybkowski (2021), 5S can serve as a valuable tool for organizations to improve their operations by reducing waste and optimizing costs. The 5S method is based on five principles, each of which starts with the letter "S" phonetically. 1) The Seiri principle involves sorting by getting rid of unneeded items. 2) The Seito principle involves putting things in order by arranging them based on their purpose. 3) The Seiso principle is centered around cleaning or shining by eliminating sources of dirt. 4) The Seiketsu principle involves standardizing a system by establishing a consistent approach. 5) Finally, the Shitsuke principle emphasizes the importance of sustaining the system to ensure its ongoing maintenance. Obulam & Rybkowski (2021) The literature related to 5S principles mainly focuses on their practical application and execution in the workplace. Various checklists are used to evaluate the 5S process, which may take the form of questionnaires, marking or numbering checklists, or evaluation sheets. These evaluations may be considered as appraisals or assessments of goal achievement or the percentage of 5S implementation achieved.

Over the past decade, automated processes have become popular in a variety of industries. Periodic and repeatable processes can be optimized via automation. Repetitive and mundane jobs in construction include monitoring progress and evaluating quality of various processes. Currently, a team of experts completes these tasks by visually inspecting the construction site and frequently making use of measuring equipment. Unfortunately, given the reliance on human factors, this can result in discrepancies between reports that are issued at different times [6].

Construction companies are adopting digital imaging to enhance their visual inspection processes [7]. With digital imaging, high-resolution images and videos of the construction site can be quickly and easily captured, enabling potential issues or problems to be identified. The use of digital images also facilitates a more thorough and consistent inspection process, as they can be easily stored and reviewed at a later time. By utilizing digital imaging, construction companies can save time, reduce the risk of human error, and improve the overall quality of their construction projects.

The construction industry has seen an increase in the use of deep learning-based object detection methods, which are primarily focused on identifying components, detecting damage and defects, classifying objects, monitoring worker activities, and detecting collisions on construction sites. These methods use computer vision algorithms to recognize and detect objects. There are several object detection models that are commonly used, including the R-CNN family [8], that includes R-CNN, Fast R-CNN, and Faster R-CNN, the YOLO family [9], and the SSD models [10]. The Yolo series of models is known for its high speed of detection. You Only Look Once (YOLO) is an advanced object detection algorithm used in computer vision. The latest YOLOv5 model has been applied in various fields such as pedestrian detection, vehicle detection and medical applications used the YOLOv5 model to develop a real-time detection method for unauthorized intrusion on construction sites [11].

3 Research Methodology

3.1 Image Classification Model

Based on studies and research, it is possible to automate the Sort principle of 5S using image classification. The Sort principle involves identifying the items that are necessary and those that are not and eliminating the unnecessary ones from the workplace. One potential application of image classification for the Sort principle is in the steel yard. Steel reinforcement bars are commonly used in construction, and they are often stored in large quantities in steel yards. However, these bars can become disorganized, making it difficult to find the required bars when needed.

To automate the sorting process in the steel yard, we can use image classification to identify and separate the steel reinforcement bars. The classification process can be achieved by training a machine learning algorithm to recognize two classes of images: one with steel bars placed in an unorganized manner, and another with steel bars placed in an organized manner, tied together in a bundle.

To train the machine learning algorithm, a large

dataset of images of steel bars in the steel yard is required. The images should include examples of both organized and unorganized placement of the steel bars. Once a sufficient dataset is obtained, deep learning algorithms can be used to train the machine to recognize the two classes of images.

Once the machine learning algorithm has been trained, it can be used to classify new images of steel bars. As new images are fed into the algorithm, it would identify the placement of the steel bars and classify them as either organized or unorganized. This would enable construction workers to quickly and easily identify the steel bars that are needed, improving the efficiency and productivity of the construction process.

A methodology was proposed to create an image processing algorithm for classifying construction images for both modal Image classification as well as Object detection model according to the principles of 5S, which consists of the following steps: (1) Collecting a dataset of construction images and annotating them to identify the objects of interest; (2) Extracting relevant features from the images, such as object shapes, textures, and colors; (3) Training a machine learning model, such as a Convolutional Neural Network (CNN) or YOLO, using the annotated data to classify objects in the images; (4) Evaluating the model's performance on a test dataset to determine its accuracy in object classification; (5) Integrating the image processing algorithm into a construction workflow, such as the 5S system, to automate the classification of objects in construction images; (6) Continuously improving the algorithm's performance by incorporating new data and adjusting the model as required.

3.2 Object Detection model

The YOLO (You Only Look Once) model is a popular object detection algorithm that has been widely used for the identification of labeled objects in images. This model is designed to recognize objects in real-time, making it an effective tool for a wide range of applications. In order to train the YOLO model to recognize labeled objects in images, it is necessary to provide a dataset of annotated images that includes both labeled and unlabeled images. The annotated dataset should include a file that specifies the location of each labeled object within the image, as well as its class label.

According to the set-in-order principle of the 5S system, visually marked, specified places are designated for everything, and visual aids such as signboards, color-coding, labeling, or outlining may be used. The research methodology adopted for Yolo model development is shown in Figure 1.

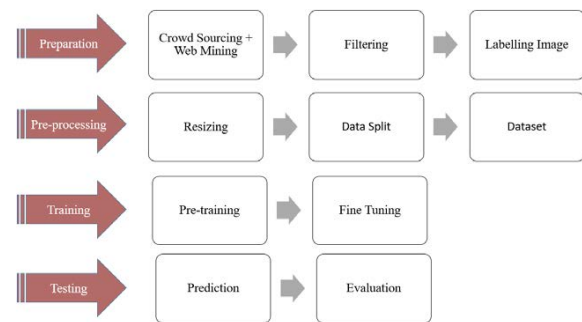


Figure 1 Methodology-Yolo model development

1. Preparation

To start building the YOLO object detection model, a dataset must be prepared, which will be utilized to train and test model. This dataset consists 50 to 60 nos of images from Web mining and construction site. This involves selecting and annotating the dataset, which includes choosing images for training and testing, and labeling the objects within the images indicating their presence and location accurately and consistently. The annotation process can be time-consuming and laborious, and there are annotation tools available like Labellmg and VGG Image Annotator (VIA) to aid in this task. Once the dataset is annotated, it is divided into the training data and testing data sets. The YOLOv5 model is trained using the training dataset, and its effectiveness is assessed using the testing set.

2. Pre-processing

Before training the YOLOv5 model, the prepared dataset must be preprocessed. This involves several steps such as resizing the images to a fixed size, normalization of the pixel values, and the application of data augmentation techniques.



Figure 2 Image annotation

The purpose of resizing the images is to make sure that they are of uniform size, which is a requirement for training the YOLOv5 model. Normalization of the pixel values is necessary to scale the pixel values within a range of 0 to 1, which increases the YOLOv5 model's

efficiency. In order to expand the training set and improve the model's generalizability, data augmentation approaches are used. Random cropping, flipping, and rotation are often used methods of data augmentation.

3. Training

To train the YOLO model with the prepared dataset, certain steps must be followed. First, the training environment is set up by installing essential software and libraries like PyTorch and CUDA and configuring hardware like GPUs. Next, the YOLO model is configured by setting hyperparameters like batch size, number of epochs, and learning rate, which significantly impact the model's performance. The preprocessed dataset is then fed into the model, and the model's parameters are iteratively changed depending on the loss function to begin the training process. The goal of this function is to minimize the difference between the expected output and the ground truth labels during training. It calculates the difference between the predicted output and the ground truth labels.

4. Testing

The performance of the YOLO model is assessed using the testing set by setting up the testing environment, configuring the YOLO model, and running the testing process. The testing environment is created by installing necessary software and libraries, like PyTorch and CUDA, and configuring the hardware, like GPUs, for testing YOLOv5 and YOLOv8 models. Configuring the YOLO model involves selecting the most suitable checkpoint, which is the saved model that performed best during training on the validation set, and then using the model to predict objects in the testing set. Various metrics like mean average precision (mAP) and intersection over union (IoU) are used to determine the accuracy and overlap of the predicted bounding boxes with the ground truth bounding boxes.

5. Fine-tuning and Deployment

Once the YOLO model has been tested, it is fine-tuned to enhance its performance. This process requires adjusting the hyperparameters and retraining the model on the complete dataset, including the validation and training sets.

Once the model is fine-tuned, it can be deployed in various applications, such as security camera systems and self-driving cars. The process involves integrating the YOLO model into a larger software system or application, enabling efficient object detection.

4 Data collection

4.1 Image Classification Dataset

Image classification is an important task in the field of computer vision, and it has numerous applications in various industries. One such application is the automated sorting of steel bars based on whether they are organized or not. To carry out this task, it is necessary to collect a significant amount of data in the form of images of steel bars. The images can be collected using various methods, such as web crawling and crowdsourcing.

Web mining is a method of collecting data from the internet. In this study, web crawling was used to collect images of steel bars from various web platforms and search engines. The images were collected using specific keywords such as "steel bars," "construction steel bars," "reinforcement steel bars," and so on. The images collected from web crawling varied in terms of quality, resolution, and orientation. Therefore, it was necessary to manually screen the images to remove any irrelevant or low-quality images. This process ensured that only high-quality images were included in the final dataset.



Figure 3 Reinforcement Bar Images for Training

In this study, crowdsourcing was used to collect images of steel bars from ongoing construction sites. The images were collected by contacting various construction companies and asking them to provide images of steel bars used in their ongoing projects. The images collected through crowdsourcing were of high quality and provided a wide variety of steel bar configurations. This method of data collection was particularly useful as it provided a more realistic representation of the steel bars used in real-life construction projects.

Once the images were collected, the next step was to clean and prepare the data. This process involved removing irrelevant and low-quality images and organizing the remaining images into relevant categories. The images were categorized based on the type of steel bars, such as organized or unorganized steel bars. The images were also resized and standardized to ensure that they were all of the same size and resolution. This step ensured that the images were processed efficiently during the image classification stage.

Techniques for data augmentation were utilized to

expand the dataset. Data augmentation entails modifying the original photos to produce new images that are similar to the originals but not exactly the same. In this study, new pictures were produced using data augmentation methods including flipping, rotating, and zooming.

4.2 Object Detection Dataset

The principle of "Set in Order" involves organizing workspaces and equipment for improved efficiency and productivity. In this study, the aim was to automate the process of identifying labels as objects that are set in order using object detection techniques. To achieve this goal, a significant amount of data was required to train and evaluate the performance of the object detection model. As before, two methods were used for data collection: web mining and crowdsourcing.

In this study, web mining was used to collect images of labels from various websites and search engines. The images were collected using specific keywords such as "labels," "organized labels," "inventory labels," and so on. The images collected through web mining were of varying quality, resolution, and orientation. Therefore, the images were manually screened to remove any irrelevant or low-quality images. This step helped to ensure that only high-quality images were included in the final dataset.

Crowdsourcing was also used to collect images of labels from various industries and workplaces. The images collected through crowdsourcing provided a more realistic representation of labels used in real-life work environments. This method of data collection provided a diverse range of labels and configurations. The images were collected by contacting various contracting companies and asking them to provide images of their labeled equipment and workspaces.

After the images were gathered, Labellmg software was used to annotate them. Annotation entails manually highlighting the subject matter of a picture. The labels were the study's items of interest, and they were marked using bounding boxes. The annotation method contributed to the creation of a labelled dataset, which was necessary for the object detection model's training and performance evaluation.

The next step was to clean and prepare the data. This process involved removing irrelevant or low-quality images and ensuring that the remaining images were appropriately labeled. The images were organized into relevant categories based on the type of labels, such as inventory labels, equipment labels, and so on. The images were also resized and standardized to ensure that they were all of the same size and resolution. This step ensured that the images were processed efficiently during the object detection stage.

Data augmentation entails modifying the original

photos to produce new images that are similar to the originals but not exactly the same. In this study, new pictures were produced using data augmentation methods including flipping, rotating, and zooming.

5 Data analysis

5.1 Image classification

This study implemented an image classification model using deep learning architectures such as TensorFlow Keras, Inception ResNet v2, and EfficientNet v2 B0. The dataset was split into training and validation sets with a validation split of 20%, and the model was trained using a batch size of 32 across 10 epochs. The history variable, which holds the loss and accuracy values for the training and validation sets for each epoch, was used to analyze the model's performance throughout training. A visual representation of the model's performance during training was made using these values. Similarly, the EfficientNetV2B0 model was tested on the dataset and the results were generated. One example of the EfficientNetV2B0 model is shown below, in Figure 4 where the accuracy graph is plotted.

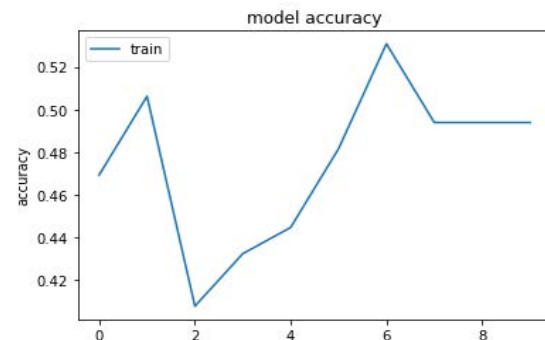


Figure 4 Model accuracy- EfficientNetV2B0

Model accuracy- EfficientNetV2B0

As shown in Figure 4, a graph displaying model accuracy is a visual representation of how well a machine learning model can predict the correct output for a given dataset. The Y-axis of the graph shows the accuracy of the model, while the X-axis indicates the number of training epochs or iterations. The training accuracy represents the model's performance on the data it was trained on, while the validation accuracy indicates how well the model is generalizing to new data.

The ideal situation is for both training accuracy and validation accuracy to increase steadily and converge over time, indicating that the model is learning and generalizing well. However, if the training accuracy

continues to increase while the validation accuracy plateaus or decreases, this indicates overfitting, where the model is too complex and has learned to fit the noise in the training data instead of the underlying patterns. Conversely, if both the training and validation accuracy are low, this suggests underfitting, where the model is too simple and cannot capture the complexity of the data.

Figure 5 presents the EfficientNetV2B0 model, and the validation loss graph generated during training. The training loss and validation loss are shown on the Y-axis, while the number of epochs or iterations is shown on the X-axis. When plotted against epochs or iterations, the training loss and validation loss graph generally depicts the fluctuation in training loss and validation loss over time. The validation loss gauges how successfully the model generalizes to fresh, untested data, whereas the training loss assesses how well the model fits the training data. When training a model, the ideal situation is for the training loss to reduce over time while the validation loss stays the same or reduces. Overfitting of the model to the training data is indicated by a rising validation loss while the training loss stays declining.

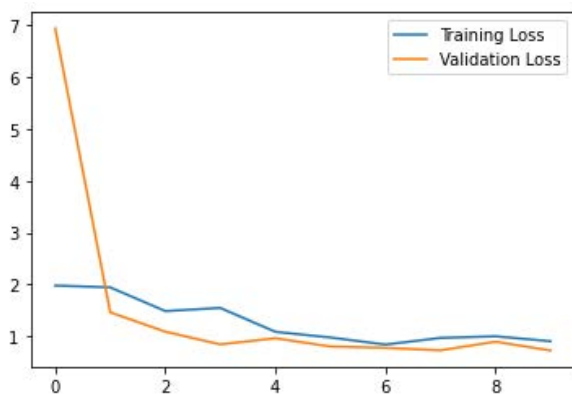


Figure 5 Train Val loss- EfficientNetV2B0

After training and validating the model on the training and validation sets respectively, the performance of each model was visualized by plotting the training and validation loss curves. The EfficientNet v2 B0 model achieved the highest accuracy of 0.8981%, followed by Inception ResNet v2 with 0.8276%, and TensorFlow Keras with 0.6969%. This indicates that the choice of architecture is crucial for achieving high accuracy in image classification tasks, as well as the quality of the dataset used for training and validation. Overall, the image classification model developed in this project successfully differentiated between organized and disorganized steel bar images.

5.2 Object Detection Model

The validation was performed using the best model

saved during training. the validation results obtained from training the YoLOv5 model on the designated dataset. Subsequent to the validation process, the model underwent testing on a distinct dataset comprising 20 images, encompassing a collective total of 184 object instances. The model was tested on 20 images, containing a total of 184 instances of objects. The precision (P) and recall (R) values for the 'all' class are 0.813 and 0.732, respectively.

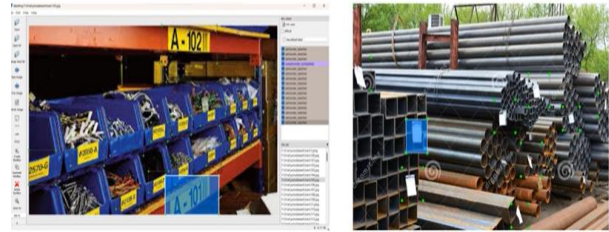


Figure 6 Image annotation for Labelled images

Furthermore, the validation results were broken down by two additional classes: 'setinorder_labelled' and 'notsetinorder & notlabelled'. These classes may have been manually labeled by the user and refer to images where the objects are either set in order or not labeled in order as shown in Figure 6.

Overall, the mAP50 value for the 'all' class was 0.751, indicating a moderate performance in object detection. The mAP50 values for the additional classes were 0.849 and 0.653, respectively. These results suggest that the model may be better at detecting objects in labeled images that are set in order.

It is important to remember that the YOLOv5 model's accuracy can vary based on a wide range of variables, including the caliber and volume of training data, hyperparameters, and network design. To assess the model's performance in various circumstances, more testing and analysis are advised.

An essential tool for assessing a classification model's performance is the confusion matrix. It provides a thorough analysis of the predictions made by the model and how well they correspond to the actual labels. A confusion matrix was created for the YOLOv5 model during the validation procedure to examine the model's performance on the validation set. For each class the model was trained on, the matrix shows the quantity of true positives, false positives, true negatives, and false negatives.

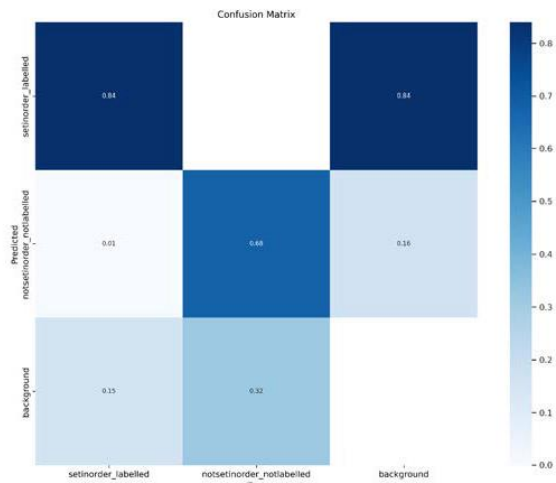


Figure 7 Confusion Matrix for YOLOv5

In the example Figure 7, the confusion matrix shows that the model correctly classified 156 of 184 instances of the "setinorder_labelled" class, resulting in a precision of 0.847 and a recall of 0.833. Similarly, the model correctly classified 22 of 28 instances of the "notsetinorder_notlabelled" class, resulting in a precision of 0.779 and a recall of 0.631. Thus, the confusion matrix generated for the YOLOv5 model provided a comprehensive overview of the performance of the model and helped to identify the areas that need improvement.

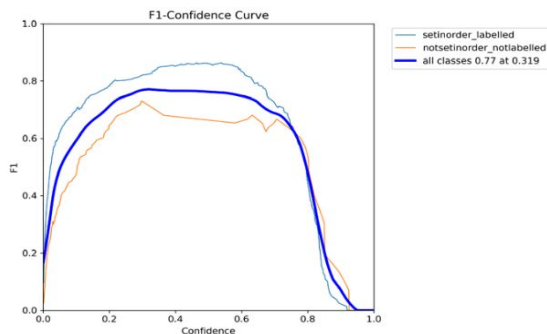


Figure 8 F1-Confidence Curve

The Figure 8 F1 confidence curve is a visual representation of the relationship between the confidence level of a machine learning model's predictions and its F1 score. The F1 score is a commonly used metric for evaluating the accuracy of binary classification models, and it takes into account both precision and recall. The F1 confidence curve is useful for identifying the optimal confidence threshold for a model, which can improve its overall performance. By analyzing the curve, we can determine at what confidence level the model achieves its maximum F1 score, allowing us to make informed decisions about how to tune the model's hyperparameters.

The YOLOv5 model generates a results.png file that gives a comprehensive overview of the model's performance on the validation dataset. The graph in the results.png file displays the precision, recall, and F1 score of the model at different confidence thresholds. Precision refers to the percentage of correctly identified objects out of all the predicted objects, while recall refers to the percentage of correctly identified objects out of all the actual objects in the image. The F1 score is a balanced measure of precision and recall and is calculated as their harmonic mean.

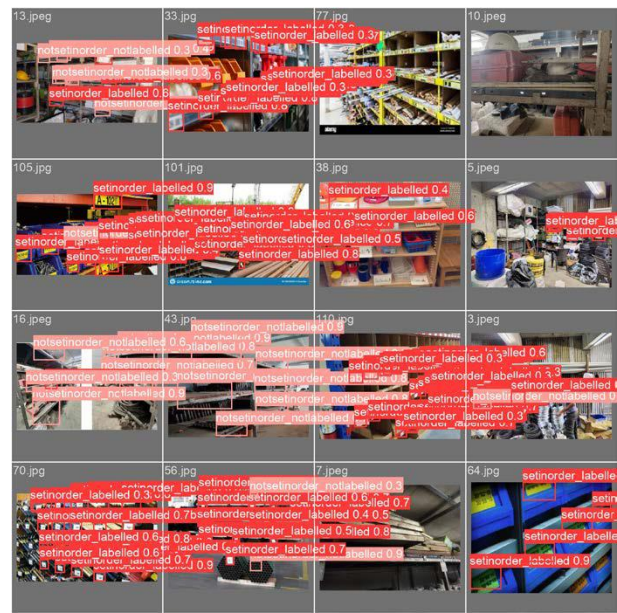


Figure 9 YOLOv8 Object Detection model

Similarly, to the yolov5 model and yolov8 dataset trained on same dataset similarly result received as shown here. The YOLOv8 model as shown in Figure 9 achieved an overall mAP50 of 0.904 and mAP50-95 of 0.688 with a precision of 0.92 and recall of 0.879 on the validation set after 300 epochs of training. The model was able to detect 384 instances of objects in 34 images, out of which 304 instances were correctly identified and labeled as "setinorder_labelled" with a precision of 0.924 and recall of 0.908. The remaining 80 instances were labeled as "notsetinorder_notlabelled" with a precision of 0.917 and recall of 0.85. The model was able to perform inference at a speed of 19.1ms per image.

The confusion matrix shows the model achieved an mAP50 of 0.904 and an mAP50-95 of 0.688 across all classes. The precision and recall for the 'setinorder_labelled' class were 0.924 and 0.908, respectively, while for the 'notsetinorder_notlabelled' class, they were 0.917 and 0.85, respectively. The model performed better for the 'setinorder_labelled' class than the 'notsetinorder_notlabelled' class, with higher

precision, recall, and mAP50 values.

Similarly, to yolov5 model and yolov8 model generated result graph which shows coming with more accuracy. The result graph generated from the yolov8 result shows the relationship between the confidence threshold and precision, recall, and F1 score for each class in the dataset.

6 Conclusion

This study underscores the transformative potential of integrating Lean principles with automation in the construction industry. By taking the first step towards integrating Lean 5S methodologies with automation.

Although the preliminary algorithm's limitations underscore the necessity for refinement, the methodology's fundamental approach, utilizing convolutional neural networks and TensorFlow/Keras libraries, demonstrates promising potential for enhancing efficiency within the construction sector. This study, conducted with a constrained dataset sourced from both web mining and on-site observations, presents a significant analysis of implementing automation alongside Lean practices in practical settings. For that further more image for training and validation as well as case studies will be require. Also considering current data This study also proves that YOLO models can integrate lean concepts with AI Tools to foster improvements in productivity, safety, and sustainability in the construction industry.

The research also suggests avenues for future work, including the verification of AI model accuracy, algorithm development for "Set in Order," comprehensive data collection, and potential revisions to optimize outcomes. Future research may address the limitations of YOLO models, such as challenges with small or distant objects and interpretability concerns.

Exploring integration of lean principles with automation in construction processes emerges as a promising avenue to further enhance the transformative capabilities of the industry.

References

- [1] Dumrak, J., & Zarghami, S. A. (2023). The role of artificial intelligence in lean construction management. *Engineering, Construction and Architectural Management*, June. <https://doi.org/10.1108/ECAM-02-2022-0153>
- [2] Koskela, L., Howell, G., Ballard, G., & Tommelein, I. (2007). The foundations of lean construction. *Design and Construction*, January 2014, 211–226. <https://doi.org/10.4324/9780080491080>
- [3] Cisterna, D., Lauble, S., Haghsheno, S., & Wolber, J. (2022). Synergies Between Lean Construction and Artificial Intelligence: AI Driven Continuous Improvement Process. *Proc. 30th Annual Conference of the International Group for Lean Construction (IGLC)*, 49, 199–210. <https://doi.org/10.24928/2022/0122>
- [4] Gil, D., Lee, G., & Jeon, K. (2018). Classification of images from construction sites using a deep-learning algorithm. *ISARC 2018 - 35th International Symposium on Automation and Robotics in Construction and International AEC/FM Hackathon: The Future of Building Things*, Isarc. <https://doi.org/10.22260/isarc2018/0024>
- [5] Obulam, R., & Rybkowski, Z. K. (2021). Development and Testing of the 5S Puzzle Game. *IGLC 2021 - 29th Annual Conference of the International Group for Lean Construction - Lean Construction in Crisis Times: Responding to the Post-Pandemic AEC Industry Challenges*, 309–319. <https://doi.org/10.24928/2021/0152>
- [6] Prieto, S. A., Giakoumidis, N., & de Soto, B. G. (2021). AutoCIS: An Automated Construction Inspection System for Quality Inspection of Buildings. *Proceedings of the International Symposium on Automation and Robotics in Construction*, 2021-Novem(November), 669–676. <https://doi.org/10.22260/isarc2021/0090>
- [7] Zhang, X., Bakis, N., Lukins, T. C., Ibrahim, Y. M., Wu, S., Kagioglou, M., Aouad, G., Kaka, A. P., & Trucco, E. (2009). Automation in Construction Automating progress measurement of construction projects. 18, 294–301. <https://doi.org/10.1016/j.autcon.2008.09.004>
- [8] Girshick, R., Donahue, J., Darrell, T., Malik, J., & Berkeley, U. C. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. 580–587. <https://doi.org/10.1109/CVPR.2014.81>
- [9] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016-Decem, 779–788. <https://doi.org/10.1109/CVPR.2016.91>
- [10] Liu, J., Jennesse, M., & Holley, P. (2016). Utilizing Light Unmanned Aerial Vehicles for the Inspection of Curtain Walls: A Case Study. 2651–2659.
- [11] Wan, H., Asce, M., Zhang, W., Ge, H., Luo, Y., Asce, M., & Todd, M. D. (2023). *Improved Vision-Based Method for Detection of Unauthorized Intrusion by Construction Sites Workers*. 149(2015), 1–12. <https://doi.org/10.1061/JCEMD4.COENG-13294>

Analysis of XES and OCEL Data Schemas: Towards Multidimensional Process Mining of Intertwined Construction Processes

Araham Jesus Martinez Lagunas¹ and Mazdak Nik-Bakht²

¹Dept. of Building, Civil & Environmental Engineering (BCEE), Concordia University, Montreal, QC, Canada

²Compleccity Lab, Dept. of BCEE, Concordia University, Montreal, QC, Canada

araham.martinez@mail.concordia.ca, mazdak.nikbakht@concordia.ca

Abstract –

Despite current digital transformation attempts, productivity in the construction industry has remained stagnant for decades. Construction organizations generate vast amounts of data from their day-to-day operations. However, these valuable digital footprints frequently remain underutilized and isolated in different IS (Information Systems). To tackle this issue, process mining, a novel and robust technology, provides organizations with the means and methods to automatically monitor and improve the efficiency of business processes by exploiting event process data. Since construction projects are process-heavy, process mining adoption becomes essential to overcome productivity stagnation. To enable process mining capabilities, process data is required to be captured in the form of event logs. In this regard, the existing body of knowledge in the construction domain yet lacks a system-agnostic framework for event log generation that considers current standard data schemas such as XES (eXtensible Event Stream) and OCEL (Object Centric Event Logs) to ensure event logs' soundness and machine readability. Therefore, to address this limitation, this work aims to facilitate the production of event logs with proper syntax and semantics by: (i) developing an ETL (Extract, Transform, Load) framework to harness process data from IS; (ii) analyzing and describing the XES and OCEL's relational data structures; (iii) providing a comparative analysis of both event log data standards. The results include a functional demonstration for constructing these data schemas applied to the Change Order Management Process (COMP) implemented in a commercial office high-rise building project. Construction change orders have the potential to either contribute to construction projects' success (if managed properly) or yield to failure otherwise.

Keywords –

Construction Management; ETL process; Event Log Generation; Data Schemas; Construction Operations; Process Mining; Process Automation

1 Introduction

The digital intensity of a country's economic sectors is an important metric in quantitatively assessing its digital economic well-being. Among the leading sectors, wholesale, finance, and manufacturing have increased their digital intensity four-fold during the last three decades by adopting cutting-edge ICT (Information and Communication Technology) developments. In this vein, there exists a strong correlation of $r \approx +0.7$ between digitization indices and productivity growth levels reported by several countries [1, 2]. In other words, the higher the digital intensity of an economic sector is, the greater its productivity and economic growth levels are. The labor productivity of digitally intensive sectors grew 22.1% in the past two decades. These leading sectors also presented a greater resilience than non-digitally intensive sectors during the COVID-19 pandemic thanks to the adoption of latest ICT developments [3-5]. Process mining, a recent technological development with exponential market growth [6], has become a crucial asset for several industries in supporting their digital transformation as its capabilities include automated process model discovery as well as monitoring, managing and improving business operations by analyzing actual operational performance behavior from event data logs[7]. To properly implement this technology, these operational event logs must be generated with formal syntax (i.e., proper notation) and clear semantics (i.e., data structures with meaningful data relationships).

In this context, the event log generation process can take up to 80% of the process mining implementation efforts [8]. Therefore, it becomes essential to understand

the event log's composition including its conceptual definitions. In this regard, an event is described by three primary components: (i) a case, which is referred to as a process instance or execution; (ii) an event activity, often referred to as an action, transaction or unit of work happening at a particular point in time; and (iii) timestamps associated with each executed activity [7]. Event logs comprise a collection of events that can be enriched with other attributes such as objects, resources, cost, etc. These enriched event logs allow to analyze specific business operational processes from different perspectives or dimensions [7].

Unlike digitally intensive sectors, construction is part of the lagging non-digitally intensive sectors facing stagnant productivity over the past three decades [3]. Most construction companies use Information Systems (IS) and software platforms to manage their construction projects and daily business operations that vary in size and complexity. These IS store vast amounts of data related to those projects and business processes. The event logs can be retrieved and reconstructed from multiple data sources including SQL relational databases [9]; Enterprise Resource Planning (ERP) systems, [10]; and Project Management Information systems (PMIS), such as Procore [11], Rhumbix [12], and Oracle NetSuite [13]. Other data sources entail Workflow Management Systems (WfMS), Application Programming Interfaces (APIs), sensors, Building Information Modeling (BIM), 3D point clouds, etc., which are enabled by the so-called 'Internet of Events' [7, 14]. However, the stored data in those IS frequently remains underexploited, in this context, the construction domain faces three major challenges that hinder successful process mining implementations: (i) need for data integration (i.e., siloed data in sparse IS); (ii) lack of suitable/quality event logs due to poor understanding of main data schemas for event log generation according to XES and OCEL; and (iii) lack of visibility and transparency over the ripple effect of intertwined processes (i.e., multidimensional process mining). The latter is particularly important given that several construction processes are by nature interrelated with one another across construction project life cycle phases [15].

Aiming at streamlining the event log generation process necessary for process mining deployment in the construction domain, this paper has the following main objectives: (i) to develop an ETL (Extract, Transform, Load) framework for data collection; (ii) to perform a comparative analysis of XES and OCEL event log data standards while analyzing and describing their relational data structures; and (iii) to apply both data schemas into a construction Change Order Management Process as a functional demonstration. This key process, if poorly managed, is one of the main root causes of cost overruns, delays, productivity decline, and legal claims. Previous

studies report expected cost growth of 10% due to change orders for most construction projects, and in some projects, the overrun far exceeds [16, 17]. Thus, it turns out essential to investigate the capabilities that more automated methods such as process mining can enable to model, monitor, audit, and manage construction processes in a more efficient and productive manner.

2 Event Log Standards - Related Works

To provide formal syntax and semantics to event logs, it is necessary to consider the most current worldwide adopted XES [18] and OCEL [19] standards for event log generation and their Event Schema Definitions (XSD).

2.1 XES Standard

The XES standard was created in 2009 by the Architecture of Information Systems (AIS) research group from the Eindhoven University of Technology (TU/E) [20] and in 2010 was adopted by the Institute of Electrical and Electronics Engineers (IEEE) [21]. Nowadays, it is a well-recognized international standard for structuring, storing, and interchange event logs in a machine-readable representation primarily based on the Extensible Markup Language (XML) suitable for process mining implementation [22]. The XES standard is founded on the concept of a single case notion, meaning that events and their related attributes belong to one process instance. Under this standard event logs are generated by selecting the perspective of interest as the event activity related to a process sequence (i.e. process control flow, organizational, cost, time, etc.). Van der Aalst (2003) applied this case-handling approach in the building industry for first time to enable automated process modeling of unstructured construction process involving a lot of uncertainty by extracting construction operational data from WfMS of semi-prefab concrete floor elements and heating, ventilation, and air-conditioning HVAC installations.

2.2 OCEL Standard

The XES standard relies on a single case notion to describe process executions, which is useful and needed to describe processes from a single perspective/angle at a time as required by the process analyst. However, for real-life processes such as those present in construction projects where the executed business operations might be composed of several cases (i.e., interrelated processes), the XES structure does not suffice to represent multiple cases within the same process model. The OCEL 1.0 standard, developed by the AIS research group in 2020, is an object-centric event log structure that empowers business analysts with the capability of structuring and storing multidimensional business processes. OCEL

provides a more realistic view of business process behavior and normally stands between the source data coming from IS and the XES event log extraction. Multidimensional OCEL logs can be flattened into XES logs as needed for further analysis on specific process perspectives/views, yet one should be aware that the flattening approach can result into discovering false process behavior due to convergence (i.e., duplicated events) and divergence (i.e., considering or omitting events that are not part of the selected perspective) problems [23].

3 Methodology

Successful process mining implementations are highly dependent on well-structured event logs. Thus, a suitable method to generate machine-readable high-quality event logs is of paramount importance.

3.1 System-agnostic ETL Framework

To this end, a system-agnostic framework for event log generation is proposed and depicted in Figure 1. Regardless of the data source system, the ETL process framework comprises (i) the extraction of the event data from source IS such as construction management systems, ERPs, and SQL databases. This data can be manually extracted in the form of comma-separated values (CSV) files. However, as this data is stored in siloed systems, the manual extraction is time-consuming and resource intensive. Thus, more automated extraction methods through representational state transfer application programming interface (RESTful API) calls or

Structured Query Language (SQL) as Java Script Object Notation (JSON) or CSV file formats are advised; (ii) the transformation step is two-fold, first, the data resulted from API calls or SQL queries is stored, and second, the extracted raw operational data is automatically transformed into event logs according to either XES or OCEL standard structure through SQL queries or coding pipelines and (iii) the loading step consist in storing the constructed event logs into a data storage service (i.e., cloud-based SQL database) for process mining enablement.

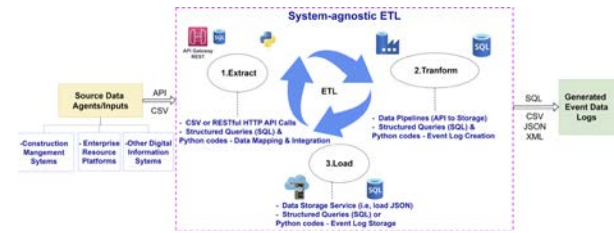


Figure 1. ETL framework for event log generation

3.2 XES Data Structure Analysis

Raw process-related data can be transformed (i.e., step 2 from ETL framework) into standard event logs that comply with XES standard. Figure 2 shows the Unified Modeling Language (UML) class diagram created based on the XES Standard [22]. The main components that should be kept in the event log schema definition include the log, extensions, global attributes, classifiers, traces, events, attributes, and data types[24].

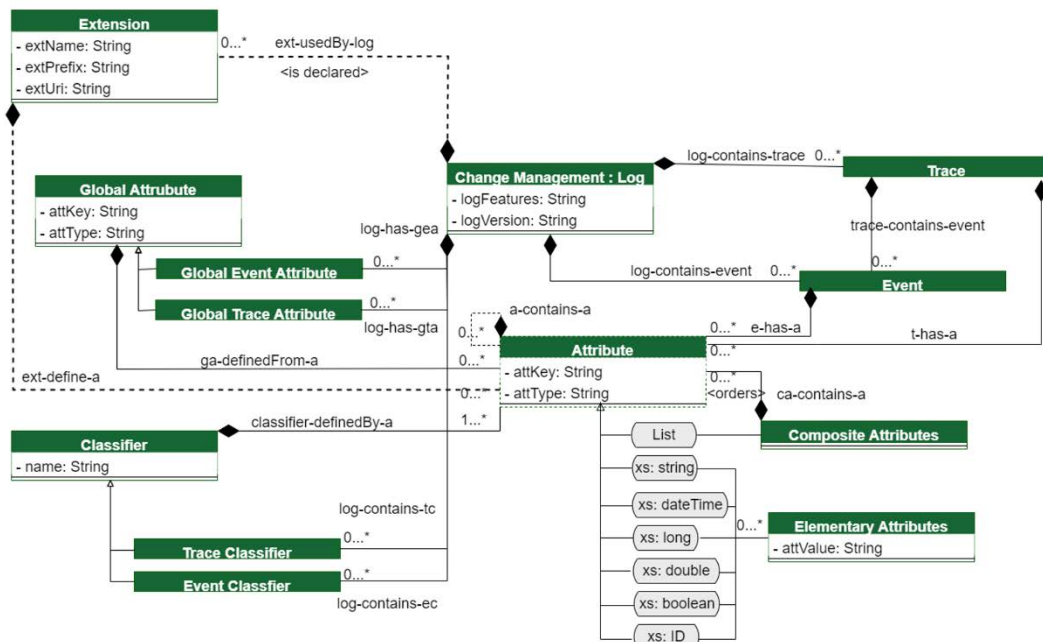


Figure 2. XES UML class diagram [22]

Event Log – An event log refers to a historical record of activities or transactions happening across processes' lifecycle. Events are labeled with execution timestamps and grouped by a unique case identifier (i.e., process instance or trace). An event log can contain one or several event traces, and it should declare/contain any required extensions to semantically describe the process.

Extensions – The extensions provide formal semantics and structure to the event log components by considering/assigning predefined attributes at various levels of the event log (i.e., log, trace, event). There are seven standard extensions in the 2016 version of the XES standard. New extensions can also be defined for domain-specific event log developments.

Global Attributes – This class refers to the declaration of global attributes used when certain predefined process-related information needs to be contained in the log. These global attributes are assigned to every trace and event within the log.

Classifiers – An identity can be assigned to each trace and event in the whole log using classifiers. These classifiers act as labels for traces and events that allow grouping them to compare against one another. An example of an event classifier shall contain two main attributes the event activity/instance name and the lifecycle transition of that activity (i.e., "Create Event – Complete").

Traces – They store the event activities related to a process instance or case. Each trace can contain several event objects. A trace can be seen as a list of activities during the execution of a process, the event activities are often chronologically ordered with the use of timestamps.

Figure 3 provides a list of the standard extensions with a description of their attributes indicating the level at which they shall be declared. The "attribute value" column in Figure 3 also shows actual attributes' values that can be included within each of the XES extension for a COMP, while the "Source Entity" column specifies the SQL table names from which the key-value pairs for each attribute can be extracted.

3.3 OCEL Data Structure Analysis

Raw digital footprints from business process operations can also be transformed (i.e., step 2 from ETL framework) into multidimensional event logs as per OCEL standard. The initial version of the Unified Modeling Language (UML) class diagram based on OCEL 1.0 event log schema definition. This schema comprises the log, global components, objects, events, elements, and data types [23].

OCEL Log – Similar to XES Standard, but without the trace and extensions class definitions.

Global Components – They are different from the XES global attributes because these global attributes are not directly assigned to events, but the class element in the log. Even when in the case that there are no global attributes defined, these classes should be defined with the "value=INVALID" as it is required for the log composition. In other words, these classes act as higher-level containers/placeholders, but the real execution process information comes from Objects, Events, and Elements/Attributes.

Objects – This is the most important component of Object Centric Event Logs. Instead of traces for single case notions. OCEL is composed of a list of object types that can be seen/defined as multi-case event logs to discover a multidimensional process model. OCEL event logs can be flattened to XES logs via the objects.

Events – Similar to XES, but they are related to objects rather than to traces.

Elements/Attributes – Similar to XES, but related to objects and events (i.e., not to traces).

Aiming to improve and simplify OCEL 1.0, OCEL 2.0 standard has been recently released to facilitate the schema definition of multidimensional event logs in the form of a Common Data Model (OCMD). Figure 4 shows its UML class diagram. The main OCEL 2.0 components are the log, events, objects, event types, object types, object-to-object relationships, event-to-object relationships and their related attribute-value pairs [25]. OCEL 2.0 data structure also sets aside the global classes previously considered in OCEL 1.0.

Once the event logs have been generated according to either XES or OCEL standards following the proposed system agnostic ETL framework, process mining capabilities can be leveraged and implemented for key processes in the construction domain.

Extension	Attribute Key	Attribute Type	Attribute Level	Description	Source Entity	Attribute Value
Concept	name	string	log, trace, event	process name; process instance; & event activities	Change_Events	Change Management Process; Change Orders; Create Event
	instance	string	event	Identifier of the executed event activity instance	Change_Events	Create Event-Created
	model	string	log	The adopted transactional model for lifecycle transitions of the events in the log	XES Standard	Standard
Lifecycle	transition	string	event	Lifecycle transition for each event in the log	Applicable to each executed activity/event	Complete (i.e., from InProgress to Closed)
	state	string	event	Lifecycle state for each event in the log	Applicable to each executed activity/event	Completed
	resource	string	event	The resource name or identifier who performed the event activity	Workers - "Workid"	Worker Id
Organizational	role	string	event	The role of the resource	ProjectRoles - "RoleLabel"	PM Approver
	group	string	event	The organizational group where the worker belongs	Workers - "UnionCode"	WorkerUnion
	timestamp	datetime	event	The date and time at which the event occurred	Change_Events	2017-01-09T09:18:00.000+01:00
Time	timestamp	datetime	event	The date and time at which the event occurred	Change_Events	2017-01-09T09:18:00.000+01:00
Semantic	model reference	string	log, trace, event, meta	It refers to model classes of a certain ontology	No currently available	Administrative Processes
ID	id	id	log, trace, event, meta	Unique identifier for the element	Change_Events	ChangeOrdersId
	total	double	trace, event	The total cost of a trace or an event	Change_Orders	\$20,000
	currency	string	trace, event	The currency of all incurred costs	Projects	CAN
Cost	drivers	list	trace, event	List of Cost Drivers	Change_Events	ChangeEventReason - "Description"
	amount	double	meta	The value amount for a cost driver	Change_Events	\$500 for Design Change - "Amount"
	driver	string	meta	Identifier of cost driver	Change_Events	ChangeEventId - "Change Id"
	type	string	log, trace, event	Type of cost	Projects - "BillingType"	Fixed

Figure 3. Sample Standard XES Extensions

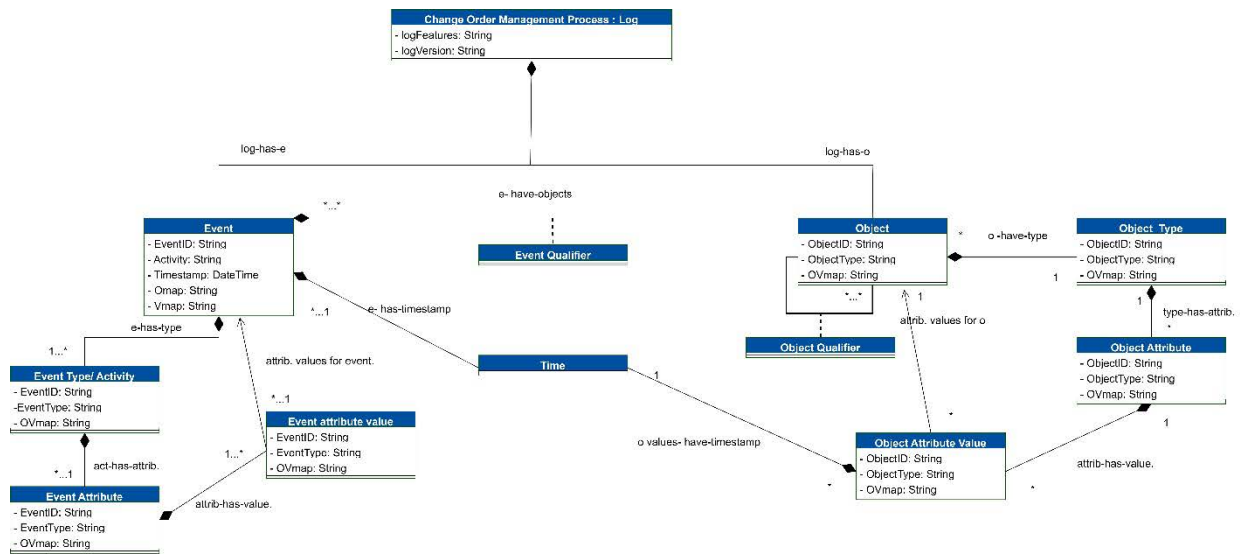


Figure 4. OCEL 2.0 UML class diagram [25]

4 Event Log Generation Process

The functional demonstration of the event log generation process mainly adopts the perspective from a General Contractor (GC) during the construction phase of a commercial office high-rise building project whose prime contracts (i.e., between owner and general contractor), budget, commitment contracts (i.e., between the general contractor and subcontractor), construction changes including Time and Materials (T&M) tracking, and their corresponding invoices are being managed on Procore construction management platform.

Data Extraction - The event log generation from digital footprints is not a trivial task, in fact it can take up to 80% of the process implementation efforts[15]. The first step consists of data extraction and collection from siloed data sources.

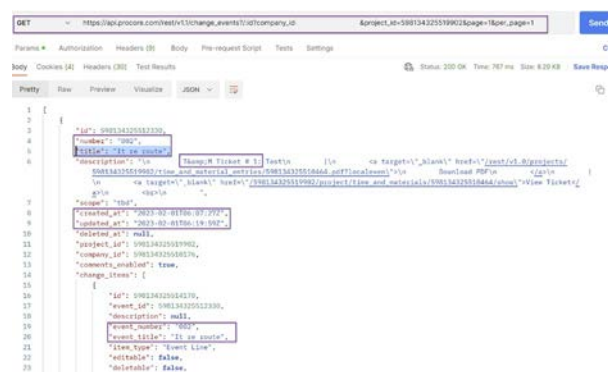


Figure 5. Data extraction Procore API - Postman [26]

In this case, the event activity timestamps, the event activity names, and their corresponding case identifiers resided on Procore, thus we extracted this information by

developing a connection data application and calling the Procore API endpoints using the RESTful framework as shown in Figure 5.

4.1 XES Event Log Generation

Data Transformation to XES - The transformation step depicted in Figure 1 consisted in changing the extracted raw process event data into XES event log format through SQL queries as depicted in Figure 6.

Figure 6. COMP Data Transformation to Event Log

A tabular excerpt of the generated event log for 2 executed traces/instances of the construction change order management process is shown in Figure 7.

CaseID	Activity Name	Timestamp	ActorRole
1	Create Project	2023-01-02 0:59	PM - GC
1	Create Contract	2023-01-02 1:15	PM - GC
1	Open Field Correspondence Extra Work	2023-01-02 1:05	Foreman - GC
1	Open Field Correspondence Extra Work	2023-01-20 17:34	Foreman - GC
1	Sign Correspondence Authorization to Proceed	2023-01-02 10:25	CM or Owner
1	Create Change Event from Correspondence	2023-01-02 11:07	PM - GC
1	Create T&M - In Progress	2023-01-02 11:08	Foreman - GC
1	Fill in T&M - Ready for Review	2023-01-02 11:12	Foreman - GC
1	GC Approval/Sign T&M	2023-02-02 11:08	Foreman - GC
1	Request Customer/Client Signature T&M	2023-02-02 11:09	Foreman - GC
1	Client Approval/Sign T&M - Field Verified	2023-02-06 9:09	CM or Owner
2	Create T&M - In Progress	2023-03-21 0:53	Foreman - GC
2	Fill in T&M - Ready for Review	2023-03-23 16:50	Foreman - GC
1	Add T&M to an existing Change Event	2023-02-07 17:12	PM - GC
1	Create Potential Change Order	2023-02-08 9:15	PM - GC
1	Create Change Order	2023-02-08 17:10	PM - GC
1	Generate Invoice	2023-02-09 10:30	Financial Manager
1	Pay Invoice	2023-02-09 9:30	Financial Manager
2	GC Approval/Sign T&M	2023-03-21 1:37	Foreman - GC
2	Client Approval/Sign T&M - Field Verified	2023-03-21 1:37	CM or Owner
2	Create Change Event from T&M	2023-03-21 1:44	PM - GC

Figure 7. COMP Tabular XES Event Log

Criteria	XES	OCEL 1.0
<i>Perspective</i>	To analyze a single perspective at a time (i.e., control flow)	Multi-perspective (i.e., event control flow & actors)
<i>Process Mining Approach</i>	Traditional Process Mining (Current state-of-the-art)	Multidimensional Process Mining (Industries shifting towards Multi-level Process Mining)
<i>Quality Aspects Suitable for Schema Basis</i>	Less transparency over events and precedence relationships Workflow executions Process Instance-based	Preserves Entities/Objects, Events, and relations E2E Process Executions Entity/object-centric based
<i>Main Schema Components</i>	Log, Traces, Events	Logs, Objects/Entities, Events
<i>Case Notion</i>	Single Case Notion (i.e., Change Events)	Multiple Case Notion (i.e., Change Events & Orders)
<i>Conversion</i>	XES to OCEL (Requires domain knowledge of the E2E Process)	OCEL to XES (One should deal with convergence and divergence problems)
<i>Conformance</i>	Enables Conformance Checking	Enables Conformance Checking
<i>Transparency</i>	Promotes Transparency	Enables Higher E2E Transparency
<i>Knowledge</i>	Knowledge of process instance (i.e., less complex networks)	Higher E2E Knowledge Capture (i.e., real complex networks)
<i>Management</i>	Less focus on interoperability	Higher focus on Interoperability

Figure 10. Comparison XES vs OCEL standards.

6 Conclusions and Future Work

Current IT data architectures of several construction companies lack process awareness in the sense that they have poor visibility and transparency over the performance behavior of actual process executions. To address this limitation, this study proposed an ETL process-oriented framework for event log generation which extracts raw construction data footprints from IS and transform it into process event logs. The reconstructed event logs enable process mining capabilities to model, monitor and manage the actual performance of their construction business operations through automated means and methods. Besides, this study contributed to the body of knowledge by shedding light on the XES and OCEL data standard schemas understanding and their application specific for the construction domain.

This research facilitates the event log generation process with proper syntax and semantics necessary for a successful process mining implementation that can support construction stakeholders during process audits and to streamline business operations that result in significant productivity gains. Despite the stated benefits, this study has the following limitations: (i) although the proposed framework aims to be system-agnostic, this study is mainly focused on Procore as the data source for event log generation, future studies should further investigate the framework's validity when considering other data sources/IS; (ii) the scope of this study is limited to the event log generation phase, thus future work should elaborate on the process analytics phase through process mining algorithms; and (iii) the functional application is limited to a single project, so future analysis should include a benchmarking comparison of process performance behavior between two or more similar projects in different location and

managed by the same or different companies.

In this context, parallel studies are underway to demonstrate process mining automating and improving capabilities in construction [32]. It is also important to note that it is necessary to use formal process modeling notations such as Business Process Modeling and Notation (BPMN) to represent the automatically discovered actual process executions, which can also enrich companies' Standard Operating Procedures (SOPs). Last but not least, future work in the construction domain should be focused on extending current descriptive and diagnostic process mining capabilities into predictive and prescriptive analytics by leveraging novel simulation and AI technologies[15]

References

- [1] J. Mollins and T. Taskin, "Digitalization: Productivity," Bank of Canada, 1914-0568, December 2023 2023. [Online]. Available: <https://doi.org/10.34989/sdp-2023-17>
- [2] A. Charles Atkins *et al.*, "Rekindling US productivity for a new era," McKinsey Global Institute, December 2023 2023. [Online]. Available: <https://www.mckinsey.com/mgi/our-research/rekindling-us-productivity-for-a-new-era>
- [3] McKinsey Global Institute, "DIGITAL AMERICA: A TALE OF THE HAVES AND HAVE-MORES," 2015. [Online]. Available: www.mckinsey.com/mgi
- [4] H. Liu, "Economic performance associated with digitalization in Canada over the past two decades Economic and Social Reports," Statistics Canada, December 2023 2021. [Online]. Available: <https://doi.org/10.25318/36280001202100200001-eng>
- [5] F. Calvino, C. Criscuolo, L. Marcolin, and M. Squicciarini, "A TAXONOMY OF DIGITAL INTENSIVE SECTORS," OECD, December 2023 2017. [Online]. Available: [https://one.oecd.org/document/DSTI/CHIE/WPI/A\(2017\)2/en/pdf](https://one.oecd.org/document/DSTI/CHIE/WPI/A(2017)2/en/pdf)
- [6] Grand View Research. "Process Mining Software Market Size Report, 2021-2028." <https://www.grandviewresearch.com/industry-analysis/process-mining-software-market-report> (accessed December, 2023).
- [7] W. M. P. Van der Aalst, *Process Mining Data Science in Action*. Berlin, Heidelberg: Springer (in en), 2016.
- [8] W. M. P. van der Aalst, "Process Mining: A 360 Degree Overview," in *Lecture Notes in Business Information Processing*, vol. 448: Springer

- Science and Business Media Deutschland GmbH, 2022, pp. 3-34.
- [9] "Cloud Computing Services | Microsoft Azure." <https://azure.microsoft.com/en-us> (accessed February 20, 2023).
- [10] SAP. "What is SAP? | Definition and Meaning." @SAP. <https://www.sap.com/about/company/what-is-sap.html> (accessed September 26, 2022).
- [11] Procore Technologies. "What is Procore? | Construction Management Software." <https://www.procore.com/en-ca/what-is-procore> (accessed September 26, 2022).
- [12] Rhumbix. "RMBX." <https://www.rhumbix.com/> (accessed September 26, 2022).
- [13] ORACLE. "ORACLE NETSUITE " Oracle. <https://www.netsuite.com/portal/products/erp.shtml> (accessed September 26, 2022).
- [14] S. Kouhestani and M. Nik-Bakht, "IFC-based process mining for design authoring," (in en), *Automation in Construction*, vol. 112, p. 103069, 2020/04//2020, [Online]. Available: <https://doi.org/10.1016/j.autcon.2019.103069>
- [15] W. M. P. Van der Aalst, J. Carmona, J. Mylopoulos, S. Ram, M. Rosemann, and C. Szyperski, *Process Mining Handbook*. Springer, 2022.
- [16] A. M. Eldeeb, M. A. M. Farag, and L. M. Abd El-hafez, "Using BIM as a lean management tool in construction processes – A case study: Using BIM as a lean management tool," *Ain Shams Engineering Journal*, vol. 13, no. 2, 2022/3// 2022, doi: 10.1016/j.asej.2021.07.009.
- [17] W. Ibbs, "Update on Quantitative Analysis of Change and Loss of Productivity," *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction*, vol. 13, no. 1, 2021/2// 2021, doi: 10.1061/(asce)la.1943-4170.0000447.
- [18] *Standard for eXtensible Event Stream (XES) for Achieving Interoperability in Event Logs and Event Streams*, 9781504424219, C. I. S. IEEE, 2016. [Online]. Available: <https://ieeexplore.ieee.org/servlet/opac?punumber=7740856>
- [19] *OCEL: A Standard for Object-Centric Event Logs*, 9783030850814, A. F. Ghahfarokhi, G. Park, A. Berti, and W. M. P. van der Aalst, 2021. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-85082-1_16
- [20] Eindhoven University of Technology. "Analytics for Information Systems - AIS." <https://www.win.tue.nl/ais/doku.php> (accessed September 29, 2022).
- [21] IEEE. "IEEE - The world's largest technical professional organization dedicated to advancing technology for the benefit of humanity." <https://www.ieee.org/> (accessed September 28, 2022).
- [22] IEEE, *IEEE Standard for eXtensible Event Stream for Achieving Interoperability in Event Logs and Event Streams*. 2023.
- [23] A. F. Ghahfarokhi, G. Park, A. Berti, and W. M. P. van der Aalst, *OCEL: A Standard for Object-Centric Event Logs* (Communications in Computer and Information Science). Springer International Publishing, 2021, pp. 169-175.
- [24] J. C. A. M. Buijs, "Mapping Data Sources to XES in a Generic Way," Master, Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, 2010. [Online]. Available: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Mapping+Data+Sources+to+XES+in+a+Generic+Way#0>
- [25] A. Berti *et al.*, "OCEL (Object-Centric Event Log) 2.0 Specification," 2023. [Online]. Available: <https://www.ocel-standard.org/2.0/ocel20>
- [26] Postman. <https://www.postman.com/>. (accessed March 14, 2024)
- [27] ProM. "ProM." http://www.promtools.org/doku.php?id=prom6_9 (accessed September 21, 2023).
- [28] J. C. A. M. Buijs, "Mapping Data Sources to XES in a Generic Way," in *Chelsea*, ed, 2010, pp. 123-123.
- [29] C. Günther and E. Verbeek, "OpenXES Developer Guide," Eindhoven University of Technology, The Netherlands, 2012. [Online]. Available: https://www.xes-standard.org/_media/openxes/openxesdeveloper-guide-1.8.pdf
- [30] FIT. "PM4py - Process Mining for Python." <https://pm4py.fit.fraunhofer.de/> (accessed September 14, 2023).
- [31] A. Berti, M. Montali, and W. M. P. van der Aalst, "Advancements and Challenges in Object-Centric Process Mining: A Systematic Literature Review," 2023/11// 2023. [Online]. Available: <http://arxiv.org/abs/2311.08795>.
- [32] A. J. Matinez Lagunas and M. Nik-Bakht, "Enabling Process Mining in the Construction Industry: An Event Log Schema for Change Management Process," in *CSCE/CRC International Construction Specialty Conference. (Manuscript accepted for publication)*, 2023, pp. 1-13.

Comparative Analysis of Cognitive Agreement between Human Analysts and Generative AI in Construction Safety Risk Assessment

Unmesa Ray¹, Cristian Arteaga² and JeeWoong Park³

^{1,2,3} Department of Civil and Environmental Engineering and Construction, University of Nevada, Las Vegas, USA

rayu1@unlv.nevada.edu, arteagas@unlv.nevada.edu, jee.park@unlv.edu

Abstract -

The construction industry struggles with safety risk assessment complexities due to evolving work environments, diverse labor forces, time constraints, regulatory intricacies, and inconsistent practices. While previous studies have highlighted the potential of Artificial Intelligence (AI) in automating processes and enhancing safety assessment, a gap exists in the convergence between human analysts and language AI models. Therefore, this study assesses the alignment in identifying risk factors by human analysts and a Language Model (LM) in Occupational Safety and Health Administration (OSHA) accident reports. Furthermore, it: 1) categorizes error types, 2) establishes an acceptance threshold for LM-generated responses, and 3) evaluates inter-rater reliability in construction accident content analysis. The test results reveal significant convergence between human and machine responses and identify potential hallucination effects in generative AI, thus paving the way for improved safety risk assessments within the construction industry.

Keywords –

Construction industry, safety risk assessment, Artificial Intelligence (AI), Occupational Safety and Health Administration (OSHA), Language Model (LM), inter-rater reliability, generative AI

1 Introduction

1.1 Motivation

The construction industry presents a significant problem concerning workplace accidents, primarily due to its inherently high-risk nature. Despite considerable

efforts over the past several years, the safety aspects of the construction industry have not witnessed enhancements at a level comparable to that observed in other industries. As a result, the industry still suffers with a substantial number of accidents [1]. In 2021, the construction industry accounted for a staggering 21% of all occupational fatalities in the United States, as reported by the Bureau of Labor Statistics [2]. 65.5% of these construction-related deaths are attributed to the "Focus Four" incidents, including falls, which accounted for 35% of the fatalities, struck-by incidents at 17%, electrocutions at 7.6%, and caught incidents at 5.8% [3].

Construction safety issues have seen a significant decline in injuries and fatalities after the implementation of OSHA's rules and standards [4]. All employers are required to notify OSHA upon death, injury, or hospitalization. The accident report form is available on OSHA's website. Employers may also contact a designated phone number to report accidents. The form contains structured data such as accident date, company name, and age/sex of injured person, as well as a narrative describing the accident. This accident narrative provides valuable information to better understand the context and sequence of events that led to the incident.

However, investigation summary, due to its unstructured nature, and possible variations of descriptive context poses several challenges in analysis and interpretation. While quantitative data is generally straightforward to handle, the unstructured and descriptive nature of narratives in accident reports presents challenges for systematic analysis. Figure 1 shows a section of OSHA Form 301 that requires investigation summary.

14)* **What was the employee doing just before the incident occurred?** Describe the activity, as well as the tools, equipment, or material the employee was using. Be specific. Examples: "climbing a ladder while carrying roofing materials"; "spraying chlorine from hand sprayer"; "daily computer key-entry."

15)* **What Happened? Tell us how the injury occurred.** Examples: "When ladder slipped on wet floor, worker fell 20 feet"; "Worker was sprayed with chlorine when gasket broke during replacement"; "Worker developed soreness in wrist over time."

Figure 1. Segment for investigation summary details in OSHA's Form 301

Consequently, the manual analysis of these narratives is a time-consuming process, which ultimately translates into substantial labor. A previous study [5] introduced framework that facilitated the extraction of standard risk factors and outcome variables from OSHA accident reports, providing a comprehensive approach to addressing the limitations of accident data analysis. However, the analysis of such databases demands highly skilled labor, making it an expensive and time-consuming administrative method [6]. Likewise, traditional content analysis has historically relied on human analysts, incorporating inter-rater reliability measures into the coding system. On the other hand, a class of neural networks known as Transformers [7] has significantly streamlined language classification tasks in recent years. However, before fully embracing their potential, it is important to assess their alignment with human coders, particularly in the realm of content analysis.

1.2 Related Work

Researchers have adopted numerous approaches to effectively handle and covert database into useful information [8]. However, one notable drawback of these methods is their lack of reliance on empirical data, coupled with a limited scope of application [9], [10]. In response to these limitations, a unified attribute-based framework has been proposed [need a citation]. This framework enables the extraction of standard risk factors and outcome variables from naturally occurring accident reports, providing a comprehensive approach to

addressing the shortcomings in accident data analysis [5]. Due to the high dimension of the injury report feature space and the diversity of construction situations, the available training data are naturally sparse. Therefore, manually analyzing tens of thousands of injury reports would have been required to put together a satisfactory training database and achieve efficient machine learning [11].

Previously, a study [12] has proposed an analytical method incorporating text mining and interpretable machine learning to discern factors influencing injury severity levels within traffic crash narratives. However, it suffered from a limitation in the text analysis since it was based on word frequency rather than meaningful semantics from the narrative. On the contrary, LMs have gained significant attention in recent years due to their ability to generate human-like text and perform a wide range of language-based tasks. In the construction industry, LMs have the potential to improve efficiency, accuracy, and communication in several different ways [13]. Regarding language representation models, different approaches have been developed in the past decade [14]. One major model is the autoregressive language model Generative Pre-trained Transformer (GPT). Developed by OpenAI, GPT is trained on a large dataset of text and can generate human-like text. GPT 3.5 has been fine-tuned for information retention during the conversation, making it suitable for activities that are solely reserved for human interaction [15]. A recent study used GPT to generate a construction schedule for a simple construction project and the outcome shows that it can generate a coherent schedule that follows a logical approach to fulfill requirements of the scope indicated [13]. Another study integrated GPT as an intervention to aid hazard recognition efforts in the curriculum of students pursuing a career in the construction industry. The results suggest that GPT can be leveraged to improve hazard recognition levels [15]. To the best of our knowledge, this study marks the first attempt to assess the alignment of Transformer models with human analysts in comprehensive understanding of accident data, offering a promising avenue for future research and applications in the field of accident analysis and prevention.

1.3 Research Objective and Scope

This study's primary objective is to assess the alignment in identifying contributing risk factors in accidents as perceived by human analysts versus an LM. By undertaking this comparative analysis, the research objectives are threefold: (1) defining a threshold for accepting LM-generated responses (2) systematically categorizing error types into Type I and Type II, and (3) evaluating the inter-rater reliability between human

analysts and the LM in the context of content analysis within construction accidents. This study implements the proposed approach by analyzing 150 accident investigation summaries related to "Fall from Heights." These summaries are sourced from the official website of OSHA.

2 Methodology

Figure 2 is an illustration of the steps adopted in this study, which include a data collection process, manual analysis of narratives, the setup of LM for generating machine responses, and the integration of a suitable statistic within the comparison of LM's responses against the manually identified contributing factors.

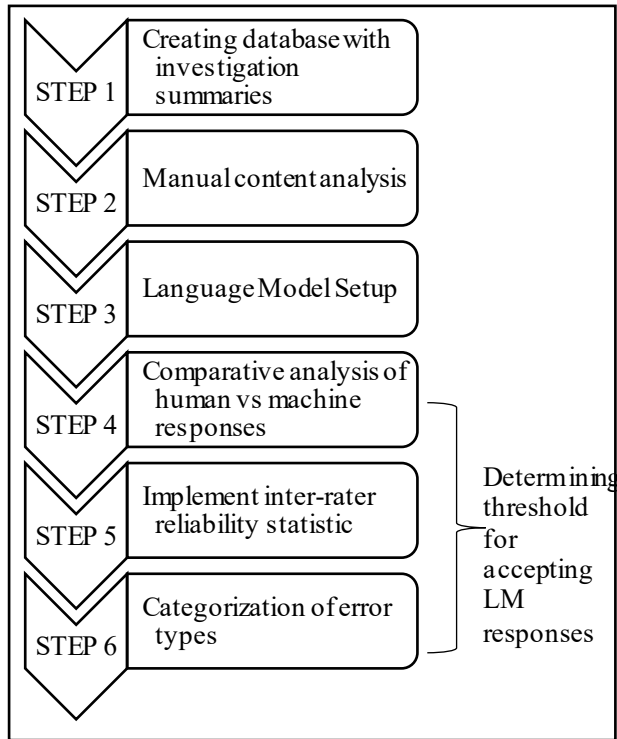


Figure 2. Proposed research methodology

2.1 Creating Database with Investigation Summaries

Initiating the proposed framework involved retrieving accident report narratives, which constitute unstructured data available on OSHA's official website, followed by the utilization of simple and easily accessible Microsoft Office tools, such as Excel, to construct a database cataloging accident identification code (IDs) alongside their respective narratives. This

ensures establishing an organized repository for subsequent content analysis.

2.2 Manual Content Analysis

Content analysis, critical to this research, involves identifying risk factors contributing to accidents within narratives. In the context of this study, a risk factor can be defined as a specific element, condition, or circumstance identified in accident reports that contributes to the occurrence or severity of accidents. These factors encompass various aspects such as environmental conditions, human behaviors, equipment usage, or any other variable that plays a role in the initiation or escalation of accidents. Recognizing these factors necessitates pattern recognition through manual labor—interpreting and inferring accident causes. We enlisted and trained a dedicated researcher for this analysis. The identified risk factors by the human analyst establish the ground truth for determining the acceptance threshold for LM-generated responses.

2.3 LM Setup and Analysis

In this phase, we integrated GPT with the collected data. The API key obtained from OpenAI facilitated interfacing Python, thereby, aiding to input prompts into the LM. The accuracy of information extraction depends on the prompt. Hence, a meticulously devised prompt must include desired outcomes in structured format, while excluding specific redundant details. We also consider fine-tuning the prompt for enhanced information retrieval.

2.4 Comparative Analysis of Responses – Human versus Machine

Upon receiving responses from the LM, we conducted a comparative analysis between contributing risk factors identified by the human analyst and the machine. A manual review of all accident reports was performed, comparing factors for each narrative. The researcher evaluating the comparative analysis considers the LM responses valid only when the LM's response aligns with the true cause leading to the accident. We treated the LM as an independent coder akin to the human analyst, and percent agreement was used to measure inter-coder reliability. Percent agreement simply represents number of agreements over total number of measures [16]. The formula for percent agreement is given below where A represents the number of agreements between the machine and the human analyst, and N is the total number of factors identified by both.

$$PA_o = A/N \quad (1)$$

2.5 Error Type Formulation and Implementing Inter-rater Reliability Statistic

In the concluding phase, we categorize the types of errors that may occur. When comparing human versus machine responses, three scenarios are possible. The first involves complete agreement, where both human and machine responses are identical, resulting in 100% agreement. However, in other cases, Table 1 shows the errors that may occur and can be classified as follows:

- **Type I Errors:** Occur when human responses include certain risk factors that are absent in the LM-generated responses.
- **Type II Errors:** Occur when the LM-generated responses include certain factors that the human analyst does not include.

Table 1. Formulation of Type I and Type II errors

Machine Coding	Manual Coding		
	Identified		Not Identified
	Identified	Pass	Type II Error
	Not Identified	Type I Error	Pass

3 Case Study

The proposed methodology leveraged OSHA accident reports as the primary dataset. Four researchers were hired and trained for data collection and manual analysis, working 6 hours a day for four weeks. This entails a thorough manual review process, where the researchers sift through accident reports to identify 150 cases resulting in fatal injuries specifically caused by falls from heights. Additionally, each researcher agrees on the contributing risk factors. In instances where discrepancies arise, the researchers conducted re-reviews to ensure accuracy until a consensus was reached. Discussions are held to establish consistent terminology. This phase is crucial as the manual coding process aims to ensure consistency and establish the ground truth for the subsequent comparative analysis with machine-generated responses. Upon examination, roof work emerges as a prevalent cause. After curating the relevant narratives, a database was constructed and integrated into Python for a comprehensive analysis using the LM. Subsequently, prompts were tailored to meet the specific requirements of this research. The fine-tuning phase is instrumental in determining the final prompt for utilization. The process of formulating the prompt involves a straightforward brute force method to refine the prompt after each response, aligning the output with

our desired outcome. These prompts are generalizable across domains, contingent on the integration of the brute force method into the process. This ensures that users can interact with the LM to extract meaningful information without requiring a background in NLP. The following is the evolution of prompts with the corresponding LM responses, and the encountered challenges throughout this iterative process until the final prompt was decided.

Prompt 1: Please give the risk factors associated with construction safety in the following narrative.

LM response for prompt 1: In the given narrative, various risk attributes associated with construction safety are evident - Working at Height: The employee was on the roof, which is a high-risk location in construction due to the potential for falls and other hazards associated with elevated work.

... seven more attributes with added explanations.

Limitations of prompt 1: Overly detailed factors, redundant OSHA data, and lengthy hazard mitigation paragraphs in tested narratives.

Prompt 2: Please give the risk factors associated with construction safety in the following narrative with no additional sentences and headers.

LM response for prompt 2: Working at Height: The employee was on the roof, which is a high-risk location in construction due to the potential for falls and other hazards associated with elevated work.

... seven more attributes with added explanations.

Limitations of prompt 2: Despite the request to remove descriptive attribute headers, they continue to be present in the responses.

Prompt 3: Please give the risk factors associated with construction safety in the following narrative excluding date, time, company name, age with no additional sentences and remove any headers the attributes begin with.

LM response for prompt 3:

Working at height

Ladder use

Dizziness

Brief rest

Returning to work

Fall from roof

Severe head, neck, and back trauma

Since implementing the final prompt involves the generation of contributing risk factors without a detailed description, this prompt is applied to the entire dataset. In the concluding phase, a comprehensive comparative analysis is conducted between LM-generated results and the manually identified contributing factors for each narrative. This qualitative analysis reveals distinct trends as follows:

- Machine responses precisely align with human-identified risk factors.
- The machine identifies all essential accident-contributing factors and introduces additional meaningful attributes, though not necessarily contributing to the accident.
- LM responses fall short in identifying one or more key accident-contributing factors.

4 Results and Discussion

In Figure 3, 104 cases out of the 150 datasets demonstrated an exact match between human-analyzed factors and machine responses. Only 1 case exhibited a Type I error, while 45 cases showed Type II errors. Further, we computed percent agreement for both error types, resulting in 66.67% for Type I errors and an average of 28% for Type II errors.

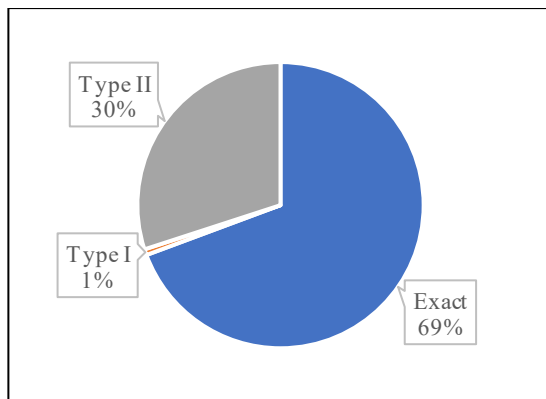


Figure 3. Distribution of alignment between manual versus machine check across the dataset

Reliable coding is achieved when the percent agreement surpasses 70% [17]. Upon examination of the statistics, it becomes evident that human and machine alignment is robust for most cases in our dataset, boasting 100% agreement. However, for error types, the percentage agreement falls below the 70% threshold. Consequently, based on the proposed methodology, it is reasonable to conclude that the inter-rater reliability threshold for human and LM is achieved at 70%.

For Type II errors, a deeper investigation into factors missing in human coding reveals that 86.67% of the

omissions pertain to a broader description of the narrative, which human analysts do not consider as important risk factors contributing to the accident. However, the remaining 13.33% encompasses factors that are not present in the narrative. This suggests a tendency for the LM to introduce additional factors, indicative of a phenomenon referred to as the "hallucination effect" of generative AI.

Example of Type II error that showcases hallucination effects of generative AI:

Investigation Summary: On December 4, 2006, Employee #1 was working for a general contractor that specialized in the construction of single-family houses. He was on a scaffold, when he fell approximately 30 feet to the ground and was killed.

LM Response:

- Employee conducting work at a height on a scaffold.
- Scaffold did not provide sufficient fall protection measures.
- Lack of proper safety equipment or personal protective gear.
- Inadequate training or lack of knowledge on scaffold safety.
- Failure to comply with safety regulations or standards.
- Absence of safety inspection or assessment prior to work.
- Potential negligence in ensuring scaffold stability.
- Insufficient communication or coordination regarding safety precautions.

For the above example, within LM responses, points 2 to 8 are absent in the original narrative.

Example of Type II error that showcases descriptive factors that do not directly contribute to the accident:

Investigation Summary: At approximately 9:00 a.m. on September 13, 2006, an employee was removing concrete forms along a floor opening, on the third floor of a building under construction. While working around the unguarded floor opening, the employee fell approximately 18 ft through the opening and landed on the first floor. The employee was not wearing fall protection. The employee was transported to a local trauma center, where he was pronounced dead on arrival.

LM Response:

- Employee removing concrete forms.
- Unguarded floor opening.
- Employee fell approximately 18 ft.
- Employee landed on the first floor.
- Employee not wearing fall protection.

- Employee transported to a local trauma center.
- Employee pronounced dead on arrival.

For the above example within LM responses, points 3,4,6, and 7 are details of the accident that do not directly contribute to it.

5 Conclusion

The study comprehensively explores the alignment between human analysts and a LM, specifically a Transformer model, in identifying risk factors contributing to construction accidents. For the proposed methodology, we established an acceptance threshold for LM-generated responses, while systematically categorizing error types into Type I and Type II discrepancies for 150 OSHA accident investigation summaries. The comparative analysis between manual and machine responses revealed that 69% of the cases demonstrated an exact match between human-analyzed factors and LM responses with only 1% showing Type I errors and 30% exhibiting Type II errors. Additionally, the calculated percent agreement for Type I and Type II errors were 66.67% and 28%, respectively. While these percentages fall slightly below the 70% threshold suggested for reliable coding, the study underscores the considerable potential of collaborative human and LM analysis, particularly in categorizing error types and establishing thresholds. Furthermore, the hallucination effect underscores the need for a nuanced interpretation of the results generated by the LM, especially in situations where extraneous factors are introduced. Potential ways to mitigate this effect could be refining prompt designs, adjusting model parameters, or incorporating additional validation steps in the AI-assisted risk assessment process. Overall, this study showcases the practical significance of collaborative human and AI analysis in construction industry risk assessments. The findings lay the groundwork for refining methodologies that optimize both human judgment and machine capabilities, offering valuable insights for industries adopting AI in risk assessment. Theoretical contributions stem from the study's categorization of Type I and Type II errors, providing insights into human-AI collaboration dynamics across industries. Future research endeavors will concentrate on enhancing collaborative models, broadening the study's scope, and delving into prompt engineering techniques to refine prompt formulation, integral to advancing this collaborative process.

References

- [1] M. M. Zaira and B. H. W. Hadikusumo, "Structural equation model of integrated safety intervention practices affecting the safety behaviour of workers in the construction industry," 2017, doi: 10.1016/j.ssci.2017.06.007.
- [2] "Fatal occupational injuries by industry and event or exposure, all United States, 2020 : U.S. Bureau of Labor Statistics." Accessed: Nov. 04, 2023. [Online]. Available: <https://www.bls.gov/iif/fatal-injuries-tables/fatal-occupational-injuries-table-a-1-2020.htm>
- [3] "CPWR | A world leader in construction safety and health research and training." Accessed: Nov. 21, 2023. [Online]. Available: <https://www.cpw.com/>
- [4] A. Chokor, H. Naganathan, W. K. Chong and M. El Asmar, "Analyzing Arizona OSHA Injury Reports Using Unsupervised Machine Learning," *Procedia Eng*, vol. 145, pp. 1588–1593, Jan. 2016, doi: 10.1016/J.PROENG.2016.04.200.
- [5] B. Esmaili and M. Hallowell, "Attribute-Based Risk Model for Measuring Safety Risk of Struck-By Accidents," *Construction Research Congress 2012: Construction Challenges in a Flat World, Proceedings of the 2012 Construction Research Congress*, pp. 289–298, 2012, doi: 10.1061/9780784412329.030.
- [6] M. Zeynalian, B. Trigunarsyah, and H. R. Ronagh, "Modification of Advanced Programmatic Risk Analysis and Management Model for the Whole Project Life Cycle's Risks," *J Constr Eng Manag*, vol. 139, no. 1, pp. 51–59, Apr. 2012, doi: 10.1061/(ASCE)CO.1943-7862.0000571.
- [7] A. Vaswani *et al.*, "Attention Is All You Need," 2017.
- [8] M. R. Hallowell and J. A. Gambatese, "Activity-Based Safety Risk Quantification for Concrete Formwork Construction," *J Constr Eng Manag*, vol. 135, no. 10, pp. 990–998, Apr. 2009, doi: 10.1061/(ASCE)CO.1943-7862.0000071.
- [9] B. Esmaili and M. Hallowell, "Attribute-Based Risk Model for Measuring Safety Risk of Struck-By Accidents," *Construction Research Congress 2012: Construction Challenges in a Flat World, Proceedings of the 2012 Construction Research Congress*, pp. 289–298, 2012, doi: 10.1061/9780784412329.030.
- [10] A. J. P. Tixier, M. R. Hallowell, B. Rajagopalan, and D. Bowman, "Automated content analysis for construction safety: A natural language

- processing system to extract precursors and outcomes from unstructured injury reports,” *Autom Constr*, vol. 62, pp. 45–56, Feb. 2016, doi: 10.1016/J.AUTCON.2015.11.001.
- [11] A. J. P. Tixier, M. R. Hallowell, B. Rajagopalan, and D. Bowman, “Application of machine learning to construction injury prediction,” *Autom Constr*, vol. 69, pp. 102–114, Sep. 2016, doi: 10.1016/J.AUTCON.2016.05.016.
 - [12] C. Arteaga, A. Paz, and J. W. Park, “Injury severity on traffic crashes: A text mining with an interpretable machine-learning approach,” *Saf Sci*, vol. 132, Dec. 2020, doi: 10.1016/j.ssci.2020.104988.
 - [13] S. A. Prieto, E. T. Mengiste, and B. García de Soto, “Investigating the Use of ChatGPT for the Scheduling of Construction Projects,” *Buildings*, vol. 13, no. 4, Apr. 2023, doi: 10.3390/buildings13040857.
 - [14] T. Schomacker and M. Tropmann-Frick, “Language representation models: An overview,” *Entropy*, vol. 23, no. 11, Nov. 2021, doi: 10.3390/e23111422.
 - [15] T. B. Brown *et al.*, “Language Models are Few-Shot Learners,” 2020. [Online]. Available: <https://commoncrawl.org/the-data/>
 - [16] Kimberly A. Neuendorf, *The content analysis guidebook*. Thousand Oaks. CA: Sage, 2002.
 - [17] Klaus H. Krippendorff, *Content analysis: An introduction to its methodology*. Thousand Oaks. CA: Sage, 2004.

Developing a Time-Cost & Storage Optimization Model for Construction project

Prasanna Venkatesan R¹, Shivaram kandasamy¹, Anshul Gupta¹, Sugeerthi M S¹

¹School of Civil Engineering, Vellore Institute of Technology, India

prasanna.venkatesan@vit.ac.in, shivaramleo@gmail.com, anshul.gupta2021a@vitstudent.ac.in,
sugeerthi.2023@vitstudent.ac.in

Abstract

Project managers in the construction industry confront challenges in managing inventories and executing projects within the scheduled time frame. In this research, a mathematical model is given for applying the Lagrange method and linear programming to tackle the time-cost and storage-related difficulties that arise in the construction project. The study's objectives are to minimize total cost escalation and ascertain the ideal order quantity for the building project while considering non-negative, start, crash, and floor space limits into account. The ideal order quantity is determined using MS Excel, and LINDO software determines the crashing duration. Program analysis produced a workable solution. The model maintains the analysis's accuracy and the outcomes' correctness. By increasing the project's cost, the linear programming approach predicts when each task will take less time to complete. The EOQ model was proposed to determine the optimal order quantity of building supplies. Adopting this technology can yield instant benefits for any real-time construction job.

Keywords:

Linear Programming Method; Lagrange method; Time-cost; Storage; Construction project; LINDO Software

1 Introduction

Project management is a study that involves the application of skills, tools, and techniques to meet the project requirements. Whenever the project gets delayed and runs behind schedules, the overall indirect cost increases. Usually, project duration can be reduced by crashing activities and assigning more resources. Construction of a project with a normal

duration will ensure specific resources and direct costs. In contrast, the same project is constructed under the crashing method, decreasing project duration and escalating the cost to the allowable percentage. A study developed an optimum solution for time and cost using ant colony optimization without dominating either function. Zhang presented a time-cost optimization problem using the ant colony method. To deal with the multi-objective problem, a modified weight approach was implemented to combine time and cost as a single objective[1]. Mondal proposed an intuitionistic fuzzy geometric programming to solve the deterministic single objective problem[2]. This study was conducted in an apartment consisting of 9 floors located in Bangalore. This paper proposes a mathematical model to solve the time-cost and storage-related problems in the construction project using linear programming and the Lagrange method, respectively[3]. The study aims to optimize the overall cost escalation and determine the construction project's optimum order quantity under start time, crash time, non-negative, and floor space constraints. To determine the crash duration and optimum order quantity, LINDO Software and MS Excel are used, respectively.

2 Background

The time and cost optimization technique decreases the total float available for non-critical activities and decreases the flexibility of the schedule. There is always a need to establish a new method for time and cost such that it can provide optimum time and cost value. The author has attempted nonlinear integer programming using the best solver technique, which can be applied to a real-time project[4]. Whenever there is a trade-off between time and cost, the duration of the project will decrease, and the cost will increase. The results obtained proved the model is significant. It helps the project manager execute

different trade-offs between time and cost[5]. A study proposed an intuitionistic fuzzy geometric programming to solve the deterministic single objective problem. Intuitionistic fuzzy geometric programming can also solve economic order quantity with a deterministic single objective model with floor space constraint. Any variable such as limited production cost, time-dependent and independent holding cost can be considered. Intuitionistic fuzzy geometric programming can be extended by existing fuzzy geometric programming to solve nonlinear and linear optimization problems. This method can minimize the total average cost of the EOQ model by applying intuitionistic fuzzy geometric programming. Intuitionistic fuzzy geometric programming is more feasible and preferable than crisp and fuzzy geometric programming[2]. The cost factor has to lie within a permissible range. The company will face huge losses if the cost exceeds the permissible range. Table 1 shows the percentage cost escalation from literature using different methods. It shows that, the average permissible cost escalation can be 1.0% - 1.3%.

Table 1. The percentage cost escalation

Author	Method	% Cost Escalation
Uroš Klansek [6]	Nonlinear Programming	1.16
Mohammed Woyeso Geda[7]	Linear Programming	1.11
Omar M. Elmabrouk [8]	Linear Programming	1.09
Michael J. Risbeck [9]	Mixed-Integer Linear Programming	1.12
Athanasios P et al [10]	Approximation method	1.96
Michael J. Risbeck at al [11]	Mixed-Integer Linear Programming	2.09
Ehsan Eshtehardian at al [12]	GA and fuzzy sets theory	1.4

Rana A. Al Haj et al [5]	Nonlinear-Integer Programming	0.99
Mohammed Nooruldeen Azeez et al [13]	Ant Colony Optimization	1.05
Yanshuai Zhang et al [1]	Ant Colony Optimization	0.75

3 Methodology

In this research, the following methodology is framed in Figure 1 to achieve the study's objectives, which are to minimize total cost escalation and ascertain the ideal order quantity for the building project while taking non-negative, start, crash, and floor space limits into account.

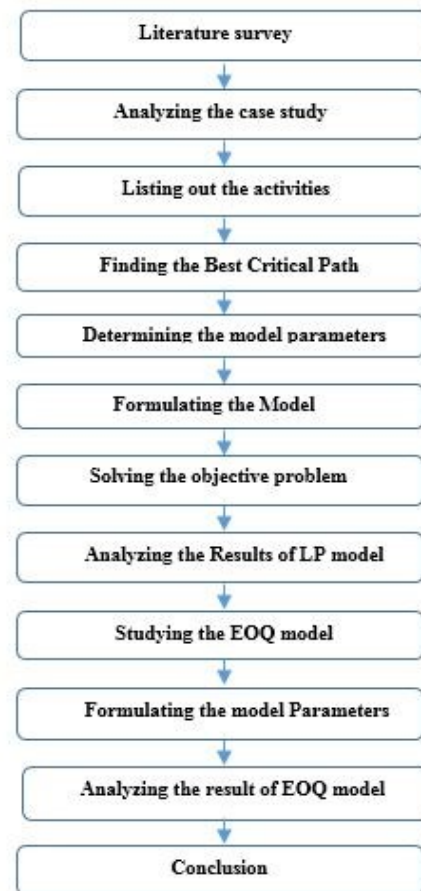


Figure 1. Methodology chart

Technical specifications in the research included:

- i. Activity on Node used to determine the network logic for the project schedule.
- ii. The cost slope is the slope of the direct cost curve, approximated as a straight line. It is defined as follows.

$$C_s = \frac{C_c - C_n}{t_n - t_c} = \frac{\Delta C}{\Delta t} \quad (1)$$

Where:

C_s	Cost slope
C_c	Crash Cost
C_n	Normal Cost
t_n	Normal Time
t_c	Crash Time
ΔC	Change in cost
Δt	Change in time

- iii. Linear programming is a technique used in mathematics to optimize processes under limitations. Maximizing or minimizing the target function is the aim of linear programming. The optimization is done using the LINDO program, which has a high degree of ease in solving complex functions.

4 PROPOSED EOQ MODEL

The best order amount a business or organization should buy to reduce inventory expenses, including holding costs, shortfall costs, and order costs, is the economic order quantity, or EOQ. Finding the ideal quantity of product units to order is the goal of the economic order quantity. If successful, a business can reduce the price of purchasing, shipping, and storing units. Production levels and order intervals are also determined using the economic order quantity model to maintain the ideal inventory level. The software can coordinate supply chain networks and logistics with the economic order quantity model. In cash flow analysis, the model is equally crucial. The approach can assist a business in managing the cash flow related to its inventory. In the EOQ model as shown in figure 2, the reorder point is defined as the point at which the

inventory is about to fall. Economic order quantity is responsible for reordering, the cost incurred while placing an order, and storing the materials.

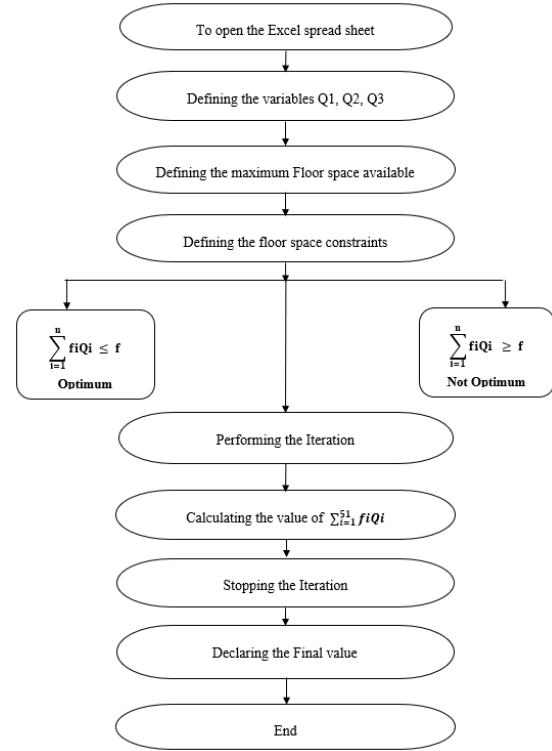


Figure 2. Algorithm To Solve the EOQ Model

In the construction industry, projects face high inventory costs because of the need for more floor space to keep the materials. To minimize the total costs of inventory, an Economic order quantity model with floor space constraints is developed using the Lagrangian function. The concept of EOQ is to determine the optimum order quantity of each material for the available floor space. This model can help construction engineers manage and control the inventory and facilitate ease of construction. The project has limited warehouse capacity, and the items compete for floor space. The available floor space in the construction project is 376 sq. m. To construct the model, the following assumptions are made: Production or supply is instantaneous with no lead time. Demand is uniform and deterministic. Shortages are not allowed. Suppose there are N-items to be stored in an inventory system; then -

The total unit cost for all the items is given by:

$$TC = \sum_{i=1}^n \left(0.5 * P_{ci} * I_{ci} * Q_i + \frac{R_{ci} * D_i}{Q_i} \right) \quad (2)$$

Where,

D_i – Demand rate for i th item

P_{ci} – Purchase cost of i th item

R_{ci} – Replenishment cost for i th item

I_{ci} – Inventory carrying cost fraction per unit per annum for i th item

Q_i – Order quantity for i th item

4.1 Application of LP Model

The case study is conducted in prestige pine wood apartments in Bangalore. Due to the extreme weather conditions, the project faced a delay in its completion. All the activities from the project have been grouped under 16 significant activities. To model, it is essential to determine the ES-EF and LS-LF in terms of starting and ending events. It is observed that the total duration for completion of the project is 2598 days based on the critical path. The list of essential activities under the critical path are: Earthwork excavation, Soil nailing, Retaining wall, Grade slab, Pedestal and column construction, Backfilling, Stitch slab, Reinforcement works, Concrete works, Brickworks, Water supply and sanitation works, Electrical works, Plastering works, Flooring works, Painting works, Wooden works.

There were seven possible ways to reach the final activity. The normal time (NT) and the crash time (CT) are calculated individually for all seven possible paths. The path with the maximum normal and full crash time is chosen for crashing. The maximum time the project can take to complete is 2511 days, and the maximum crashing time allowable is 2288 days.

$$2511 \text{ days} \geq X_i \geq 2288 \text{ days}$$

Where X_i = Amount of time that each activity i will be crashed

Y_i = Start time of activity i

U_i = Change in cost by change in time for activity i

Table 2 shows the cost slope for all activities. The value of U_i can be written as shown in Table 3.

Table 2. Determining the cost slope

Activity ID	Normal Time (days)	Crash Time (days)	Normal Cost (Rs)	Crash Cost (Rs)	Cost Slope (C_s)
1	56	37	1,16,61,840	13032640	72147.36
2	37.5	30	735000	750750	2100
3	15	13	1056873	1064873	4000
4	51	46	18297190	18312190	3000
5	34	34	1470547	1470547	0
6	11.5	6	1711304	1743804	5909.09
7	50	45	18297190	18312190	3000
8	417	371	14064681	14170481	2300
9	365	332	222709837	222815437	3200
10	292	271	4,57,26,135	4,57,66,135	1904.76
11	80	80	88,41,000	88,41,000	0
12	73	73	88,41,000	88,41,000	0
13	486	459	24823281	24875281	1925.92
14	331	303	19663020	19758220	3400
15	226	201	27180189	27230189	2000
16	73	73	29982541.38	29982541.4	0

Table 3. Determining the U_i value

U_1	U_2	U_3	U_4
72147.3	2100	4000	3000
U_5	U_6	U_7	U_8
0	5909.09	3000	2300
U_9	U_{10}	U_{11}	U_{12}
3200	1904.76	0	0
U_{13}	U_{14}	U_{15}	U_{16}
1925.92	3400	2000	0

To Determine the Objective Function and Constraints of the LP Model:

Minimize $Z = 72147.3 X_1 + 2100 X_2 + 4000 X_3 + 3000 X_4 + 0 X_5 + 5909 X_6 + 3000 X_7 + 2300 X_8 + 3200 X_9 + 1904.7 X_{10} + 0 X_{11} + 0 X_{12} + 1925 X_{13} + 3400 X_{14} + 2000 X_{15} + 0 X_{16}$

Subject to the conditions mentioned in Table 4, all the activities' crash time constraints are shown.

Table 4. Determining Crash Time Constraints

Crash Time Constraints			
$X_1 \leq 19$	$X_5 \leq 0$	$X_9 \leq 33$	$X_{13} \leq 27$
$X_2 \leq 7.5$	$X_6 \leq 5.5$	$X_{10} \leq 21$	$X_{14} \leq 28$
$X_3 \leq 2$	$X_4 \leq 5$	$X_{11} \leq 0$	$X_{15} \leq 25$
$X_4 \leq 5$	$X_8 \leq 46$	$X_{12} \leq 0$	$X_{16} \leq 0$

Table 5 shows the Start time constraints of all the activities.

Table 5. Determining Start Time Constraints

Start time Constraints				
Y2- Y1+X 1>=56	Y6- Y5+X5 >=34	Y10- Y9+X9 >=365	Y13- Y12+X 12>=73	YF- Y16+X 16>=73
Y3- Y1+X 1>=56	Y7- Y6+X6 >=11.5	Y11- Y10+X 10>=29 2	Y14- Y13+X 13>=48 6	YF<=2 367
Y4- Y3+X 3>=15	Y8- Y7+X7 >=50	Y12- Y10+X 10>=29 2	Y15- Y14+X 14>=33 1	-
Y5- Y4+X 4>=51	Y9- Y8+X8 >=417	Y13- Y11+X 11>=73	Y16- Y15+X 15>=22 6	-

Table 6 shows the non-negative constraints of all the activities.

Table 6. Determining Non-Negative Constraints

Non-Negative Constraints							
X1 >=0	X5 >=0	X9 >=0	X1 3>=0	Y1 >=0	Y5 >=0	Y9 >=0	Y13 >=0
X2 >=0	X6 >=0	X1 0>=0	X1 4>=0	Y2 >=0	Y6 >=0	Y1 0>=0	Y14 >=0
X3 >=0	X7 >=0	X1 1>=0	X1 5>=0	Y3 >=0	Y7 >=0	Y1 1>=0	Y15 >=0
X4 >=0	X8 >=0	X1 2>=0	X1 6>=0	Y4 >=0	Y8 >=0	Y1 2>=0	Y16 >=0

4.2 Application of EOQ Model

The EOQ model is demonstrated using three different materials: Steel (Q1), Granite flooring (Q2), and Solid concrete block (Q3). The available floor

space 'f' in the construction project is 376 sq.m. The project data is given in Table 7 below.

Table 7. Determining Material Parameters

Parameters	Q1	Q2	Q3
D_i	1900	2000	10000
P_{ci}	34	32	10
O_{ct}	100	150	180
f (sq. m)	12	11.9	1.2

The values of Q1, Q2, Q3 are determined by varying the lambda function. Solving the material Parameters:

Table 8 shows the value of $\sum_{i=1}^{51} fiQi$:- 375.98, where fi values is mentioned in table 7.

Table 8. Solving Material Parameters

λ	Q1	Q2	Q3	$\sum_{i=1}^{51} fiQi$
1	114.27	144.84	960.76	4247.81
2	84.59	42.73	79.30	1618.80
3	70.20	36.92	76.47	1373.57
4	61.30	32.97	73.92	1216.83
5	55.11	30.07	71.61	1105.21
6	50.48	27.82	69.50	1020.35
7	46.85	26.01	67.57	952.93
8	43.90	24.51	65.79	897.64
9	41.45	23.25	64.15	851.19
10	39.37	22.16	62.62	811.43
11	37.57	21.21	61.19	776.87
12	36.00	20.38	59.86	746.47
13	34.61	19.63	58.62	719.44
14	33.37	18.96	57.44	695.20
15	32.26	18.36	56.34	673.29
16	31.25	17.81	55.30	653.37
17	30.32	17.31	54.31	635.13
18	29.48	16.84	53.37	618.37
19	28.70	16.42	52.48	602.88
20	27.98	16.02	51.63	588.51
21	27.32	15.65	50.83	575.13
22	26.69	15.30	50.06	562.64
23	26.11	14.98	49.32	550.93
24	25.57	14.68	48.62	539.94
25	25.05	14.39	47.94	529.58
26	24.57	14.12	47.29	519.80
27	24.12	13.87	46.67	510.55

28	23.69	13.63	46.07	501.78
29	23.28	13.40	45.50	493.45
30	22.89	13.18	44.94	485.52
31	22.52	12.97	44.41	477.97
32	22.17	12.77	43.89	470.76
33	21.83	12.58	43.39	463.87
34	21.51	12.40	42.91	457.27
35	21.20	12.23	42.44	450.95
36	20.91	12.06	41.99	444.89
37	20.62	11.90	41.55	439.06
38	20.35	11.75	41.13	433.46
39	20.09	11.60	40.72	428.07
40	19.84	11.46	40.32	422.88
41	19.60	11.32	39.93	417.87
42	19.36	11.19	39.55	413.04
43	19.14	11.06	39.19	408.37
44	18.92	10.93	38.83	403.86
45	18.71	10.81	38.49	399.49
46	18.51	10.70	38.15	395.27
47	18.31	10.59	37.82	391.17
48	18.12	10.48	37.50	387.20
49	17.93	10.37	37.19	383.35
50	17.75	10.27	36.88	379.61
51	17.58	10.17	36.58	375.98

5 Discussions

The proposed models have been applied to a construction project to demonstrate their practicality. The main aim of this study is to mitigate the time-cost and storage-related problems occurring in the construction industry. The linear programming solution in Table 9, indicates the crashing activities to reduce the project duration to 2485 days from 2598 days, which increased the overall cost to Rs. 45,52,96,751 from Rs. 45,50,61,628.

Table 9. Determining the Crashed Duration and Crashed Cost

Variable	Value	Reduced Cost
X ₁	0	69847
X ₂	0	2100
X ₃	0	1700
X ₄	0	700
X ₅	0	0
X ₆	0	3609
X ₇	0	700

X ₈	40	0
X ₉	0	900
X ₁₀	21	0
X ₁₁	0	0
X ₁₂	0	0
X ₁₃	27	0
X ₁₄	0	1100
X ₁₅	25	0
X ₁₆	0	0

The objective found at the 40th iteration and its value is given by,

Minimum value of Z = Rs. 235123.7

Table 10. Comparing the Obtained result with the Standard value

Parameters	Permissible range	Obtained result
Cost		
Escalation	1.0% - 1.3%	1.0005%
Crash	Between 2511	
Duration	and 2285 days	2485 days

Table 10 shows that to crash the total construction time for 113 days, Rs 235123.7 crash cost is needed. Thus, an additional 235123.7 Rs is required to crash the total duration of the construction. From Table 9, it can be inferred that Activity 8,10, 13, and 15 have been hit to 40, 21, 27, and 25 days, respectively. For time-related problems, the CPM method is used to identify the critical path. The model indicates that about a 4.36% decrease in time can be achieved by increasing cost by 1.0005%, which is satisfactory, as shown in Table 10.

The Single objective EOQ model with limited floor space is solved for the floor space constraints using the lagrangian function. The Economic order quantity value was found at the 51st iteration; its values are given in Table 11.

Table 11. Determining the value of material parameters

Items	Optimum Value
Q1	18
Q2	10
Q3	36

The available floor space to accommodate the materials in the construction site is 376 sq. m. The $\sum_{i=1}^{51} fiQi$ value is 375.989 sq. m. Hence, the floor space constraints lie within the range. The optimum order quantity of steel, Granite, and Solid concrete blocks is 18, 10 and, 36 respectively, which can be accommodated in the area of 376 sq. m.

6 Conclusion

The linear programming model offers the best solution to time and cost constraints concerns. After optimization, it was discovered that the whole crash lasted 113 days. Therefore, the extra expense incurred due to the time reduction is Rs. 235123.70. According to the model, it is possible to obtain a sound time reduction of around 4.36% by raising the cost by 1.0005%. The model accurately analyses while maintaining the correctness of the outcomes. The linear programming methodology effectively ascertains a reduction in the length of every task. By using software techniques, the strategy may readily address complicated crashing situations and is very versatile. The approach works well for large-scale building projects with plenty of moving parts. It is challenging to cycle through many activities manually using the trial-and-error method. Therefore, using the linear programming method, the construction manager may quickly estimate the crash cost required to crash the complete project for a given set duration. The project manager can efficiently organize all activities thanks to the model's ease of use. It is limited to solving linear constraints and single-objective problems. Therefore, several techniques, such as fuzzy multi-objective linear programming, mixed integer linear programming, and particle swarm optimization, can also be utilized to handle multi-objective optimization problems.

Using the economic order quantity is an efficient approach to figuring out inventory control. The ideal order quantities for steel, Granite, and solid concrete blocks are 18, 10, and 36 lots respectively, and they may all fit within a 376-square-meter space. To maintain a good flow of production and prevent overinvestment in stocks, the model can keep an eye on the acquisition and storage of materials in the inventory. Any real-time project can directly benefit from the application of this technology. Regarding the building business, the model helps determine the order quantity at various process phases based on space constraints and demand. The model is easy to use and effective. Different programming techniques, including geometric, nonlinear, and Newton-Geometric programming, can be applied to determine the solution. Therefore, by employing this technique, the construction industry's inventory management quality can be raised. By introducing different uncertainties, additional research on the time-cost programming model, nonlinear discrete optimization of project schedules, and multi-project scheduling with resource constraints can be conducted. In the economic order quantity model, modified geometric programming in a neutrosophic environment can be used to accomplish multiple product optimization.

References

- [1] Y. Zhang and S. T. Ng, "An ant colony system based decision support system for construction time-cost optimization," *J. Civ. Eng. Manag.*, 2012, doi: 10.3846/13923730.2012.704164.
- [2] B. Mondal, A. Garai, and T. K. Roy, "Optimization of EOQ model with space constraint: An intuitionistic fuzzy geometric programming approach," *Notes Intuitionistic Fuzzy Sets*, 2018, doi: 10.7546/nifs.2018.24.4.172-189.
- [3] O. M. Elmabrouk and F. Aljebali, "Crashing Project Activities Using Linear Programming Technique," *Proc. Int. Conf. Ind. Eng. Oper. Manag. Istanbul, Turkey*, 2012.
- [4] S. B. Kurhade and A. R. Patel, "Optimization in Construction Management," *Int. Res. J. Eng. Technol.*, 2008.
- [5] R. A. Al Haj and S. M. El-Sayegh, "Time-Cost Optimization Model Considering Float-Consumption Impact," *J. Constr. Eng. Manag.*, 2015, doi: 10.1061/(asce)co.1943-7862.0000966.
- [6] U. Klanšek and M. Pšunder, "Cost Optimization of Time Schedules for Project Management," *Econ. Res. Istraživanja*, 2010, doi: 10.1080/1331677x.2010.11517431.
- [7] M. W. Geda, "A Linear Programming Approach for Optimum Project Scheduling Taking Into Account Overhead Expenses and Tardiness Penalty Function," *Int. J. Eng. Res. Technol.*, 2014.
- [8] O. M. Elmabrouk, "A Linear Programming Technique for the Optimization of the Activities in Maintenance Projects," 2011.
- [9] M. J. Risbeck, C. T. Maravelias, J. B. Rawlings, and R. D. Turney, "Mixed-integer optimization methods for online scheduling in large-scale HVAC systems," *Optim. Lett.*, 2020, doi: 10.1007/s11590-018-01383-9.
- [10] A. P. Chassiakos and S. P. Sakellariopoulos, "Time-Cost Optimization of Construction Projects with Generalized Activity Constraints," *J. Constr. Eng. Manag.*, 2005, doi: 10.1061/(asce)0733-9364(2005)131:10(1115).
- [11] M. J. Risbeck, C. T. Maravelias, J. B. Rawlings, and R. D. Turney, "Cost optimization of combined building heating/cooling equipment via mixed-integer linear programming," in *Proceedings of the American Control Conference*, 2015. doi: 10.1109/ACC.2015.7170976.
- [12] E. Eshtehardian, A. Afshar, and R. Abbasnia, "Time-cost optimization: Using GA and fuzzy sets theory for uncertainties in cost," *Constr. Manag. Econ.*, 2008, doi: 10.1080/01446190802036128.
- [13] M. N. Azeez and A. Alsaffar, "Construction Time-Cost Optimization Modeling Using Ant Colony Optimization," *J. Eng.*, 2023, doi: 10.31026/j.eng.2014.01.09.

Prototype development of an automated 3D shop drawing generation tool for reinforcement work

Kanae Miyaoka¹, Fumiya Matsushita¹ and Kazumasa Ozawa¹

¹School of Engineering, University of Tokyo, Japan

miyaoka@i-con.t.u-tokyo.ac.jp, matsushita@i-con.t.u-tokyo.ac.jp, ozawa@i-con.t.u-tokyo.ac.jp

Abstract –

In concrete structure projects, activities such as drawing, fabrication, and assembly related to the reinforcement substantially impact the schedules and costs. It is necessary to apply building information modeling (BIM) to smoothly share and distribute information regarding reinforcement among the diverse range of suppliers to streamline work.

This study aims to establish a BIM-based rebar information linkage system applicable across the reinforcement work supply chain. The paper reports on the prototype development and trial of an automated 3D shop drawing generation tool, which is one of the essential tools in the system. Through a trial conducted in a subway station construction project, it was validated that the tool led significant reductions in labor costs and timesaving compared with the conventional 2D-based drawing method. As future work, it will be necessary to define an information model for rebar in order to build a system that can promote information sharing and coordination across the supplier network during the construction phase.

Keywords –

BIM; Concrete reinforcement; Parametric modeling; Shop drawing

1 Introduction

Building information modeling (BIM) is a system designed to enhance the efficiency and sophistication of construction production and management. It achieves this by introducing 3D models from the planning and design stages and linking and developing these models throughout the construction and maintenance. This facilitates information sharing among the project stakeholders. BIM adoption for concrete structures is increasing rapidly, extending to large-scale infrastructure in civil engineering (e.g. railroads, highways, and bridges) [1] [2] [3]. In infrastructure projects, activities related to reinforcement (such as

drawing, fabrication, and assembly) significantly impact the schedules and costs. Given the diverse range of suppliers involved, including contractors, subcontractors, and manufacturers, coordination is crucial for project execution. The use of BIM is likely to streamline this process.

However, notwithstanding the increasing adoption of BIM, only approximately 30% of users frequently model reinforcement elements [4]. This leaves scope for improvement in 3D modeling for reinforcement. Past applications of 3D rebar models include evaluations of interference, enhancement visibility for client–contractor agreements [5], quantity calculations [6] and generation of input data for rebar processing machines [7]. However, these applications are task specific. Moreover, research focused on improving efficiency throughout the supply chain is insufficient.

This study aimed to establish a BIM-based rebar information linkage system applicable across the rebar construction supply chain. This paper reports the development and trial of an automated 3D shop drawing generation tool.

2 Overview of BIM-based rebar information linkage system

2.1 Processes of conventional reinforcement work

Figure 1 illustrates conventional reinforcement work processes and the suppliers involved in the construction phase. The initial step in this phase is the generation of shop drawings that provide detailed information regarding the rebar, such as its shape, specifications, joint locations, and placement. These drawings also include section/plan drawings and bar bending schedules. Shop drawings are generated from technical drawings, considering construction plan elements such as temporary openings and beams. During the procurement stage, the contractor produces a rebar order sheet based on the bar bending schedules and sends it to the manufacturer. The manufacturer inputs this

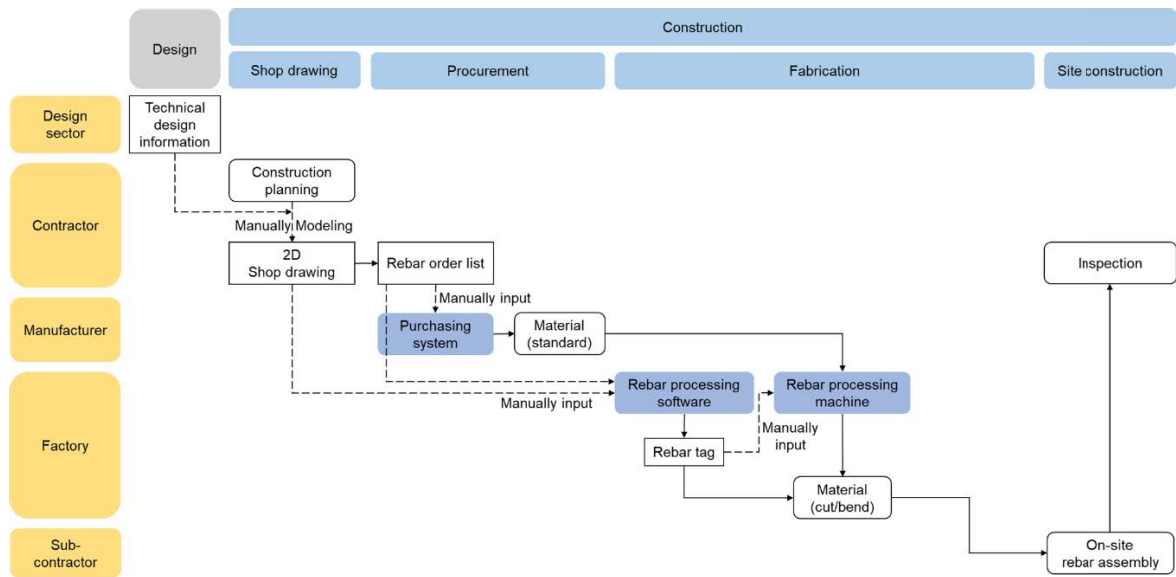


Figure 1. Conventional reinforcement work processes and the suppliers involved.

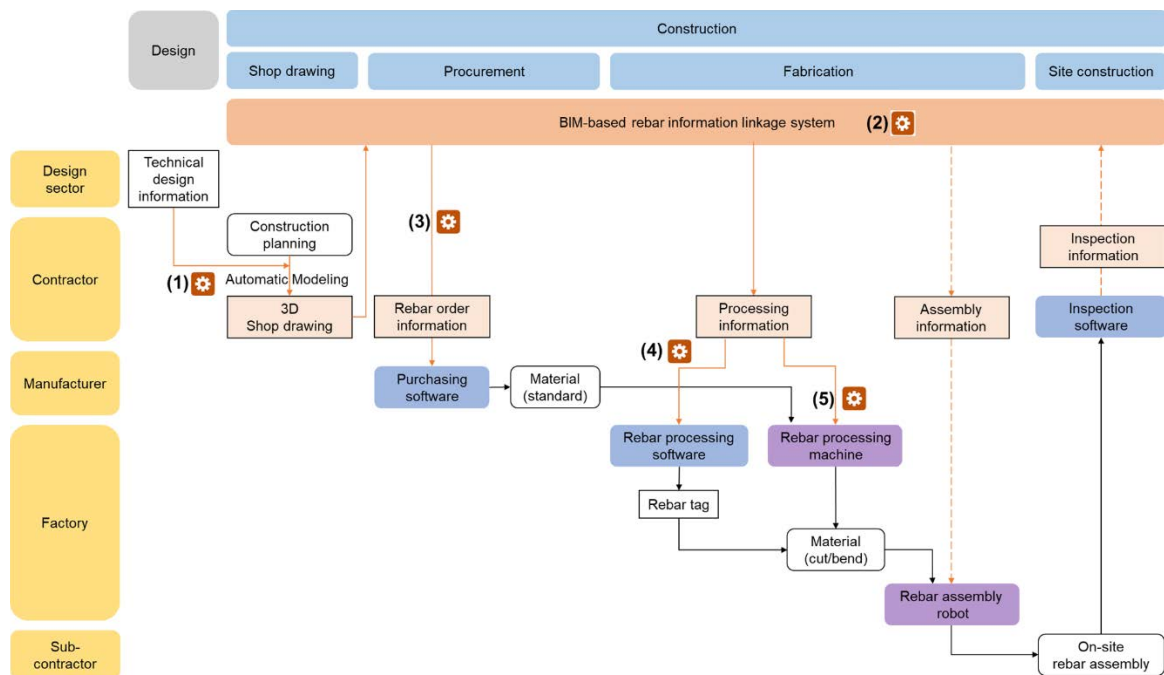


Figure 2. Overview of the BIM-based rebar information linkage system

information into the purchase system. In the fabrication stage, the rebar is cut and fabricated at the factory of the manufacturer or subcontractor. The processing-information management software at the factory receives details from the rebar order sheet and shop drawings and generates rebar processing tags for each processing line. Machine operators use these tags to process rebars. The recent introduction of automatic

processing machines requires manual data input through touch panels. After processing, the materials are delivered to the construction site, inspected, and accepted. Subsequently, a subcontractor assembles the rebar. This is followed by an inspection by the prime contractor.

Conventional workflows encounter several challenges. First, generating and revising shop drawings

is labor-intensive and vulnerable to errors owing to the frequent revisions based on modifications in the construction plan. Second, the communication between suppliers relies significantly on documents such as 2D drawings and order sheets. This results in inefficient information transfer between processes. Manually entering data into the purchasing and processing information management software further contributes to time-consuming processes. Third, the significant personnel requirements for rebar processing at factories and assembly sites, particularly in the large civil engineering field, pose a key obstruction in the overall process.

2.2 Overview of the proposed BIM-based rebar information linkage system

Figure 2 presents an overview of the BIM-based rebar information linkage system proposed in this study. Table 1 outlines the challenges in the conventional work process, in conjunction with the corresponding requirements and development items essential for resolving these issues.

The proposed system provides a user-friendly tool for generating 3D shop drawings (development item (1)). To effectively distribute and utilize information from these drawings in subsequent operations, it is necessary to define an information model to manage the rebar information (development item (2)). In the construction sector, information is organized along project timelines to facilitate the exchange of

information between stakeholders and applications. A representative standard, ISO 12006-2, defines a framework for construction-sector classification systems and identifies a set of recommended classification tables and their titles for a range of construction object classes [8]. Based on ISO 12006-2, countries have introduced classification systems such as Uniclass-2015 and OmniClass [9]. There is various modeling software that handle rebar models and different applications used depending on the timeline of the work. However, by defining the information model for rebar, it is possible to exchange information among the parties and applications involved. Additionally, tools would be developed to extract and link the required information to various software applications such as purchasing systems and rebar processing information management software (development items (3) and (4)). Furthermore, enabling data linkage with machines such as rebar bending machines (development item (5)) and rebar assembly robots would expedite prefabrication in factories. This would reduce the labor required for processing and assembly. This research is also working actively on the data linkage represented by the solid line in Figure 2, with a forward-looking perspective on future connections between the rebar assembly robots and inspection systems indicated by the dotted line.

3 Prototype development of an automated 3D shop drawing generation tool

This paper presents a report on an automated 3D shop drawing generation tool. It is essential in this system. This chapter details the development of the tool. The next chapter presents its implementation and a verification of its effectiveness.

3.1 Previous studies and issues

In the conventional approach to generating shop drawings, designers convey the necessary information by incorporating it into basic design drawings. The CAD operators manually produce 2D shop drawings based on this information. These drawings encompass various elements such as the section/plan view drawings and bar-bending schedules for each construction block or pattern (see Figure 3). Frequent revisions are necessary because of the influence of construction plans, such as the placement of temporary structures and construction methods. Moreover, because shop drawings convey intricate rebar details, errors are likely to occur during drafting. The adoption of 3D models for rebar shop drawings effectively mitigates these challenges.

Existing technology enables the extraction of 2D

Table 1. Requirements and development items

Challenges	Requirements	Dev. item
Preparation of shop drawing	Convenient generation and modification of 3D shop drawing.	(1)
Information sharing	Sharing of rebar information across suppliers and distribution among software programs.	(2)
	Extraction and linkage of input data of each software program from 3D shop drawings.	(3) (4)
Processing and assembly work	Extraction and linkage of input data of machines such as processing machines to achieve pre-fabrication.	(5)

drawings from 3D models, a feature implemented in software programs such as Revit [10] and Rhinoceros [11] (see Figure 4 (1)). When the construction plan undergoes alteration, modification of the 3D models automatically updates the linked 2D drawings.

However, the generation of 3D shop drawings poses several challenges. The complexity of drawings makes the modeling process time-consuming. Moreover, the quality depends on the skills of the BIM operator. Parametric modeling has emerged as an efficient technique for addressing these challenges. The method

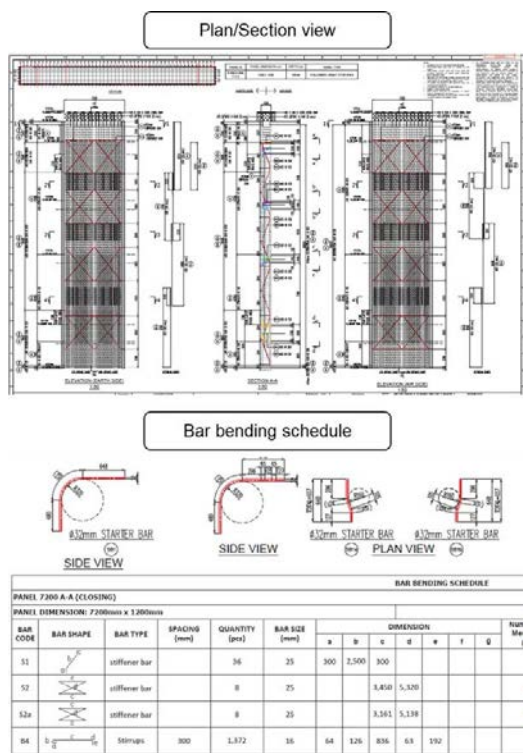


Figure 3. Examples of shop drawing

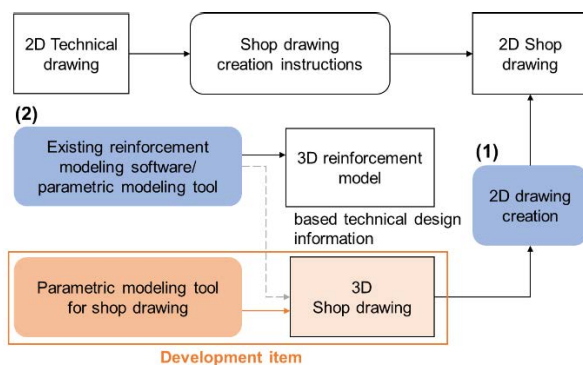


Figure 4. Existing technology and development item

involves defining the dimensions and parameters in advance. This enables the generation and modification of 3D models by entering variable values. Unlike direct modeling, in which the operator manually draws the model irrespective of the operator's skills, the parameters and constraints should be considered effectively. Although certain automatic rebar placement functions exist in 3D modeling software such as Revit that use parametric modeling [12] [13], their user-configurable parameters are limited to basic design information (rebar diameter, length, rebar shape, spacing). Although there are examples of the application of parametric modeling to building structures and the rebar in previous studies [14] [15], it is difficult to model rebar to reflect the detailed information contained in the shop drawings, such as the type and location of rebar joints, changes in spacing, and arrangement of setup bars (see Figure 4 (2)). Given these considerations, there is a need to develop parametric modeling tools capable of handling the intricate reinforcement information available in shop drawings. This would ultimately facilitate the efficient generation of 3D shop drawings.

3.2 Strategy to develop the tool

Figure 5 presents an overview of the automated 3D shop drawing generation tool. The tool leverages Revit for 3D modeling and Grasshopper (a plug-in compatible with Rhinoceros) for parametric modeling. The initial step involves preparing a 3D model outlining the basic concrete structure frame (referred to as the “concrete body model”) using basic design information. In step 2, the parameter data for the shop drawings of the targeted

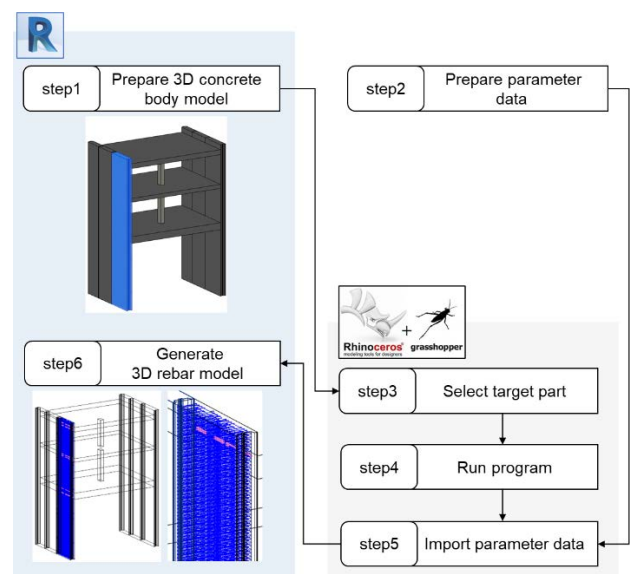


Figure 5. Overview of the developed tool

part is created. Details of the parameter data are explained in the next section. After that, a program for automated 3D shop drawing generation is opened on Grasshopper, and the target part is selected (step 3). Then, when the program is executed, the prepared parameter data is read (step 4 and 5), and a 3D rebar model is generated in Revit according to the parameter data (step6).

3.3 Prototype development

In this study, a prototype tool was developed for a subway station. The development process began by referencing 2D shop drawings from previous subway station projects and identifying the necessary parameters for modeling. An example of the parameter data for a wall (specifically a diaphragm wall) is presented in Table 2. The first line represents the fixed parameter, whereas the numerical values are entered by the user in the second line and beyond. As emphasized in Section 3.1, effective parameter settings are crucial for tool development. The parameters were configured to enable the model to convey shop drawing-specific information including the rebar cover in planes and sections, rebar shapes, joint locations and specifications, and arrangements of the set-up bars. Separate parameter data were established for each rebar type (main bar, reinforcement bar, shear reinforcement bar, etc.). The data were formatted as a CSV file.

Subsequently, a program for generating a 3D rebar model in Revit was constructed using Grasshopper. An example of the Grasshopper program is shown in Figure 6. The program includes buttons for generating models for each rebar type. When a button is pressed and the program executed, it reads the parameter data stored in the specified path folder and generates a 3D model in Revit. Testing of the developed tool using sample data verified the seamless generation of the 3D model in Revit, as depicted in Figure 7.

Additionally, a comparison between the parameters of the generated 3D rebar model and the input values from the parameter data in the CSV verified that the rebar was modeled accurately based on the specified parameter data values, as shown in Figure 8.

Table 2. Example of parameter data

name	z	xoffset	yoffset	shape code	Dia.	...
mainbar1	100	50	100	00	22	...
mainbar2	100	1600	100	01	19	...
mainbar3

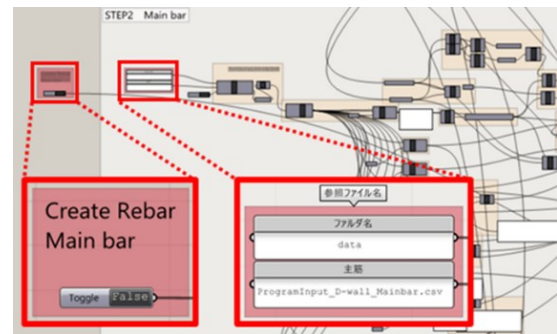


Figure 6. Example of Grasshopper program

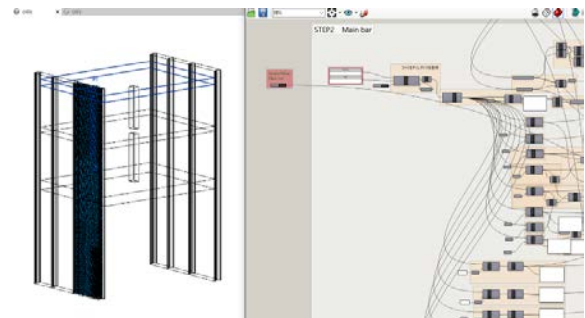


Figure 7. Screen during model creation

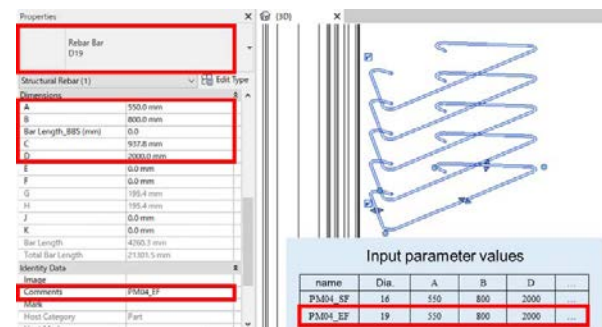


Figure 8. 3D model parameters and CSV input values

4 Trial of the developed tool and verification on effect

4.1 Trial of the developed tool

The developed tool underwent a trial run in a project involving the construction of three subway stations and connecting tunnels. The aim here was to elucidate the workflow involved in using the tool and validate its effectiveness in the context of the project. Specifically, the tool was employed to generate shop drawings of a diaphragm wall.

The workflow during the implementation of the development tool is shown in Figure 9. This sequence

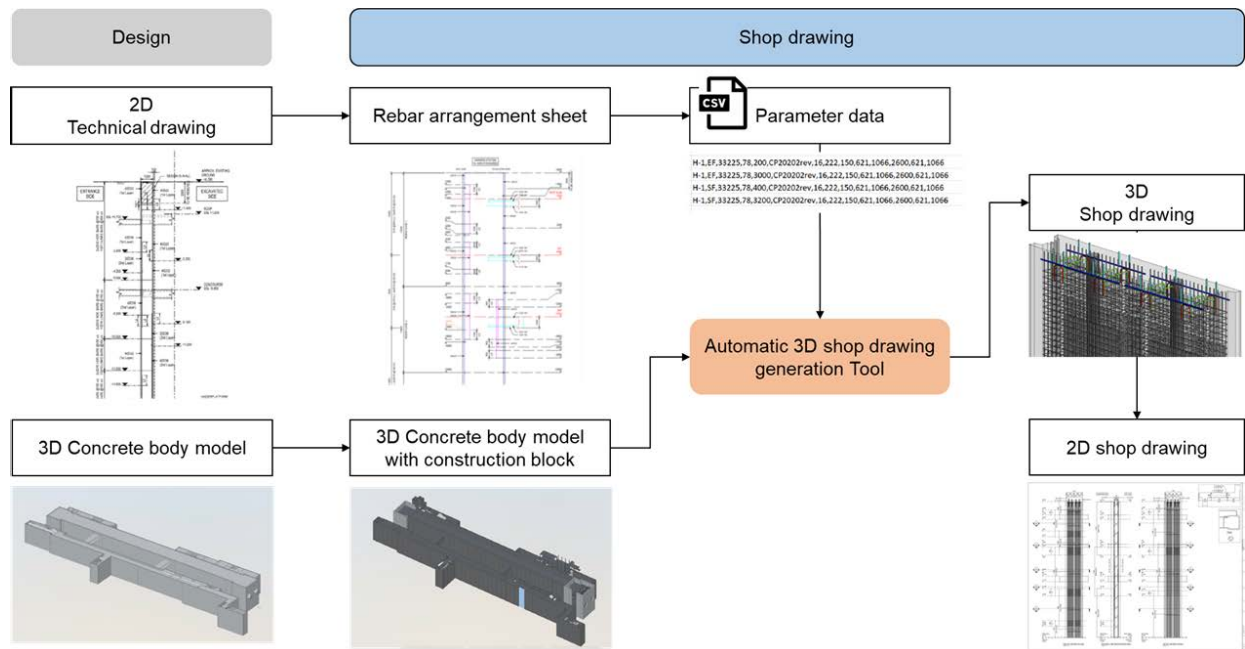


Figure 9. The workflow during the implementation of the development tool

was developed based on the activities of the shop drawing generation team involved in the subject project. It reveals the various steps preceding the preparation of the input data for the development tool (comprising the concrete body model and parameter data). The shop drawing generation team initially received 2D technical drawings from the design department. It was challenging to translate these directly into parametric data. To address this issue, a rebar arrangement sheet for each panel was first generated by extracting pertinent information from technical drawings. The rebar arrangement sheet includes essential details for shop drawings, such as the diameter, spacing, number of rebars, rebar shape, and joint locations and specifications. Subsequently, the parameter data to be imported into the tool were obtained from the information within the rebar arrangement sheet and cross-sectional drawings. With regard to the concrete body model, the 3D model obtained from the design department integrates various members such as walls and slabs. Consequently, during the shop drawing stage, it is necessary to segment the model into construction blocks. Following this process, 3D shop drawings were generated using the developed tool, and 2D shop drawings were extracted manually from the 3D model.

4.2 Validation on effects

Table 3 compares the work required to generate conventional 2D shop drawings and the utilization of the developed tool. The table reveals a labor reduction

of 1.5/h/panel for the underground continuous wall when employing the developed tool, compared with the conventional method. With the construction project involving modeling of approximately 100 panels per station, the estimated cost savings amounted to approximately USD 52,500.00 per station (refer to Table 4). It is noteworthy that although the 2D drawings were extracted manually from the 3D model in this trial, the potential introduction of software functions to this process is likely to further decrease the workload.

Table 3. Comparison of the work required

for a panel of diaphragm wall	
Conventional 2D method	Method using the developed tool
Instructions for CAD operator 2D drawing	Generating rebar arrangement sheet 1 day
	Generating CSV parameter data 1 h
	3D model generation 5 min
	Generating 2D drawing from 3D model 4 h
total 3 days	total 1.5 days

Table 4. Cost effectiveness of the tool

Items	Effects
(1) Productivity-effect	1.5 day/panel
(2) Number of panels	100 panels/station
(3) Average cost of BIM operator	USD 350/day
Cost-effective: (1) × (2) × (3)	USD 52,500/station

5 Conclusion

This paper presents the development and trial of an automated 3D shop drawing generation tool as a component of a BIM-based rebar information linkage system during the construction phase. By incorporating the detailed information of the shop drawings into the parameter data, the tool succeeded in parametrically and automatically modeling the 3D shop drawings. A demonstration on a subway station construction project revealed that applying this tool can significantly reduce labor costs compared to traditional 2D-based drafting method. The ultimate goal is to establish a comprehensive system that seamlessly connects information on reinforcement arrangements during the construction stage. As future work, the linkage of rebar information will be focused. Particularly crucial in this endeavor is the consideration of an information model that is essential for facilitating the sharing and coordination of information across an entire supplier network. As noted in 2.2, the development of an information model allows for the integration of various BIM software and information processing applications. In order to increase the robustness of the proposed system, an information model will be developed with reference to the existing information classification system.

Acknowledgement

This study was conducted under a joint research agreement between School of Engineering, the University of Tokyo, and Shimizu Corporation.

I wish to express my deepest appreciation to the members of the shop drawing team of the subject project for their cooperation during the trial.

References

- [1] Bensalah M. et al. BIM integration into railway

- projects—Case study. *Contemporary Engineering Sciences*, 11(44):2189–2199, 2018.
- [2] Fanning B. et al. Implementing BIM on infrastructure: Comparison of two bridge construction projects. *Practice Periodical on Structural Design and Construction*, 20(4):2189–2199, 2018.
- [3] E Halim. et al. Building Information Modelling (BIM) Implementation for Highway Project from Consultant’s Perspectives in Malaysia. IOP Conference Series: Earth and Environmental Science, 971, 2022.
- [4] McGraw-Hill Construction. *McGraw Hill Construction Smart Market Report*, McGraw-Hill Construction, New York, 2008.
- [5] National Institute for Land and Infrastructure Management. BIM/CIM case study Vol. 1. On-line: <https://www.nilim.go.jp/lab/qbg/BIM/BIMExamplesR1.pdf>, Accessed: 10/12/2023.
- [6] Woobin K. et al. Analysis of BIM-Based Quantity Take-Off in Simplification of the Length of Processed Rebar. *applied sciences*, 13, 2468, 2023.
- [7] Liu Y. et al. BIM-BVBS integration with openBIM standards for automatic prefabrication of steel reinforcement. *Automation and Construction*, 125, 2021.
- [8] ISO12006-2:2015: <https://www.iso.org/standard/61753.html>, Accessed: 10/3/2024.
- [9] Veronica R. et al. Analysis of classification systems for the built environment: Historical perspective, comprehensive review and discussion. *Journal of Building Engineering*, Volume 67, 2023.
- [10] Revit: <https://www.autodesk.co.jp/>, Accessed: 10/12/2023.
- [11] Rhinoceros: <https://www.rhino3d.co.jp/>, Accessed: 10/12/2023.
- [12] Lightning BIM: <https://lightningbim.com/>, Accessed: 10/12/2023.
- [13] SOFiSTiK: <https://www.sofistik.com/>, Accessed: 10/12/2023.
- [14] A. Aramburu. et al. Parametric modelling of 3D printed concrete segmented beams with rebars under bending moments. *Journal of Building Engineering*, Volume 18, 2023.
- [15] Weilong Z. et al. Development and application of rebar information table module based on Revit platform. IOP Conference Series: Earth and Environmental Science , 567, 2020

The Effect of Material Handling Strategies on Time and Labour Fatigue in Window Manufacturing

Omar Azakir¹, Fatima Alsakka¹, and Mohamed Al-Hussein¹

¹ University of Alberta, Canada

azakir@ualberta.ca, falsakka@ualberta.ca, malhussein@ualberta.ca

Abstract –

Manual material handling in the workplace is a leading cause of musculoskeletal disorders among workers and a key driver of workers' compensation costs. As such, the primary objectives of this study are to assess the impact of manual material handling on both the time spent manually handling materials and the well-being of workers within the context of glass window manufacturing, as well as to examine potential strategies to reduce the adverse impact of manual material handling. The study focuses on the process of manually pushing carts fully loaded with 35 glass units over a 100-meter-long route at a case window-manufacturing facility. Two alternative strategies are examined: (1) reducing the weight of the carts by partially loading them, and (2) using power jacks to transfer fully loaded carts. The study analyzes the effects of these strategies on parameters such as total transfer time and the corresponding labour hours, as well as worker fatigue levels, measured on a scale ranging from 1 (low fatigue) to 5 (severe fatigue). The results reveal that the task of transferring the fully loaded carts along the route costs the company about \$1,800 in labour every month, consumes more than 10% of the 8-h work shift of three workers, and causes high levels of fatigue among the workers engaged in this task. Reducing the cart load size from 35 to 25 units is found to reduce the total time needed to transfer 700 glass units by 7%, the corresponding total labour hours by 38%, and the average fatigue score (as reported by the workers) from 5.0 to 4.0. Further reducing the load size to 20 units reduces the average fatigue score to 2.3 but increases the requisite transfer time and labour hours compared to the load size of 25 units. Introducing the use of a power jack as an alternative to reducing the load size, meanwhile, reduces the total transfer time by 55%, the total labour hours by 85%, and the average fatigue score from 5.0 to 2.0 when comparing it to the current practice. Moreover, a cost analysis of the different alternatives reveals that introducing the power jack is the alternative with the lowest net present value when considering labour and purchase

costs, and that the case company could realize a return on its investment within five months following the purchase of a power jack. As such, this alternative is deemed to be the most attractive among the alternatives considered in terms of time efficiency, cost efficiency, and the well-being of workers.

Keywords –

Material Handling; MMH; Time Efficiency; Cost Efficiency; Labour Fatigue; Cart System; Power Jack; Window Manufacturing.

1 Introduction

Despite the versatility and ubiquity of material handling equipment, which includes different types of conveyors, cranes, automated guided vehicles, and others [1], manual material handling (MMH) that does not involve the use of any equipment is still common in many workplaces [2]. MMH is defined as the process of “moving or handling things by lifting, lowering, pushing, pulling, carrying, holding, or restraining” [3]. It often requires workers to bend and stretch their bodies when carrying heavy loads [4]. As such, MMH is recognized as a leading cause of occupational fatigue and musculoskeletal disorders, leaving approximately three out of four Canadians whose job involves MMH with back pain resulting from workplace injuries at some point in their lives [3]. In fact, the largest portion of workers' compensation claims are associated with MMH [5]. In Canada, MMH-related back injuries contribute to roughly one-third of all lost work and an even higher percentage of total workers' compensation costs [3]. Furthermore, the impacts of MMH are not limited to physical injuries of workers and the corresponding compensation claims. In this regard, improving work design rooted in ergonomics has been shown to have many benefits, such as increased worker productivity, reduced physical demand, and reduced employee turnover [6]. Moreover, besides the ergonomic risks it poses, material handling is also generally considered a non-value-added task, from a lean perspective, as it does not directly contribute to the transformation of raw

materials into a saleable commodity [7]. These non-value-added tasks can account for a significant portion of the total manufacturing time. For instance, in a recent case study undertaken in a tannery facility, non-value-added tasks were found to account for 17.42% of the manufacturing cycle time, with material handling tasks being the primary contributor [8]. Because these non-value-added tasks consume time and resources, they drive up operating costs. In fact, material handling alone has been found to be responsible for 20–50% of the operating costs in a manufacturing environment [9].

Therefore, effective material handling strategies can play a pivotal role in enhancing the overall performance of manufacturing. For instance, a case study in the automotive manufacturing sector demonstrated that implementing hand trolleys or carts for material transportation, as opposed to manual methods, led to a productivity increase of 19.6% [10]. However, as illustrated in this manuscript's case study, even a hand cart system can encounter critical issues if the load handled during each transfer trip is not properly managed (as elaborated in Section 2). Currently, there is a lack of standardized numerical guidelines regarding acceptable load limits or force exertions for hand cart operations, as stated by the Canadian Centre for Occupational Health and Safety [11]. While general recommendations suggest a load limit of 450 lb for four-wheel hand carts [11], management may be inclined to increase the load per trip with the intent of minimizing the number of trips and total transfer time, thereby boosting productivity. However, it is arguable that heavier carts might prolong transfer times and elevate worker fatigue, potentially counteracting the intended productivity gains. This hypothesis regarding cart systems has not been thoroughly explored in previous research conducted on MMH [12–15]. Therefore, this study aims to assess how varying cart loads affect total transfer time and worker fatigue levels, focusing on a glass window manufacturer where glass units are manually transferred between workstations using rolling carts. The study also evaluates the use of power jacks for cart transfer compared to manual handling and includes a cost analysis of different material handling strategies.

2 Problem Description

The case company specializes in the manufacture of double- and triple-glazed window units. The process of manufacturing a sealed glass unit entails various activities, beginning with loading the glass onto a cutting table. The glass is then cut into desired sizes, and workers load the glass pieces onto rolling carts located in Area A, as shown in Figure 1. They are then transferred to a production line, which starts at Location B (see Figure 1), where double-glazed and triple-glazed glass units are

formed. The formed glass units are then loaded onto rolling carts at the end of the production line positioned at Location C, then transferred to another workstation located in Area D to begin the window framing process.

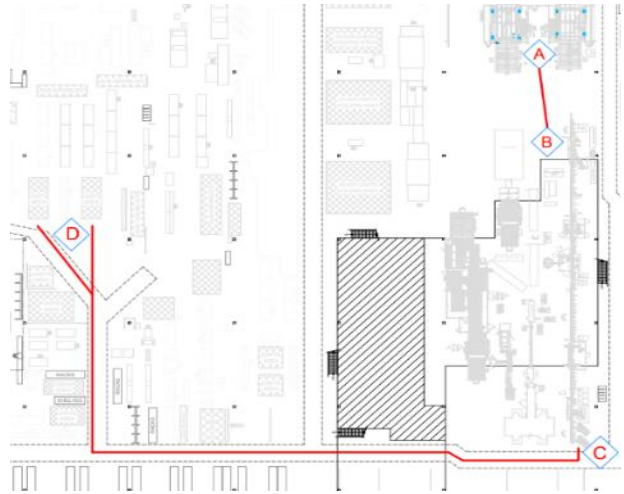


Figure 1. Material flow

The rolling carts used at the facility, shown in Figure 2, are manually transferred by workers between workstations. This manual transport of material was flagged as a notable material handling issue due to the substantial weight of the carts. This issue was highlighted through a message conveyed by the workers, as seen in Figure 3 where it required the efforts of multiple workers to handle the carts. For instance, the trip from Location C to Location D requires two to three workers to manually push fully loaded carts (35 double- and triple-glazed glass units) over a 100-meter-long route. The weight of a fully loaded cart can reach 1,102 lb, and they are manually transferred between these locations more than twenty times a day. The high physical demand associated with this task increases the risk of injury to workers and has an adverse effect on worker well-being. Moreover, this material handling practice also consumes a significant amount of labour hours, as described later in this paper.

As such, we set out to quantify the effect of the current material handling practice on transfer time and labour costs and to investigate alternative strategies that could mitigate the inefficiencies associated with the current practice. We focused in particular on the trip from Location C to Location D in the case facility, as this is material handling trip that involves the longest travel distance and the heaviest carts.

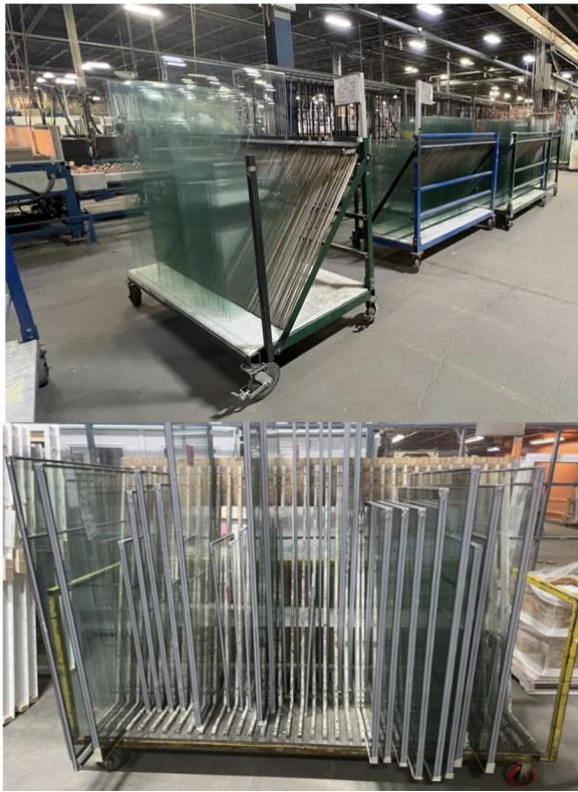


Figure 2. Rolling carts



Figure 3. Worker's note

3 Methods

This study involved four main research stages, as described in the following subsections.

3.1 Stage I: Evaluating the current practice of using fully-loaded rolling carts

The first stage consisted of identifying the different activities involved in the manufacturing process and

observing the material handling system during operation. The number of glass units that need to be transferred between workstations on a daily basis and the number of workers required to transfer them were first recorded. Next, a time study was conducted to measure the transfer time required to move the fully loaded carts along the 100 m-long route. A total of 24 data points on transfer times were collected, and these were found to accurately represent the actual durations of the recorded variables. The dataset was deemed to be sufficient in size due to the low variability in the recorded data (as indicated by the low standard deviation of 0.5 min obtained). The trip duration was measured using a stopwatch which, for each trip, was started when the cart left Location C and stopped when it arrived at Location D.

To quantify the fatigue experienced by different workers while transferring the carts, the workers were asked to indicate a fatigue score for each trip on a scale ranging from 1, representing the lowest fatigue level, to 5, representing a severe fatigue level. In this stage, the workers consistently reported the highest fatigue score due to the heavy weight of the carts fully loaded with double- and triple-glazed glass units.

3.2 Stage II: Experiments with reduced cart loads

The first alternative material handling strategy examined was to reduce the number of units loaded onto the cart to reduce its weight. Although reducing the load increases the number of trips required in order to transfer the same number of units, we hypothesized that it may also reduce (1) the number of workers needed to transfer a cart, (2) the adverse effect of manual material transfer in terms of worker fatigue, and (3) trip durations to transfer the carts. To test these hypotheses, three experiments were run with the number of glass units loaded onto the 35-slot carts reduced by increments of five for each experiment (i.e., reducing the load from 35 units to 30, 25, and 20 units for the three respective experiments). To minimize interruptions to the facility operations, each experiment was restricted to only four repetitions, thereby generating four data points per experiment (i.e., per load size), which is a limitation of the present study as discussed in the Conclusions section of this manuscript. In each experiment, the trip durations, the number of trips required, and the level of fatigue reported by workers were recorded.

3.3 Stage III: Experiments with use of power jack

Following the first two stages of analysis, the use of a motorized power jack, shown in Figure 4, to transfer glass units between workstations was tested. The power jack can load up to 4,500 lb of glass and is operated by a

single worker. The durations of four trips completed using the power jack were recorded, and the worker operating it was asked to report their fatigue level.



Figure 4. Power jack

3.4 Stage IV: Comparative cost analysis

Once the data had been collected for the four material handling alternatives (i.e., the three reduced cart load sizes and the use of a power jack), the costs associated with each strategy were calculated in order to gain a better understanding of the financial implications of each. First, the total labour hours spent on handling the same total quantity of glass units using each reduced cart load size was computed and multiplied by a sample hourly rate of \$25/h in order to compute labour costs. Then, since the introduction of a power jack entails a purchase cost, the net present value equivalent to the labour and capital costs incurred was computed for each of the four strategies (i.e., the three reduced card load scenarios and the power jack scenario) considering a minimum attractive rate of return (MARR) of 20%. This threshold was chosen because the case company does not consider any investment that does not yield at least 20% return. This analysis was mainly intended to determine whether the decision to invest in a power jack was financially sound.

4 Results and discussion

4.1 Fully loaded carts

Data was collected on 24 randomly selected, fully loaded carts transferred from Location C to Location D,

as summarized in Table 1 below. During the data collection process, four instances of transfer times (highlighted in red in Table 1) were identified as outliers due to events interrupting the transfer of the corresponding carts. Specifically, workers occasionally had to navigate slowly while pushing the cart because other workers were simultaneously pushing other materials along the same path. Consequently, they had to pause and wait until the pathway cleared. Workers also had to pause the cart transfer in some instances due to congestion of carts in the designated area, necessitating a rearrangement of the area in order to accommodate all the carts.

Most of the fully loaded carts were transferred by three workers, as they were deemed too heavy to be handled by only two workers in most cases. It took, on average, a total of 3.3 min for workers to transfer a loaded cart to Location D. Moreover, the workers consistently reported a fatigue score of 5 after transferring the carts.

Table 1. Current state results

Cart #	Transfer Time per Cart (minutes)	Cart #	Transfer Time per Cart (minutes)
1	3.5	13	3.0
2	4.4	14	3.3
3	3.3	15	3.2
4	3.1	16	3.1
5	3.2	17	4.5
6	3.2	18	3.4
7	4.5	19	3.1
8	3.4	20	5.1
9	3.4	21	3.1
10	3.5	22	3.3
11	4.2	23	3.0
12	3.0	24	4.4

Although the time spent on a given material handling trip from one location to another may be insignificant, the cumulative time spent on material handling, and the potential savings associated with strategies to reduce the trip duration, can be significant considering how many material handling trips occur in a given day. To demonstrate this, let us consider the total time and labour hours needed to transfer twenty carts, containing 35 glass units each, which is a commonly sized glass batch scheduled for production on a given day. The average total time and labour hours can be calculated by using Eq (1) below. The total time and labour hours needed to transfer twenty carts amount to 1.1 h and 3.3 h, respectively. A total transfer time of about one hour per day is equivalent to more than 10% of an 8-h work shift

spent by each of the three workers on manually transferring heavy carts—a significant amount of non-value-added work.

$$\begin{aligned}
 & \text{Estimate of total transfer time per day} \\
 & = \text{average transfer time per cart} \\
 & \times \text{carts per day} = 3.3 \frac{\text{minutes}}{\text{cart}} \times 20 \text{ carts} \\
 & = 66 \text{ minutes} \\
 & = 1.1 \text{ hours} \times (\sim 3 \text{ workers per trip}) \\
 & \approx 3.3 \text{ labour hours}
 \end{aligned} \quad (1)$$

4.2 Partially loaded carts

To evaluate the effect of the load size on the number of carts required in order to transfer a given batch of glass from Location C to Location D, a batch size of 700 glass units was considered. The total number of carts necessary to transfer this batch was then computed satisfying Eq. (2). The total transfer time and labour hours were also computed, satisfying Eq. (1). It was found that, for a load size of 30 units, the number of required carts would total 23.3 based on Eq. (2). In this case, 23 carts would be loaded with 30 units each and one cart would be loaded with the remaining ten units. The total transfer time corresponding to the 23 carts was calculated satisfying Eq. (1) using the average transfer time per cart loaded with 30 units. As for the time needed to transfer the remaining cart containing ten units, since the relevant data was missing, it was assumed to be equal to the minimum time recorded for pushing a cart containing 20 units during the present time study, which was measured at 1.7 min as shown in Table 2. As such, the total time needed to transfer the 24 carts was computed by adding 1.7 min to the transfer time computed using Eq. (1). The results are summarized in Table 2.

$$\begin{aligned}
 & \text{Number of required carts} \\
 & = \frac{700 \text{ glass units/day}}{\text{Load size (units/cart)}}
 \end{aligned} \quad (2)$$

As the presented results show, decreasing the cart load size reduced the number of required workers from about three workers for the carts with 35 glass units to two for carts with 30, 25, and 20 glass units. The average transfer time also decreased, with the lowest total transfer time of 61.6 min and total labour hours of 2.1 h recorded for carts with 25 glass units. Reducing the load size from 35 units per cart to 25 units per cart reduced the total transfer time by about 7% (i.e., increased daily productivity by about 7%) and the total labour hours by

about 38%. Further reducing the load size to 20 units per cart, on the other hand, increased the total transfer time and total labour hours, as the increase in the number of required trips outweighed the reduction in transfer time per cart. These results are reasonable, considering that walking freely (without pushing a cart) from Location C to Location D at the factory takes about 1.5 min, meaning that further reducing the load size will not have a significant effect on the transfer time, as it approaches the average walking time needed to travel between the two locations.

Table 2. Results from reducing cart weights

# Glass units per cart	35	30	25	20
Transfer time per cart (min)	-	2.8	2.3	1.8
		2.7	2.3	1.7
		2.9	2.2	2.2
		2.8	2.1	2.1
Average transfer time per cart (min)	3.3	2.8	2.2	2
# Needed workers	3	2	2	2
Average fatigue score	5	5	4	2.3
# Trips per day	20	24	28	35
Total transfer time per day (min)	66	66.1	61.6	70
Labour hours per day	3.3	2.2	2.1	2.3

The average fatigue score dropped from 5.0 for carts loaded with 35 and 30 units to 4.0 for carts loaded with 25 units, and further dropped to 2.3 for carts loaded with 20 units. The fatigue score remained high when the load size was reduced to 30 units because there were two workers transferring the carts rather than the original three. Hence, while this initial reduction in load size was beneficial in reducing the total labour hours, there was no discernible benefits in terms of alleviating worker fatigue. Decreasing the load size by another five units, though, did result in a reduction the average fatigue score to 4.0. While further reducing the load size to 20 units adversely affected transfer time and labour hours, it significantly reduced the average fatigue score. In this regard, a reduction in load size that may adversely affect productivity may nevertheless be considered preferable as a means of further mitigating the adverse effects that manually transferring heavy carts has on the health of workers.

4.3 Power jack

Following experimentation with the different partial

load sizes, the use of a power jack to transfer fully loaded carts, each containing 35 glass units, from Location C to Location D, was introduced. Although the purpose of this strategy was to reduce manual input in material handling, it was more feasible to remove the cart from the power jack and manually maneuver it to place it at its intended position, as the drop-off area is often congested with other carts. Despite this limitation, the data presented in Table 3 shows the promising results obtained for this material handling strategy. On average, using a power jack reduced the transfer time for each cart by about half. The effect on the total labour hours was also significant, since only a single worker is needed to operate the power jack. The total labour hours decreased by about 85%, from 3.3 h in the case of manually handled carts to 0.5 h in the case of carts transferred using the power jack. The average fatigue score also significantly decreased from 5.0 to 2.0, but it did not drop to 1 since the worker must manually maneuver the cart at the drop-off location (as noted above). Overall, the use of a power jack was found to be the most attractive material handling strategy in terms of time efficiency and the well-being of workers. Further analysis was conducted in order to better understand the financial implications of investing in this equipment, as described in the following section.

Table 3. Power jack results

Cart handling strategy	Manual		Power Jack			
# Glass units per cart	35		35			
Transfer time per cart (min)	-	1.3	1.5	1.4	1.7	
Average transfer time per cart (min)	3.3	1.5				
# Needed workers	3	1				
Average fatigue score	5	2				
# Trips per day		20				
Total transfer time per day (min)	66	30				
Labour hours per day	3.3	0.5				

4.4 Cost analysis

The labour costs associated with each material handling strategy and the respective net present value equivalent to twelve months of labour costs were computed. The results obtained for a MARR of 20% are presented in Table 4. Notably, the expenditures on transferring carts from Location C to Location D alone amount to \$1,815 each month. It is also remarkable that

a mere adjustment in the load size of these carts could lead to monthly labour cost savings of over \$800 for the company. The power jack strategy was found to have the lowest net present value among the tested strategies, with a recorded value of \$10,177, resulting in cost savings of 48% when compared to current practice. The strategy of using a cart load size of 25 units was also found to be financially attractive, where a 36% reduction in costs could be realized without investing in new equipment. Still, despite requiring initial capital investment, the power jack strategy stands out as the preferable choice due to its significant time and cost savings coupled with the reduction in the level of fatigue experienced by workers.

Table 4. Cost results

Cart handling strategy	Ma-nual	Manual			Power Jack
# glass units per cart	35	30	25	20	35
Capital investment	-	0	0	0	\$7,210
Monthly labour costs	\$1,815	\$1,200	\$1,100	\$1,200	\$275
Net present value	\$19,584	\$13,056	\$12,462	\$13,649	\$10,177
Cost savings	-	33%	36%	30%	48%

To better judge the financial benefits of the different strategies, though, the capital investment cost of the power jack needed to be taken into consideration. Using the power jack results in monthly labour cost savings of \$1,540. Considering the initial purchase price of \$7,210, the number of periods (N) needed to recover the cost of the power jack can be calculated using Eq. (3) as follows.

$$PV = C + \frac{1 - ((1 + r)^{-N})}{r} \times S \quad (3)$$

$$0 = -7,210 + \frac{1 - ((1 + 0.0167)^{-N})}{0.0167} \times 1,540$$

where:

- PV is the present value, set to 0 in order to determine the breakeven point at which the case company would recover the purchase cost of the power jack;
- C is the initial capital investment, which is the purchase price of the power jack;
- r is the monthly discount rate; and

- S is the monthly savings in labour costs realized using the power jack.

Based on Eq. (3), it would take the case company approximately five months to recoup the initial investment of \$7,210. Beyond the breakeven point, the company would begin realizing monthly savings of \$1,540 relative to the base scenario cost of \$1,815 (i.e., manually pushing fully loaded carts from Location C to Location D).

5 Conclusions

This study examined the influence of MMH on time and worker fatigue in glass manufacturing. Various strategies of transferring glass between workstations were investigated to evaluate the influence MMH has on these factors. These strategies included reducing the number of glass units per cart and utilizing a power jack.

The current practice involved workers dedicating a considerable amount of time to transferring carts, resulting in fatigue and increased labour costs. Among the alternatives examined, three demonstrated promising results: using a cart with 25 glass units instead of 35, using a cart with 20 glass units, and using a power jack. Among the three reduced load size scenarios considered, the 25-unit option proved to be the most economical, while the 20-glass option led to the least fatigue for workers. As for the power jack, it significantly reduced both the time spent on transferring carts and worker fatigue. However, the decision to adopt this solution ultimately hinges on the company's willingness to introduce new equipment and the anticipated return on investing in it. In this regard, the cost analysis results show that the case company would recoup the purchase costs within five months, while also saving costs, improving productivity, and improving the well-being of workers in the long term.

Perhaps the most notable limitation of the present study was the limited size of the dataset used to evaluate the alternative material handling strategies. To address this limitation, average transfer time per cart was used to estimate the total transfer time corresponding to a batch of glass units. In future research, alternative strategies could be tested over the course of days in order to obtain more accurate transfer times and labour cost data. For prolonged data acquisition periods, automated technologies like radio frequency identification (RFID) may prove more efficient compared to manual data collection methods. Further exploration could encompass additional analysis factors such as (1) occupational safety, (2) ergonomics of worker movement, and (3) training needs for adopting new technology in material handling processes. Finally, given the labour costs expended on material handling, future research could examine the prospect of implementing fully automated material

handling systems such as automated guided vehicles to eliminate the reliance on human labour.

6 References

- [1] M. G. Kay, "Material Handling Equipment," 2012. doi: https://mgkay.github.io/Material_Handling_Equipment.pdf.
- [2] Ontario.ca, "Manual materials handling," Ontario.ca. Accessed: May 15, 2023. [Online]. Available: <https://www.ontario.ca/document/manual-materials-handling#>
- [3] Canadian Center of Occupational Health and Safety, "Manual Materials Handling (MMH)." Accessed: May 15, 2023. [Online]. Available: <https://www.ccohs.ca/oshanswers/ergonomics/mmh/>
- [4] D. Health and Safety Authority 10 Hogan Place, "Guidance on the Management of Manual Material Handling in the Workplace, Health and Safety Authority," 2005.
- [5] P. G. Dempsey and L. Hashemi, "Analysis of workers' compensation claims associated with manual materials handling," *Ergonomics*, vol. 42, no. 1, pp. 183–195, Jan. 1999, doi: 10.1080/001401399185883.
- [6] H. W. Hendrick, "Determining the cost-benefits of ergonomics projects and factors that lead to their success," *Appl Ergon*, vol. 34, no. 5, pp. 419–427, 2003, doi: [https://doi.org/10.1016/S0003-6870\(03\)00062-0](https://doi.org/10.1016/S0003-6870(03)00062-0).
- [7] J. Liker, *The Toyota Way*, 14 Management Principles from the World's Greatest Manufacturer, 1st ed., vol. 11. New York: McGraw-Hill, 2004.
- [8] N. P. Wangari, N. P. Muchiri, and D. O. Onyancha, "Impact Assessment of a Facility Layout on Manufacturing Cycle Time and Throughput: A Case Study of a Tannery in Central Kenya," 2018.
- [9] J. A. Tompkins, J. A. White, Y. A. Bozer, and J. M. A. Tanchoco, *Facilities planning*, 4th ed. John Wiley & Sons, Inc., 2014.
- [10] S. T. Yang, M. H. Park, and B. Y. Jeong, "Types of manual materials handling (MMH) and occupational incidents and musculoskeletal disorders (MSDs) in motor vehicle parts manufacturing (MVPM) industry," *Int J Ind Ergon*, vol. 77, p. 102954, May 2020, doi: 10.1016/J.ERGON.2020.102954.
- [11] Canadian Center of Occupational Health and Safety, "Manual Materials Handling (MMH)." Accessed: May 14, 2023. [Online]. Available: <https://www.ccohs.ca/oshanswers/ergonomics/mmh/>

- [12] A. D. Nimbarte, Y. Sun, M. Jaridi, and H. Hsiao, "Biomechanical loading of the shoulder complex and lumbosacral joints during dynamic cart pushing task," *Appl Ergon*, vol. 44, no. 5, pp. 841–849, 2013, doi: <https://doi.org/10.1016/j.apergo.2013.02.008>.
- [13] A. Dormohammadi, H. Amjad-Sardrudi, M. Motamedzade, R. Dormohammadi, and S. Musavi, "Ergonomics intervention in a tile industry: A case of manual material handling," *J Res Health Sci*, vol. 12, pp. 109–113, Mar. 2012.
- [14] T. Ramsey, "The Effects of Load-Positioning Material Handling Equipment on Spinal Loading During Manual Handling of Bulk Bags," 2013. [Online]. Available: <https://api.semanticscholar.org/CorpusID:115404026>
- [15] B. Dagnew, "Design of manual material handling system through computer aided ergonomics : A case study at BDTSC textile firm" *International Journal for Quality Research*, vol. 8, no. 4, pp. 557–568, 2014.

Impact of Technology Use on Time Needed for Information Retrieval for Frontline Supervisors in the Construction Industry

Bassam Ramadan¹, Hala Nassereddine¹, Timothy Taylor¹, Paul Goodrum²

¹Department of Civil Engineering, University of Kentucky, USA

²Department of Construction Management, Colorado State University, USA

bara235@uky.edu, hala.nassereddine@uky.edu, tim.taylor@uky.edu, paul.goodrum@colostate.edu

Abstract –

In the construction industry, information is considered to be the lifeblood of modern construction projects. As a data-dependent industry, information retrieval and technological innovations are essential for optimizing construction performance. The construction industry is currently experiencing rapid adoption of new technologies and innovations. Existing research highlights the positive impact of technological use on construction sites, particularly on performance and information access. However, no research has yet evaluated how technological adoption impacts the time needed to receive or gain access to requested information. The objective of this paper is to analyze and understand the impact of technological adoption on the time needed to receive or gain access to requested information among frontline supervisors in the construction industry. To achieve the research objective, 1138 construction frontline supervisors were surveyed using an online questionnaire. The participants in this survey were asked to specify which technologies are being used on their construction sites; and evaluate the average time needed to receive or gain access to requested information. Key findings show that the adoption of six out of 13 technologies had a statistically significant positive impact in decreasing the time needed to receive or gain access to requested information among frontline supervisors in the construction industry.

Keywords –

Information Access; Technology Use; Frontline Supervisors.

1 Introduction & Background

The United States' aging infrastructure is urgently in need of renovations amid a significant shortage of labor and skilled craft workers [1]. Recent trends highlight a rapid aging of the construction workforce, contributing to

an industry-wide labor shortage [2,3]. One study found that labor-related challenges are among the top challenges contractors currently face [4]. Forecasts by industry leaders and experts still suggest a substantial transformation in the construction sector over the next two decades [5]. Despite the industry's need for advancements, construction technologies are not being implemented at a pace matching their development, preventing the sector from fully realizing their potential [1]. According to a McKinsey & Company analysis of Venture Capitalist (VC) investments data found that VC investment growth in construction technologies has surged approximately 15-fold compared to other industries, with no apparent signs of this momentum slowing down [6].

In the contemporary construction industry, information is regarded as the lifeblood of any construction project [7]. Given the extensive amount of knowledge generated across various phases of the construction project lifecycle, effective information management becomes indispensable, particularly due to the compartmentalized nature of the industry [8,9]. The industry is experiencing a noteworthy shift from traditional paper-based systems to digital information and e-construction [10,11]. Recognizing the substantial benefits, adequate information access has been identified as a catalyst for improvement in construction safety and performance [12,13]. Moreover, the integration of information access and technologies has been shown to enhance both external and internal collaboration, communication, and employee satisfaction [14]. In addition to these advantages, effective information access and management contribute to elevated work quality, simplification of challenging tasks, and increased worker productivity [15]. Organizational information management further streamlines the dissemination of critical information across related projects, potentially facilitating on-time project completion. Within construction teams, it fosters communication between workers and supervisors, promoting coordination,

collaboration, and effective decision-making [16].

Extensive research has been conducted on Construction 4.0 technologies [17], including recent advancements in blockchain technologies [18–20]. The existing body of literature underscores the versatile benefits of these construction technologies. Notably, a study highlighted that the adoption of diverse information technologies enhances information access for the construction workforce [7]. For instance, the efficiency of object identification and information collection is significantly improved through barcode scanning technology [21]. Augmented or virtual reality has proven instrumental in real-time project visualization and data collection, contributing to heightened productivity, improved quality, and enhanced communication and collaboration [22]. Radio Frequency Identification (RFID) chip tracking optimizes information collection and management throughout different project phases by efficiently tracking construction materials, resources, equipment, components, and systems [23]. The utilization of Building Information Modeling (BIM) facilitates faster, more efficient, and more accurate planning and construction. It enables simultaneous contributions from all project stakeholders and promotes information sharing [10,24–26]. Artificial intelligence has furthered the advancement of services and business operations, enhancing the automation processes of companies to gain a competitive advantage [27]. Drones, which have been proven to be contributors to overall cost reduction and fewer project delays, have improved safety records and provided high-resolution aerial imagery for more accurate data collection and surveying [28]. The utilization of robotics and autonomous machinery in construction processes has streamlined operations, leading to significant reductions in costs and time spent on tasks. Additionally, these technologies have proven effective in enhancing the overall quality of the executed product [29]. Prefabrication/ modularization techniques enable faster on-site construction, diminish on-site work requirements, enhance quality, lower energy consumption and emissions, and ultimately reduce overall construction costs [30]. Finally, quick connection systems have been identified as effective in reducing work hours, minimizing the number of required work packages, improving quality, and facilitating construction processes [31].

Within the construction industry, there persists a prevailing notion that insufficient productivity stems from inadequate performance by craft workers. In reality, the root of such issues is more likely attributed to inadequate supervision and the failure of frontline supervisors to provide essential planning, information, support, and motivation [32]. Frontline supervisors bear the responsibility of ensuring a safe and healthy

workplace while serving as a crucial communication link between management and the craft workforce [33–35]. Numerous studies extensively underscore the pivotal role and influence that frontline supervisors wield in determining construction efficiency [36–39].

While the existing literature has assessed and emphasized the benefits of technology, and research has studied the impact of several factors on performance and the difficulty of information access on construction sites, including workforce skills [40], workforce training [41], and crew diversity [42], no research has yet studied how technology use on construction sites impacts the time needed to retrieve or gain access to needed information. Construction frontline supervisors' prompt access to project information is critical for their ability to do their job. Equipping construction frontline supervisors with tools that help decrease the time needed to receive or gain access to information would improve overall construction performance and efficiency. The goal of this study is to bridge this gap of knowledge using a questionnaire of construction frontline supervisors that investigates technology use and the time needed to retrieve or gain access to needed information. This paper aims to analyze and understand the impact of construction on-site technology adoption on the time needed to retrieve or gain access to needed information among frontline supervisors in the construction industry.

2 Methodology

The objective of this paper is to analyze and understand the impact of construction on-site technology adoption on the time needed to retrieve or gain access to needed information among frontline supervisors in the construction industry. To accomplish the research objective, an online questionnaire was created using "Qualtrics" and disseminated to construction frontline supervisors in the United States. The survey includes questions about the time needed to receive or gain access to needed information; and enquires whether specific technologies are employed on their construction sites. The survey was developed based on a thorough review of the literature, where gaps in research were identified and addressed. The questions were reviewed and approved by the Internal Review Board (IRB). A total of 1138 responses were received from participants across all 50 states. States including New York, California, Texas, Pennsylvania, and Illinois were among the biggest contributors to the survey in terms of respondents. The geographic distribution of the respondents is presented in Figure 1.

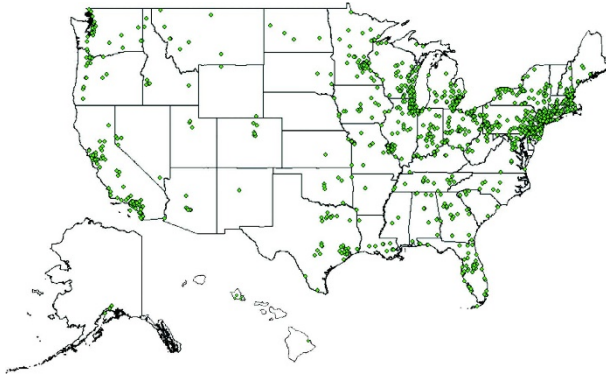


Figure 1. Geographic distribution of the survey respondents

The responses have an overall gender distribution of 97.5% male and 2.0% female. Among the supervisors, 31% hold the title of foremen, 29% are superintendents, 17% are general foremen, 2% are craft superintendents, 2% are assistant superintendents, and 19% indicated having another title. The breakdown of the survey respondents by job title is presented in Figure 2.

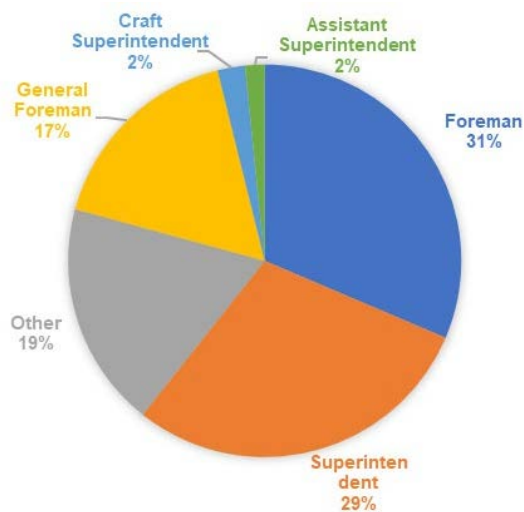


Figure 2. Survey respondent's breakdown by job title

Regarding age distribution, 33.2% of respondents are aged over 55, 33.2% fall within the 45-54 age group, 25.3% are in the 35-44 age range, 8.0% are in the 25-34 age group, and 0.3% are below the age of 25. The breakdown of the survey respondents by age is presented in Figure 3.

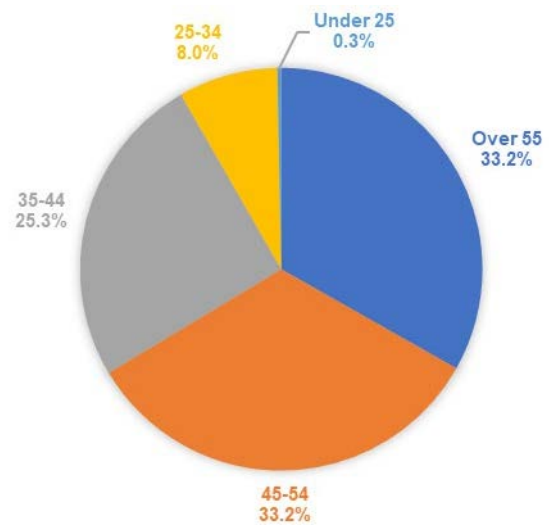


Figure 3. Survey respondent's breakdown by age

In this research, participants were prompted to choose from a list of information technologies employed on their construction sites. The available options for selection included RFID Chip Tracking, Virtual or Augmented Reality (VR or AR), Artificial Intelligence, Barcode Scanning, Building Information Modeling (BIM), Drones, and Robotics, Prefabrication/Modularization, Quick Connection systems, Autonomous Machinery, Battery Powered Tools, and New types of Hand Tools or Construction Machinery developed in the last five years. Respondents also had the option to specify any additional "Other" technologies they utilize on their construction sites.

Finally, the survey asked participants the two following questions relating to the level of difficulty to gain access to information and their performance record:

- If you need access to project information for your work, how long does it typically take (in Days and Hours) to receive this information or get access to it?

For this question, a lower numerical value on the response to the time needed to receive or get access to required information indicates a more positive outcome.

To compare the time needed to receive or get access to required information for frontline supervisors based on whether technologies are used on construction sites, the data was grouped based on whether each specific technology is or not. The corresponding average of the time needed to receive or get access to the required information of each of these groups is calculated and compared.

To assess the statistical significance of the difference between the two groups, the non-parametric Mann-Whitney U-test was employed to obtain the p-value. This

choice was made to adopt a conservative approach in statistical analysis, given substantial variations in the sample sizes of the two compared groups [43]. The student's t-test was not utilized due to the significant differences in sample sizes between the groups. When sample sizes in both conditions are equal, the t-test is very robust against unequal variances. However, if sample sizes are unequal, unequal variances can influence the Type 1 error rate of the students' t-test by either increasing or decreasing the Type 1 error rate from the nominal alpha significance level. In such instances, the Mann-Whitney U-test demonstrates better performance and behavior than the t-test [43]. Hence, the Mann-Whitney U-test was chosen for this statistical analysis. A significance level, α , of 0.1 was considered for statistical significance, corresponding to a 90% confidence level.

3 Results & Analysis

3.1 Impact of Technology Use on Time Needed to Receive or Access Needed Information

The impact of technology use on the time needed to receive or access needed information based on whether the frontline supervisors indicated each specific technology is used on their construction is shown in Table 1. The table presents the average responses for the difficulty of information access and the average time needed (in Days) to receive or access needed information for respondents who specified whether each technology is used or not. The table additionally shows the p-value resulting from the Mann-Whitney U-test, aiming to assess the statistical significance of any observed differences in averages.

The results in Table 1 show that the on-site use of only four out of the 13 technologies resulted in a higher average time needed to receive or access needed information for construction frontline supervisors. However, the results are statistically significant for only two of those technologies, including *Artificial Intelligence*, and *New Hand Tools developed in the last five years*. While artificial intelligence is a new emerging technology with significant potential, it remains a sophisticated technology that is rarely used on construction. Potentially, lack of training on such complicated technologies can be a contributing factor as to why it has not caused a positive contribution to the average time to access information.

On the other hand, the on-site use of nine out of the 13 analyzed technologies resulted in a decrease in the average time needed to receive or access needed information for construction frontline supervisors. However, only six of those technologies showed results that were statistically significant. The technologies that

have statistically improved the average time needed to receive or access needed information include *Barcode Scanning*, *RFID Chip Tracking*, *Building Information Modeling*, *Drones*, *Prefabrication/ Modularization*, and *Quick connection systems*. While *Virtual or Augmented Reality* and *Battery-Powered tools* did have a slight positive impact, the difference in the results was not statistically significant between those two groups.

Table 1. Average time needed to receive or access needed information based on technology use among construction frontline supervisors (in Days)

Technology	Used	Not Used	P-value
Barcode Scanning	0.86	1.25	0.00*
Virtual or Augmented Reality	1.07	1.16	0.31
RFID Chip Tracking	1.06	1.16	0.02*
Building Information Modelling	1.09	1.18	0.01*
Artificial Intelligence	1.61	1.14	0.09*
Robotics	1.15	1.15	0.43
Drones	0.88	1.19	0.02*
Prefabrication/ Modularization	1.13	1.17	0.09*
Quick Connection Systems	1.01	1.18	0.09*
Autonomous Machinery	1.54	1.14	0.16
New Hand Tools	1.27	1.12	0.03*
Battery Powered Tools	1.13	1.30	0.23
New Type of Construction Machinery	1.34	1.12	0.12

*Difference in averages is statistically significant at the 90% level

4 Conclusion, Limitations, and Future Studies

Over the last decade, technological advancements have played a crucial role in fostering growth and progress within a rapidly transforming construction industry. While the existing body of literature has extensively examined construction technologies and information access, no existing work has assessed its direct impact on the time needed to receive or gain access to needed information among construction frontline supervisors. The objective of this paper is to examine the impact of on-site technology use on the average time needed to receive or gain access to needed information. This research used data from a survey of 1138 construction frontline supervisors. The statistical analysis of this data showed that there are substantial benefits for on-site technology use in terms of information access. This study found that for the construction frontline supervisors, on average, on-site technology use has a positive impact in decreasing the time needed to receive or access needed information when eight out of the 13 technologies are used, six of which had results that were statistically significant.

While this research presents results of a robust statistical analysis of the impact of technology use on the average time needed to receive or gain access to needed information, the study does have certain limitations. The survey of construction frontline supervisors doesn't ask any open-ended or multiple-choice questions that discuss specific benefits, challenges, or factors that resulted in enhanced access to information. Consequently, although this analysis empirically measures a positive impact on performance and information access, it cannot address the "why" or "how" behind the improvements resulting from technology use. Future research endeavors can build upon this study, aiming to answer these specific questions using structured interviews and focus groups with construction workers and industry experts, thus, providing a roadmap for construction industry leaders. This roadmap could help identify paths that maximize potential benefits tailored to the unique needs of specific construction projects.

5 Acknowledgments

The authors acknowledge and extend their gratitude to the Construction Industry Institute (CII) for their invaluable support and partial funding that facilitated the data collection for this research project. Continuous support from the College of Engineering at the University of Kentucky is also acknowledged. The authors express their appreciation to all the survey participants, whose contribution is indispensable to the success of this research. Any opinions, findings, conclusions, and

recommendations presented by the authors in this paper do not necessarily represent the views of the University of Kentucky or the Construction Industry Institute.

References

- [1] Construction Industry Institute: *Workforce 2030: What You Need to Know Now About Your Future Workforce*. Austin, TX, US: Construction Industry Institute
- [2] Construction Industry Institute: *Is There a Demographic Labor Cliff that Will Affect Project Performance*. Austin, TX: Construction Industry Institute
- [3] Ramadan BA.; Taylor TRB.; Real KJ.; Goodrum P: The Construction Industry from the Perspective of the Worker's Social Experience. *Construction Research Congress 2022*, Arlington, Virginia: American Society of Civil Engineers, 611–21, 2022
- [4] Ramadan B.; Nassereddine H.; Taylor T: Workforce Challenges and Strategies of Top Contractors in the United States. *Proceedings of the Creative Construction Conference 2023*, Keszthely, Hungary, 385–90, 2023
- [5] Ribeirinho MJ.; Mischke J.; Strube G.; Sjödin E.; Blanco JL.; Palter R.; Biörck J.; Rockhill D.; Andersson T: *The Next Normal in Construction: How disruption is reshaping the world's largest ecosystem*. McKinsey & Company
- [6] Bartlett K.; Blanco JL.; Fitzgerald B.; Johnson J.; Ribeirinho MJ: Rise of the platform era: The next chapter in construction technology. *McKinsey & Company*, 2020. Available at: <https://www.mckinsey.com/industries/private-equity-and-principal-investors/our-insights/rise-of-the-platform-era-the-next-chapter-in-construction-technology>
- [7] Ramadan B.; Nassereddine H.; Taylor TRB.; Goodrum P: Impact of Technology Use on Workforce Performance and Information Access in the Construction Industry. *Frontiers in Built Environment*, 2023. Doi: 10.3389/fbuil.2023.1079203
- [8] Dave B.; Koskela L: Collaborative knowledge management—A construction case study. *Automation in Construction* 18(7):894–902, 2009. Doi: 10.1016/j.autcon.2009.03.015
- [9] Kazi AS.; Koivuniemi A: Sharing Through Social Interaction: The Case of YIT Construction Ltd. *Real-Life Knowledge Management: Lessons from the Field* (952-5004-72-4):63–80, 2006
- [10] Dadi GB.; Nassereddine H.; Taylor TR.; Griffith R.; Ramadan B: *Technological Capabilities of Departments of Transportation for Digital Project*

- Management and Delivery*. Washington, D.C.: Transportation Research Board
- [11] Nassereddine H.; Ramadan B.; Li Y.; Sturgill R.; Patel P: *Practices for the Collection, Use, and Management of Utility As-Built Information*. Washington, D.C.: Transportation Research Board
- [12] Al-Shabbani Z.; Ammar A.; Dadi G: Preventative Safety Metrics with Highway Maintenance Crews. *Construction Research Congress 2022*, Arlington, Virginia: American Society of Civil Engineers, 510–9, 2022
- [13] Ammar A.; Al-Shabbani Z.; Dadi G: Assessing the Safety Climate of a State Department of Transportation through Its Highway Maintenance Crews. *Construction Research Congress 2022*, Arlington, Virginia: American Society of Civil Engineers, 345–55, 2022
- [14] Kline R.; Dolenc M.; Turk Ž: Possible Benefits of WEB 2.0 to Construction Industry:10, 2008
- [15] Vasista TG.; Abone A: Benefits, Barriers and Applications of Information Communication Technology in Construction Industry: a Contemporary Study. *IJET* 7(3.27):492–9, 2018. Doi: 10.14419/ijet.v7i3.27.18004
- [16] Prasanna SVSNDL.; Raja Ramanna T: Application of ICT benefits for building project management using ISM Model. *International Journal of Research in Engineering and Technology* 3(6):68–78, 2014
- [17] Forcael E.; Ferrari I.; Opazo-Vega A.; Pulido-Arcas JA: Construction 4.0: A Literature Review. *Sustainability* 12(22):9755, 2020. Doi: 10.3390/su12229755
- [18] Li J.; Greenwood D.; Kassem M: Blockchain in the built environment and construction industry: A systematic review, conceptual models and practical use cases. *Automation in Construction* 102:288–307, 2019. Doi: 10.1016/j.autcon.2019.02.005
- [19] Sadeghi M.; Mahmoudi A.; Deng X: Adopting distributed ledger technology for the sustainable construction industry: evaluating the barriers using Ordinal Priority Approach. *Environ Sci Pollut Res* 29(7):10495–520, 2022. Doi: 10.1007/s11356-021-16376-y
- [20] Sadeghi M.; Mahmoudi A.; Deng X.; Luo X: Prioritizing requirements for implementing blockchain technology in construction supply chain based on circular economy: Fuzzy Ordinal Priority Approach. *International Journal of Environmental Science and Technology*:1–22, 2022
- [21] Yan Y.; Li Q.; Cao M.; Chen H.; Xue J: Application research of two-dimensional barcode in information construction of colleges. *Proceedings of the 2012 International Conference on Information Technology and Software Engineering*, Springer, 71–80, 2013
- [22] Nassereddine.; Schranz C.; Bou Hatoum M.; Urban H: A Comprehensive Map for Integrating Augmented Reality During the Construction Phase. *Proceedings of the Creative Construction e-Conference 2020*, Online: Budapest University of Technology and Economics, 56–64, 2020
- [23] Valero E.; Adán A.; Cerrada C: Evolution of RFID Applications in Construction: A Literature Review. *Sensors* 15(7):15988–6008, 2015. Doi: 10.3390/s150715988
- [24] Guo F.; Jähren CT.; Turkan Y.; David Jeong H: Civil Integrated Management: An Emerging Paradigm for Civil Infrastructure Project Delivery and Management. *J Manage Eng* 33(2):04016044, 2017. Doi: 10.1061/(ASCE)ME.1943-5479.0000491
- [25] Patel P.; Sturgill R.; Nassereddine H.; Ramadan B.; Li Y: Current Benefits of ASCE 75 and its Potential to Affect Digital As-Built Initiatives at State Departments of Transportation. *Transportation Research Record: Journal of the Transportation Research Board*, 2023. Doi: 10.1177/03611981231182961
- [26] Torres HN.; Ruiz JM.; Chang GK.; Anderson JL.; Garber SI.; others *Automation in highway construction part I: Implementation challenges at state transportation departments and success stories*. United States. Federal Highway Administration. Office of Infrastructure ...
- [27] Abioye SO.; Oyedele LO.; Akanbi L.; Ajayi A.; Davila Delgado JM.; Bilal M.; Akinade OO.; Ahmed A: Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges. *Journal of Building Engineering* 44:103299, 2021. Doi: 10.1016/j.jobe.2021.103299
- [28] McGuire M.; Rys MJ.; Rys A.; others A study of how unmanned aircraft systems can support the Kansas Department of Transportation's efforts to improve efficiency, safety, and cost reduction, 2016
- [29] Prasath Kumar VR.; Balasubramanian M.; Jagadish Raj S: Robotics in Construction Industry. *Indian Journal of Science and Technology* 9(23), 2016. Doi: 10.17485/ijst/2016/v9i23/95974
- [30] Lopez D.; Froese TM: Analysis of Costs and Benefits of Panelized and Modular Prefabricated Homes. *International Conference on Sustainable Design, Engineering and Construction* 145:1291–7, 2016. Doi: 10.1016/j.proeng.2016.04.166
- [31] Shan Y.; Goodrum P.; Haas C.; Caldas C: Assessing Productivity Improvement of Quick

- Connection Systems in the Steel Construction Industry Using Building Information Modeling (BIM). *Construction Research Congress 2012*, West Lafayette, Indiana, United States: American Society of Civil Engineers, 1135–44, 2012
- [32] Howell GA: What is Lean Construction? IGLC7. *University of California, Berkeley, CA, USA*, 1999
- [33] Oswald D.; Lingard H: Development of a frontline H&S leadership maturity model in the construction industry. *Safety Science* 118:674–86, 2019. Doi: <https://doi.org/10.1016/j.ssci.2019.06.005>
- [34] Ramadan B.; Nassereddine H.; Taylor TRB.; Real K.; Goodrum P: Impact of Skill Proficiencies on Frontline Supervision Practices in the Construction Industry. *Proceedings of the Creative Construction e-Conference (2022)*, 2022
- [35] Uwakweh BO: Effect of Foremen on Construction Apprentice. *J Constr Eng Manage* 131(12):1320–7, 2005. Doi: 10.1061/(ASCE)0733-9364(2005)131:12(1320)
- [36] Gerami Seresht N.; Fayek AR: Factors influencing multifactor productivity of equipment-intensive activities. *IJPPM* 69(9):2021–45, 2019. Doi: 10.1108/IJPPM-07-2018-0250
- [37] Hewage KN.; Gannoruwa A.; Ruwanpura JY: Current status of factors leading to team performance of on-site construction professionals in Alberta building construction projects. *Can J Civ Eng* 38(6):679–89, 2011. Doi: 10.1139/111-038
- [38] Liberda M.; Ruwanpura J.; Jergeas G: Construction Productivity Improvement: A Study of Human, Management and External Issues. *Construction Research Congress*, Honolulu, Hawaii, United States: American Society of Civil Engineers, 1–8, 2003
- [39] Ramadan B.; Nassereddine H.; Taylor TRB.; Goodrum P: Impact of Technology Use on Frontline Supervision Practices in the Construction Industry. *2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, Istanbul, Turkiye: IEEE, 2023
- [40] Ramadan B.; Nassereddine H.; Taylor TRB.; Goodrum P: Impact of Administrative and Computer Skill Proficiency on Workforce Performance and Information Access in the Construction Industry. *2023 Canadian Society of Civil Engineers Annual Conference*, 2023
- [41] Ramadan B.; Nassereddine H.; Taylor T.; Goodrum P: Impact of Workforce Training on Worker Performance in the Construction Industry. *Proceedings of the Creative Construction Conference 2023*, Keszthely, Hungary, 319–24, 2023
- [42] Ramadan B.; Nassereddine H.; Taylor T.; Goodrum P: Impact of Crew Diversity on Worker Information Access and Performance. *Proceedings of the 40th International Symposium on Automation and Robotics in Construction*, Chennai, India: International Association for Automation and Robotics in Construction (IAARC), 309–16, 2023
- [43] Gibbons JD.; Chakraborti S: Comparisons of the Mann-Whitney, Student's *t*, and Alternate *t* Tests for Means of Normal Distributions. *The Journal of Experimental Education* 59(3):258–67, 1991. Doi: 10.1080/00220973.1991.10806565

Digital Info Screen - A Visual Management Tool for Construction Workers

Markku Kiviniemi¹, V. Paul C. Charlesraj², Marja Liinasuo¹, Pekka Siltanen¹ and Kari Rainio¹

¹ VTT Technical Research Centre of Finland, Finland

² School of Arts and Humanities, University of Huddersfield, UK

Markku.Kiviniemi@vtt.fi, P.C.Vasantharaj@hud.ac.uk

Abstract

The Info creen development is part of wider research project where automated data collection of site operations and processing the data for provision of role-specific and contextualized information to targeted professions to fulfil the predicted information needs are being studied. The specific objective of the research work reported in this article is to arrive at a suitable user interface design that is easy to use and provides contextualized useful information for the construction workers. A Digital Visual Management method is proposed to share dedicated information to workers at construction sites with digital Info screens. The concept development has focused on information needs in drywall installation and the coordination of needed work phases. The information needs and barriers have been studied earlier with interviews of drywall installers to select and define information contents for Info screens. The Visual Management methods have been applied for this purpose and digital tools aspects are investigated especially. In visualization on the screen, the 3D-view of Building Information Model (BIM) has been used as a basis to present the information. In the user interface most information types are presented in relation to the geometry of the building and visualized with BIM 3D-view. The preliminary results from the user testing with the help of structured interviews on usefulness & ease of use of the prototype are encouraging.

Keywords

Construction operations management; construction workers; digital tools; user interface; visual management.

1 Introduction

The methods for sharing the design and planning information to construction workers are quite traditional ones; document deliveries and verbal communication

between the site managers and workers are most used. The site managers and trade contractors' foremen are responsible for detailed planning and controlling of activities at site, and to inform worker about schedules and correct work sequences. In practice, besides this management level input the workers must organize and plan their activities in the short term, and more formal decentralized planning approaches have been proposed [1]. In any case, better information exchange is expected to improve productivity with increasing the share of value-adding time of work and reducing waste [2].

Our earlier interviews among construction workers [3] have shown that the challenges and the demands that the workers have at the construction site partly originate from the issues outside the site and partly from the site itself. These interviews were conducted with 11 male Finnish drywall installers (average age & experience are 49 & 26 years respectively). Some of the demands, i.e., the ones originating from the site, would be relatively easily manageable. For instance, the need to consume time in searching for the tools or materials could be tackled by more efficient site management and better information flow to workers. Other demands, that is, the ones with roots in the work of the construction planning team, seem to require more profound changes in the construction "system"; extremely pressing timetable and problems with the floor plans and construction drawings are harder to meet or mitigate by simple orders or recommendations.

Many information needs could be met by providing workers with updated information about the situation at the site; in the following, the solutions found relevant, based in drywall installers' interviews, are briefly described in [3]. Pertaining to the demand to consider work proactively, an illustration of the finished space would support work, especially if presented at the location in question. To support flexible work sharing, one demand identified in the interview material, some digital means to inform other workers when personal share of work is done would support collaboration. Regarding lost tools and material, guiding the way to these items would diminish the time to find them. One

issue with construction plans and drawings is that they may be obsolete without the worker knowing it: thus, an efficient, updated document deliverer would mitigate this problem. Finally, there are many roles working with the same target. A shared illustration about the target, with information for all worker roles, would support collaboration at the construction site.

One possibility of directly delivering information to workers is through digital screens on site. Such screens are usually described as kiosks which typically consist of one or more large screens in fixed or movable positions at site with the protective casing needed in the physically demanding environment. Information contents varies from wide collection of design and management information [4] to more dedicated contents, like BIM-kiosks (BIM, Building Information Modelling) [5]. There are also research examples based on a screen at site for more specific BIM-based visual management use cases (see e.g., [6]).

The proposed Digital Visual Management (DVM) user interface is developed as part of a wider research project, ACTOR (Automatic Coordination of Construction Actors), involving the digital situational awareness of construction workers that empowers them in decision making. The prototype of the proposed system is demonstrated in the context of work phases needed for drywall installation. The drywall installation activity is divided in following work phases: a) installation of drywall frame and first gypsum board, b) electrical installations inside the wall, as well as HVAC installation if needed, c) installation of soundproof insulation, if needed, and installation of second sides board, and d) installation of joint tapes and plastering.

The anticipated information needs in drywall installation were derived from the interview result analysis and those are widely related in work preconditions of work defined in Lean Construction principles [3]. Workers need information on free workspace where previous activities are finished and correct external conditions exist, as well as information of availability of the design, materials, and equipment. Essential information is also the requirements and expectations for the work. Professional workers have skills to fulfill the core requirements, but project specific detailed requirements must be provided.

Accordingly, the following use cases are identified for the proposed design & development of the Digital Info screen - a Visual Management Tool for construction workers involved in drywall installation.

- Readiness assessment - checklist for a specific task to commence; availability of work front, resources such as materials, equipment, tools, etc.
- Suggestion for next work area to optimize the workflow at site.
- Design information - present the design information

such as plans and specifications in 3D/4D.

- Incidence reporting - interface for reporting incidents and seeking assistance from managers.

Hence, the objective of this research work is to arrive at a suitable user interface design that is easy to use and provides contextualized useful information for the construction workers. Structured interviews with the construction workers were used to capture their demands & information needs as well as to test the ease of use & usefulness of the proposed user interface. It is targeted to provide workers with near real time information supporting their daily work. The collected information can also be utilized for site managers' views, like showing the work accomplished over a period against the plan or creating short term forecasts to identify potential issues.

An overview of Visual Management is presented in the following section and subsequently the details of the development of the tool. Later, the proposed interface for the tools is elaborated with the discussions along with the user testing.

2 Overview of Visual Management

There is inherent waste in the construction processes and [7] and the construction productivity is historically low compared to other non-farm sectors [8]. Lean production principles have been attempted in construction projects to improve productivity, reduce waste, and create value to stakeholders [9]. Visual Management (VM) is one of the fundamental elements of the lean production system [10]. A simple definition of VM as found in [11]:

VM is a managerial strategy that attempts to improve organizational performance through connecting and aligning organizational vision, core values, goals and culture with other management systems, work processes, workplace elements, and stakeholders, by means of sensory stimuli (information), which directly address one or more of the human sensory modalities (visual, auditory, tactile, olfactory, and gustatory).

VM is also a management strategy that emphasizes using sensory information systems to increase process transparency or the communication ability of process elements [12-13] and an approach to manage and control information. VM can support the identification of issues, reduction of waste and detection of inconsistencies [14], contributing to cognitive, emotional, and social benefits [15]. VM can also serve a broad range of functions in an organization such as transparency, job facilitation, simplification, and unification [16]. VM and its tools are used to realize, communicate, and coordinate the lean

production system's targets in the workplace [17]. A range of tools, from the traditional analog to digital and hybrid tools are commonly used in VM [18]. VM using digital tools/platforms can be called Digital Visual Management. There are challenges reported in the adoption of DVM in construction sites [19-20]. It can be observed that there is greater emphasis on and the need for use of digital visual tools for automated provisioning of information to construction workers on situational awareness for their independent decision-making. Hence, it is proposed to design an effective user interface for DVM to promote situational awareness and independent decision making.

3 Application Development

3.1 DVM Info Screen System Requirements

The target of the proposed DVM system is to provide workers role-specific, context and location aware information in visual form. This would be achieved with automated data collection of site operations to provide up to date information for automatically coordinating actors and operations at site. The DVM Info screens are the information sharing component of ACTOR (Automatic Coordination of Construction Actors) research project. The shared information is processed in the backend systems and methodology for those are developed in other works packages of the ACTOR project.

This first version of the Info screen implementation is based on the idea of installation of touch sensitive info screens at ground floor in each staircase or beside of a door of an elevator in use during construction. Especially in high-rise buildings, the worker has some time to browse the screen when waiting for the elevator. In all cases the screens would be along centric movement routes at site. The information needs to be focused on the expected needs to provide instant value to the worker and be easily consumed. To reach these goals, the following requirements have been set for the DVM Info screen information sharing.

3.1.1 Role-specific Information

It is proposed that a worker's profession/trade can be identified automatically to provide dedicated information to a person near the info screen. Identification can be based on different methods but in this case, it would be based on indoor positioning and tracking system as this information is needed for other features in ACTOR project. It is noted that there are attitude barriers for tracking personnel at site and there is European regulation on how personal data can be used and managed.

3.1.2 Location-specific Information

As Info screens are installed along the central routes used by workers, the default content covers the areas inward from the screen location. However, this depends on the direction of the worker's movement as it affects the expected need for information.

3.1.3 Context-specific Information

The Info screen shall have the same kind of features for information selection as current public web services to predict what information user needs based on user's behavior. Some of this context sensitive information can be determined by the direction of the movement (inwards/outwards) and time in relation to work shift. Basic events are when the worker enters for the first time during the shift to the Info screen and then outward or inward around the breaks. In these basic use cases, and especially when user is entering/leaving the Info screen on any other time, there is a need to inspect the user's movement patterns recorded with indoor positioning, to make reasoning of the possible information needs. This reasoning will be executed in the ACTOR backend systems where the methods are based on results of ongoing and previous research, e.g. [21], [22], [23].

The shared information may contain uncertainties as the provided view is based on expected need and the information contents based on analysis of automatically collected datasets. The system needs to evaluate the reliability of shared information and notify users of uncertainties. This would improve the system's credibility by providing accurate information. The user can be asked if the detailed information is correct, which will also encourage the user to interact with the system but, in general, the user is not required to input data to the Info screen system. However, users will have the possibility with the touch screen to browse different views, change to another role (profession) or change language as some visual information is supported with text.

The limited number and the locations of Info screens at site affects the design of the shared information contents. Within this general information the system shows some QR-codes leading to more detailed information which could be picked up with mobile devices for usage at other locations. The suitability and acceptability of mobile devices for these use cases need to be evaluated in later user tests.

3.2 Choice of Technology

Main selection for Info screen implementation was using BIM 3D-view as a basis in user interface views. BIM provides geometrical visual representation of buildings, but also other model information can be shared with Info screen users. There are plenty of different tools

for showing and interacting with 3D IFC data, of which Trimble Connect [24] was chosen because it has good 3D visualization capabilities, as well as a programming API for interacting with the BIM data. The BIM data is uploaded in Trimble Connect in IFC-format and all IFC information content is available for retrieval by API. It is also possible to add new models in Trimble Connect and show those objects in addition to other models. This functionality is used for showing e.g., material storage in the UI.

In practice, the info screen would be implemented using a large touch screen. For development purposes, a Windows laptop with a 14-inch touch screen foldable in tablet mode was selected. Since Info screen is implemented as browser/server application, there are no specific requirements for the user terminal: any touch screen capable of running modern browser is enough. Same applies to the server used: since all the demanding BIM model handling is implemented by Trimble Connect, a basic Linux Ubuntu server, with 6GB memory and Intel Xeon 2.90GHz processor, was used for DVM Info screen backend.

3.3 System Architecture

Info screen was implemented as a web-based application (Figure 1) where user interface was implemented as JavaScript application using React programming framework. Server was implemented as Java Spring Boot application, with embedded Tomcat server. Server is used for serving user interface code, as well as offering REST and WebSocket application interfaces. Data used in the user interface is stored in a relational database (PostgreSQL).

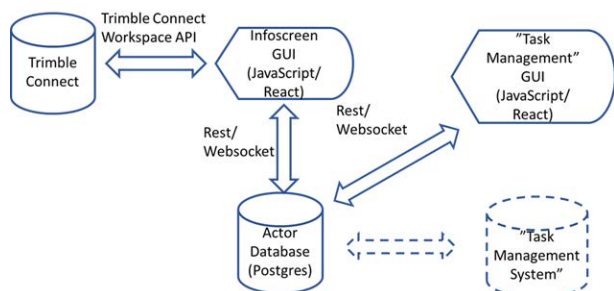


Figure 1. Info screen System Architecture

As described earlier, the Info screen is user interface for automated work management system investigated in ACTOR project. In Figure 1, the “Task Management System” describes this backend system, but integration has not yet been implemented due to parallel developments. Meanwhile, temporary UI is developed to input data in Info screen for testing purposes.

4 Proposed DVM User Interface

The basic features of the proposed DVM Info screen user interface is presented in Figure 2. On the header row are selections for language and profession/role of the user. These are set automatically to default values of user’s profile if the user’s profession can be identified automatically by indoor positioning system. If it is not possible or there are several professions nearby, the values can be selected manually with the touch screen.

Buttons on the left represent floors and those open 3D-view of the selected floor. The icons on the buttons summarize the work status of the selected profession in different floors. The blue check-box icon (☐) indicates where the work is going on currently. The green “checked” check-box icon (☒) visualize that all work is done in the floor. The thump (👊) and pointing finger (☞) icons express next possible work areas for the profession. The thump means the proposed next work location and the pointing finger indicates other possible free location.

For managing the activities at detail level, the DVM Info screen UI is following Location-Based Management System concepts (LBMS, [25]): the construction site is divided in work locations and there shall be just one activity going on in a work location at same time. Configuration of the DVM Info screen system requires definition of the activities / work phases as well as work locations in relation to IFC-objects in the BIM. In figure 2, the work locations are defined as sets of individual drywall objects. By default, in the UI the work locations where the work is going on are shown (blue) and the next proposed work location (light green) for this profession.

The work phases and locations shall be the same as are used in weekly planning of site operations, and DVM Info screen system is a channel to share the same information in visual format to the workers. The term “work phase” is used here to indicate that installation of a drywall requires a few work phases by different professions in correct order. The automated coordination of these is expected to improve productivity by increasing direct installation work time of workers total time at site.

Key objective in ACTOR approach is that the readiness of work phases can be recorded with sensing systems and no user input to update statuses is needed. Target is that on daily bases the work statuses will be checked with machine vision-based analysis [21] and at start of next work shift the statuses are expected to be known and will be showed to users at screen. During the work shift this baseline information will be compared to workers’ indoor movement to detect whether the work has progressed to the next work location. In screen is shown question mark (?) to indicate that status of the work location may have been changed and user may update it manually.



Figure 2. DVM Info screen View for a Selected Floor with Visualization of Work Status of Drywall Installer

Other features in the UI are that user may highlight other information about work location occupancy, like other possible free work locations (yellow) and all finished locations (green) (Figure 2), and the locations that are restricted no-go areas (red) or reserved by some

other activity (light yellow) (Figure 3). Users can also select to show the material storage locations (Figure 3). These materials and locations are targeted to identify with machine vision.

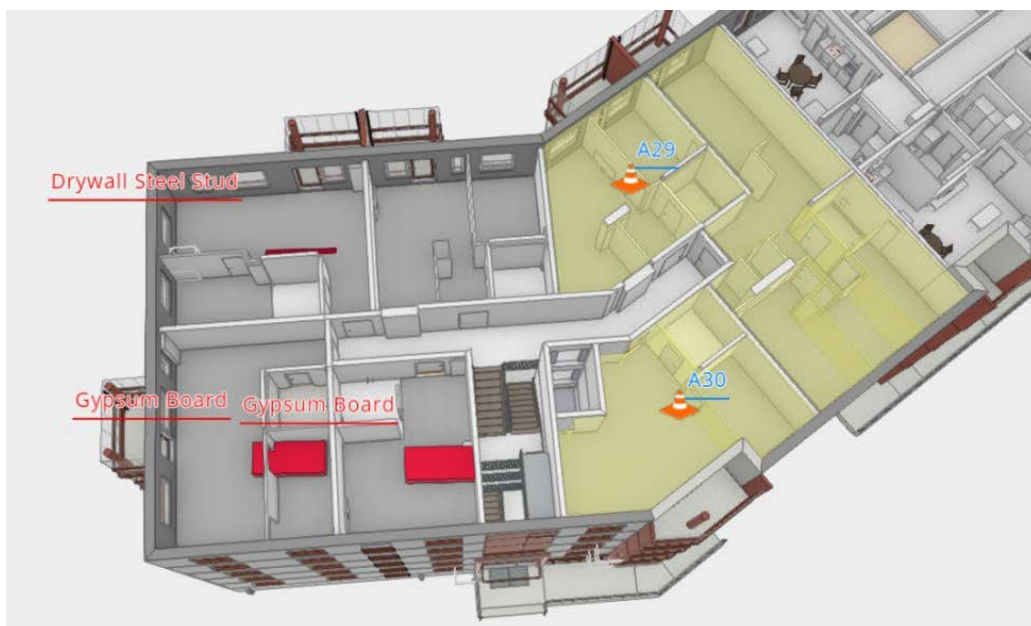


Figure 3. Examples of Visualisation of Restricted and Reserved Areas, and Material Storages

The information on design objects in BIM and drawings is one of the key information types for Info screen. Basic feature is providing links to drawings with QR-codes or open a drawing in info screen or showing detailed object data (Figure 4).

An important feature is to notify users of new versions of the drawing needed in user's work. The IFC data of individual wall can be shown e.g., wall types or the quantities in the work locations (Figure 4). Basic problem with drywalls is that structures are not modelled at detail level in BIM and specific requirements for wall installation e.g., additional braces in drywall frame, must be found out from related surrounding installations. However, this kind of detailed information is more needed at the work location and is not so important at the Info screen.

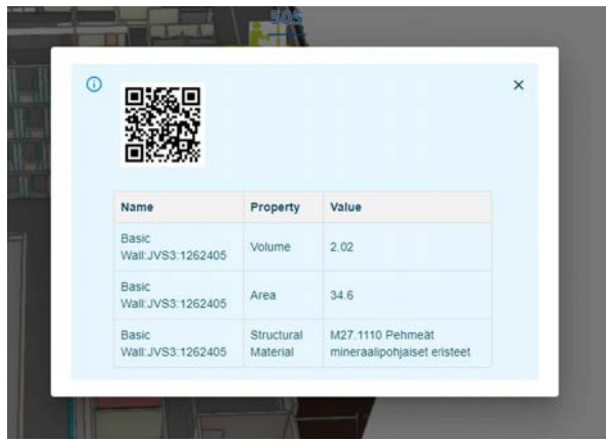


Figure 4. First example for showing drywall object data and QR-code to fetch data with mobile device.

The Info screen can also be used to push some notifications, even alarms, to give this information more visibility at site. Another idea is to add or read location specific notes ("virtual post-it notes") with Info screen UI.

5 User Testing

The potentiality of DVM Info Screen for future use was tested by workers at a construction site. Testing was performed at the office by the site. One researcher presented the user interface of Info Screen, conducted the interview, and wrote down the responses. The responses were then analyzed using narrative analysis to understand the perceptions of the users.

5.1 Method

Four carpenters performing drywall installation have been interviewed on the usefulness & ease of use of the Info Screen. Their average age was 48,5 years (min 39

years, max 63 years). Work experience was about 30 years on average (min 20 years, max 47 years).

Each interview lasted for one hour. During the interview, all main functionalities from the user perspective were presented and interviewees' opinions were asked about them.

5.2 Results

All testers except one found Info Screen easy to use. The only person expressing difficulties was the oldest one, not accustomed to digital technology. Despite this, all testers appeared to navigate through the 3D model effortlessly and had no difficulties in manipulating the device. As a grade, all testers found the usefulness of the information provided by Info Screen as good (4) and all, but one tester considered the usability of the device as good (4, average 3.25; the scale is 1-5 so that 1 is poor and 5 is excellent). One commented that it would be easy to learn to use the device, only brief instructions could be provided next to the device for the first use of it.

The information of the device can be located on a large screen at the site and/or in a mobile phone, as an application. Two testers would have the information in both ways, one considered the large screen as the best option, and one would only use the mobile phone application.

Testers were asked how much of the users' identity could be revealed to the device so that the device could provide tailored information when standing next to it. Options were the revealing of the person or the work role. Somewhat surprisingly, three testers found no obstacle in the ability of Info Screen to detect the personal identity of the worker when standing next to the device. One commented that the site is already full of cameras and there is such a hurry anyway that you cannot do anything but focus on the work. One tester would only allow the identification of a person's work role and the preferred language setting for the device.

No functionality provided by Info Screen was found unnecessary. The most useful information was found to be the statuses of the working locations (2 responses), easy access to design documents (1 response), or both statuses and documents (1 response).

The statuses of the working locations provoked discussions the most. Among the responses, only the symbols indicating a finished area and a forbidden area were found clear by all testers; the symbol indicating a possible work location was commented confusing the most often. Two testers stated that the status information about the work phase of the drywall installation is important to electrician, plumber, and painter; one noted that the information about the presence of other workers in the area would be useful so that when their work is about to be finished, the drywall installer can go there. This is in accordance with the fact that the construction

work is performed in harmony among various work roles and collaboration is important.

6 Discussion

This Info screen deviates from other information kiosks at construction sites, providing to workers role-specific, context and location aware information in visual form. However, the DVM Info screen development is a concept study. Practical implantation would require extensive automatic information collection from site and interaction with external planning and management systems. Sharing information with Digital Info screen would be one value-adding proposition for implementing this kind of emerging automated data collection and system interoperability. In addition to data collection, it is necessary to develop simulation and optimization methods to automate decisions about what are optimal next operations at site.

There were various information needs identified in the Finnish drywall installation interviews [3]. Some of these needs can be fulfilled with Info screen. Electricians, for instance, are involved in drywall installation as electricity related devices are installed in the walls so both drywall installers and electricians need the status information of the drywall. Regarding the need to conduct flexible work sharing, Info screen provides a means to inform all collaborating workers when some specific location is available for the work of the specific role(s). It can be concluded that Info screen would be beneficial also when work is done in solitude. The start of any work at a construction site usually requires that the previous work phase in some specific location is finished before the next phase can be started. Info screens can be used in these instances. Other information needs reported by [3] call for a different type of solution or solutions.

The results of the user testing revealed the challenges faced by some of the users, specifically, the older workers in using the proposed system. It is also reinforced during the tests that necessary training is required for effective use. There have been varied opinions on the fixed versus mobile and small versus large screens. It is also observed that there could be wider acceptance to such systems that can track the location (and personal work-related information) among the construction workers that is encouraging for the greater adoption of technology and collaborative work at construction sites. However, a large-scale validation is necessary to generalize these findings.

The Info screen system is the first attempt towards automatic coordination of construction actors at site environment. Beyond the first user tests further research is needed to improve the usability of the system and information contents, as well as adaptability to different approaches in workers' ways on organizing and planning

their work recognized in [26].

Main visual element in the Info screen is the usage of BIM 3D-view as basis for visualizing work-related information. Current BIM models are produced for design purposes, even though those are also enriched for production use cases, like 4D-scheduling. The BIMs could also be enriched with information directed at workers e.g. including work requirements, detailing of structures or safety aspects. This information could be easily shared to workers in the Info screen UI. The detailed work planning tool used at site should also be linked to BIM in order exchange information between systems and making it easier to configure the Info screen system for a construction project. It can be noted that the management views are not in the scope of this paper.

7 Conclusions

It has been attempted to propose a suitable user interface for a digital visual management tool for construction workers. The design of the Info screen is based on the interviews conducted with the construction workers involving drywall installations on their demands & information requirements. Appropriate technical solutions have been chosen in the development of prototype with which it was possible to demonstrate the utility of the system under various use case scenarios. The initial testing of the prototype with the users revealed that the proposed system is easy to use and provides all the contextual information that is useful in their independent decision making. A wider study is required to generalize this proposed solution across other construction processes and locations.

References

- [1] Lehtovaara, J., Seppänen, O., and Peltokorpi, A. Improving construction management with decentralized production planning and control: exploring the production crew and manager perspectives through a multimethod approach. *Construction Management and Economics*, 40(4): 254-277, 2022,
- [2] Kärkkäinen, R., Lavikka, R., Seppänen, O., and Peltokorpi, A. Situation picture through construction information management. *In the Proceedings of the 10th Nordic conference on construction economics and organization*, pages 155-161, 2019.
- [3] Liinasuo, M., Salonen, T. T., and Görsch, C. Drywall Installers' Work Demands-Tackling between Normal Duties and Absurd Challenges. *In Proceedings of the European Conference on Cognitive Ergonomics 2023*, pages 1-8, 2023.
- [4] Ruwanpura, J. & Hewage, K. and Silva, L.

- Evolution of the i-Booth® onsite information management kiosk. *Automation in Construction*, 21(): 52-63, 2012.
- [5] Bråthen, K. and Moum, A. Bridging the gap: bringing BIM to construction workers. *Engineering, Construction and Architectural Management*, 23(6): 751-764, 2016.
- [6] Sacks, R., Barak, R., Belaciano, B., Gurevich, U. and Pikas, E. KanBIM Workflow Management System: Prototype implementation and field testing. *Lean Construction Journal*, 2013:19-35, 2013.
- [7] Koskela, L., Bølviken, T., and Rooke, J. Which are the wastes of construction? In *Proceedings of the 21st Annual Conference of the International Group for Lean Construction 2013, IGLC 2013*, pages 905–914, Fortaleza, Brazil, 2013.
- [8] Office for National Statistics. Productivity in the construction industry, UK: 2021. On-line: <https://www.ons.gov.uk/economy/economicoutputandproductivity/productivitymeasures/articles/productivityintheconstructionindustryuk2021/2021-10-19#understanding-construction-productivity>, Accessed: 25/12/2023.
- [9] Koskela, L. Lean Production in Construction. Automation and Robotics in Construction X: Proceedings of the 10th International Symposium on Automation and Robotics in Construction (ISARC), pages 47-54, Houston, USA, 1993.
- [10] Liker J.K. The Toyota way: 14 management principles from the world's greatest manufacturer, McGraw-Hill, USA. 2004.
- [11] Steenkamp, L. P., Hagedorn-Hansen, D., and Oosthuizen, G. A. Visual Management System to Manage Manufacturing Resources. *Procedia Manufacturing*, 8, pages 455–462. 2017.
- [12] Tezel A., Koskela L., Tzortzopoulos P., Formoso C.T. and Alves T. Visual management in Brazilian construction companies: taxonomy and guidelines for implementation, *Journal of management in engineering*, 31(6):05015001-1-14, 2015.
- [13] Pedó, B., Tezel, A., Koskela, L., Tzortzopoulos, P., Formoso, C. T., Vrabie, E., and Robinson, S. Visual Management Implementation Strategy: An Analysis of Digital Whiteboards. In *the Proceedings of the 31st Annual Conference of the International Group for Lean Construction (IGLC31)*, Berkeley, California, USA, pages 608–619. 2023.
- [14] Tezel, A., Koskela, L., and Tzortzopoulos, P. Visual management in production management: A literature synthesis. *Journal of Manufacturing Technology Management*, 27(6): 766-799, 2016.
- [15] Bresciani, S., and Eppler, M. J. The risks of visualization - a classification of disadvantages associated with graphic representations of information. *ICA Working Paper*, 1-22, 2008.
- [16] Tezel, B. A., Koskela, L. J., and Tzortzopoulos, P. The functions of visual management. *International Research Symposium*, pages 201–219, Salford, UK, 2009.
- [17] Galsworth G.D. *Visual workplace: visual thinking*, Visual-Lean enterprise press, USA, 2005.
- [18] Tezel, A., and Aziz, Z. From conventional to it based visual management: A conceptual discussion for lean construction. *Journal of Information Technology in Construction*, 22(May 2016): 220–246, 2017.
- [19] Reinbold, A., Lappalainen, E., Seppänen, O., Peltokorpi, A., and Singh, V. Current Challenges in the Adoption of Digital Visual Management at Construction Sites: Exploratory Case Studies. *Sustainability*, 14(21): 14395-1-16, 2022.
- [20] Tezel, A., Koskela, L., Tzortzopoulos, P., Koskenvesa, A., and Sahlstedt, S. An examination of visual management on Finnish construction sites. In *the proceedings of the 19th Annual Conference of the International Group for Lean Construction 2011, IGLC 2011*, Lima, Peru, pages 115–124, 2011.
- [21] Chauhan, I. and Seppänen, O. Automatic indoor construction progress monitoring: challenges and solutions. In *Proceedings of the 2023 European Conference on Computing in Construction and the 40th International CIB W78 Conference*, Crete, Greece, DOI: [10.35490/EC3.2023.225](https://doi.org/10.35490/EC3.2023.225), 2023.
- [22] Reinbold, A., Seppänen, O., Peltokorpi, A., Singh, V. and Dror, E. Integrating Indoor Positioning System and BIM to Improve Situational Awareness. In *the Proceedings of the 27th Annual Conference of the International Group for Lean Construction (IGLC)*, Dublin, Ireland, pages 1141-1150, 2019.
- [23] Zhao, J., Pikas, E., Seppänen, O. and Peltokorpi, A. Using Real-Time Indoor Resource Positioning to Track the Progress of Tasks in Construction Sites. *Frontiers in Built Environment*. 7: <https://doi.org/10.3389/fbuil.2021.661166>, 2021.
- [24] Trimble Connect. Online: <https://connect.trimble.com/>, Accessed: 19.12.2023.
- [25] Kenley, R., and Seppänen, O. *Location-Based Management for Construction: Planning, scheduling and control*, (1st ed.). Routledge. <https://doi.org/10.4324/9780203030417>, 2009.
- [26] Görsch, C., Seppänen, O., Peltokorpi, A. and Lavikka, R. Task planning and control in construction: revealing workers as early and late planners. *Construction Management and Economics*, DOI: 10.1080/01446193.2023.2270080, 2023.

Assessing the Impact of AR-Assisted Warnings on Roadway Worker Stress Under Different Workload Conditions

Fatemeh Banani Ardecani¹ and Amit Kumar¹ and Omidreza Shoghli¹

¹William State Lee College of Engineering, University of North Carolina at Charlotte, USA.

fbanania@charlotte.edu, akumar28@charlotte.edu, oshoghli@charlotte.edu

Abstract -

Recent data from the Federal Highway Administration highlights an alarming increase in fatal crashes in roadway work zones, underscoring the need for enhanced worker safety measures. This study addresses this concern by evaluating stress levels among roadway workers exposed to multi-sensory AR-assisted warnings during varying work intensities, using a high-fidelity Virtual Reality environment for simulation. Unlike previous studies that mainly concentrated on external factors, this study investigates the internal impact of these factors on workers. Our findings from 18 participants indicate significant physiological differences between light- and medium-intensity activities in terms of heart rate variability, mean heart rate, NN50, pNN50, and HF-HRV, though SCR peaks showed no considerable variation. The study's significant contributions include insights into higher stress levels in workers performing moderate-intensity tasks, aiding in the development of improved warning systems. Additionally, it offers valuable data for optimizing resource allocation in construction settings. Ultimately, this research bridges a gap and provides insights for future research on improving both safety and productivity in roadway work zones through informed stress management and effective hazard warning systems.

Keywords -

Worker's Stress; Physical Activity; Roadway Work Zone; Virtual Reality

1 Introduction

Fatal traffic crashes at roadway work zones have been on the rise in recent years. According to the Federal Highway Administration (FHWA) work zone facts and statistics report, there was a significant increase in fatal crashes at work zones from 863 in 2020 to 956 in 2021 [1]. Research has shown that excessive stress in workers heightens the risk of errors, injuries, and various health issues among workers while concurrently being associated with reduced productivity levels [2]. In roadway work zones, lane closures, proximity to moving vehicles, night shifts, and the presence of construction vehicles and employees can all cause stress in workers. A construction worker's behavior can be significantly affected by stress [3]. The hormones released by the brain under stress, including cortisol and adrenaline, may have an impact on cognitive processes, judgment, and response times.

In the wake of rising highway worker fatalities at road construction sites, there's an immediate need to create ad-

vanced safety systems that safeguard workers. Augmented Reality (AR) shows promising potential in alerting workers, but its implementation in road work zones and its impact on stress and attention have not been thoroughly investigated. Currently, most of the research is focused on external factors like lane closures, traffic congestion, warnings, determining safe work radius, and traffic management. While this is important, the body of knowledge is missing an integrated effort to analyze how these factors affect the worker's stress. Therefore, this study aims to assess the effects of two categories of work (light- or medium-intensity activity) on stress levels experienced by roadway workers when they were receiving multi-sensory warnings during their routine tasks.

This study aims to assess the stress levels experienced by roadway workers as they receive multi-sensory AR-enabled warnings during their routine tasks, examining the effects within two categories of work as either light- or medium-intensity activity. Furthermore, it leverages a high-fidelity virtual reality environment for safe evaluation and testing of rare high-risk scenarios. Additionally, it introduces a model that enables continuous and non-invasive monitoring of stress levels among roadway workers based on work activity. The proposed methodology can be leveraged for active monitoring of stress in the field, thereby enhancing safety at construction sites and promoting the well-being and productivity of workers.

2 Literature Review

2.1 Mental Stress in Construction Workers

The number of mortality and disability cases involving construction workers was the highest amongst the major industrial sectors [4]. Current approaches investigating the physical demands of various tasks provide valuable information for evaluating certain construction activities. However, these approaches often focus on specific individual traits, such as physiological characteristics, and environmental factors, such as ambient temperature and humidity. In simpler terms, this means that employees may exert varying levels of effort when performing the same task due to these individual and environmental differences [5]. In individuals working under continuously demanding and

stressful conditions, emotional stress manifests in chronic fatigue, emotional drain, and a loss of devotion to job duties. Stress can lead to reduced attention on work tasks and subsequently cause one to ignore safety behaviors, thereby increasing injury incident rates [6].

Molen et al.[7] presented a qualitative study on the nature and feasibility of measures to reduce work stress. Construction work conditions can induce anxiety that may be high enough to elicit a robust fear response. Fear and anxiety are known to have different neuro-anatomical substrates and physiological outcomes and may lead to differences in behavior and action tendencies [8].

Because they are at the bottom of the organizational hierarchy, construction workers have little authority over their work. These organizational stressors not only stress out the construction workers but also affect how they behave in terms of safety. The primary factor behind most construction workers' injury events is a deterioration in safety behaviors [6]. In the case of roadway construction and maintenance, dynamic worksite conditions and proximity to live traffic frequently expose highway workers to unsafe work zone proximity, resulting in accidents [9, 10, 11].

2.2 Stress Monitoring Using Wearable Sensors

Everyone reacts to stress differently; thus, measuring and tracking stress can be difficult. Psychoanalysis, human sensing, and medical examinations could be used to identify the signs and symptoms of stress. Basic signs of elevated stress include headaches, stiff muscles, sleeplessness, and rapid heartbeat [12]. In the clinical setting, physiological indicators of biochemical reactions, such as stress hormones, have been widely used as accurate stress markers. Because stress-related hormones, such as cortisol and glucocorticoids, alter in response to stress, monitoring these hormone changes can give us useful information about how stressed people are [13]. Though this strategy is effective and desirable, it is impractical for continuous stress monitoring at an active construction site because serum, saliva, urine, or hair samples must be taken repeatedly to measure stress-related hormones. Continuous monitoring of stress is particularly important on construction sites due to the constantly changing nature and challenges at the workplace. Additionally, the analysis of the collected biological samples necessitates laboratory processing, which is challenging to implement in the field. To bypass this issue, recent studies [14, 15] have utilized wearable sensors for the assessment of stress levels. For this purpose, physiological signals, which are generated by the body's processes, can be collected. These signals have traditionally been categorized into two groups. The first group includes electrical physiological signals such as electrodermal activity (EDA) [16], electroencephalog-

raphy (EEG) [17], heart rate (HR), heart rate variability (HRV), electrocardiography ECG [18], photoplethysmography (PPG) [19], peripheral skin temperature (ST) [20]. The second group comprises non-electrical physiological signals, like inertial measurement units (IMU) and potential of hydrogen (pH). Variations in electrical physiological signals are closely linked to stress levels [5]. Accurately determining workers' stress levels in the field can lead to early recognition of stress, thereby enhancing both safety and productivity at construction sites.

A smart sensor in a wristband can provide the signal known as EDA. By measuring the fluctuation in skin conductance, EDA provides information about the electrical characteristics of the skin. An EDA sensor inserts a low, steady voltage into a wristband-style sensor. Next, the voltage variations brought on by sweat gland activity are monitored. A wristband-style sensor contains an infrared thermopile that measures the ST signal. An infrared thermopile uses the skin's infrared energy to measure temperature. A higher temperature is correlated with higher infrared energy. [21, 22]. Heart signal analysis allows for the computation of heart activity characteristics known as heart rate variability (HRV). These characteristics result from processing the heart signal across time, frequency, and nonlinear domains. For instance, in the time domain, an HRV metric is the root mean square of successive heart-beat intervals (R-R intervals), termed RMSSD. Frequency domain HRV metrics include the power within specific frequency bands (such as high frequency - 0.15 to 0.4 Hz and low frequency - 0.04 to 0.15 Hz) denoted as HF and LF within the HRV spectrum. Lastly, in the nonlinear domain, an HRV feature example is the entropy calculated from beat-to-beat intervals.

Earlier research has demonstrated that these characteristics are associated with specific human conditions. For instance, stress levels have been linked to a reduction in RMSSD, while an increase in HF (high frequency) is associated with cognitive load [23]. Electrodermal activity (EDA), also known as Galvanic Skin Response (GSR), has similarly exhibited correlations with human conditions like stress and workload. EDA signals are typically broken down into two primary components: tonic and phasic. Tonic represents long-term changes in the signal, while phasic accounts for momentary shifts in the EDA signal. The tonic aspect helps establish the skin conductance level (SCL), while the phasic component defines the skin conductance responses (SCR). Both SCL and SCR have been found to be linked with heightened cognitive load and stress levels.

3 Method

To achieve the objective of the study, we designed an experiment in a high-fidelity Virtual Reality environment

replicating a real-world roadway work zone. The study protocol (21-0357) was reviewed and approved by the Institutional Review Board (IRB) at the University of North Carolina at Charlotte. Prior to data collection, all individuals were briefed about the data-collection procedure, and had the option of choosing to opt out of participation at any time without giving any explanation. None of the participants mentioned any physiological or physical issues impacting their ability to function at work. The participants were asked to perform two routine roadway maintenance tasks in the VR environment while wearing the sensing technologies and warning delivery devices. Three members of the research team recorded the activities over a half-hour session.

AR-assisted safety protocols within roadwork zones combined with warnings can extensively impact the response to danger and safety of workers as shown in Figure 1. Thus, it is indispensable to determine the worker's response to stress in a roadway work zone to gain a better understanding of the usability and effectiveness of the AR-assisted safety technology.

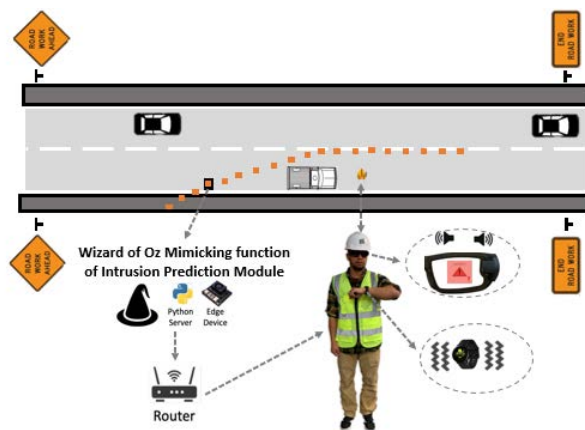


Figure 1. Outline of the AR-assisted Warning System

In this study, we evaluated the metrics related to the skin and cardiovascular system. To achieve this, we employed a wristband-type wearable equipped with embedded sensors, including photoplethysmography (PPG), skin temperature (ST), and electrodermal activity (EDA). The PPG sensor analyzes the difference between transmitted and reflected light and quantifies the variations in blood pulse within the arteries. This difference served as an indicator of heart activity. Additionally, we investigated cardiovascular parameters such as heart rate (HR) and heart rate variability (HRV). These parameters were derived by further analysis of the obtained data from the PPG sensor. It should be noted that PPG technology is commonly integrated into smart wearables and can be effectively utilized in both naturalistic and experimental settings.

3.1 Experiment Design

The experiment design consisted of participants triggered by warnings at certain time intervals while carrying out maintenance activities categorized into light and medium, as follows:

Warning Trigger: This consisted of three warning stimuli that were administered to the participants while they carried out the task in the VR environment. All three multi-sensory warnings were intended to communicate predicted risk to the workers by delivering haptic, audio, and visual cues at the same time. Tizen Native framework was used to administer the haptic stimulus, which utilizes a predefined pattern available in the API [24]. The audio warning was a high-pitched beep sound with a frequency of 44,100 Hz and a duration of 0.2 milliseconds. The warnings were designed to function simultaneously, without any delay, as soon as the back-end system activated them. Moreover, the visual warnings about potential hazards were delivered using AR simulation in the VR environment.

Activity Types: The construction tasks were categorized as low and medium-intensity activities. Similar separation of tasks has been done previously based on energy-expenditure prediction program (EEEP) [5] and [25].

Light Activity: These work scenarios include tasks like standing, briefing, and other tasks involving little movement of the body parts. Assembly, reading the construction drawings, and inventory work are a few examples of typical low-intensity tasks. For the purpose of low-intensity activity, an inspection task was chosen for this study, which included taking a picture of a clogged stormwater inlet by the roadway shoulder. Such incidents are a common maintenance task at roadway work zones where curbs are often blocked by overgrowth of vegetation, small debris, and fallen leaves.

Moderate Activity: These tasks include tasks like cleaning up the site, locating tools, moving light materials, and measuring and cutting sheetrock. For the purpose of the study, cleaning the clogged inlet drainage by using a leaf blower was selected as a medium-intensity task. Cleaning jobs, as such, are standard upkeep at roadway work zones.

3.2 Apparatus

Roadway Work Zone in Virtual Reality. The study used the guidelines provided in the Manual on Uniform Traffic Control Devices (MUTCD) [26] as a reference to create a virtual highway work zone environment. Figure 2(c) shows the simulated environment, including live traffic and highly detailed 3D models, closely resembling the real roadway work zone environment. Using the VR environment developed by the research team [27], participants were equipped with a VR headset while holding a real-life

tablet (light activity) or a leaf blower (medium activity). Also, they wore a smartwatch for delivering the haptic warning and a wristband for collecting physiological data. The figure also depicts the immersive and interactive VR environment captured from the participants' viewpoints while they performed the tasks of the experiment.

The modeling utilized the Oculus Quest 2 VR headset in the Virtual Reality simulation as shown in Figure 2(a). Furthermore, the Unity 3D game engine was used to develop the stream VR software. To allow the exact simulation interaction and user experience in the VR interface, the participants were able to simultaneously observe the gadgets and equipment (smartwatch, tablet, and leaf blower) in the VR while wearing the smartwatch and handling the tablet/leaf blower during the actual participation.

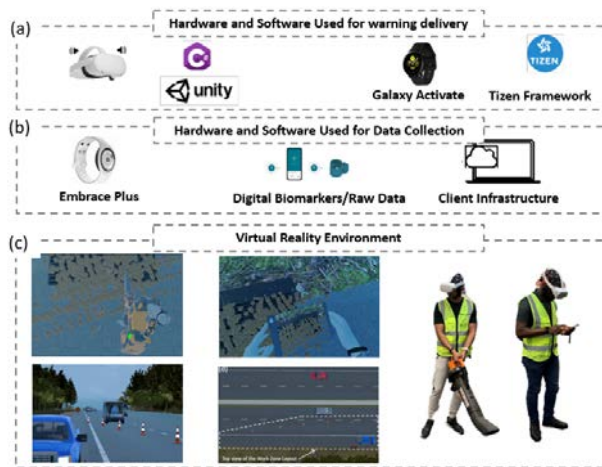


Figure 2. Apparatus of the Experiments (a) Hardware and Software Used for warning delivery, (b) Hardware and Software Used for Data Collection, and (c) Virtual Reality Environment

Wristband Sensing Device to Capture Physiological Data. To measure workers' physiological data in the roadway work zone environment in VR, a wristband-type sensing technology was used. An off-the-shelf wearable sensing wristband, Embrace Plus, was used to collect workers' PPG, EDA, and ST signals [28]. These were recorded simultaneously at the highest recording rate (PPG at 64 Hz and EDA and ST at 4 Hz).

Leaf Blower, Tablet, Warning Watch, and Camera for Video recording. Participants wore a Samsung Galaxy Watch, as shown in Figure 2(a), that was used to deliver haptic warnings. The warnings were administered using the Tizen Native framework. Additionally, the participants used an Amazon Fire Tablet, as shown in Figure 2(c), to take pictures of a clogged storm water inlet (light level activity) and a WORX WG509 12Amp leaf blower to clear the leaves (medium level activity).

3.3 Experiment Procedure

The experiment involved a carefully structured sequence of activities to ensure a comprehensive and ethical approach to the research. First, the nature of the study, its goals, and the participants' roles were explained to the volunteers. Following the consent process, administrators of the experiment explained the objectives and ensured that the participants clearly understood what was expected of them throughout the study. The experiment was designed to be completed within a reasonable time frame, with a maximum duration of up to half an hour per session.

The participants were tasked to replicate real-world scenarios commonly encountered in roadway work zones. The focus was on taking pictures of the clogged inlet and removing obstructions from the same. To design the study, existing literature [29, 24] that discusses the influence of physical activity, intensity, and cognitive load on worker's stress was utilized. Based on this knowledge, an obstruction removal task requiring participants to engage in higher levels of physical exertion than routine inspection of roadways was developed. The task began with participants capturing pictures of the clogged drain and then, as a subsequent task, activating the leaf blower and directing it towards the obstructed drop inlet within the virtual environment. As they did so, the virtual reality environment featured a carefully designed photograph capture with sound on the tablet. On the subsequent task, the blowing effect with a leaf blower sound effectively cleared the leaves positioned on top of the drop inlet. This dynamic and interactive task continued until all necessary warnings were delivered, and the administrator signaled the completion of the task.

The developed virtual work zone is depicted in Figure 2(c). To enhance the realism and interactivity of the study, the model uses a mixed-reality interaction for this experiment. This approach allows the participants to simultaneously engage with physical objects in the real world and virtual objects within the virtual environment.

3.4 Participants

The study utilized data from 18 out of the 22 participants. The removal of the data for 4 participants was due to a technical challenge in recording and synchronizing physiological data from the study. Out of the final count ($N=18$), the average age of the participants was 28.27 years. Five participants did not have work experience, while the remaining 13 had an average work experience of 3.92 years ($SD=4.78$). The average duration for all the participants to complete the medium and light-intensity tasks was 1 minute 42 seconds and 1 minute 59 seconds, respectively.

3.5 Physiological Data Analysis

The Heart Rate Variability (HRV) features were computed from the interbeat interval data (IBI) as depicted in Figure 3 captured through the Empatica Embrace Plus device [28]. The features span across different time, frequency, and nonlinear domains. Using the pyHRV [30] package in Python, features such as HF and RMSSD were calculated across these domains for each participant during the study. It is important to note that certain computed features, like LF, might not be suitable for short-term data collection, prompting a focus only on features applicable to shorter timeframes. Additionally, for Electrodermal Activity (EDA) analysis, denoising of EDA signals from the Empatica Embrace Plus involved using a Butterworth filter with a low cutoff frequency of 1.5 Hz [5].

Furthermore, a Hampel filter was utilized to eliminate outliers in the physiological data, following Allen's method [31], which replaces spikes with the median value of neighboring signals. Subsequently, the modified signal was processed using the Neurokit 2 package [32]. This package allows for signal decomposition into tonic and phasic components and facilitates the computation of various Electrodermal Activity (EDA) features, including the count of phasic skin conductance responses and the skin conductance level (SCL). The study then compared the number of SCR peaks among different participants and conditions.

A paired t-test was used to compare heavy and light activity across all participants. Additionally, parameters like mean-NNI, mean-HR, std-HR, SDNN, RMSSD, NN50, PNN50, and the number of SCR Peaks were analyzed to draw inferences and conclusions from the model. RMSSD, calculated as the square root of the mean squared differences between successive NN intervals, stands as one of the most frequently utilized measures in the time domain. NN50 signifies the count of interval differences greater than 50 milliseconds between successive NN intervals, while pNN50 is the proportion calculated by dividing NN50 by the total count of NN intervals. Additionally, the Blood Volume Pulse (BVP) captured by the PPG sensor stands as a primary output. This BVP data shares a strong correlation with Interbeat Interval (IBI) data, which represents the time lapse between individual heartbeats. It is to be noted that the IBI information stems from processed PPG/BVP signals from the Embrace Plus wristband, utilizing an algorithm that effectively eliminates erroneous peaks caused by noise in the BVP signal.

4 Result and Discussion

HR, HRV, and HF-HRV: The comparison of the mean heart rate over the two activities (light and moderate) was conducted using a paired t-test to assess the potential dif-

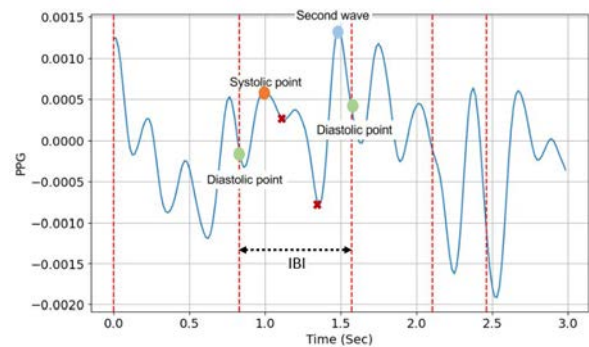


Figure 3. Interbeat Interval (IBI) defined as the time interval between two fiducial points on the diastolic pulse wave for the BVP signal

ferences between the two work scenarios. The paired t-test revealed a p-value of 0.0729 with a confidence level of 90 percent, suggesting a significant difference between these two activities. Participants' mean HR (beat per minute) when performing the light and moderate activities were 100.35 and 104.29 bpm, respectively. Figure 4(a) shows the distribution of the mean HR. The lower mean heart rate in light activity can be associated with decreased distractions due to external stimuli. Since the leaf blower is heavier and noisier than the tablet, the workers could not pay attention to the environment. Similarly, we com-

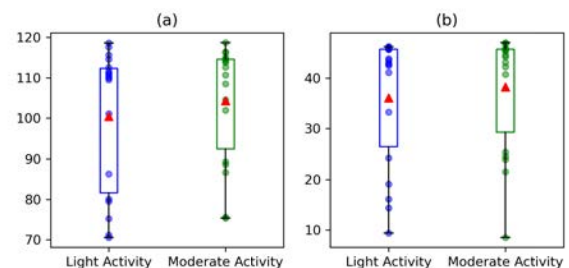


Figure 4. (a) Mean Heart Rate and (b) Heart Rate Variability of Moderate and Light Activity

pared the heart rate variability over the two activities using a paired t-test to assess the potential differences between the two activities. The paired t-test revealed a p-value of 0.1070, suggesting a confidence level of approximately 90 percent. The mean HRV of the light and moderate activity level were 36.06 and 38.22 bpm, respectively. Figure 4(b) shows the distribution of the HRV. A decrease in heart rate variability (HRV) indicates a reduction in the adaptability and responsiveness of the autonomic nervous system. This can be associated with an increase in stress when participants performed moderate activity.

Another interesting finding from HRV is in the pNN50

and NN50 features. The paired t-test revealed a p-value of 0.7349 and 0.0257, respectively. The NN50 with a confidence level of 95 percent, suggesting a significant difference between these two activities. The moderate activity recorded a lower mean NN50 compared to the light activity. Figure 5(a) and (b) shows the distribution of the HRV. Both pNN50 and NN50 are used in HRV analysis to assess the balance between the sympathetic and parasympathetic (PNS) branches of the autonomic nervous system [33]. Although in a low-intensity scenario, the pNN50 has no significant difference as compared to the moderate intensity, lower pnni50 shows lower PNS activity during moderate activity, revealing a higher workload level [34]. Comparing the HF-HRV across the light and moderate

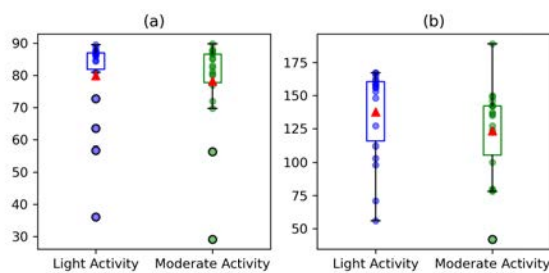


Figure 5. (a) pNN50-HRV and (b) NN50-HRV of Moderate and Light Activity

activity, the paired t-test shows a significant difference between them with a confidence interval of approximately 90 percent (P-value = 0.1025). The distribution of the HF-HRV is depicted in the Figure 6(a). This parameter is also associated with the parasympathetic nervous system. Lower HF-HRV values are generally associated with more stress [34] during moderate activity.

It should be noted that in this study, we used the Normal-to-Normal Intervals (NNI) and Inter beat Intervals (IBI) for the calculation of the HRV, which is a measure of the variation in time between each heartbeat. However, the intensity of physical workloads can also affect stress indicators, and distinguishing between the effects of physical workloads and mental stress on these indicators presents a challenge due to their overlapping impacts. When the body is subjected to varying levels of physical demand, the HR increases, but it does not necessarily mean that the body is under mental stress. A higher heart rate, generally resulting from physical activity or stress, is associated with lower HRV. This is because a faster heart rate reduces the time between individual heartbeats, leaving less room for variability. In this study, none of the workload scenarios involved intense physical activities. The results were presented with the assumption that the variations between low and medium physical activities will not have a significant impact on stress indicators. This assumption was based on

the understanding that the influence of physical activity on heart rate (HR) and heart rate variability (HRV) becomes more pronounced under conditions of intense physical exertion.

EDA: The mean of SCR peaks associated with the cognitive load of participants for the low-intensity and moderate-intensity scenarios were 36.61 (SD=15.43) and 36.5 (SD=15.22), respectively. Also, the paired t-test showed a p-value of 0.9778, suggesting that in a low-intensity scenario, the mean of SCR peaks has no significant difference as compared to the moderate intensity, as shown in Figure 6(b).

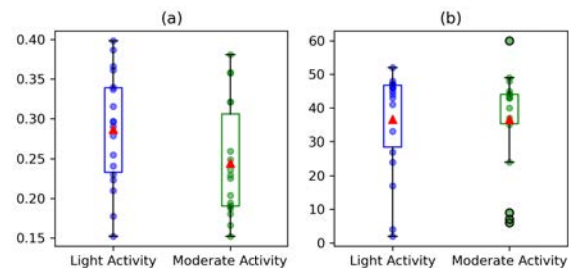


Figure 6. (a) HF-HRV and (b) Number SCR Peaks in Moderate and Light Activity

The findings revealed significant physiological differences between light and moderate activities, particularly in HRV measures of HF, RMSSD, NN50, and pNN50, which are used for assessing the autonomic nervous system's balance. The moderate activity demonstrated a reduction in HRV and a lower mean NN50, indicating increased stress levels and a potential decrease in parasympathetic activity. Furthermore, the analysis did not find significant differences in the low and moderate scenarios as measured by the number of SCR peaks, suggesting that the type of physical activity may not impact cognitive stress indicators in controlled conditions.

5 Conclusion

This study aimed to assess the stress levels experienced by construction workers as they received warnings during regular roadway work zone tasks, examining the effects within the categories of light or moderate-intensity tasks. The experiment model used physiological signals collected from a wristband-type sensing technology device while they carried out routine highway maintenance work in a VR environment. Results from 18 participants in the simulated work zone through a virtual environment indicate that (1) workers had a significantly lower heart rate variability and mean heart rate when they did medium-intensity activity than low-intensity activity; (2) moderate activity recorded a lower mean NN50 and pNN50 as com-

pared to the light activity; (3) the HF-HRV across these two activities show a significant difference between them; (4) there are no significant differences between the SCR peaks for the low and moderate activity scenarios. The human-sensing stress recognition model introduced in this study offers two significant contributions: (1) It incorporates workers' physiological characteristics to assess their stress levels while receiving an AR-enabled warning through a combination of audio, visual, and haptic signals. (2) It further clarifies the relationship between worker stress, workload, and activities that demand physical effort.

This study acknowledges a few limitations that should be addressed in future research. Expanding the sample size of participants to include a more diverse and larger population would enhance the generalizability of the findings. Moreover, the study did not explore the effects of gender, race, age, and disability, areas which future studies could consider to deepen the understanding of stress recognition in varying demographics. Additionally, the use of head-mounted VR displays may introduce motion sickness and dizziness, potentially confounding stress measurements. The varied backgrounds of participants, including their AR/VR experience and onsite experience, could also influence the results, suggesting a need for subgroup analysis to reveal these effects. Furthermore, the experiment was limited to tasks categorized as light and medium in terms of activity level. Future research could benefit from incorporating a broader range of tasks, including those classified as heavy, to more thoroughly assess the impact of workload and physical demand on stress levels.

References

- [1] Federal Highway Administration. Work Zone Safety and Mobility Data, Facts, and Statistics. https://ops.fhwa.dot.gov/wz/resources/facts_stats.htm. Accessed on Dec 4, 2023.
- [2] Houtan Jebelli, Sungjoo Hwang, and SangHyun Lee. Eeg-based workers' stress recognition at construction sites. *Automation in Construction*, 93:315–324, 2018.
- [3] Soram Lim, Seokho Chi, Joon Deuk Lee, Hoon-Jin Lee, and Hyunjung Choi. Analyzing psychological conditions of field-workers in the construction industry. *International Journal of Occupational and Environmental Health*, 23(4):261–281, 2017.
- [4] Yue Pan and Limao Zhang. Roles of artificial intelligence in construction engineering and management: A critical review and future trends. *Automation in Construction*, 122:103517, 2021.
- [5] Houtan Jebelli, Byungjoo Choi, and SangHyun Lee. Application of wearable biosensors to construction sites. ii: Assessing workers' physical demand. *Journal of construction engineering and management*, 145(12):04019080, 2019.
- [6] Mei-yung Leung, Isabelle Yee Shan Chan, and Jingyu Yu. Preventing construction worker injury incidents through the management of personal stress and organizational stressors. *Accident Analysis & Prevention*, 48:156–166, 2012.
- [7] HF Van der Molen and PLT Hoonakker. Work stress in the construction industry: causes and measures. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 44, pages 5–651, 2000.
- [8] Justin R Davis, Adam D Campbell, Allan L Adkin, and Mark G Carpenter. The relationship between fear of falling and human postural control. *Gait & posture*, 29(2):275–279, 2009.
- [9] Sepehr Sabeti, Nichole Morris, and Omidreza Shoghli. Mixed-method usability investigation of arrows: Augmented reality for roadway work zone safety. *International Journal of Occupational Safety and Ergonomics*, 0(ja):1–40, 2023. doi:10.1080/10803548.2023.2295660.
- [10] Sepehr Sabeti, Omidreza Shoghli, Mohammadreza Baharani, and Hamed Tabkhi. Toward ai-enabled augmented reality to enhance the safety of highway work zones: Feasibility, requirements, and challenges. *Advanced Engineering Informatics*, 50: 101429, 2021.
- [11] Sepehr Sabeti, Omidreza Shoghli, and Hamed Tabkhi. Toward wi-fi-enabled real-time communication for proactive safety systems in highway work zones: A case study. In *Construction Research Congress 2022*, pages 1166–1173, 2022.
- [12] Xiyuan Hou, Yisi Liu, Olga Sourina, Yun Rui Eileen Tan, Lipo Wang, and Wolfgang Mueller-Wittig. Eeg based stress monitoring. In *2015 IEEE International Conference on Systems, Man, and Cybernetics*, pages 3110–3115. IEEE, 2015.
- [13] Salam Ranabir and K Reetu. Stress and hormones. *Indian journal of endocrinology and metabolism*, 15 (1):18, 2011.
- [14] Mona Arabshahi, Di Wang, Junbo Sun, Payam Rahnamayiezekavat, Weichen Tang, Yufei Wang, and Xianguyu Wang. Review on sensing technology adoption in the construction industry. *Sensors*, 21(24): 8307, 2021.

- [15] Chukwuma Nnaji, Ibukun Awolusi, JeeWoong Park, and Alex Albert. Wearable sensing devices: towards the development of a personalized system for construction safety and health risk mitigation. *Sensors*, 21(3):682, 2021.
- [16] Shahrad Shakerian, Mahmoud Habibnezhad, Amit Ojha, Gaang Lee, Yizhi Liu, Houtan Jebelli, and SangHyun Lee. Assessing occupational risk of heat stress at construction: A worker-centric sensor-based approach. *Safety science*, 142:105395, 2021.
- [17] Yimin Qin and Tanyel Bulbul. Electroencephalogram-based mental workload prediction for using augmented reality head mounted display in construction assembly: A deep learning approach. *Automation in Construction*, 152:104892, 2023.
- [18] Sajjad Rostamzadeh, Alireza Abouhossein, Shahram Vosoughi, Saeid Bahramzadeh Gendeshmin, and Rasoul Yarahmadi. Stress influence on real-world driving identified by monitoring heart rate variability and morphologic variability of ecg signals: The case of intercity roads. *International Journal of Occupational Safety and Ergonomics*, pages 1–35, 2023.
- [19] Gaang Lee, Byungjoo Choi, Houtan Jebelli, and SangHyun Lee. Assessment of construction workers' perceived risk using physiological data from wearable sensors: A machine learning approach. *Journal of Building Engineering*, 42:102824, 2021.
- [20] Waleed Umer. Simultaneous monitoring of physical and mental stress for construction tasks using physiological measures. *Journal of Building Engineering*, 46:103777, 2022.
- [21] Stefano Betti, Raffaele Molino Lova, Erika Rovini, Giorgia Acerbi, Luca Santarelli, Manuela Cabiati, Silvia Del Ry, and Filippo Cavallo. Evaluation of an integrated system of wearable physiological sensors for stress monitoring in working environments by using biological markers. *IEEE Transactions on Biomedical Engineering*, 65(8):1748–1758, 2017.
- [22] Martin Gjoreski, Mitja Luštrek, Matjaž Gams, and Hristijan Gjoreski. Monitoring stress with a wrist device using context. *Journal of biomedical informatics*, 73:159–170, 2017.
- [23] Monika Lohani, Brennan R Payne, and David L Strayer. A review of psychophysiological measures to assess cognitive states in real-world driving. *Frontiers in human neuroscience*, 13:57, 2019.
- [24] Tizen IoT Documentation. https://docs.tizen.org/iot/api/5.0/tizen-iot-headed/group__CAPI__SYSTEM__FEEDBACK__MODULE.html. Accessed: [April 6, 2024].
- [25] Hannu Virokannas. Thermal responses to light, moderate and heavy daily outdoor work in cold weather. *European journal of applied physiology and occupational physiology*, 72:483–489, 1996.
- [26] Texas MUTCD. Manual on uniform traffic control devices. *Texas Department of Transportation, Austin*, 2006.
- [27] Sepehr Sabeti, Fatemeh Banani Ardecani, and Omidreza Shoghli. Augmented reality warnings in roadway work zones: Evaluating the effect of modality on worker reaction times. *arXiv preprint arXiv:2403.15571*, 2024.
- [28] Embraceplus. <https://www.empatica.com/embraceplus/>. Retrieved from <https://www.empatica.com/embraceplus/>.
- [29] Natalie A Snyder and Michael E Cinelli. Aperture crossing in virtual reality: physical fatigue delays response time. *Journal of Motor Behavior*, 54(4): 429–437, 2022.
- [30] PMAHPDSP Gomes, Petra Margaritoff, and Hugo Silva. pyhrv: Development and evaluation of an open-source python toolbox for heart rate variability (hrv). In *Proc. Int'l conf. On electrical, electronic and computing engineering (icetran)*, pages 822–828, 2019.
- [31] David P. Allen. A frequency domain hampel filter for blind rejection of sinusoidal interference from electromyograms. *Journal of Neuroscience Methods*, 177(2):303–310, 2009. ISSN 0165-0270.
- [32] Dominique Makowski, Tam Pham, Zen J Lau, Jan C Brammer, François Lespinasse, Hung Pham, Christopher Schölzel, and SH Annabel Chen. Neurokit2: A python toolbox for neurophysiological signal processing. *Behavior research methods*, pages 1–8, 2021.
- [33] Fred Shaffer and Jay P Ginsberg. An overview of heart rate variability metrics and norms. *Frontiers in public health*, page 258, 2017.
- [34] Hye-Geum Kim, Eun-Jin Cheon, Dai-Seg Bai, Young Hwan Lee, and Bon-Hoon Koo. Stress and heart rate variability: a meta-analysis and review of the literature. *Psychiatry investigation*, 15(3):235, 2018.

Worker-Centric Path Planning: Simulating Hazardous Energy to Control Construction Safety Using Graph Theory

Kilian Speiser and Jochen Teizer

Department for Civil and Mechanical Engineering, Technical University of Denmark

kilsp@dtu.dk, teizerj@dtu.dk

Abstract

Occupational accidents in the European Union continue to pose a significant threat to construction workers, with skill-based errors contributing substantially to these incidents. Virtual training gains prominence in improving skills, but evaluating trainee performance based on safety behavior is challenging to quantify. This paper introduces a pioneering worker-centric simulation that assesses the hazardous energy that a worker may be exposed to while navigating a construction site. The result of the simulation is a Safety Graph, aiding in determining the safest routes for workers. The graph is generated based on a known construction site geometry, with each node representing a one-square-meter area. The simulation developed in the game engine Unity calculates hazardous energy associated with falls and trips that a worker is exposed to when moving between nodes. The evaluation demonstrates a 97% accuracy in estimating hazardous energy. A practical application in virtual training demonstrates how the approach allows for quantifying the safety performances of workers. The study, however, reveals minor shortcomings in the simulation, such as considering an energy threshold or incorporating more hazard types. The results also indicate further applications of the Safety Graph, hinting at its potential in hazard detection or forwarding the safest paths to construction workers using smart glasses.

Keywords –

Occupational Construction Safety, Health, and Well-being; Hazard Simulation; Safe Path Planning; Alternative Routes; Game Engine; Graph Theory; Virtual Education and Training.

1 Introduction

Occupational accidents remain high in the European Union (EU). Approximately 6% of construction workers in the EU face accidents each year [1]. These numbers indicate that, on average, a construction worker

undergoes 2.4 accidents in a 40-year career. While various factors contribute to the occurrence of incidents within the construction industry, it is noteworthy that a significant proportion of these incidents are attributed to skill-based errors made by workers [2].

Efficient training is a crucial method to improve the skills of the workforce, with virtual training methods gaining more interest over the last years. Several studies propose virtual training to improve hazard identification [3], train tool handling [4], or collaboration with equipment based on real-world location data [5]. However, the assessment of the trainees in virtual training experiences is challenging. While some studies rely on manual and subjective evaluation methods [6], others analyze collected data [7–9]. Nevertheless, it is challenging to compare the performance of trainees and assess the collected data in a meaningful manner [6,8].

For instance, three workers navigate a construction site: Worker A chooses a safe but longer path, Worker B takes the shorter path with several minor hazards on the way, and Worker C chooses a path with one high-energy hazard that would most likely result in a fatal accident. While subjective evaluation may somewhat evaluate the paths and conclude that Worker A performed the best, an objective data-driven assessment is quantifiable and unbiased. Nevertheless, we would not know if any worker chose the safest available path. To facilitate such assessment, path planning algorithms may allow safety trainers to compare a worker's chosen route with the optimal route. Such optimal routes can be determined using path-planning algorithms.

Path planning is widely adopted in various industries. With the emergence of robots and autonomous vehicles, path planning has also become more relevant in construction. Path planning algorithms can utilize graph theory to find a feasible or optimal route from a starting point to a target. Graph theory is a mathematical approach to model relationships between entities. A graph comprises a set of nodes and a set of edges connecting these nodes [10]. Nodes and edges can represent a wide range of entities and relationships in various fields. Among others, graph theory has been

applied to solve assignment problems, transportation problems, knowledge representation, or path planning [10]. To find the optimal route, the edges of the graph, whether directed or undirected, connect the nodes with associated costs. The cost indicates the expenses occurring when using this specific edge. While there is no constraint as to what such cost may represent, standard measures are the traveling time or distance [10]. The path producing the lowest cost represents the optimal path.

In construction, various studies investigated path planning problems. Several studies plan operation paths for construction vehicles [11–13]. Other studies determine optimal routes for crane operations [14] using Building Information Modelling (BIM) [15,16]. Notably, Hu and Fang include safety aspects [15], while the other studies purely focus on productivity aspects. Fewer studies address workers' path planning. Cheng et al. determine paths using workers' trajectory data [11]. Wang and Qin determine safe paths by assessing fall hazards using BIM models [17]. It is noteworthy that all these studies determine path planning using environmental conditions without addressing decisions made by the humans involved when operating. Kim et al. address this shortcoming by integrating deep reinforcement learning to mimic workers' decision-making processes [18].

Building upon the existing literature, the absence of a worker-centric path-finding approach that focuses explicitly on determining the safest paths becomes evident. No previous approach provided a graph that allows for an objective assessment of chosen paths. Such a solution, however, could be integrated into virtual training, as indicated above, but it also allows for the comparison of safety behavior in real-world settings. Hence, the objective of this research is to create a worker-centric algorithm to generate a *Safety Graph* based on the known geometry of a construction site. The generated graph facilitates the risks for a worker to travel between locations within the construction site using the hazardous energy of two of the four most common accidents in the EU [1], namely falls and trips. This *Safety Graph* contains valuable information and can function as input to determine the safest routes for construction workers when navigating through construction sites.

The path planning simulation will be integrated into a virtual training environment to ease the evaluation of trainees' performances. A brief experiment will showcase this application at the end of this paper. A different practical application of the *Safety Graph* could forward the safest path to a construction worker wearing smart glasses and help them navigate the actual construction site. By the end of this paper, readers will gain insights into a cutting-edge approach that combines graph theory, safety assessment, and path planning, which could be used for several practical applications.

2 Research Method

This study employed a mixed research approach to create and assess the proposed simulation. First, a comprehensive literature review investigated worker-centric path planning in construction. Subsequently, a research gap was identified, and the requirements and goals to address this gap were specified. Based on the requirements, a simulation was developed using the Unity game engine. This simulation generates a *Safety Graph* facilitating hazardous energies for a worker.

The simulation was evaluated in an artificial testing environment with fall and trip hazards to validate the approach. Based on the evaluation, the authors could refine the algorithm. To validate the accuracy of the simulation, the authors compared the results from the simulation to a manually created validation graph. Finally, the resulting graph was integrated into a virtual training environment to demonstrate one application of such a *Safety Graph*.

3 Simulation Generating the Graph

The *Safety Graph* assesses the potential hazards faced by human workers while navigating a construction site. This paper presents an algorithm embedded in a simulation framework designed to create such a graph. The following section first overviews the *Safety Graph* before describing the algorithm for creating the graph.

3.1 Algorithm Creating the Safety Graph

As described before, a graph consists of nodes and edges. To generate the nodes, we distributed the construction site into one-meter squares. Each square represents one node in the *Safety Graph*, and adjacent nodes, which a worker can reach, are connected by edges (see Figure 1a). Each edge has a cost, representing the hazardous energy a worker faces when traveling from one node to another. While this cost should encompass all potential hazards like falling, tripping, being struck by objects, or electrocution, this work only integrates falls and trips as they account for more than 40 percent of accidents in construction [1]. Wang and Qin have considered fall hazards in path planning [17], but no previous work included trip hazards in path planning for workers. The novelty of this work lies in generating the *Safety Graph* by simulating the hazardous energy impact on the worker for each possible movement in the construction site. The resulting graph can then function as the base for finding the safest route in the graph (see the blue graph in Figure 1) using search algorithms like Dijkstra [19]. While Figure 1 depicts the shortest path between a particular start and end node, it is crucial to note that graphs enable finding the shortest path for any combination of start and end nodes [10].

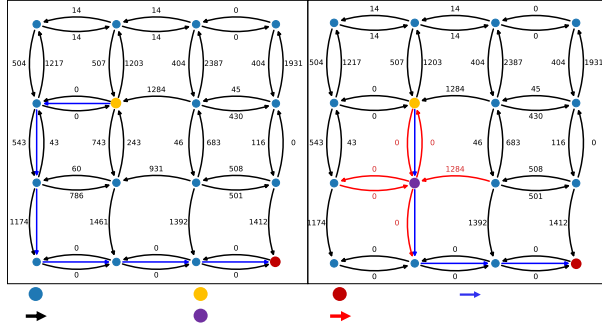


Figure 1. (a) Brief example of a Safety Graph and (b) updating after changes in the construction site.

The benefits of having such a graph are twofold. First, in a static environment, the algorithm runs once. After that, path planning algorithms can efficiently use the graph to identify optimal routes and do not require continuously re-evaluating the site layout. Second, in dynamic environments like construction sites, our graph-based approach allows for updating specific areas of the map without repeating the entire simulation. Figure 1b illustrates the concept: The construction site has changed as, for instance, an object was moved by machinery. Now, the simulations must only update edges connected to this node (highlighted in red in Figure 1b). The update of the node and connecting edges also changes the optimal path from the starting node to the target (see blue edges in Figure 1b).

This section has described the concept of the *Safety Graph* and the motivation behind creating it. The following parts will delve into how our algorithm generates this graph and how the Unity simulation we developed calculates the costs associated with the edges.

3.2 Simulation Development in Unity

The described *Safety Graph* is generated using the algorithm implemented in a Unity simulation. The developed algorithm simulates each possible motion of a worker from one node to an adjacent node and measures the hazardous energy impact on the worker during this motion. For instance, if a worker jumps one meter down during this motion, the hazardous energy of this jump will be considered the cost of this particular movement. Similarly, the energy can be determined for other kinds of hazards such as trips, electrocution, struck-by, or caught-in-between.

The underlying idea of utilizing hazard energy as the edge cost goes back to the findings that the energy of a hazard correlates with the result of an accident [20]. According to Hallowell et al., hazards between 500 Joule and 1,500 Joule are likely to result in medical treatment, while hazards with more than 1,500 Joule most likely cause a severe injury or fatality [21]. The energy intensity would provide more insight, but it would be more

challenging to obtain [21]. Thus, in this work, we compute the hazardous energy that a human worker is exposed to when traveling from one location in the construction site to adjacent locations.

To determine the hazardous energy, we developed a simulation framework in the game engine Unity. As Unity facilitates real-world physics, the hazardous energy can be calculated. While the authors expect this algorithm to be generally feasible, obtaining the hazardous energy through the game engine is only one proposed approach. Other methods may utilize camera footage from a site to detect hazards and corresponding safety potential. Nevertheless, this approach has several advantages: It not only includes the safety aspect but also allows for the determination of other parameters, such as accessibility for agents or expected time of traveling.

Listing 1 shows the algorithm to create the *Safety Graph*. The algorithm first distributes the construction site in nodes, adds the agent, and instantiates an empty set of edges (lines 1-3). Then, for each of the nodes, the agent is moved to the neighboring squares. If the motion is possible (no obstacles), the energy potential is calculated, and an edge is added to the graph (lines 5-13). Eventually, the algorithm returns the completed graph in line 14. The following section will first describe the setup of the Unity scene before illustrating the underlying concept of calculating the hazardous energy.

Listing 1. Algorithm generating the *Safety Graph*.

```

1  Function CreateSafetyGraph
2    nodes = SegrateSite()
3    edges = empty collection
4    agent = CreateAgent()
5    For each node in nodes:
6      If edge = SimMotion(agent, node, 1, 0) is not null
7        Then add edge to edges
8      If edge = SimMotion(agent, node, -1, 0) is not null
9        Then add edge to edges
10     If edge = SimMotion(agent, node, 0, 1) is not null
11       Then add edge to edges
12     If edge = SimMotion(agent, node, 0, -1) is not null
13       Then add edge to edges
14  Return Graph G with nodes and edges

```

The simulation undergoes testing in an artificial environment, depicted in Figure 2. This setting comprises several objects that could cause a trip, such as cement bags, a fuel canister, stairs, and a rebar laying area. The objects in the scene were added from different Unity assets. Moreover, the scene features elevated areas from where the worker may fall to a lower level, with some sections safeguarded by guardrails while others remain unprotected. To ensure a thorough evaluation of the algorithm, the authors included several configurations to ensure that the environment is functioning for testing purposes. The algorithm is implemented in Unity, and all related scripts and results are available in a data repository [22].

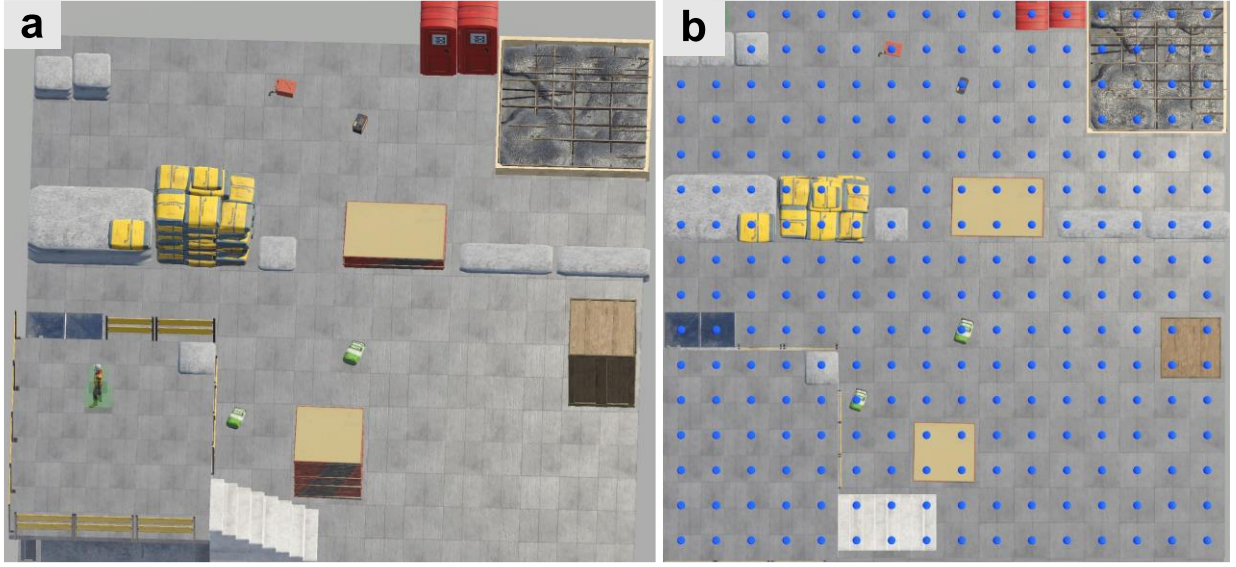


Figure 2. (a) Construction scene in Unity and (b) segregated into nodes (in blue).

3.3 Hazardous Energy Simulation

The proposed algorithm initializes the Graph G , utilizes the geometry of the environment, segregates it into one-meter squares, places a node at each center point (see blue dots in Figure 2b), and adds the nodes to G . After generating the nodes, the algorithm iterates through the nodes and assesses the hazardous energy associated with traveling from each node to its adjacent tiles. As indicated in Listing 1, the simulation of the motions occurs four times for each node. Figure 3 illustrates the process for simulating one motion, and the following paragraphs will describe each of the steps.

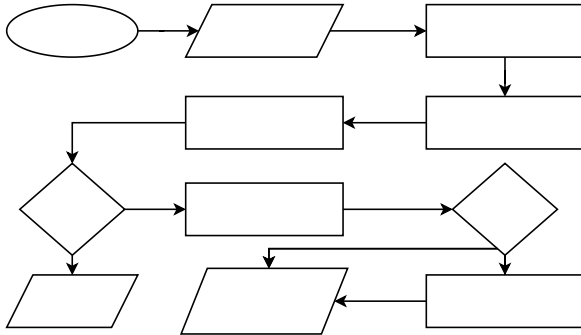


Figure 3. Flow chart for risk assessment of one individual movement of the agent.

The simulation of an individual motion to a neighboring node begins with placing the avatar on the current node n . To ensure the correct vertical placement, we cast a ray downwards from a 100m offset using Unity's *RayCastHit* structure. Then, we place the agent in the position of the first hit.

After placing the avatar, the simulation waits for 100

milliseconds to ensure that the avatar is placed correctly. Then, the algorithm queries the neighboring node n_{adj} from G , and moves the agent towards n_{adj} using the method *Rigidbody.MovePosition* in Unity. This method moves the agent, but only if the path is free. After moving the agent, the simulation stops for 200ms to evaluate the impact of the motions. For instance, if the worker moves from an unprotected leading edge, the simulation needs to wait until the agent touches the ground. The vertical gravity in Unity is set to -300 to accelerate the simulation.

In case of colliding objects during the movement from n to n_{adj} , the motion will fail, and the simulation continues with the next neighboring node. However, if the motion is successful, our algorithm adds an edge e to the G and calculates the hazardous energy $E_{Hazardous}$ during the motion as the cost of e .

The energy for the fall hazards is computed with Equation 1 where m corresponds to the mass of the worker, g represents the gravity, and Δh corresponds to the difference in height before and after moving the agent.

$$E_{Fall} = m \times g \times \Delta h \quad (1)$$

$$E_{Trip} = \frac{1}{2}mv^2 = \frac{78kg \times 10km^2}{2h^2} \approx 300J \quad (2)$$

$$E_{Hazardous} = E_{Fall} + E_{Trip} \quad (3)$$

Equation 2 regulates the hazardous energy for trips, assuming a constant value of 300 Joule. This value corresponds to the kinetic energy of an average human, assuming a mass of 78kg and a velocity of 10km/h. The velocity is selected rather high, integrating a safety factor. The trip hazard is detected using a collider object at the bottom of the worker (see Figure 4). When this object collides with any other object while moving the avatar, the trip hazard energy is added to the cost of e .

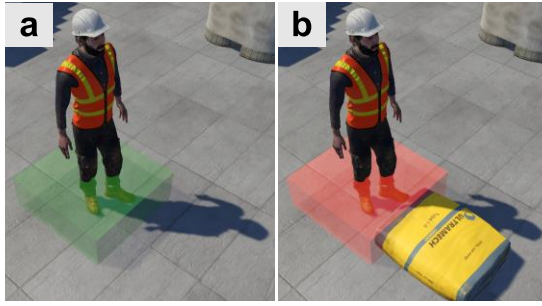


Figure 4. The concept for detecting tripping hazards with (a) a safe path without hazards and (b) a tripping hazard being present.

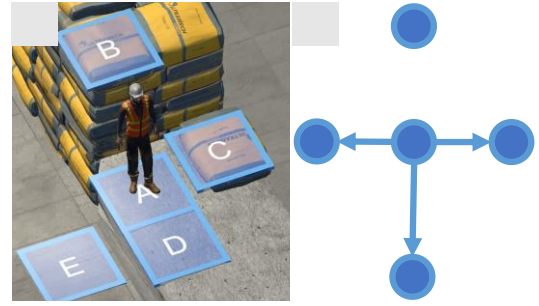


Figure 5. Generating the edges for one node (assuming no diagonal movement is possible).

Figure 5 illustrates the simulation for one node, repeating the previously demonstrated simulation for each adjacent node. There is no edge between node *A* and *B*, as the avatar cannot reach *B* from *A*. Moving to *D* is associated with no risk as there are neither objects causing trips nor an elevation. When moving to *C*, a trip hazard is detected, and the cost of the motion is set to 300 Joule. Lastly, moving to *E* includes a potential fall of 0.6m (hazardous energy of 400 Joule).

4 Results and Discussion

The simulation was executed, and the resulting graph was exported into a JSON file for further processing. To validate the approach, the graph was first evaluated regarding its accuracy and then applied in a training scenario with four trainees to find the safest path. Figure 6 depicts the resulting graph highlighting additional information, which the following sections describe and discuss in detail.

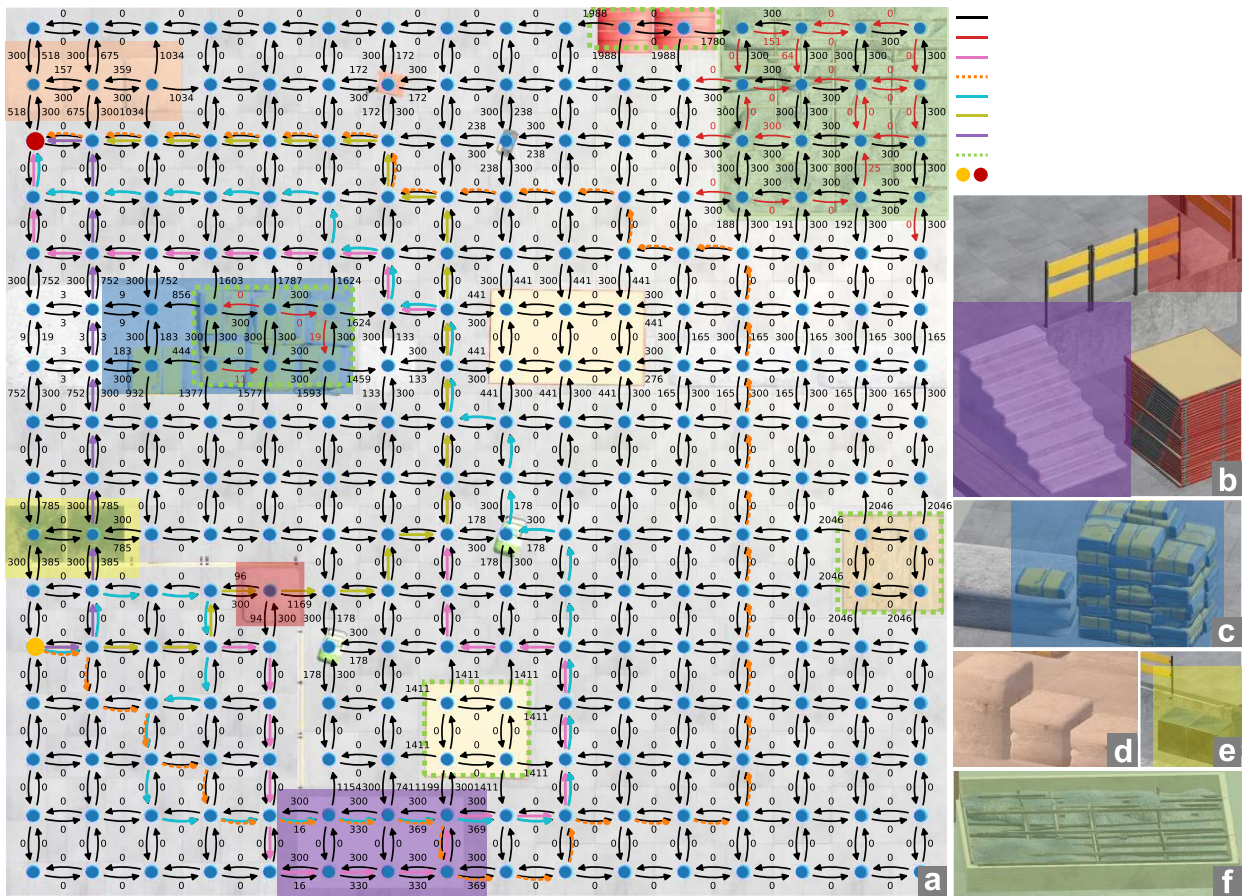


Figure 6. The safety graph includes the hazardous energy from the simulation, critical areas, and the paths taken by the trainees when applying the concept within a virtual training environment.

4.1 Accuracy Evaluation

The simulation's accuracy was evaluated by manually assessing each of the worker's motions. We categorized each of the 1024 edges into one of the five categories shown in Table 1 and created a validation graph (available at [22]). The categories were selected to make sure to detect both falls and trips without losing too much accuracy. Then, we converted the results from the simulation into the same categories based on the hazardous energy. This method allows for the evaluation of the simulation without calculating the exact expected hazardous energy for each of the edges. Still, we manually assessed the exact calculation of the hazardous energy using the area with differently elevated spots shown in Figure 6d.

Table 1. Categorized edges in the validation set.

Category	Category description	Simulated hazardous energy [J]
0	No hazard	Not accessible
1	$Fall(\Delta h < 0.4m)$	$0 < cost < 300$
2	Trip	$cost = 300$
3	$Fall(0.4m < \Delta h < 1m)$ or $[Fall(\Delta h < 0.4m) \& Trip]$	$300 < cost < 785$
4	$Fall(1m < \Delta h < 2m)$ or $[Fall(1m < \Delta h < 1.6m) \& Trip]$	$785 < cost < 1500$
5	High energy hazard	$cost > 1500$

Using the Python package NetworkX, we compared the edges from the validation graph with the simulation graph. Figure 6a highlights the incorrect simulated edges in red. The comparison of the graphs suggests that the simulation estimates for 97% of the edges the hazardous energy correct. 28 connections are incorrect, of which 25 relate to an underestimate of the hazardous energy and 3 to an overestimate. In 20 cases, the simulation did not detect the expected trip hazard. All these cases occurred in the rebar laying area shown in Figure 6f. Here, the surface is uneven but not uneven enough for the simulation to detect the trip hazard. Similarly, in this area, the simulation detected a trip in five cases, while hazardous energy was expected due to a fall hazard.

Four errors relate to the cement bags in the middle of the scene (Figure 6c), where the surface is also uneven. The hazardous energy while walking on this pile could be considered as a trip or a fall. Nevertheless, it is improbable that a worker would walk up there.

4.2 Safe Path-Planning in Virtual Training

The second step in the validation represents a short case study by integrating the results into a virtual training environment. Four researchers were asked to choose a safe path to navigate from a starting node to a target node in the virtual training environment, as shown in Figure 2.

They navigated an avatar in a desktop-based first-person view using a keyboard and mouse (Figure 7). During the training, a Unity script records the trajectory, which we superimpose with the *Safety Graph*. Figure 6a illustrates the chosen paths of the four trainees and the optimal path determined with the Dijkstra algorithm using NetworkX.



Figure 7. The trainee navigates in a desktop-based first-person view using a keyboard and mouse.

Table 2 summarizes the results and suggests that Trainee 4 chose the shortest but most hazardous path. Trainee 4 did not take the stairs but instead jumped down at an unsecured area (Figure 6e) and later crossed the elevated area in Figure 6c. The paths of Trainees 1, 2 and 3 result in a similar hazardous energy. However, Trainee 3 chose a shorter path than Trainees 1 and 2. Trainee 3 jumps down at the missing guardrail while Trainees 1 and 3 take the stairs. This situation reveals a shortcoming of our approach. As we aggregate the hazardous energies for each edge, several minor hazards may lead to a worse route than a path with one severe hazard, like Trainee 3. Future improvements should weigh higher energy hazards more than low energy hazards. Another approach would be the removal of edges higher than a certain energy threshold, as the likelihood of an injury is too high.

Table 2. Comparison of the paths of the four trainees.

Trainee	Hazardous Energy [Joule]	Distance [number of edges]	Duration [seconds]
Trainee 1	1,549	41	24
Trainee 2	1,562	37	25
Trainee 3	1,469	22	23
Trainee 4	2,225	11	10
Optimal	1,045	32	-

Another observation relating to the tripping hazards is that Trainees 1 and 2 chose the stairs but later exposed themselves to hazardous energy relating to tripping. In our study, the hazardous energy for tripping hazards is assumed to have a static value. This value should be reconsidered, as a trip close to another hazard could result in a more severe injury. For instance, if the tripping object is at a location where a worker could fall, this edge should be considered with a higher risk.

4.3 Summary

The simulation generally estimates the risk correctly. Further improvements, however, must address uneven surfaces, classify high-energy hazards as unreachable, and consider the secondary impact of detected tripping hazards. In addition, moving elements have not been considered in the current implementation. A potential approach to facilitate dynamic elements is to update the parts of the graph where the changes occurred, as demonstrated in Figure 1.

By connecting the simulation to other data sources, such as a digital twin, the location of other hazard types could be integrated [23]. For instance, detecting hazard zones for struck-by hazards [24] would allow the *Safety Graph* to remove the components within such zones. The simulation would only require the geometry of the construction site from the digital twin. For instance, a BIM model or a point cloud collected by a laser scanner should be sufficient when converted into a mesh. However, further research must validate this expectation.

Additionally, the edges could be accessible to some workers depending on their role. For instance, while the rebar layer should access the area highlighted in Figure 6f, the electrician must not go there. Another approach to declaring such hazard zones and integrating them is the detection based on the actual worker's path [25]. Integrating the trajectory would also allow us to compare how often the workers choose the safe path and help identify tailored training for workers [26].

5 Conclusion and Outlook

While existing studies have explored path planning for construction vehicles and crane operations, few studies focus on workers. A primary contribution of our work is the introduction of a worker-centric path-finding approach, creating a *Safety Graph*.

By representing the construction site as a graph of nodes and edges and assigning costs to edges based on hazardous energy simulations, this research provides a systematic and structured approach to safety evaluation. This novel application of graph theory contributes to a more objective, quantifiable, and data-driven safety analysis and eventually allows for evaluating the safety performance of workers in safety training.

The evaluation of the simulation accuracy, involving the categorization and comparison of edges, adds a layer of robustness to this work. Applying the concept in a virtual training scenario highlights the potential benefits of the approach. Lastly, the study indicates that the *Safety Graph* can demonstrate additional benefits in other applications, such as hazard detection.

Future research can expand the capabilities of the agent by allowing a broader range of motions, thus enhancing the simulation's realism. Additionally,

incorporating more types of hazards and consideration of secondary impacts could further refine the graph's ability to assess hazards in dynamic construction environments accurately. Including other agents and dynamic hazards such as a compactor or forklift represents another challenge that future research should investigate. An in-depth study within a virtual training environment is recommended to comprehensively evaluate the impact of this approach on evaluating workers' safety performance. Here, interviews with the trainees could reveal interesting findings on why they chose different routes.

This exploration can contribute to the development of advanced methods for next-generation construction safety training, providing active and personalized feedback to trainees.

Furthermore, testing the proposed approach in a real-world setting, such as forwarding optimal routes to construction workers through an augmented reality (AR) headset, presents an exciting opportunity to validate the practicality and effectiveness of the *Safety Graph*. Such tests would also allow us to investigate the sufficiency of informing workers about the location of hazards or if other measures need to be implemented, e.g., stopping machines or blocking paths.

Acknowledgments

This research has been funded by the European Union Horizon 2020 research and innovation program under grant agreement no. 95398.

References

- [1] Eurostat. Accidents at work by NACE Rev. 2 activity and size of enterprise. https://ec.europa.eu/eurostat/databrowser/view/H SW_N2_05/Default/Table?Lang=en, 2023.
- [2] Garrett JW, Teizer J. Human Factors Analysis Classification System Relating to Human Error Awareness Taxonomy in Construction Safety. *J Constr Eng Manag*, 135, 2009. [https://doi.org/10.1061/\(asce\)co.1943-7862.0000034](https://doi.org/10.1061/(asce)co.1943-7862.0000034).
- [3] Wolf M, Teizer J, Wolf B, Bükür S, Solberg A. Investigating hazard recognition in augmented virtuality for personalized feedback in construction safety education and training. *Advanced Engineering Informatics*, 51, 2022. <https://doi.org/10.1016/j.aei.2021.101469>.
- [4] Bükür S, Wolf M, Böhm B, König M, Teizer J. Augmented virtuality in construction safety education and training. *EG-ICE 2020 Workshop on Intelligent Computing in Engineering*, 2020.
- [5] Speiser K, Hong K, Teizer J. Enhancing the realism of virtual construction safety training: integration of real-time location systems for real-world hazard

- simulations. *23rd International Conference on Construction Applications of VR*, 2023. <https://doi.org/10.36253/979-12-215-0289-3.15>.
- [6] Salinas D, Muñoz-La Rivera F, Mora-Serrano J. Critical Analysis of the Evaluation Methods of Extended Reality (XR) Experiences for Construction Safety. *Int J Env Res Public Health*, 19, 2022. <https://doi.org/10.3390/ijerph192215272>.
- [7] Speiser K, Teizer J. An Efficient Approach for Generating Training Environments in Virtual Reality Using a Digital Twin for Construction Safety. *Proceedings of CIBW099W123*, Porto, Portugal, 481-490, 2023. ISBN: 978-972-752-309-2, <https://doi.org/10.24840/978-972-752-309-2>.
- [8] Speiser K, Teizer J. An Ontology-Based Data Model to Create Virtual Training Environments for Construction Safety Using BIM and Digital Twins. *30th European Group for Intelligent Computing in Engineering Conference (EG-ICE)*, 2023.
- [9] Golovina O, Kazanci C, Teizer J, König M. Using Serious Games in Virtual Reality for Automated Close Call and Contact Collision Analysis in Construction Safety. *36th International Symposium on Automation and Robotics in Construction*, 2019. <https://doi.org/10.22260/ISARC2019/0129>.
- [10] Bondy JA, Murty USR. Graph Theory with Applications. 5th ed. New York: *Elsevier Science Publishing*, 1982.
- [11] Cheng T, Mantripragada U, Teizer J, Vela PA. Automated Trajectory and Path Planning Analysis Based on Ultra Wideband Data. *Comp in Civil Eng*, 26, 2012. [https://doi.org/10.1061/\(asce\)cp.1943-5487.0000115](https://doi.org/10.1061/(asce)cp.1943-5487.0000115).
- [12] Hammad A, Vahdatikhaki F, Zhang C, Mawlana M, Doriani A. Towards the smart construction site: Improving productivity and safety of construction projects using multi-agent systems, real-time simulation and automated machine control. *Proc. of the Winter Simulation Conference*, Berlin, Germany, 2012. <https://doi.org/10.1109/WSC.2012.6465160>.
- [13] Akegawa T et al. Loading an Autonomous Large-Scale Dump Truck: Path Planning Based on Motion Data from Human-Operated Construction Vehicles. *IEEE Intl. Conf. Intelligent Robots & Systems*, 2022. <https://doi.org/10.1109/IROS47612.2022.9981828>.
- [14] Kang SC, Miranda E. Planning and visualization for automated robotic crane erection processes in construction. *Automation in Construction*, 15, 2006. <https://doi.org/10.1016/j.autcon.2005.06.008>.
- [15] Hu S, Fang Y. Automating Crane Lift Path through Integration of BIM and Path Finding Algorithm. *37th ISARC*, 2020. <https://doi.org/10.22260/isarc2020/0072>.
- [16] Lin Z, Petzold F, Hsieh SH. Automatic Tower Crane Lifting Path Planning Based on 4D Building Information Modeling. *Computer Applications - Construction Research Congress*, 2020. <https://doi.org/10.1061/9780784482865.089>.
- [17] Wang TK, Qin C. Integration of BIM, Bayesian belief network, and ant colony algorithm for assessing fall risk and route planning. *Safety and Disaster Management - Selected Papers from the Construction Research Congress 2018*, 2018. <https://doi.org/10.1061/9780784481288.021>.
- [18] Kim M, Ham Y, Koo C, Kim TW. Simulating travel paths of construction site workers via deep reinforcement learning considering their spatial cognition and wayfinding behavior. *Autom Constr*, 147, 2023. <https://doi.org/10.1016/j.autcon.2022.104715>.
- [19] Dijkstra EW. A note on two problems in connexion with graphs. *Numer Math (Heidelb)* 1959;1. <https://doi.org/10.1007/BF01386390>.
- [20] Alexander D, Hallowell M, Gambatese J. Precursors of Construction Fatalities. II: Predictive Modeling and Empirical Validation. *J Constr Eng Manag*, 143, 2017. [https://doi.org/10.1061/\(asce\)co.1943-7862.0001297](https://doi.org/10.1061/(asce)co.1943-7862.0001297).
- [21] Hallowell MR, Alexander D, Gambatese JA. Energy-based safety risk assessment: does magnitude and intensity of energy predict injury severity? *Constr Mngmt and Econ*, 35, 2017. <https://doi.org/10.1080/01446193.2016.1274418>.
- [22] Speiser K, Teizer J. SafetyGraph: Worker-Centric Path Planning for Construction Safety Simulating Hazardous Energy [Data set]. DTU Data, 2024. <https://doi.org/10.11583/DTU.24754278>.
- [23] Teizer J, Johansen KW, Schultz CL, Speiser K, Hong K, Golovina O. A Digital Twin Model for Advancing Construction Safety. *Construction Logistics, Equipment, and Robotics, CLEaR Conference*, vol. 390, p. 201–12, 2024. <https://doi.org/10.1007/978-3-031-44021-2>.
- [24] Johansen KW, Schultz C, Teizer J. Automated spatiotemporal identification and dissemination of work crews' exposure to struck-by hazards. *Proceedings of CIBW099W123*, Porto, Portugal: 2023, p. 1–10. ISBN: 978-972-752-309-2, <https://doi.org/10.24840/978-972-752-309-2>.
- [25] Hong K, Teizer J. A data-driven method for hazard zone identification in construction sites with wearable sensors. *Proceedings of CIBW099W123*, 2023, p. 41–8. ISBN: 978-972-752-309-2, <https://doi.org/10.24840/978-972-752-309-2>.
- [26] Speiser K, Teizer J. Automatic creation of personalised virtual construction safety training using digital twins. *Proceedings of the Institution of Civil Engineers - Management, Procurement and Law*, 2024.

Experimental Evaluation of Exoskeletons for Rebar Workers Using a Realistic Controlled Test

Malcolm Dunson-Todd¹, Mazdak Nik-Bakht¹, Amin Hammad^{1,2}

¹Gina Cody School of Engineering and Computer Science, Concordia University, Canada

²Corresponding author: amin.hammad@concordia.ca

Abstract – Workers in the construction industry are exposed to the risk factors of high forces, repetitive tasks, and awkward postures, and consequently suffer from work-related musculoskeletal disorders (WMSDs). Occupational exoskeletons (OEs) are promising interventions to reduce WMSD rates in the construction trades. Previous work has evaluated the efficacy of OEs in controlled laboratory conditions and semi-realistic test courses; however, there are no standard evaluation methods for the construction industry. Standardizing the methods for evaluating and comparing the efficacy of OEs for construction workers is essential for the adoption of effective OEs in the construction industry. Toward this goal, a realistic, controlled, and repeatable test was implemented to evaluate the efficacy of back-support exoskeletons (BSEs) for the trade of rebar installation. The test was implemented at a steel trades school where nine experienced student participants performed the test course with and without wearing an OE. Objective effects on lower-back muscle activity and subjective effects on discomfort, effectiveness, obstruction, and usability were measured. The study demonstrates the initial implementation of the test, and the results show objective and subjective evidence that the OE reduces loads in the lower back during realistic rebar installation tasks.

Keywords –

Occupational Exoskeleton; Construction; Standard Testing

1 Introduction

Construction workers face considerable health and safety risks on construction sites. For example, reinforcing ironworkers (rebar workers) are exposed to heavy loads, prolonged awkward postures, and repetitive tasks, increasing the risk of significant musculoskeletal stress that leads to work-related musculoskeletal disorders (WMSDs), such as lower back pain and herniated discs [1–3]. These risks can be mitigated by replacing human workers with robots in certain cases, or by supporting the workers with ergonomic tools and

wearable technologies. In the case of the rebar trade, workers can be supported by existing devices, such as standing rebar tying tools to reduce the need for bending. The emerging technology of occupational exoskeletons (OEs) has the potential to provide more effective and widely adopted ergonomic tools in the construction industry, compared to existing interventions. OEs have been adopted as tools to reduce WMSD rates in many industries, including construction, with back-support exoskeletons (BSEs) and shoulder-support exoskeletons (SSEs) being the most common designs. OEs can be evaluated in terms of immediate effects in laboratory studies (efficacy) and long-term effectiveness in real-world conditions, but evidence of their impact on WMSD rates is scarce [4–6]. The research community faces the challenge of establishing evidence-based guidelines for the construction industry, necessitating standard testing to compare OE performance and inform large-scale adoption.

To help address this need, and building on prior research [7], the present paper reports the outcomes of implementing a realistic, controlled, and repeatable test to assess BSEs for rebar workers. The main objectives involve implementing a controlled, practical test course and evaluation method designed for a distinct construction trade (rebar workers), a specific type of OE (back-support), and a designated project scenario (horizontal rebar installation). Additionally, the study seeks to extract insights for refining the test in future iterations. The realistic test course and evaluation method is aimed at contributing to future standard tests which will be necessary for large-scale adoption of OEs in the construction industry.

2 Literature Review

In recent years, research on the application of OEs in the construction industry has garnered attention. Guidelines from [8] recommend OE types for specific construction trades based on WMSD statistics from the United States.

Laboratory efficacy studies, such as [9], have assessed the impact of OEs on worker safety, acceptance,

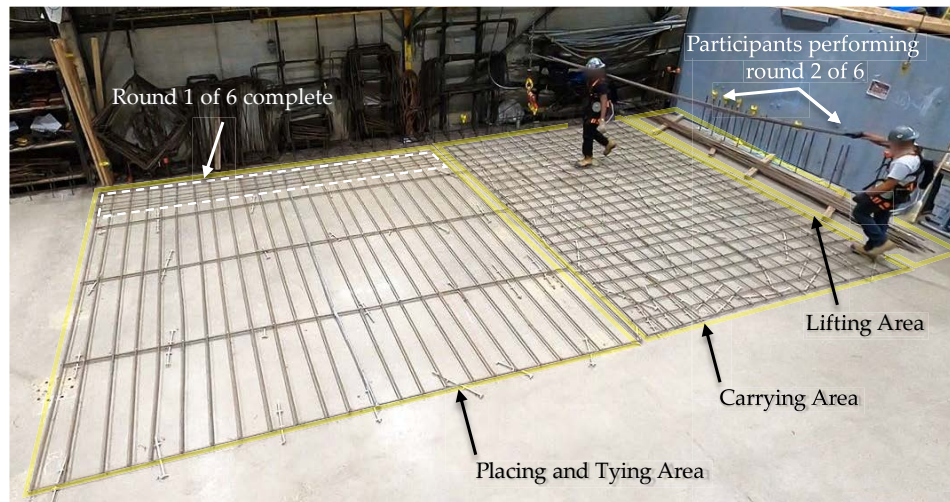


Figure 1. Implementation of the test

and productivity for generic tasks. Measures previously used to assess an OE's effect on health and safety include reductions in muscle activity measured using surface electromyography (sEMG) sensors, biomechanical simulation, and concepts such as lower back disorder risk. Measures previously used to assess OE acceptability include surveys on perceived discomfort, obstruction, and usability. A framework proposed by [10] for evaluating OE efficacy emphasized subjective and objective measures of efficacy, specified target tasks, and specified body postures. De Bock et al. [11] emphasized that the evaluation of the effects of an OE should include realistic tasks and test conditions. Previous work has also investigated the efficacy of OEs for specific construction trades, including rebar installation [12].

However, standard evaluation methods are lacking, as noted by [13] and [14]. Various efforts to standardize OE evaluation methods include the ASTM F48 standards [15], the standard test course for material handling [16], and the *Exoworkathlon* [17]. Ralfs et al. [18] also contributed to the development of generic methodologies and standard tests for OEs.

Crea et al. [19] outlined a roadmap for OE adoption, delineating laboratory assessment, field assessment, and large-scale adoption eras. Within this framework, the authors believe that standard versions of realistic test courses, such as those proposed by [17], will serve as a crucial bridge between controlled laboratory efficacy tests and the large, long-term field studies that will be necessary for the large-scale adoption of effective OEs in the construction industry.

3 The Test and Its Implementation

Our proposed test [7] was designed in consultation with ergonomic experts and implemented in collaboration with the Steel Trades Training Center

(STTC) in Montréal, Canada. This collaboration was in line with the recommendations suggested by [20] of working with trade schools and apprenticeship training programs as a potential solution to the barriers to OE adoption.

Figure 1 shows this initial implementation of the test. The test course has three designated areas for the four main tasks of the test, i.e., *lifting* a load of rebars, *carrying* the rebars over an existing matrix, and *placing* and *tying* the rebars. Six *rounds* of the four *tasks* complete the test course and are defined as one *course*, which takes approximately one hour.

Within each *course*, half of the test course (three *rounds*) was completed in the control condition (No OE) and half was completed in the intervention condition (OE). Each participant was asked to complete four *courses* over four consecutive days, defined as one *test*, resulting in four hours of testing, with and without OEs. Figure 1 shows two participants performing the *carrying task* during the second *round* of that *course* in the intervention condition (OE).

A test duration of several hours was recommended by [21] to ensure the stabilization of measures related to back kinematics, pressure perception, and task performance when using a BSE. Similarly, [22] highlighted the advantageous outcomes of training in an experiment involving an ankle exoskeleton, noting that significant reductions in metabolic energy consumption were observed after several hours of training, in contrast to the results following only 12 minutes of exposure to the device. Therefore, each participant in the proposed test completed a one-hour training session and four one-hour test sessions, for a total of three hours of exposure to the OE and two more hours without the OE.

The aim of the test is to evaluate the short-term efficacy of OEs for the purpose of predicting long-term OE effectiveness. Based on the literature review, a

repeatable test that balances experimental control with realistic conditions reflecting site conditions, is well suited for predicting OE effectiveness. The authors are not aware of any similar tests in the area of construction. The following sections will explain about the objective and subjective measures of OE efficacy to predict the potential effectiveness of the OE considering its effects on participants' health and safety, and the acceptability of the OE.

3.1 Lower Back Muscle Activity

Muscle activity of the lower back was recorded throughout the test using sEMG sensors and synchronized to *task* conditions using video reference. Muscle activity was chosen as an indirect but objective measure of the OE's effects on the forces in the lower back during rebar work. Specifically, muscle activity of the erector spinae longissimus thoracis (LT) was measured by placing two sensors on either side of the spinous process of the first lumbar vertebra (L1), as recommended by [23], and as seen in Figure 2.

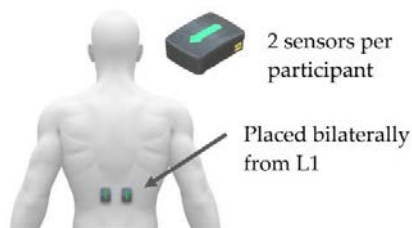


Figure 2. sEMG sensor placement locations

In previous work, different participants experienced reduced fatigue in different lower back muscles, namely the lumbar multifidus (LM), after an OE intervention [24]. Measuring LM activity was not possible in this study because the back pad of the OE covers this muscle group.

Maximum voluntary isometric contraction (MVIC) tests of the LT muscles were used to normalize the data and to compare effects between participants. As recommended by [25], MVIC tests were performed using a roman chair in a 45° position with the hips flexed so that the trunk is parallel to the ground. Participants were asked to contract for three seconds during which non-threatening verbal encouragement was provided. MVIC tests were performed for the LT muscle group two times with at least a 30 s rest in between.

The Delsys Trigno Wireless System was used as sEMG sensor hardware, and data was collected using EMGworks. A minimum sampling rate of 1259 Hz was used for all recordings. Recordings were band pass filtered (20-450 Hz, 4th-order Butterworth), rectified, and a low pass filtered (3-450 Hz, 4th-order Butterworth). During several *courses* in the OE condition, the OE briefly touched one or both sEMG sensors, inducing

noise in the sensor data. When OE interference was infrequent, the noise appeared as outliers in the muscle activity data and were successfully removed using a z-score (threshold = 6). *Courses* where sensor noise was not successfully removed were excluded from the analysis.

3.2 Subjective Measures

After participants finished each half of the test course (three *rounds* of *tasks*) in either the control or intervention condition, participants were asked to complete the Discomfort Survey. The survey asks the participants to quantify their perceived discomfort for each area of the body that is in contact with the OE by using the Rating of Perceived Discomfort (RPD) scale [26]. For passive BSEs, there is a concern that the assistive torque of the OE will induce unwanted forces in non-targeted body areas (i.e., in areas other than the extensor muscles of the hip and back). The Discomfort Survey was used as a subjective measure of these potential undesirable effects.

After a participant finishes the *test* (four *courses* over four days), they are asked to complete a Health, Acceptability, and Safety (HAS) Survey developed by the authors. The questions of the survey are shown in Figure 5. This survey uses a 5-point Likert scale allowing participants to agree or disagree with phrases by choosing from five responses: *Strongly Disagree*, *Somewhat Disagree*, *Neutral*, *Somewhat Agree*, and *Strongly Agree*. As the realism of OE evaluations approaches that of field conditions, and the level of experimental control decreases, subjective measures of an OE's effects become more important. The HAS Survey was used as a subjective measure of the OE's effects regarding perceived support, balance, range-of-motion, thermal comfort, obstruction, and usability.

In both the Discomfort and HAS Survey, space for comments is provided after each question. The HAS Survey also asks for comments regarding how the design of the OE can be improved.

3.3 OE Characteristics

The OE evaluated during the test implementation was the Biolift [27] passive BSE. Figure 3 shows the OE worn by one of the participants, as well as the integrated tool pouch and tie-wire reel necessary for rebar work. The OE weighs 3.5 kg and the time to put on and take off the OE is approximately 10 seconds each. The time to initially adjust the fit of the OE was approximately 1 minute, depending on the user's familiarity with the device. The OE has three levels of support and was set to the second support level for all participants, which has a maximum torque of 50 Nm during hip and back extension, accounting for hysteresis (energy lost due to friction).



Figure 3. The Biolift passive BSE

3.4 Participant Characteristics

Nine students from the STTC participated in the test. The participants were contacted via their instructor and all interested students signed the provided consent forms after a thorough explanation of the test during a one-hour training session. Eight participants were male, and one participant was non-binary and male at birth. The average age of the participants was 27 yr. with a standard deviation of 10 yr., ranging from 19 yr. to 49 yr.

Larger sample sizes are desirable in order to have statistically significant results regarding interaction effects between measures of efficacy and human attributes, including age, gender, and experience level. The limited sample size during the initial implementation was the result of time limitations and the class size of the school. Despite the small sample size, the authors strived to, as much as possible, have diversity with respect to participant gender and age.

All participants had the same level of experience, as they were all in the process of taking their exams at the end of a seven-month (735-hour) professional degree. As recommended by De Bock et al. [28], the participants had a near professional level of experience in regards to the four tasks of the test course, approximately at the same level as junior rebar workers. During their degree, the participants learned the fundamentals of the steel reinforcing trade, including assembly for reinforced concrete slab construction. Participants were also frequently monitored by their instructor throughout the test. This insured that the participants applied the best practices of the trade, and the authors believed that this improved the realism of the tasks they performed, in addition to the realism provided by the participants' experience levels.

Eight of the participants were divided into four groups of two (G1 to G4), and each participant was named according to their group (e.g., PA1 and PB1 belong to G1). A ninth participant was designated as a replacement in the case of absences (PR). The quantity of data from the test is not consistent among participants due to technical issues with the sensors; however, on

average, participants performed each *task* 11 times in the control condition and 10 times in the intervention condition.

4 Test Results

4.1 Muscle Activity Results

Figure 4 compares the average peak (95th percentile) normalized LT muscle activity for each participant in the control (No OE) and intervention (OE) conditions for each *task*. Right and left LT muscle activity is averaged. Error bars show 95% confidence intervals and * marks significant differences ($p < 0.05$) from the control condition (i.e., No OE).

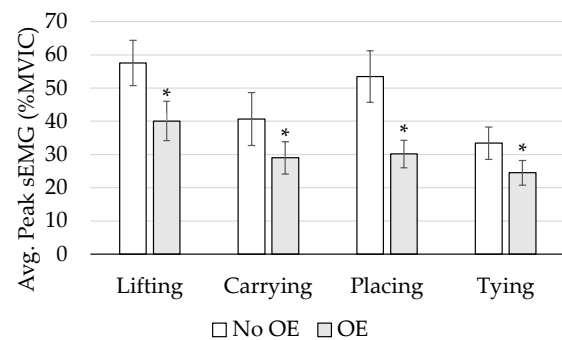


Figure 4. The effect of the OE on average peak normalized LT muscle activity

The average peak (95th percentile) LT muscle activity decreased by 30.4%, 28.7%, 43.6%, and 26.6% for the *lifting*, *carrying*, *placing*, and *tying* tasks, respectively. Average LT muscle activity decreased by 17.1%, 14.2%, 19.5%, and 20.7% for the *lifting*, *carrying*, *placing*, and *tying* tasks, respectively.

During the test, it was observed that some participants took on a narrower stance while tying as compared to the traditional wide stance. Rebar workers use a wide stance to tie the maximum number of intersections before having to move to a new area. Moving to a new area is done either by standing up to move or by shifting positions while maintaining a forward leaning position. Both ways of changing position engage the muscles of the lower back. A narrower stance therefore necessitates more frequent use of the lower-back muscles for tying the same number of intersections. In addition, when a worker's hands are in a working position near the ground, the wider stance will allow him or her to have less back flexion compared to the narrow stance.

Using video reference, each of the eight main participants was classified into using either a wide or narrow stance while tying. This classification was done subjectively and did not rely on objective joint angle data. The average decrease in peak (95th percentile) muscle

activity for the two groups of wide and narrow stance was -18.6% and -38.4% respectively. The larger decrease in muscle activity in the narrow stance group compared to the wide stance group may suggest, as explained above, that the traditional wide stance allows workers to rely less on their LT muscles while tying. As recommended by Golabchi et al. [10], posture must be considered while assessing OE efficacy, and future iterations of the test can record joint angles to objectively measure tying posture.

4.2 Subjective Survey Results

Table 1 shows the results of the discomfort survey as mean levels of discomfort at the OE contact areas for all participants. The minimum and maximum values of the Rating of Perceived Discomfort (RPD) scale are 0, and 10, respectively [26]. The mean discomfort for all contact points was at or below level 1, defined as “Minor Discomfort.” Although the standard deviations for each contact area are large compared to the mean, the deviations are relatively small compared to the range of the RPD scale.

Figure 5 shows the results of the HAS Survey. Questions three and nine were originally written “While carrying rebar, the OE made me feel unbalanced” and “The heat retained by the OE was uncomfortable.” However, these questions have been rewritten in the positive sense to visually unify the data. Green responses represent positive opinions regarding the OE and orange and red responses represent negative opinions. The survey results on perceived support, OE heat retention, obstruction, and usability were generally positive.

Table 1. Discomfort at the OE contact points

Contact Area	Discomfort [Mean (SD)]
OE Shoulder Straps	0.5 (0.4)
OE Chest Strap	0.6 (0.6)
OE Belt	1.0 (1.2)
OE Thigh Straps	0.7 (0.8)

The two oldest participants PA3 and PA4 (49 and 36 years old, respectively) both strongly agreed that the OE made them feel unbalanced while carrying rebars. PA3 commented that he also felt more unbalanced while tying. All other participants strongly disagreed or were neutral regarding the OE making them feel unbalanced. The average age of all other participants, excluding PA3 and PA4, was 23.1 yr. (SD 3.3 yr.). The third oldest participant, PB4 (27 years old) commented that the OE “... did not interfere with my balance.” PB4 commented that at times, his movements were restrained while carrying rebars with the OE activated and that this was not a problem when the OE was deactivated while walking during breaks. The OE designers recommend deactivating the OE if the wearer is walking more than 100 meters. The distance of the carrying task is small enough that all participants left the OE activated while carrying. For PB4, the average and average peak (95th percentile) LT muscle activity showed no significant change ($p < 0.05$) between the No OE and OE conditions during the *carrying task*, suggesting that the activated OE had little to no effect on the LT muscles during the *carrying task*.

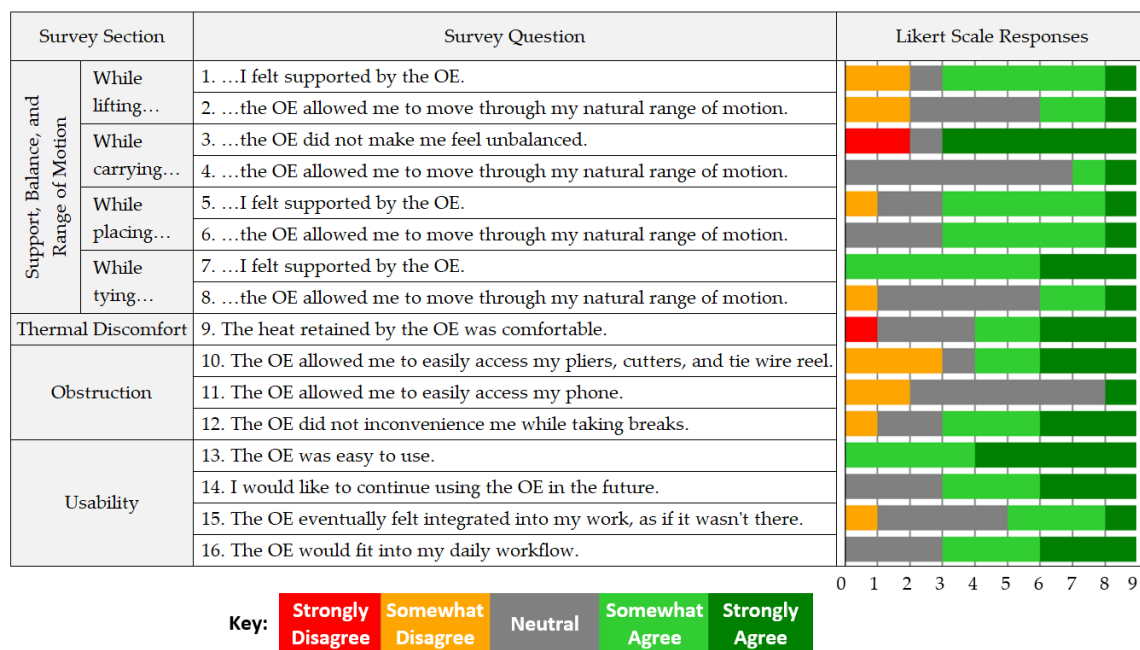


Figure 5. Responses of the nine participants on a 5-point Likert scale from the HAS Survey

5 Conclusions and Future Work

Standard testing of OE efficacy for the construction trades is an integral step from laboratory efficacy evaluation to field effectiveness evaluation, leading to large-scale adoption of effective OEs in the construction industry. This study demonstrated the initial implementation of our previously proposed test for evaluating the effects of BSEs on workers installing horizontal rebars [7]. The potential benefits of the Biolift BSE were evaluated using realistic testing conditions and tasks. The nine participants from the collaborating professional school were at the end of their programs and had the level of experience of junior rebar workers. The study revealed significant reductions in lower-back muscle activity for the main tasks of *lifting*, *carrying*, *placing*, and *tying*. Suggestive differences in muscle activity were found when investigating and comparing participants' tying postures. Subjective results from the surveys showed low levels of discomfort from using the OE and positive opinions regarding OE support, obstruction, and usability.

Future work using the proposed test involves comparing multiple OEs for the trade of rebar assembly. However, the test has limitations that can be addressed in future work. First, conducting the test with a larger sample size with diversity regarding participant age and experience level, as well as having a gender balanced sample, will allow for statistically significant analyses of interaction effects between measures of OE efficacy and human attributes. The manual process of labelling sEMG data led to an unexpected reduction in muscle activity during the *carrying task*, as peaks in muscle activity belonging to the *lifting* and *placing tasks* may have been incorrectly labelled. Future work will involve inertial measurement units (IMUs) for full-body motion capture. Full-body motion capture can ensure accurate task labelling of the muscle activity data and gives an objective measure of an OE's effect on posture. Some loss of data occurred due to technical difficulties with the sEMG sensors. As novel methods of estimating OE assistance mature, for example, combining IMUs and pressure-sensing insoles [29] and simulating spine loading [30], the test may benefit from measures of efficacy that are more practical than sEMG-based measures. Introducing practical measures would be beneficial in two ways: (1) the test would be easier to reproduce, and (2) long-term field studies could implement the same measures of performance as the test, allowing for the results of both types of studies to inform the other, and iteratively improve. The authors hope this research works towards standard testing of OEs for construction trades that balances the factors of implementation repeatability, realism of the testing conditions, and experimental control.

Acknowledgments

This research is in collaboration with the STTC and Biolift Technologies Inc. We thank the STTC instructors Mr. Jean-François Bussi eres and Mr. Alix Lundy for making the project possible with their expertise and generosity, as well as the rest of the STTC team and student participants for their supportive involvement. We thank the Biolift team for their technical support and valuable expertise. We also thank Ms. Annie Boisclair and Mr. Chad Roy-Gauthier from Hydro Qu ebec for their insightful feedback.

References

- [1] W. Umer, M. F. Antwi-Afari, H. Li, G. P. Y. Szeto, and A. Y. L. Wong, "The prevalence of musculoskeletal symptoms in the construction industry: a systematic review and meta-analysis," *Int. Arch. Occup. Environ. Health*, vol. 91, no. 2, pp. 125–144, Feb. 2018, doi: 10.1007/s00420-017-1273-4.
- [2] X. Wang, X. S. Dong, S. D. Choi, and J. Dement, "Work-related musculoskeletal disorders among construction workers in the United States from 1992 to 2014," *Occup. Environ. Med.*, vol. 74, no. 5, pp. 374–380, 2017.
- [3] M. F. Antwi-Afari, H. Li, D. J. Edwards, E. A. P arn, O.-M. De-Graft, J. Seo, *et al.*, "Identification of potential biomechanical risk factors for low back disorders during repetitive rebar lifting," *Constr. Innov.*, vol. 18, no. 2, 2018, doi: 10.1108/CI-05-2017-0048.
- [4] S. Kim, M. A. Nussbaum, M. Smets, and S. Ranganathan, "Effects of an arm-support exoskeleton on perceived work intensity and musculoskeletal discomfort: An 18-month field study in automotive assembly," *Am. J. Ind. Med.*, vol. 64, no. 11, pp. 905–914, 2021, doi: 10.1002/ajim.23282.
- [5] M. Smets, "A Field Evaluation of Arm-Support Exoskeletons for Overhead Work Applications in Automotive Assembly," *IIE Trans. Occup. Ergon. Hum. Factors*, vol. 7, no. 3–4, pp. 192–198, Oct. 2019, doi: 10.1080/24725838.2018.1563010.
- [6] G. Ferreira, J. Gaspar, C. Fуж o, and I. L. Nunes, "Piloting the Use of an Upper Limb Passive Exoskeleton in Automotive Industry: Assessing User Acceptance and Intention of Use," In *Advances in Human Factors and Systems Interaction*, Cham, 2020, pp. 342–349, doi: 10.1007/978-3-030-51369-6_46.
- [7] M. Dunson-Todd, M. Nik-Bakht, and A. Hammad, "Evaluating Occupational Exoskeleton Performance in Construction: Towards Standard Testing for the

- Construction Industry,” Moncton, NB, 2023, [Online]. Available: https://www.researchgate.net/publication/378521774_Evaluating_Occupational_Exoskeleton_Performance_in_Construction_Towards_Standard_Testing_for_the_Construction_Industry.
- [8] Z. Zhu, A. Dutta, and F. Dai, “Exoskeletons for manual material handling – A review and implication for construction applications,” *Autom. Constr.*, vol. 122, Feb. 2021, doi: 10.1016/j.autcon.2020.103493.
 - [9] M. M. Alemi, S. Madinei, S. Kim, D. Srinivasan, and M. A. Nussbaum, “Effects of Two Passive Back-Support Exoskeletons on Muscle Activity, Energy Expenditure, and Subjective Assessments During Repetitive Lifting,” *Hum. Factors*, vol. 62, no. 3, pp. 458–474, May 2020, doi: 10.1177/0018720819897669.
 - [10] A. Golabchi, A. Chao, and M. Tavakoli, “A Systematic Review of Industrial Exoskeletons for Injury Prevention: Efficacy Evaluation Metrics, Target Tasks, and Supported Body Postures,” *Sensors*, vol. 22, Apr. 2022, doi: 10.3390/s22072714.
 - [11] S. De Bock, J. Ghillebert, R. Govaerts, S. A. Elprama, U. Marusic, B. Serrien, *et al.*, “Passive Shoulder Exoskeletons: More Effective in the Lab Than in the Field?,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 173–183, 2021, doi: 10.1109/TNSRE.2020.3041906.
 - [12] N. Gonsalves, O. R. Ogunseiju, O. Ogunseiju, A. Akanmu, and C. Nnaji, “Assessment of a passive wearable robot for reducing low back disorders during rebar work,” *J. Inf. Technol. Constr.*, vol. 26, pp. 936–952, Nov. 2021, doi: 10.36680/j.itcon.2021.050.
 - [13] N. Hoffmann, G. Prokop, and R. Weidner, “Methodologies for evaluating exoskeletons with industrial applications,” *Ergonomics*, vol. 65, no. 2, pp. 276–295, Feb. 2022, doi: 10.1080/00140139.2021.1970823.
 - [14] M. Bär, B. Steinhilber, M. A. Rieger, and T. Luger, “The influence of using exoskeletons during occupational tasks on acute physical stress and strain compared to no exoskeleton – A systematic review and meta-analysis,” *Appl. Ergon.*, vol. 94, Jul. 2021, doi: 10.1016/j.apergo.2021.103385.
 - [15] B. D. Lowe, W. G. Billotte, and D. R. Peterson, “ASTM F48 Formation and Standards for Industrial Exoskeletons and Exosuits,” *IIEE Trans. Occup. Ergon. Hum. Factors*, vol. 7, no. 3–4, pp. 230–236, Oct. 2019, doi: 10.1080/24725838.2019.1579769.
 - [16] R. Bostelman, Y.-S. Li-Baboud, A. Virts, S. Yoon, and M. Shah, “Towards Standard Exoskeleton Test Methods for Load Handling,” In *2019 Wearable Robotics Association Conference (WearRAcon)*, Scottsdale, AZ, USA, Mar. 2019, pp. 21–27, doi: 10.1109/WEARRACON.2019.8719403.
 - [17] V. Kopp, M. Holl, M. Schalk, U. Daub, E. Bances, B. García, *et al.*, “Exoworkathlon: A prospective study approach for the evaluation of industrial exoskeletons,” *Wearable Technol.*, vol. 3, ed 2022, doi: 10.1017/wtc.2022.17.
 - [18] L. Ralfs, N. Hoffmann, U. Glitsch, K. Heinrich, J. Johns, and R. Weidner, “Insights into evaluating and using industrial exoskeletons: Summary report, guideline, and lessons learned from the interdisciplinary project ‘Exo@Work,’” *Int. J. Ind. Ergon.*, vol. 97, Sep. 2023, doi: 10.1016/j.ergon.2023.103494.
 - [19] S. Crea, P. Beckerle, M. D. Looze, K. D. Pauw, L. Grazi, T. Kermavnar, *et al.*, “Occupational exoskeletons: A roadmap toward large-scale adoption. Methodology and challenges of bringing exoskeletons to workplaces,” *Wearable Technol.*, vol. 2, 2021, doi: 10.1017/wtc.2021.11.
 - [20] D. Mahmud, S. T. Bennett, Z. Zhu, P. G. Adamczyk, M. Wehner, D. Veeramani, *et al.*, “Identifying Facilitators, Barriers, and Potential Solutions of Adopting Exoskeletons and Exosuits in Construction Workplaces,” *Sensors*, vol. 22, no. 24, Art. no. 24, Jan. 2022, doi: 10.3390/s22249987.
 - [21] A. Favennec, J. Frère, and G. Mornieux, “Changes in Human Motor Behavior during the Familiarization with a Soft Back-Support Occupational Exoskeleton,” *Appl. Sci.*, vol. 14, no. 3, Art. no. 3, Jan. 2024, doi: 10.3390/app14031160.
 - [22] K. L. Poggensee and S. H. Collins, “How adaptation, training, and customization contribute to benefits from exoskeleton assistance,” *Sci. Robot.*, vol. 6, no. 58, p. eabf1078, Sep. 2021, doi: 10.1126/scirobotics.abf1078.
 - [23] H. J. Hermens, B. Freriks, C. Disselhorst-Klug, and G. Rau, “Development of recommendations for SEMG sensors and sensor placement procedures,” *J. Electromyogr. Kinesiol. Off. J. Int. Soc. Electrophysiol. Kinesiol.*, vol. 10, no. 5, pp. 361–374, Oct. 2000, doi: 10.1016/s1050-6411(00)00027-4.
 - [24] E. P. Lamers, J. C. Soltys, K. L. Scherpereel, A. J. Yang, and K. E. Zelik, “Low-profile elastic exosuit reduces back muscle fatigue,” *Sci. Rep.*, vol. 10, no. 1, Art. no. 1, Sep. 2020, doi: 10.1038/s41598-020-72531-4.
 - [25] G. Biviá-Roig, J. F. Lisón, and D. Sánchez-Zuriaga, “Determining the optimal maximal and submaximal voluntary contraction tests for normalizing the erector spinae muscles,” *PeerJ*, vol. 7, Oct. 2019, doi: 10.7717/peerj.7824.

- [26] J. Steele, J. Fisher, S. McKinnon, and P. McKinnon, "Differentiation between perceived effort and discomfort during resistance training in older adults: Reliability of trainee ratings of effort and discomfort, and reliability and validity of trainer ratings of trainee effort," *J. Trainology*, vol. 6, Jan. 2017, doi: 10.17338/trainology.6.1_1.
- [27] "Biolift." <https://biolift.co/en/home/> (accessed Mar. 28, 2023).
- [28] S. De Bock, J. Ghillebert, R. Govaerts, B. Tassignon, C. Rodriguez-Guerrero, S. Crea, *et al.*, "Benchmarking occupational exoskeletons: An evidence mapping systematic review," *Appl. Ergon.*, vol. 98, Jan. 2022, doi: 10.1016/j.apergo.2021.103582.
- [29] E. S. Matijevich, P. Volgyesi, and K. E. Zelik, "A Promising Wearable Solution for the Practical and Accurate Monitoring of Low Back Loading in Manual Material Handling," *Sensors*, vol. 21, no. 2, Art. no. 2, Jan. 2021, doi: 10.3390/s21020340.
- [30] S. Madinei and M. A. Nussbaum, "Estimating lumbar spine loading when using back-support exoskeletons in lifting tasks," *J. Biomech.*, vol. 147, Jan. 2023, doi: 10.1016/j.jbiomech.2023.111439.

Digital Twin-based Instructor Support System for Excavator Training

Faridaddin Vahdatikhaki¹, Leon olde Scholtenhuis¹, and Andre Doree¹

¹Department of Construction Management and Engineering, University of Twente, the Netherlands
f.vahdatikhaki@utwente.nl, l.l.oldscholtenhuis@utwente.nl, a.g.doree@utwente.nl

Abstract

Given the severity and magnitude of accidents caused by excavators all over the world, the training of excavator operators plays an important role in ensuring the safety of construction operations. New training modes, such as Virtual Reality-based (VR) training simulators, have started to transform training in the construction industry. Although these new developments have been proven to be very effective, it is unanimously agreed by instructors that they do not substitute on-equipment training. During the on-equipment training, an instructor needs to monitor several novice trainees and provide feedback to ensure a safe learning environment. However, being outside the cabin, having to focus on multiple trainees at the same time, and having to stay at a safe distance, instructors are very susceptible to missing important details about the performance of the students. This oversight, in the long term, can result in the institutionalization of wrong behavior in the trainees, which is then very difficult to unlearn. In recent years, with the advent of cyber-physical systems and digital twins, it has become possible to support instructors by providing them with a digital replica of trainees' operations. However, to the best of the authors' knowledge, there has been no systematic research on the use of digital twin as an education/training support tool in excavation training. To this end, this research proposes a comprehensive instructor support system that utilizes a digital twinning approach to help instructors circumvent the limitations of traditional training. A prototype is used in a case study to indicate the potential of the proposed approach. It is shown that the proposed system offers great potential in supporting instructors to provide more in-depth feedback to the trainees.

Keywords –

Digital Twin, Excavator, Training, Support System

1 Introduction

Statistics suggest that jobsite injuries and fatal accidents are significantly high in the construction industry [1]. Among others, inadequate knowledge and skills of the workforce account for a large portion of these accidents[2] That is why effective training is especially important in this industry. This is particularly the case for complex heavy equipment such as excavators [3,4] However, the large equipment size, proximity hazards, complex kinematics, high costs, and limited human capital render training or operators especially challenging. To address these challenges, current excavator operator training includes theoretical and practical components, which focus on equipment mechanics, rules, and safety during theoretical sessions [5]. The practical training can be delivered in two forms: (1) simulator training using virtual reality scenes [6], and (2) on-equipment training using actual excavators for hands-on tasks.

While valuable, simulator-based training, where trainees use VR-based devices to experience near-real working conditions, has several limitations. These include limited contextual interaction, potential cybersickness, and insufficient physics simulation [7,8]. Besides, the extensive integration of VR simulators in curricula demands substantial investment and instructor training. This can be a significant adoption barrier [9]. Consequently, VR-based simulators are not commonly perceived as a complete substitution for on-equipment training, which remains preferred for safety instruction [10].

On the other hand, on-equipment training also has limitations pertaining to safety risks and costs [11]. This is because inexperienced trainees may struggle with hazardous situations. Limited access to expensive excavators puts a cap on training time, and budget constraints result in an imbalanced instructor-to-trainee ratio [11]. These limitations impact feedback quality during practical sessions. Addressing these challenges is crucial to optimize the effectiveness of on-equipment training.

Effective practical training heavily relies on

instructors providing content-rich, relevant, specific, and timely feedback [12–14]. Lack of feedback during psychomotor skill training can lead to the institutionalization of incorrect behaviour[15]. While instant feedback enhances temporary performance [16], conventional methods have limitations such as disrupting workflow continuity, providing feedback from an outsider's perspective, safety concerns leading to a safe distance, and challenges in assessing multiple trainees simultaneously[17,18]. These issues call for a new feedback approach in excavator training that assures high-frequency, relatable, and non-intrusive feedback. Besides, current excavator training lacks clear skill indicators and it relies on instructors' subjective assessments. While some quantitative indicators exist[19,20], they mainly measure performance rather than skill. Proficiency-based indicators related to smoothness, joint velocity, force distribution, and motion consistency are considered more suitable for skill assessment, but their effectiveness is unclear.

In recent years, sensor-based monitoring of construction equipment has become popular [21,22]. With all these systems in place, it is nowadays possible to accurately capture the fine details of excavator operations using sensory data. On the other hand, an influx of research on construction digital twinning provides a solid base for using sensory data to rebuild the operation in the virtual world [23]. However, the research in the development of a digital twin as a means to support the training process is unprecedented.

This research hypothesizes that an instructor support system that records trainees' performance in a non-intrusive manner can address these issues. In other words, building a digital replica of the training process (i.e., digital twin) can help instructors address many of the existing limitations in their work. This system aims to offer specific, spatiotemporally referenced feedback through a VR environment, allowing instructors and trainees to review performance with clarity and detail. While the previous work of the authors has presented the technical specifications of such a digital-twin-based system [24], a thorough analysis of system requirements from the instructors' view was not presented. Given that the success of such an instructor support system depends heavily on the perfect alignment between system functionalities and instructors' needs, it is important to analyze the user perspective on such a system. This paper outlines the high-level requirements of such a system and presents a case study to indicate its feasibility.

2 Research Methodology

To achieve the research objective, we employed the System Development Lifecycle (SDLC) V-model design methodology proposed by Balaji and Murugaiyan [25].

Initially, user requirements were identified through a two-fold process: (1) a thorough review of pertinent literature from scholarly sources covering topics such as equipment safety, equipment operator training, VR-based training simulators, and equipment tracking and visualization, and (2) conducting intake meetings with system users. These high-level user requirements served as the foundation for the entire system development and later served as assessment criteria at the project's conclusion to evaluate the instructor support system's alignment with these requirements.

The subsequent step involved translating the client's overarching requirements into the functional requirements of the system. This transformation was achieved through stakeholder analysis and multiple meetings with instructors and managers at SOMA College. The goal was to delineate the anticipated functions of the instructor support system and conduct an analysis of feedback types. Understanding commonly provided feedback to trainees and prioritizing different feedback types was crucial for the system's ability to furnish automated cues for POAs.

Following this, the functional requirements guided the creation of a conceptual model, outlining the type and quantity of modules (e.g., motion tracking, head tracking, VR environment) essential to deliver the identified functions.

The subsequent phase involved crafting the technical system architecture. By evaluating various hardware and software alternatives identified in the previous step, the most efficient options for each module were determined. Throughout this process, alternatives were assessed against the high-level system requirements defined by users. For example, a comparison between GPS and Ultra-Wideband considered factors like accuracy, reliability, and cost to ascertain the optimal choice for the system. The outcome was a comprehensive system architecture detailing the structure of the instructor support system.

In the final stage, a prototype was developed and subjected to testing in a case study. User feedback was collected to evaluate the prototype's efficacy in meeting user requirements.

3 Requirement Analysis

3.1 Stakeholder Analysis

Table 1 shows the stakeholders involved in this research, their viewpoints, and their needs. Each view involves specific stakeholders that have certain goals. Based on their goals, several needs can be extracted. Using (INCOSE, 2015) guidelines, these needs can be converted to high-level requirements which can be verified during the system design cycle.

Knowledge institutions, such as universities, are actively engaged in the project from a corporate perspective, aiming to assist the industry in minimizing excavation damages. Their primary objective is to initially identify shortcomings in the current operator training program and subsequently address them by proposing and showcasing innovative technological and organizational solutions.

The initiative introduces a novel approach to training construction equipment operators, which impacts excavation contractors and asset owners, such as road agencies. Accidents on construction sites are undesirable for both contractors and clients as they lead to project delays and financial costs. The majority of these accidents result from a failure to adhere to safety instructions while operating construction equipment. Consequently, contractors look for operators with market-ready skills acquired through an optimal training program. Likewise, project clients prefer to collaborate with contractors who have highly skilled operators.

Table 1 Manuscript margins

Stakeholder	Need
Knowledge Institutions	Helping industry to reduce excavation damages and injuries by identifying the limitations of the current practice and trying to tackle them
Asset Owners	Less damage and safety hazards on the construction site
Training Schools	Use limited resources at their disposal in a more efficient way and train operators with higher sets of skills
Excavation Contractors	Hiring skilled operators who can perform safe and productive excavation operations
Instructors	Use their limited time efficiently to make sure trainees make good progress with their training
Trainee	Use their limited time efficiently to make sure trainees make good progress with their training

The project has a direct impact on instructors and trainees at training schools, who are very important figures in vocational training programs. On-equipment training sessions create an environment for instructors to supervise trainees directly, offer feedback, and observe trainees' skill improvement based on the feedback. Instructors aim to maximize their time during these sessions to provide more valuable feedback to trainees efficiently. Simultaneously, trainees are eager to comprehend this feedback more clearly to acquire the skills essential for their future careers as construction equipment operators as fast as possible.

3.2 Functional Requirements

After conducting the stakeholder analysis outlined earlier, a workshop involving four instructors and two educational support staff was conducted, alongside informal interviews with instructors and training school managers. The goal was to discern the system's functional requirements, a task approached indirectly due to the limited technical insights on the part of interviewees and workshop participants. These interviews were deliberately unstructured and informal, fostering creativity in generating valuable requirements. To guide the discussion, a set of questions was devised and is detailed in Table 2.

Table 2 illustrates that functional requirements were derived from the responses of SOMA instructors and management representatives. According to these requirements, the system should:

1. Capture the movement across all degrees of freedom of construction equipment, as illustrated in Figure 1.
2. Monitor the head pose of trainees, particularly rotation inside the cabin, to track their shoulder-check tendencies.
3. Provide visualization with centimetre-level accuracy (measured at the bucket tip in Figure 1) and a frame rate of at least 60 Hz. Rotation accuracy should be around ± 1 degree with a drifting error of no more than ± 1 degree/hour. Positioning accuracy (i.e., translation of the excavator in Figure 1) should be 3 meters.
4. Represent the pose (i.e., location and orientation) of other equipment in the vicinity.
5. Supply automated cues/pointers to instructors, indicating POAs for feedback.
6. Offer offline visualization while being fast enough to support feedback within 24 hours.
7. Feature a user-friendly GUI for instructors to interact with the virtualized training site, encompassing 3D navigation and the ability to traverse training sessions in time.
8. Include a feature for instructors to annotate feedback in timestamped text.
9. Incorporate a feature allowing the replay of annotated visualizations for trainees.

An additional desirable functional requirement was also identified: the system should enable trainees to use the same virtual environment to practice operations based on provided feedback using joysticks.

Among the outlined functional requirements, requirement 5 required further clarification. This was necessary because instructors offer a diverse range of feedback to trainees, and for the system to generate automated cues/pointers to POAs, understanding the typology of feedback provided to trainees is crucial. Given the time constraints of the project, implementing automated cues/pointers for all feedback types was not

feasible. Consequently, it became essential to establish the priorities of various feedback types, leading to the consideration of only the top 5 types for automated cues/pointers.

Table 2 Functional requirements of the system

Questions	Functional Requirement
What aspects of equipment need to be represented in the visual representation?	The system must capture all degrees of freedom of the excavator. This includes the full motion of the arm, the rotation of the superstructure, the rotation of the tracks, and translation of the excavator, as shown in Figure 2
What aspects of trainees need to be represented in the visual representation?	The system must capture the rotation of the head of the trainee. This is important to make sure trainees perform the shoulder check and blind-spot control when needed during the operation
What level of detail and accuracy of the visual representation of trainees' performance is considered sufficient for providing feedback?	Given that the focus of the accuracy is on the position of the bucket tip with respect to the ground, the system is expected to have an accuracy of about a few centimetres. Additionally, the rotation accuracy should be around ± 1 degree with a drifting error of ± 1 degree/hour. The positioning accuracy needs to be in the order of 3 meters. The visualization of the performance must also not be lagging. This can be translated to requirements for the data capture frequency of at least 60 Hz
What aspects of the context of the operation (i.e., surrounding environment) need to be represented for the feedback?	The system, at this stage, is not expected to track the changes in the terrain (i.e., soil tracking). But, the movements of other equipment need to be captured in the visualization.
During the feedback sessions, what aspects of the trainees' performance are being analyzed?	The system should help analyze certain aspects of the trainee's performance in terms of automated cues/pointers to POAs for the instructors. These aspects are explained in detail in Table 3
When is the feedback to each trainee expected to be delivered?	The system is not required to be real-time. The instructors would want to give feedback to students within 24 hours
How much time do instructors envision to spend on reviewing the feedback of each trainee?	The system should be able to speed up and down the reply of performance visualization. Instructors mentioned that the system should potentially be able to reply to the entire operation of about 1 hour in 5 minutes (approximately fast forward x12)
How would instructors expect to provide feedback?	The system should enable at the very least timestamped textual feedback
How are trainees expected to interact with the feedback system?	The system should enable, at the very least, the reply of annotated visualization (i.e., visualization + timestamped textual feedback) at different playback rates. It might be also beneficial for trainees to be able to switch to interactive mode of the feedback system to practice the operation in the virtual environment in the actual context of the work

To identify and prioritize these feedback types, workshop participants were queried about the typical

feedback they usually provide to trainees. To facilitate discussion and guide participants in thinking about feedback types, a set of ten predefined feedback types was presented to instructors, drawing from literature reviews and observations of on-equipment training sessions. Throughout the workshop, instructors also proposed new feedback types they deemed necessary. The suggested feedback types are detailed in Table 3.

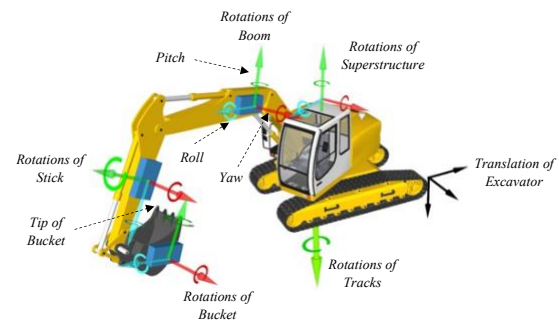


Figure 1. Required degrees of freedom for visualization of an excavator

The instructors who participated in the workshop played a key role in determining the priority of feedback types. They were tasked with assigning a score to each feedback type on a five-point scale, where 1 indicated "Not Useful," 2 represented "Partially Useful," 3 denoted "Useful," 4 signified "Very Useful," and 5 indicated "Crucial." Subsequently, the mean of the scores provided by instructors was computed to establish the ranking of feedback types based on their priorities. The outcomes of this ranking are illustrated in Figure 2.

Considering the project's time constraints and the ranking depicted in Figure 2, six feedback types were chosen for implementation in the instructor support system. From the top 8 feedback types, specifically operator concentration points and excavator vibration, required additional sensors beyond those necessary for capturing the motion of the excavator and trainee. Consequently, it was collectively decided (in consultation with the instructors) to exclude these two feedback types. Consequently, the final set of feedback types supported by automated cues/pointers for instructors' POAs comprises:

1. Shoulder check
2. Bucket movement smoothness
3. Bucket loading distance
4. Simultaneous axes movement
5. Axes movement speed

Additionally, an extra type, stability check, was later introduced during the project by experts. Since this feedback type wasn't part of the workshop discussion, it is treated as an additional aspect.

Table 3 Functional requirements of the system

Feedback Type	Description
Shoulder check	This feedback type targets situational awareness of operators by detecting the gaze point of the operators while they are moving the excavator backward. In this situation, the operators should look back in that direction of movement otherwise the maneuver would be unsafe
Operator concertation point	This metric also targets the situational awareness skill of equipment operators. While performing a swing action on the excavator, the operator should focus on the bucket destination rather than following the bucket position itself
Scenario evaluation	Each trainee is responsible for practicing specific tasks (e.g. dumping a truck) in a collaborative training scenario with other trainees. This feedback type concerns how well the trainee performed the assigned task
Simultaneous axes movement	An efficient maneuver from the equipment fuel consumption perspective is defined as moving the joints of the excavator at the same time as much as possible. This feedback type indicates that the operator is moving the joints simultaneously or not
Bucket movement smoothness	The excavation performance of operators depends on how smooth they move the bucket. This feedback type concerns the movement smoothness
Bucket load	This feedback type focuses on the number of soil operators load during each maneuver. If this amount exceeds from a certain threshold which depends on the excavator specifications, the maneuver is not efficient
Trench geometry	The geometry of the trench dug by operators is an important indicator of excavator performance. This feedback type evaluates the trainees' performance based on geometrical parameters of the trench, e.g., trench depth
Operator drowsiness level	This metric determines the drowsiness level of operators while they are working with the excavator
Operator stress level	This feedback type addresses the stress level of trainees while working with the excavator
Axes movement speed	This feedback deals with how fast operators move the excavator arm in each manoeuvre
Excavator vibration	This feedback concentrates on the extent of vibration induced to the excavator tracks during the operation. If trainees are not experienced enough, lots of vibrations are generated over the excavator tracks
Bucket loading distance	This feedback concerns the amount of pressure on the hydraulic jacks while trainees dig. When trainees lift the loaded bucket, if the bucket is too far or close to the tracks, more fuel is consumed. Appendix 1 provides more detail about this feedback

4 Architecture of the Proposed System

Figure 3 shows an overview of the proposed

instructor support system. This figure also illustrates the hardware and software components required to deliver these functions. The initial module of the system, as depicted in the figure, focuses on collecting motion data for the excavator and trainees, including translation, rotation, and head rotation. Tracking technologies are utilized to capture excavator location, pose, and trainee head pose. The module runs in real time. Module 2 integrates, time-stamps, and synchronizes the collected data, necessitating a processor and storage means (e.g., cloud space or local drive). These steps are also performed in real time. Module 3 visualizes stored data in a virtual environment post-training session. This requires a platform for visualization, linking collected data to 3D models of equipment and trainees. In the subsequent module, the system analyzes trainee performance, automatically detecting POAs and providing cues to instructors. Algorithms for POA identification are integrated into an analyzer module within the visualizer.

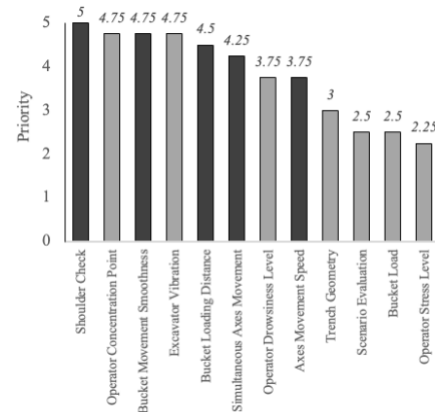


Figure 2. The feedback type priority

In the subsequent module, the instructor assesses trainee performances offline, identifying areas for improvement based on provided POAs and personal visual reviews. A custom GUI facilitates navigation and performance review from various points of view. The instructor can add notes, exporting the annotated virtual scene for trainees. Trainees then receive these annotated VR scenes, gaining insights for future sessions.

In the system's final module, the visualizer transitions to interactive mode, allowing trainees to use control units (e.g., joysticks) for task practice in the virtual environment. This module aligns with functional requirement

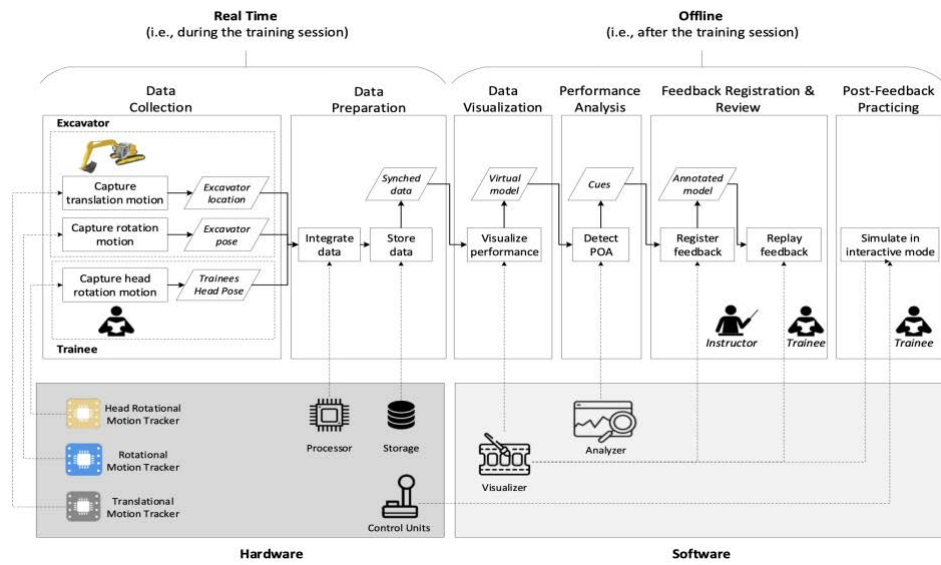


Figure 3. The architecture of the instructor support system [24]

5 Architecture of the Proposed System

The design outlined in Section 4 has been translated into a prototype system. The objective of this prototype is to showcase the viability of the proposed design and evaluate its suitability for the intended purpose. It is crucial to note that the prototype is not intended to function as a fully operational system. This chapter provides a concise overview of the prototype.

5.1 Hardware Components

Figure 4 provides a summary of the prototype's hardware components.

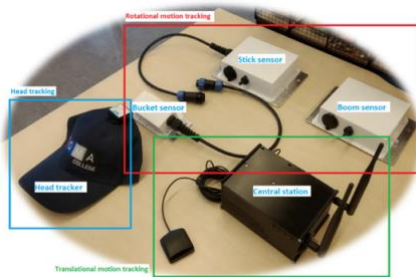


Figure 4. The hardware components of the Motion Capture Module

The system hardware includes 4 IMU sensors, a GPS antenna, and a single-board computer. A cap is employed to monitor the head movement of trainees within the cabin. The connection between the IMUs of the bucket and stick is wired, eliminating the need for WiFi transmission of bucket IMU data and, consequently, the need for a Raspberry Pi in the bucket IMU. As a result, the casing of the bucket sensor is smaller compared to the

other two, which rely on Raspberry Pi for WiFi data transmission. Figure 5 shows the configuration of sensors when installed on an excavator. Unity Game Engine [26] is used as a platform for data visualization, performance analysis, feedback registration and review, and post-feedback practice. An interface is designed inside Unity as shown in Figure 6. Figure 6 also shows the main interface of the system for visualization playback and feedback registration.

5.2 Implementation and Case Study

To validate the efficiency of the proposed system, a case study was conducted at SOMA College, i.e., the largest construction equipment training institution in the Netherlands, using the prototype. In this study, an excavator was outfitted with the implemented data collection module, and the trainee's performance was captured on a USB stick. The task assigned to the trainee involved excavating a trench approximately 70 cm deep within 20 minutes. After the experiment, a quantitative assessment via a workshop involving instructors and students was used to assess the system. The assessment form includes a total of 7 general metrics (1 to 7) applicable to both instructors and trainees. Additionally, instructors are evaluated based on 3 specific metrics (8 to 10), while trainees are assessed using 3 distinct metrics (11 to 13). The outcomes of this workshop are depicted in Figure 7.

Figure 7 presents the results of the assessment. Both from the perspective of instructors and trainees, the User Interaction and Real-video Superimposition features received the highest ratings. The 3D navigation feature

allows both instructors and trainees to observe and assess performance from various angles. Participants particularly valued the Real-video Superimposition feature for reducing miscommunication by displaying real video alongside the visualization. Moreover, instructors found the Feedback Insertion feature to be an effective means of providing textual feedback. The lowest rating during the evaluation was assigned to Versatility. Trainees expressed comfort with wearing the head tracker during operations and agreed with instructors on the helpfulness of text feedback for performance improvement. However, they noted that the simulator mode did not offer significant added value compared to their existing simulator training program.



Figure 5. The configuration of the system's hardware on an excavator



Figure 6. The interface of the prototype system (adapted from [24])

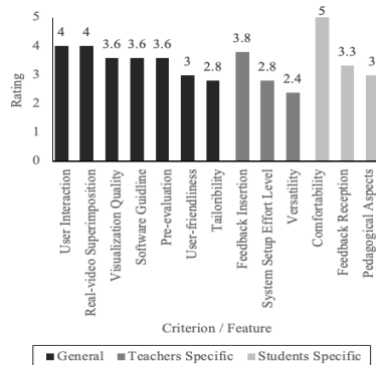


Figure 7. The interface of the prototype system

6 Conclusions

The instructor support system is a cost-effective sensing kit designed to monitor the movements of both the equipment and the operator, offering an interactive interface for educational use. This system empowers

instructors to closely monitor trainee performance during on-equipment training sessions from diverse viewpoints, minimizing the likelihood of overlooking cues indicating mistakes. Additionally, it allows trainees to thoroughly review their performance post-training session, ensuring they don't forget instructors' feedback and can maintain focus without interruptions.

Participating experts in the case study affirm that the instructor support system introduces three notable advantages to the existing education system: (1) It effectively addresses the issue of "transience" in the current feedback approach; (2) The system's automatic pre-evaluation feature proves highly valuable, saving instructors time during trainee assessments; and (3) The head tracker provides insights into the trainees' situational awareness.

However, the experts also identify three main drawbacks of the instructor support system: (1) The current state lacks sufficient accuracy in visualization to analyze the quality of trainee performance; (2) The system is currently tailored for a single type of excavator and requires calibration for use with other types; and (3) Integration of real-video recordings with the system is not yet realized, making synchronization with visualization challenging.

There are two points that can be considered in the future. First, in its current form, the system operates as an open-loop system with a unidirectional flow of feedback. The authors have already implemented a simulation mode in the system, where the trainees can practice their improved operation in the VR and send it back to the trainees for re-evaluation. However, this feature is not systematically checked for its usefulness. This will be done in the future. Another direction to pursue in the future pertains to capturing the operation of equipment by tracking the movement of the control units inside the cabin. This tracking allows instructors to better understand the root causes of bad operational habits of the trainees.

7 Acknowledgements

The authors would like to thank Mr. Armin Langroodi for his invaluable contributions to the development of the prototype and data collection campaigns.

8 References

- [1] Ale BJM, Bellamy LJ, Baksteen H, Damen M, Goossens LHJ, Hale AR, et al. Accidents in the construction industry in the Netherlands: An analysis of accident reports using Storybuilder. *Reliability Engineering and Safety System*, 93(10):1523–1533, 2008.

- [2] Haslam RA, Hide SA, Gibb AGF, Gyi DE, Pavitt T, Atkinson S, et al. Contributing factors in construction accidents. *Applied Ergonomics*, 36(4):401–415, 2005.
- [3] Hinze J, Huang X, Terry L. The Nature of Struck-by Accidents. *Journal of Construction Engineering and Management*, 131(2):262–268, 2008.
- [4] Hinze JW, Teizer J. Visibility-related fatalities related to construction equipment. *Safety Science*, 49(5):709–718, 2011.
- [5] Wang X, Dunston PS. Design, strategies, and issues towards an Augmented Reality-based construction training platform. *Electronic Journal of Information Technology in Construction*, 12:363–380, 2007.
- [6] Oliveira DM, Cao SC, Fernández Hermida X, Martín-Rodríguez F, Oliveira DM, Martín Rodríguez F. Virtual Reality System for Industrial Training, in *Proceedings of IEEE International Symposium on Industrial Electronics*, 1715–1720, 2007.
- [7] Burke MJ, Sarpy SA, Smith-Crowe K, Chan-Serafin S, Salvador RO, Islam G. Relative effectiveness of worker safety and health training methods. *American Journal of Public Health*, 96(2):315–324, 2006.
- [8] Sacks R, Perlman A, Barak R. Construction safety training using immersive virtual reality. *Construction Management and Economics* 31(9):1005–1017, 2013.
- [9] Abdel-Wahab M. Rethinking apprenticeship training in the British construction industry. *Journal of Vocational Education & Training*, 64(2):145–154, 2012.
- [10] Psotka J. Immersive training systems: Virtual reality and education and training. *Instructional Science*, 23(5):405–431, 1995.
- [11] Dunston PS, Proctor RW, Wang X. Challenges in evaluating skill transfer from construction equipment simulators. *Theoretical Issues in Ergonomics Science*, 15(4):354–375, 2014.
- [12] Wu WFW, Young DE, Schandler SL, Meir G, Judy RLM, Perez J, et al. Contextual interference and augmented feedback: Is there an additive effect for motor learning? *Human Movement Science*, 30(6):1092–1101, 2011.
- [13] Nesbitt CI, Phillips AW, Searle RF, Stansby G. Randomized trial to assess the effect of supervised and unsupervised video feedback on teaching practical skills. *Journal of Surgical Education*, 72(4):697–703, 2015.
- [14] West J, Turner W. Enhancing the assessment experience: improving student perceptions, engagement and understanding using online video feedback. *Innovations in Education and Teaching International*, 53(4):400–410, 2016.
- [15] Shea CH, Shebilske W, Worchel S. *Motor learning and control*. Prentice Hall, 1993.
- [16] Shea CH, Wulf G. Enhancing motor learning through external-focus instructions and feedback. *Human Movement Science*, 18(4):553–571, 1999.
- [17] Cassidy T, Stanley S, Bartlett R. Reflecting on Video Feedback as a Tool for Learning Skilled Movement. *International Journal of Sports Science & Coaching*, 1(3):279–288, 2006.
- [18] Wulf G, Shea CH, Matschiner S. Frequent feedback enhances complex motor skill learning. *Journal of Motor Behavior*, 30(2):180–192, 1998.
- [19] Bernold LE. Quantitative Assessment of Backhoe Operator Skill. *Journal of construction engineering and management*, 133(11):889–99, 2007.
- [20] Sakaida Y, Chugo D, Yamamoto H, Asama H. The analysis of excavator operation by skillful operator - Extraction of common skills. In *Proceedings of the SICE Annual Conference*, 538–542, 2008.
- [21] Akhavian R, Behzadan AH. Construction equipment activity recognition for simulation input modeling using mobile sensors and machine learning classifiers. *Advanced Engineering Informatics*, 29(4):867–877, 2015.
- [22] Ryu J, Seo J, Jebelli H, Asce AM, Lee S, Asce M. Automated Action Recognition Using an Accelerometer-Embedded Wristband-Type Activity Tracker, *Journal of construction engineering and management*, 145(1), 2018.
- [23] Fu T, Zhang T, Lv Y, Song X, Li G, Yue H. Digital twin-based excavation trajectory generation of Uncrewed excavators for autonomous mining. *Automation in Construction*, 151:104855, 2023.
- [24] Vahdatikhaki F, Langroodi AK, Scholtenhuis L olde, Dorée A. Feedback support system for training of excavator operators. *Automation in Construction*, 136:104188, 2022.
- [25] Balaji S, Murugaiyan MS. Waterfall vs. V-Model vs. Agile: A comparative study on SDLC. *International Journal of Information Technology and Business Management*, 2(1):26–30, 2012.
- [26] Unity 3D. <https://unity.com/> (accessed February 17, 2021).

Fit Islands: Designing a Multifunctional Exergame to Promote Healthy Aging for Chinese Older Adults

Zixin Shen¹, Rongbo Hu², Dong Wan⁴, Ami Ogawa³ and Thomas Bock¹

¹Chair of Building Realization and Robotics, Technical University of Munich, Germany (www.rod.de)

²Kajima Technical Research Institute Singapore, Kajima Corporation, Singapore

³Department of System Design Engineering, Keio University, Japan

⁴Civil Design Department, Sinoma International Engineering Co., Ltd, Nanjing, Jiangsu, China
zixin.shen@tum.de; r.hu@kajima.com.sg; wandong@sinoma-ncdri.cn; ami_ogawa@keio.jp; bockrobotics@web.de

Abstract –

Emerging economies are ill-prepared to deal with population aging. Uncontrollable events such as the COVID-19 pandemic and regional conflicts have exacerbated the plight of older adults. The consequent development of smart aging offers endless possibilities for improving the quality of life and health of older adults. In this paper, the authors propose a universal approach to designing a health-promoting Exergame system in the format of a virtual village for emerging economies to cope with aging populations. To verify the feasibility, the authors designed an expandable Exergame called "Fit Islands", using China as a case study. Based on the initial demonstration the authors conducted functional tests. The result is that Fit Islands can meet the development objective of motivating Chinese older people to increase their physical activity, providing initial evidence of the feasibility of an Exergame system to promote healthy aging in emerging economies. The application of Fit Islands demonstrates the feasibility of the universal Exergame development method, which can, in principle, provide comprehensive and practical guidance for other countries.

Keywords –

Aging in China; COVID-19 Pandemic; Digital Technology; Exergame; Health; Population Aging; Serious Game; Silver Economy; Smart Aging

1 Introduction

Compared to developed countries, the growth rate and magnitude of the elderly population in developing regions are much higher [1]. The large population base, strong economic growth, abundant resources, and enormous market present both opportunities and challenges [2]. With the rapid ageing of the population, problems have arisen such as a huge gap between the rich

and the poor in urban and rural areas, an unfriendly society for the older adults, empty nesters, a rapid increase in the number of chronically ill patients, and an underprepared health-care system [3]. The COVID-19 pandemic and regional conflicts have undoubtedly exacerbated the physical and mental ill health of the older adults. Lockdowns and other restrictions on travel make older adults, who are already socially isolated due to family relationships, social connections, marital status, Digital gap, etc., more likely to develop a sense of loneliness, which in turn triggers psychological problems such as depression and anxiety. These psychological problems can easily lead to a decline in the quality of sleep, increased incidence of cardiovascular and cerebrovascular diseases, and can lead to cognitive decline in the older adults, etc., which directly and negatively affects the quality of life of the older adults [4,5].

Smart ageing offers new possibilities to deal with the problem of ageing population. By utilising advanced Internet, cloud computing, wearable and other new-generation information technology tools, it is possible to build urban Internet of Things systems and urban information platforms for home-based, community-based and institutional elderly care, to maximise the satisfaction of the needs of the older adults in terms of social participation, leisure and culture [6]. In smart cities, digital entertainment interactive applications such as age-friendly games have great potential to enhance the quality of life and physical and mental health of the older adults. In recent years, a new digital technology has been introduced to motivate people to engage in physical exercise, and is known as Exergame [7]. Exercise has potential benefits for improving the physical health, cognitive abilities, mood, personality and overall well-being of older adults [8,9]. In the context of ageing societies, Exergame is no longer just for teenagers or game enthusiasts, but offers new ways of coping with the problems of ageing as a tool for disease prevention and health education for older people.

The aim of this study is therefore to try to incorporate Exergame's game format to design a virtual village that encourages physical activity, using China as an example, breaks down spatial isolation and crosses the digital gap for older people in emerging economies (Figure 1). By proposing a design methodology for the health promoting Exergame system, new solution ideas are proposed for different emerging economies to cope with the problem of population ageing.

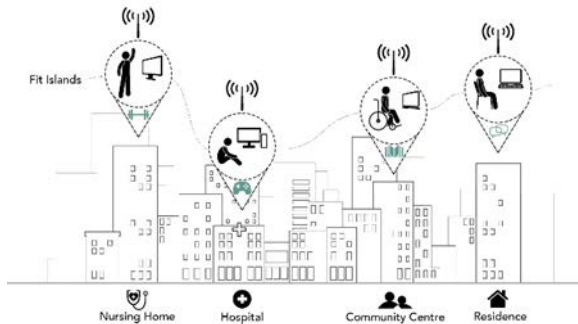


Figure 1. Fit Islands breaks spatial isolation and provides a health promoting virtual platform for Chinese older adults

2 Background

Exergames typically utilize various types of sensors as input devices. Cameras, as input device, along with corresponding algorithms, can enable motion capture functionality, allowing the possibility of playing Exergames without holding or wearing input devices [10].

2.1 Existing projects

Gerling et al. in 2010 explored design guidelines for sports game design for older adults with a game called SliverBalance [11]. Chartomatsidis et al. 2019 developed an Exergame called Fruit Collector, based on the Kinect sensor and developed using Unity for 3D game world design, demonstrating the positive impact of Exergames on improving physical health in older adults [12]. In 2017 a virtual therapeutic gaming system called SilverFit 3D further demonstrates that treating older adults through Exergames is an accepted, effective and potentially cost-efficient approach through a system based on scientific research and clinical practice that can develop rehabilitation protocols and targeted exercise programmes for older adults [13]. The REACH research project conducted by Hu et al. included an Exergame system called the ActivLife Gaming Platform, which senses and assists with a smart device called the PI2U-iStander, designed to prevent falls in older adults through physical activity and cognitive training [14].

SliverBalance takes into account the physical conditions and the cognitive limitations of older adults. Fruit Collector interviewed orthopaedic and physiotherapists to assess the game's movements and focused on building older adults' confidence in the game. SilverFit 3D can create over 1.2 million exercises based on scientific research and clinical practice. The flexibility and adaptability of the PI2U-iStander smart device in REACH also ensure the safety of older adults during play. However, these games also have some drawbacks, for example, SliverBalance and Fruit Collector have a single content and do not take into account the long-term motivation of the players. The SilverFit 3D and REACH systems are very expensive, require a medical resource system or a smart device and are not widely used in everyday life.

2.2 Inclusive design

In the 1980s and 1990s, scholars began to make proposals for age-friendly games. Weisman and Whitcomb, after studying the effects of age, suggested that gameplay should include adjustable game speed and difficulty and easy-to-use interface design [15]. Ijsselsteijn et al. also suggested in their study of digital games in the early 21st century that visually adjustable game interfaces should be considered. For example, fonts, colours and contrast, providing a presentation that incorporates both dynamic resources and multimodal feedback in addition to text [16]. Flores et al. expand on this by arguing that it makes sense to combine appropriate cognitive challenges, simple user interfaces and motivational feedback in the design [17]. Gerling et al. also suggest that older players should be given the possibility to adjust the difficulty of the game, the speed of the game and the sensitivity of the input device [11]. A large proportion of today's older adults have limited experience with digital games and therefore Exergames should be designed to minimise the number of steps required to complete tasks and reduce cognitive load, enhance the scientific, safety and practicality of the game, thereby increasing the user adaptability of the game system.

The above content of the background research generated guidance and insights into the following research methodology. The special situation of countries with emerging economies requires researchers to analyse the current situation in the country with a large amount of data before designing, which in turn determines the main concepts of the game, its purpose and users, its functionality, interfaces and interactions. The development and the test of Demonstration can further ensure the scientific validity and credibility of the study.

3 A universal method for health promoting Exergame system design

The process of Exergame development for emerging economies is proposed in Figure 2.

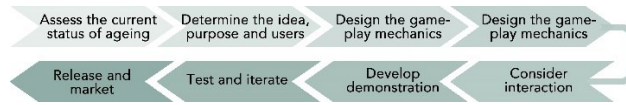


Figure 2. A universal development flow for health promoting Exergame systems for emerging economies

4 Fit Islands

4.1 The current status of ageing in China

China, as a representative emerging economy, has a huge population base. The improvement in medical conditions and the one-child policy have directly contributed to population aging in China [18,19]. Generally, individuals aged 65 and above are considered older adults. However, in China, based on the retirement age of 60 for men (55 for women), individuals aged 60 and above are classified as older adults. Data shows that even using the age of 65 as the threshold, China's population aging exhibits significant scale, depth, and speed. The speed of population aging in China has surpassed that of any other country in modern history. By the end of 2022, China's population decreased by 850,000 compared to the previous year, marking the first negative growth. The rapid decline in the number of women of childbearing age and the fertility rate suggests that China is entering a period of normalized population decline [20]. It is projected that by 2057, the peak of China's elderly population aged 65 and above will reach 425 million, accounting for 32.9%-37.6% of the total population. The current status of population ageing in China has the following main aspects:

- Age structure and health: The age structure of China's elderly population is young. The health status is at a low level, and the number of people suffering from chronic diseases is increasing rapidly [21,22].
- Education and informatization levels: as education levels rise, adaptation to digital technologies shows a polarized phenomenon [21,23].
- Economic status: China's older adults are growing old before they get rich. The wealth disparity in different regions is very large. Older adults rely mainly on savings, children, continued work and property [24].
- Psychological health: In China, cultural traditions

emphasize family systems and collectivism. The phenomenon of elderly people who have lost their only child and empty nesters has intensified, shaking the Chinese tradition of family social support for the older adults and exacerbating the social isolation and loneliness [25].

- Leisure and recreation: Chinese older people's leisure activities are mostly sedentary, and their interest in sports is generally at a low level. In terms of physical activity, square dancing is the preferred choice for them [26].
- Living space: Most Chinese older adults currently have their own private living space. However, the construction of public facilities for the older adults, such as activity centers, physical exercise centers, barrier-free facilities, is still incomplete [27].
- Policy and Smart Aging: China is promoting a three-tiered elderly care system, with plans for 90 per cent of older people to stay at home, 7 per cent in community centers and 3 per cent in institutional care centers. China's elderly care products and ageing services have not yet developed to the extent that they can meet the needs of the ageing population, and are not yet an alternative to 'family pension' [28].

4.2 Determine the main idea, objectives, and users of the game

"Fit" represents physical and mental health. The game design will need to meet the following objectives:

- Adaptable: All older adults can use the game. Membership will be set up to provide extended games for various numbers of players to ensure the games can be adapted to different scenarios.
- Affordable: The game will be played without external wearable devices. It will only require a computer with a camera to play, aiming to minimize game costs.
- Elderly-friendly: The design will be inclusive of different levels of cultural, linguistic, visual, and auditory abilities, creating a user-friendly interface and game experience for the older adults.
- Enjoyability: To increase the motivation, the game will have many functions, but will also provide practical assistance to the target users, such as adding valuable health knowledge in the game.

4.3 Design the gameplay mechanics

In Fit Islands, players will assume the role of lively animated characters living on their own floating village (see Figure 3). Players can visit shops, cultivate crops, interact with friends on their contact lists at the post office, and play games at the game centre.



Figure 3. Fit Islands' Poster

The specific process can be summarized as follows: Initial game installation - Game instructions - Account creation - Main interface - Game map - Gameplay - Points and health knowledge. Elderly players who may be less familiar with information technology can receive assistance from family members or caregivers during the initial game installation (see Figure 4).

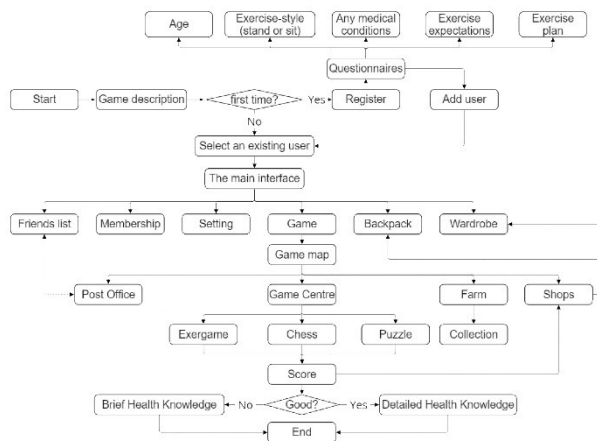


Figure 4. Functional flowchart of Fit Islands

The game will be designed to include two forms of exercise, full-body standing exercise and upper-body exercise (see Figure 5). The points-based incentive system, allowing players to use points to redeem various items in the game store. At the end of each gaming session, players will receive different health knowledge. To simplify the preference settings for older adults, the game will employ individual user profiles and automatically match game content. The game can accommodate two users for free in total, and after each game session, the progress and user data will be automatically saved. In the settings, older adults can freely adjust game content, language, colour brightness and contrast, font size, and more. Furthermore, Fit Islands is expandable, allowing different membership levels to enjoy larger game maps, more game varieties, and accommodate additional independent user profiles and services.

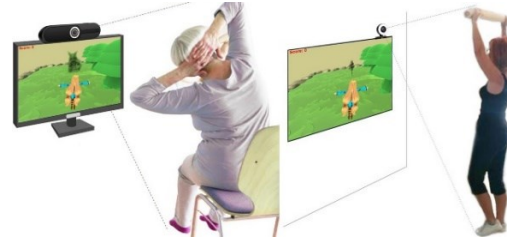


Figure 5. Two forms of games: (a) full body standing movement; (b) upper body movement

4.3.1 Exergame design

In Fit Islands, the shadow representing the action will gradually move closer to the game character from behind. The degree of overlap between the shadow and the game character will determine the player's score. The movements in Fit Islands will primarily focus on preventing cardiovascular diseases [29], enhancing muscle strength [30], and improving flexibility and balance among the older adults [31]. Fit Islands will categorize and assess the difficulty of all expert-evaluated movements, and these movements will be combined in a default sequence to form a set of progressively challenging game levels. Each level will last approximately 10-15 minutes by default and follow the sequence of warm-up, exercise, and stretching for the movements. An option to adjust the difficulty level will be added at the end of each game, which allows the game to maintain a basic level of challenge while providing inclusivity, ensuring the autonomy and safety.

4.3.2 Health knowledge

Authors have collected knowledge on diet, exercise, living habits, self-monitoring, disease prevention, psychology, etc., such as: knowledge of nutrients and proper dietary combinations [32], exercise and exercise precautions, how to maintain good sleep, how to prevent cardiovascular and cerebral vascular diseases [33], how to maintain a good state of mind. When presenting health knowledge to older adults, it is important not to excessively emphasize the authority of the knowledge but to consider ways to make it easier and quicker for them to comprehend the information. The image on the right uses a friendly tone and presents the health knowledge through visuals and audio (see Figure 6).



Figure 6. Comparison of health knowledge

4.3.3 Business

Fit Islands has added a product point exchange and membership system, where points can be earned through farm cultivation, sports games, cognitive games, and interaction with friends. These points can be exchanged in the shop for virtual game character skins as well as some real elderly goods, such as electronic blood pressure monitors. Appropriate adverts or videos are played at the end of the game. This will attract brands, communities, nursing homes, etc. with the same target audience to place commercial ads or offer real products. Fit Islands offers three levels of membership for different number of players.

4.4 Plan the user interface and visuals

The design of user interfaces for older adults needs to take into account their habits and preferences in terms of style and aesthetics (see Figure 7). Fit Islands has been designed to meet the inclusive design principle.



Figure 7. the interface in the demonstration: (a) initial interface; (b) game description; (c) user registration; (d) interface for selecting gender; (e) main interface; (f) settings interface; (g) membership interface; (h) game map; (i) shop interface; (j) post office interface; (k) farm interface; (l) game centre; (m) Exergame interface; (n) health knowledge interface; (o) multiplayer interface.

4.5 Consider the gameplay interaction

Fit Islands employs an audio-visual sensory experience to simplify data and information feedback, making it easier for older adults to take in different information through voice-reading and pictures. The interface switching throughout the game mainly relies on direct mouse clicks, while the motion capture in Fit Islands avoids any wearable devices, and the game can be played simply by standing in front of a computer with a camera. Immersive interaction is an important aspect of the interaction design process in Fit Islands. In addition to creating natural scenes through ambient sounds and dynamic effects, Fit Islands also adds multiplayer online. Players can play against each other in cognitive games or complete challenges in two-player mode in exergame.

4.6 Develop demonstration

The development of the Fit Islands demonstration consisted of interface design, a motion capture program, motion pose image recognition and the creation of a 3D game world. For the motion capture part, the demonstration uses the ThreeD Pose Tracker algorithm. Motion pose recognition uses a skeleton recognition algorithm to identify imported motion pose images and display these skeletons in the image of the game character. The 3D game world for the Fit Islands demonstration was created using an open-source resource called 'Top-down island' from the Unity shop. The game is played from a bird's eye view and the character is controlled by running around with a right click of the mouse.

4.7 Test and iterate

4.7.1 Test

The initial testing of the demonstration for Fit Island was split into two phases. In the first phase, a 62-year-old German woman was invited to test the functionality and operation of the game after the initial demonstration was completed. After receiving feedback, the initial game demonstration was refined. The second stage was to invite 9 Chinese older adults to test the functionality and useability of the game after the completion of the refined game demonstration. The nine seniors, aged 51-77, included three women and six men. B-F, were tested in the computer room of a community centre in Yancheng, Jiangsu Province, China, G and H were tested in the activity room of a community centre in Yancheng, Jiangsu Province, China, and the other two, I and J, were tested in their homes. These older adults were asked to operate and use the game and to provide feedback on the corresponding issues (see Figure 8).



Figure 8. Field photos of nine older Chinese adults test subjects

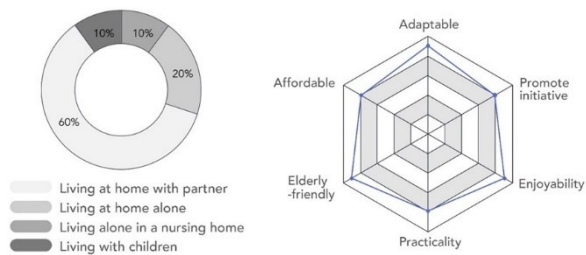


Figure 9. Analysis of test results. (a) Residency (b) Evaluation scores of the game

Once individuals' data is involved, protecting their data privacy will be critical. Therefore, the good practice for protecting individual privacy in the test is reported as follows. First and foremost, the test was conducted in a fully anonymous manner, meaning that information such as names, birth dates, addresses, and resident identity card numbers were not recorded. This approach excludes any possibility to identify any individual test participant. Furthermore, to participate in the test, all the participants were informed of the transparent purpose of the study and needed to give consent to provide their basic demographic information such as age, gender, residential address, and level of education. These measures also go in line with Guide on the Good Data Protection Practice in Research [34].

The test was conducted in the form of a questionnaire. The questions mainly included evaluation of the game interface, game content and features, and whether it is useful for health. Based on the results of ten older adults in the initial test. (see Figure 9), the following findings can be concluded:

- This system can effectively motivate older Chinese adults to exercise and increase their amount of

physical activity. The initial test participants were able to complete the exercises relatively easily, and the music and scoring format allowed them to concentrate on the exercises. Many older adults said they look forward to the health knowledge.

- This system provides a free platform for Chinese seniors to communicate and play multiplayer games. The affordability and anonymous and fun way can be used as a complement to family pensions, allowing older adults to socialize more with people of their age, mentally reducing over-dependence on their children and alleviating feelings of social isolation and loneliness. Several older adults in the test expressed their enjoyment of the post office function and their expectations of the online multiplayer function.
- The age-friendly and inclusive design of this system can attract older adults to actively use the technology. Most of the older adults responded positively to the game's functions and interface, found the game relatively easy to operate and use, and expressed acceptance of the game's format.
- The system is adaptable to different devices and different scenarios. Desktop computers were used in the computer room, while laptops were used in the activity room and the testers' homes. Older adults can play the game in different building scenarios using existing equipment.

4.7.2 Limitations and Future Work

The demonstration of Fit Islands has basically realized most of the contents and functions of the game design, and in the future, it will further complete more accurate motion capture, improve the whole game's motion library and realize the function of multi-person online. The above test as an initial test of Fit Islands has already obtained positive results, but the sample is small the demonstration lacks further refinement and is only a functional test. Further testing and evaluation will require larger participant sizes and long-term follow-up observations to ensure the long-term sustainability and effectiveness of Fit Islands.

After further refinement of the game demonstration, several older adults will be invited for further testing ($n \geq 30$). A uniform site will be chosen and the test subjects will be brought together in a uniform site for testing. The future test will be divided into four main phases. The first phase of the test will be the distribution of information about Fit Islands' research objectives and planned tasks to the test participants, who will be introduced to the Exergames equipment and procedures. The second phase will divide the test subjects into a control group ($n \geq 10$) and an intervention group ($n \geq 20$). The physical and psychological conditions of the two groups will be assessed separately to ensure that the subjects are not

statistically significantly different. There are many methods of testing specifically for older adults. For example, a research team assessed the functional status of older adults using the Activities of Daily Living Scale (ADL) and the Instrumental Activities of Daily Living Scale (IADL) [35]. Mental health status can be tested using the UCLA Loneliness Scale or subjective questions [36]. The third phase will have the intervention group ($n \geq 20$) play the game three times a week for five weeks, either at home or at this uniform venue of their choice. Various parameters will be recorded during the game, including the time of entry into the game, the time of user creation, the time of exercise, the score of the game and whether the participant completed the game in its entirety or was launched early, etc. The control group ($n \geq 10$) will not be intervened and will live their lives according to their own way of life. In the fourth phase, the physical and psychological well-being of the two groups will be assessed again after five weeks in the same way as in the second phase. Effective proof can be derived by comparing the data of the control group ($n \geq 10$) and the intervention group ($n \geq 20$) at Phase 4.

After the validation, the project team plans to initiate an internet-based start-up to further promote the project and its products to as many people as possible through multiple channels in order to make a real impact. Furthermore, the start-up (i.e., online store) can potentially create an economy around elderly-focused educational and health games, where profits can in turn support future product iterations and new product development, and potentially create new jobs around the business (e.g., customer service, manufacturing, etc.).

5 Conclusion

In conclusion, Fit Islands strictly follows the generic development methodology presented in Section 3 and meets the development goal of motivating Chinese older adults to increase their physical activity. As the problem of aging continues to grow in emerging economies, the development and application of Exergame will play an effective role in the future to create a smarter, healthier, more active and fun lifestyle for the older adults. Furthermore, due to its advantages such as high flexibility, affordability, and ease of use, the proposed system can be easily integrated into larger smart city frameworks such as Dynamic Vertical Urbanism [37] to generate a comprehensive smart city experience for the aging population and beyond.

Acknowledgment

The paper was written during the time when the co-author Rongbo Hu was at Department of System Design

Engineering, Keio University, Japan.

References

- [1] United Nations. World Population Prospects: The 2017 Revision. Online: https://population.un.org/wpp/Publications/Files/WPP2017_KeyFindings.pdf, Accessed: 26/12/2022.
- [2] Hoskisson, R., Eden, L., Lau, C., & Wright, M. Strategy in Emerging Economies. *Academy of Management Journal* 2017.
- [3] Shetty, P. Grey Matter: Ageing in Developing Countries. *The Lancet* 2012, 379, 1285–1287.
- [4] Cacioppo, J., Hawkley, L., & Thisted, R. Perceived Social Isolation Makes Me Sad: 5-Year Cross-Lagged Analyses of Loneliness and Depressive Symptomatology in the Chicago Health, Aging, and Social Relations Study. *Psychol Aging*, 25, 453–463, 2010.
- [5] Steptoe, A., Owen, N., Kunz-Ebrecht, S., & Brydon, L. Loneliness and Neuroendocrine, Cardiovascular, and Inflammatory Stress Responses in Middle-Aged Men and Women. *Psychoneuroendocrinology*, 29, 593–611, 2004.
- [6] Skouby, K., Kivimäki, A., Haukipuro, L., Lynggaard, P., & Windekilde, I. Smart Cities and the Ageing Population. *The 32nd Meeting of WWRP*, 2014.
- [7] Larsen, L., Schou, L., Lund, H., & Langberg, H. The Physical Effect of Exergames in Healthy Elderly—a Systematic Review. *GAMES FOR HEALTH: Research, Development, and Clinical Applications*, 2, 205–212, 2013.
- [8] Chen, A., Yin, H., & Yan, J. The Relationship between Physical Activity and Well-Being in Older Adults: The Mediating Role of Loneliness. *China Sports Technology*, 135–139, 2010.
- [9] Li, J. A Study of the Effects of Physical Activity on the Mental Health, Well-Being and Quality of Life of Older People and Its Associated Influencing Factors, 2005.
- [10] Studenski, S., Perera, S., Hile, E., Keller, V., Spadola-Bogard, J., & Garcia, J. Interactive Video Dance Games for Healthy Older Adults. *The journal of nutrition, health & aging*, 14, 850–852, 2010.
- [11] Gerling, K., Schild, J., & Masuch, M. Exergame Design for Elderly Users: The Case Study of SilverBalance. In Proceedings of the Proceedings of the 7th International Conference on Advances in Computer Entertainment Technology, 66–69, 2010.
- [12] Chartomatsidis, M.; Goumopoulos, C. A Balance Training Game Tool for Seniors Using Microsoft Kinect and 3D Worlds. In Proceedings of the ICT4AWE, 135–145, 2019.

- [13] SilverFit 3D Interactive Rehabilitation Training System. Online: <https://gies.hk/zh-hk/expo/exhibition-products/detail/1223>, Accessed: 27/12/2022.
- [14] Hu, R., Kabouteh, A., Pawlitza, K., Güttler, J., Linner, T., & Bock, T. Developing Personalized Intelligent Interior Units to Promote Activity and Customized Healthcare for Aging Society. Proceedings of 36th International Symposium on Automation and Robotics in Construction (ISARC 2019), 234–241, 2019.
- [15] Weisman, S. Computer Games for the Frail Elderly. *Computers in Human Services*, 11, 229–234, 1995.
- [16] Ijsselstein, W., Nap, H.H., de Kort, Y., & Poels, K. Digital Game Design for Elderly Users. In Proceedings of the Proceedings of the 2007 conference on Future Play, 17–22, 2007.
- [17] Flores, E., Tobon, G., Cavallaro, E., Cavallaro, F., Perry, J., & Keller, T. Improving Patient Motivation in Game Development for Motor Deficit Rehabilitation. In Proceedings of the Proceedings of the 2008 international conference on advances in computer entertainment technology, 381–384, 2008.
- [18] Mao, G., Lu, F., Fan, X., & Wu, D. China's Ageing Population: The Present Situation and Prospects. In *Population Change and Impacts in Asia and the Pacific*, Poot, J., Roskruege, M., Eds., New Frontiers in Regional Science: Asian Perspectives. Springer: Singapore, 269–287, 2020
- [19] Woo, J., Kwok, T., Sze, F., & Yuan, H. Ageing in China: Health and Social Consequences and Responses. *International Journal of Epidemiology*, 31, 772–775, 2002
- [20] Zhang, A., & Master, F. China's First Population Drop in Six Decades Sounds Alarm on Demographic Crisis, Online: <https://www.reuters.com/world/china/chinas-population-shrinks-first-time-since-1961-2023-01-17/>, Accessed: 18/02/2023.
- [21] Fang, Y., Wang, H., & Ouyang, Z. An Overview Report on the Situation of Elderly People in Urban and Rural China: 2000–2015. *China popul. dev. stud.* 2019, 2, 323–345.
- [22] World Health Organization. China Country Assessment Report on Ageing and Health. Online: https://iris.who.int/bitstream/handle/10665/194271/9789241509312_eng.pdf, Accessed: 27/12/2022
- [23] Hu, R., Linner, T., Wang, S., Cheng, W., Liu, X., Güttler, J., Zhao, C., Lu, Y., & Bock, T. Towards a Distributed Intelligent Home Based on Smart Furniture for China's Aging Population: A Survey. In Proceedings of the 38th International Symposium on Automation and Robotics in Construction (ISARC 2021), 41–48, 2021
- [24] World Economic Outlook Database. Online: <https://www.imf.org/en/Publications/WEO/weo-database/2022/April>, Accessed: 16/02/2023.
- [25] Qiang, L., Gang, X., & Zhen, Z. Loneliness and Its Correlates Among the Aged People Living Alone in Urban Areas. *Journal of East China Normal University*, 51, 160, 2019.
- [26] Wang, X., Luo, W., & Chen, X. Influence of Square Dance on Physical and Psychological Health of the Elderly. *Chin J Gerontol*, 34, 477–478, 2014.
- [27] Li, B. Divergence and Distinction: Living Arrangements of Older People in China. *Population Sciences in China*, 2010.
- [28] National Ageing Development Bulletin 2020. Online: <http://www.nhc.gov.cn/cms-search/xxgk/getManuscriptXxgk.htm?id=c794a6b1a2084964a7ef45f69bef5423>, Accessed: 27/12/2022.
- [29] 10 Best Chair Gym Exercises Printable - Printablee.Com Available online: https://www.printablee.com/post_chair-gym-exercises-printable_178691/, Accessed: 18/02/2023..
- [30] 21 Chair Exercises for Seniors: Complete Visual Guide Available online: <https://californiamobility.com/21-chair-exercises-for-seniors-visual-guide/>, Accessed: 05/11 /2023.
- [31] 10 Best Printable Chair Yoga Exercises For Seniors - Printablee.Com Available online: https://www.printablee.com/post_printable-chair-yoga-exercises-for-seniors_178538/, Accessed: 05/11 /2023.
- [32] Alzheimer's Association Available. Online: https://act.alz.org/site/SPageServer/?jsessionid=00000000.app20115b?NONCE_TOKEN=9C725D55D38175E0DC73CF979C87911E, Accessed: 05/11 /2023.
- [33] Herrschaft, H. Prevention of Cerebrovascular Circulatory Disorders. *Fortschritte der Neurologie-Psychiatrie*, 53, 1985.
- [34] Wilms, G. Guide on good data protection practice in research. *European University Institute*, 2019.
- [35] Guo, L., An, L., Luo, F., & Yu, B. Social Isolation, Loneliness and Functional Disability in Chinese Older Women and Men: A Longitudinal Study. *Age and Ageing*, 50, 1222–1228, 2021.
- [36] Russell, D., Peplau, L., & Cutrona, C. The Revised UCLA Loneliness Scale: Concurrent and Discriminant Validity Evidence. *Journal of Personality and Social Psychology*, 39, 472–480, 1980.
- [37] Hu, R., Pan, W., & Bock, T. Towards dynamic vertical urbanism: A novel conceptual framework to develop vertical city based on construction automation, open building principles, and industrialized prefabrication. *International Journal of Industrialized Construction*, 1(1), 34–47, 2020.

Use of robotics to coordinate health and safety on the construction site

Fabian Dornik¹ and Niels Bartels^{1,2}

¹Faculty of Process Engineering, Energy and Mechanical Systems, University of Applied Sciences Köln, Germany

²Faculty of Civil Engineering and Environmental Technology, University of Applied Sciences Köln, Germany
fabian_michael.dornik@smail.th-koeln.de, niels.bartels@th-koeln.de

Abstract

The construction industry has the highest number of accidents compared to other sectors of the economy. Minimising the risk of accidents requires compliance with safety standards and constant monitoring by responsible persons. The use of robots for high-risk work can increase safety and prevent accidents on construction sites. On the other hand, the use of robots can create new health risks. Especially the use of UAVs is therefore highly regulated, whereas this is not the case for legged robots. This paper identifies the challenges and opportunities for coordinating health and safety on construction sites with robots. It is based on a literature analysis that looks at the status of the general area of application of robots and health and safety coordination services. Based on the literature analysis, guided workshops were held with various stakeholders (site managers, health and safety coordinators, German employer's liability insurance association (BG Bau), clients and planners). To validate the results, a quantitative survey was conducted. The interviewees believe that robots can assist them with activities on the construction site or take over some of the activities independently. Challenges identified the interviewees were mainly related to purchase costs, acceptance by construction workers an unclear legal framework. The results show that robots should work autonomously. To create synergies, robots should not only coordinate health and safety, but also carry out target-performance comparisons with the building model, produce daily construction reports with photos and take measurements of the trades.

Keywords – Unmanned Aerial Vehicles (UAV); Legged Robot; Safety and Health Protection; Construction Site; Safety Management

1 Introduction

Statistically, one person dies on a construction site in Germany every three and a half working days [1].

Around 20 % of accidents at work in Germany and the European Union occur in the construction sector [2]. Falls are the main cause of fatal accidents on construction sites [3]. Compared with other sectors of the economy, the construction industry therefore has the highest accident rate and a high sickness rates [3]. Unlike other sectors, the construction industry is characterised by daily changing conditions and responsibilities for those carrying out the work [4]. Daily changes in construction progress create dynamic workplace situations that are particularly hazardous. Frequent changes in personnel, a high proportion of foreign employees on site, time and cost pressure and weather conditions can also lead to additional hazards [5].

In order to prevent accidents, there are various regulations, laws and standards designed to improve occupational safety [6]. In Germany, the Construction Site Ordinance in particular, is designed to improve safety and health protection on construction sites. In addition, building owners are obliged to comply with health and safety measures.

On construction sites where employees of several employers are working, at least one coordinator must be appointed to coordinate health and safety on the construction site [7]. This means that the risk of accidents can be minimised through safety standards and continuous monitoring, such as inspections by a safety and health coordinator. Because of their moderating role, safety and health coordinators must be able to deal with conflict situations, have organisational talent, a high level of social competence and be able to mediate between the parties involved.

As part of his or her duties, the safety and health coordinator is required to carry out regular site inspections during the execution of the work and document and report any defects or malfunctions. They must also influence those carrying out the work to rectify any errors. [8,9] In practice, the resulting regular inspections mean that safety and health coordinator have to travel long distances between individual construction sites.

At the same time, robots are becoming increasingly

important for construction sites. There are various forms of robots, such as exoskeletons, 3D printers, unmanned aerial vehicle (UAV), robots, humanoids or stationary robots that can support activities on construction sites. Robots could improve safety on construction sites and preventing accidents, especially by using robots for activities where humans are exposed to a high-risk potential. Figure 1 shows these robots and their main area of application.

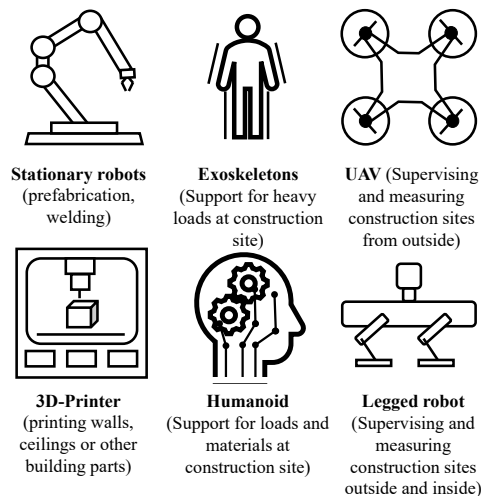


Figure 1: Types of robots on construction sites

Regarding safety managements, especially drones and legged robots could support humans on construction sites by supervising construction sites autonomously and documenting findings with photos or videos. Therefore, this paper examines strategies for coordinating health and safety on construction sites using these two forms of robots by taking practical application, regulatory requirements, and the human-robot-interaction into account.

2 State of the art

The paper is based on a literature analysis that examines the current state of robotic applications on construction sites in general, but also regarding health and safety management services. The keywords "unmanned aerial system / vehicle (UAS / UAV)", "unmanned ground vehicle (UGV)", "building", "construction site", "legged robot", "spot boston dynamics" and "safety" were searched in different variations on scopus, google scholar and springerlink. Figure 2 illustrates the number of publications in the period from 2013 to 2023.

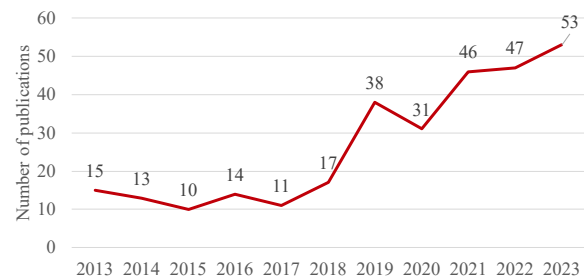


Figure 2: Publications in the period 2013 to 2023

The results show that, particularly in recent years, many papers have been published, that deal with monitoring the construction process, inspecting buildings, and surveying the site using UAVs or legged robots. It can also be seen that the focus of the papers is certainly on the use of UAVs and not legged robots. It should also be noted that the use of robots to coordinate health and safety measures is currently not fully researched. Only few papers deals with the use of UAVs to monitor safety on construction sites by analysing video recordings (e.g. [10]). Several papers also show that there are currently no guidelines for the safe use of robots on construction sites to protect people, especially for legged robots (e.g. [11,12]).

The next sections show the characteristics and possibilities to use legged robots and UAV for safety management on construction sites.

2.1 Legged robot

Legged robots mimic human or animal movements, which means that they can overcome different ground conditions such as stairs or obstacles. In addition, this type of robot requires little to no preparation and setup of the construction site in order to move around. [13]

The stereo camera installed in the legged robot on site can be used to create 3D point clouds for mapping the environment [13]. In addition to the stereo camera, the legged robot can be equipped with other modular configurations, such as panoramic cameras, LiDAR sensors and a with six degrees of freedom robot arm for opening doors [13]. The battery life varies depending on the product. For example, it can be approx. 90 minutes in operation and 180 minutes in standby mode [14].

There is currently a wide range of applications where construction workers could be assisted by a legged robot inside the building, e.g. in construction documentation or quality control. This is currently the main use case and there are several studies in which construction progress has been recorded using a 360° camera [11,15]. These studies are already providing initial insights into the use of legged robots for health and safety measures. In particular, the studies show that the following technical measures apply to the legged robot (1) an illumination level of more than 2 lux is required for operation, (2)

glass or transparent objects (e.g. a glass door is difficult for the legged robot to detect), (3) objects less than 30 cm high and objects than 3 cm on the ground are not detected by the legged robot, (4), no use on scaffolding steps and (5) to avoid the risk of collision with the legged robot during operation, people should keep a distance of 2 metres. [11]

A key aspect of using the robot as an aid to health and safety coordinator is that current studies show that applications where the robot frequently encounters people are excluded. In concrete terms, this means that the use of legged robots during working hours could be excluded, or that an operator would need to monitor the legged robot.

2.2 Unmanned aerial vehicles (UAV)

An unmanned aerial vehicle (UAV) is an aircraft without a pilot on board. Thanks to the lift generated by their rotors, UAVs can stay aloft without changing their position or altitude. They are therefore well suited for use on construction sites [16].

A distinction can also be made between remotely piloted and fully automated flying UAVs [17]. Under European law, only remotely piloted UAVs can currently be licensed. Fully automated UAVs do not currently comply with either European law or International Civil Aviation Organisation regulations. [17] In order to operate a UAV in Germany, both national and European legislation must be observed. These include the German Air Traffic Act (LuftVG) and the Easy Access Rules for Unmanned Aircraft Systems apply. [18]

The operation of UAVs may require an operating licence. This depends on the UAV's take-off mass, the flight altitude and flying within visual range. If the flight altitude of the UAV is more than 120 metres or if it is operated in geographical areas, e.g. near hospitals or airports, a risk assessment and approval by the aviation authorities is required for operation. In addition, the pilot will need an EU certificate of competence or an EU remote pilot licence. [19] The flight time of a UAV depends on the take-off weight and the payload, e.g. for cameras or measuring instruments, as well as on wind, weather and flight altitude, and can be between 7 and 30 minutes [17]. UAVs offer the possibility of inspecting buildings or terrain at high altitudes.

Although UAVs can be used efficiently and in a variety of ways outdoors, their use indoors is limited [20]. One of the reasons for this is the difficulty of collision avoidance for UAVs in the air compared to other robots on the ground [21].

The possibility of using UAVs on construction sites has already been investigated in several studies. Inspections of all types of structures, such as bridges or wind turbines, are carried out using UAVs equipped with a camera [17]. Surveys of the terrain using single images

and photogrammetric three-dimensional point clouds have also already been carried out [17]. In addition, the use of a UAV to improve site safety has been investigated. The videos recorded by the UAV were analysed using an image processing programme to determine whether the people on the construction site were wearing their personal protective equipment (PPE) [10].

2.3 Conclusions for the integration of robots and safety management

As seen in sections 2.1 and 2.2, the use of legged robots and UAV have different use cases in the construction industry. The resulting technical requirements offer corresponding advantages and disadvantages for use on the construction site. These in turn have an impact on the use of the legged robot and the UAV for the safety management on construction sites. Table 1 shows the advantages and disadvantages of the individual use cases for legged robots and UAVs.

Table 1: Advantages and Disadvantages of the usage of Legged Robots and UAV for safety management on construction sites

Robot	Advantages	Disadvantages
Legged Robot	Is able to climb stairs; Full view of construction site due to 360° camera; Could be used inside buildings	Technical barriers, such as glass or low objects; Is not able to speak or interact with humans; Risk of collision with humans; Must be obeyed by humans
UAV	Could supervise tall buildings; Photos and documentation of high places; Fewer obstacles than on the ground; Larger operating radius	License needed; Partial authorisations required; Could not open doors (not suitable inside buildings; Influence of the weather on operations; Risk of collisions with aircraft; Noise pollution from UAVs

The Table 1 shows, that neither the legged robot nor the UAV fits for all tasks, that need be done by the safety management on construction sites. Especially the human-robot-interaction and the documentation can actually not be done automatically by the robots.

3 Material and methods

Based on the literature analysis (step 1), workshops were initially held with various stakeholders in step 2. These were based on guidelines and aimed to evaluate the current areas of application of robots in health and safety management and requirements for implementing robots for health and safety management in practice. A total of 9 workshops was held. The documentation was done by transcribing the video and audio recordings. The workshops were held via videoconference.

Based on the literature review and the workshops, a quantitative survey was carried out to validate the theses more broadly in step 3. The survey "Strategies for the use of robotics to coordinate health and safety on the construction site" was divided into five subject areas with a total of 14 questions. It was carried out between 29/09/2023 and 23/10/2023 and disseminated through social networks, professional associations (e.g., BG Bau) and health and safety companies. A total of 40 people from different companies and professional associations participated in the survey. Based on the number of members of the German Association of Health and Safety Coordinators, this sample size is sufficient at a confidence level of 80%. The association has 190 members [22].

In step 4 the results of the qualitative and quantitative survey were analysed and discussed by the authors. The aforementioned methodology is shown in Figure 3.

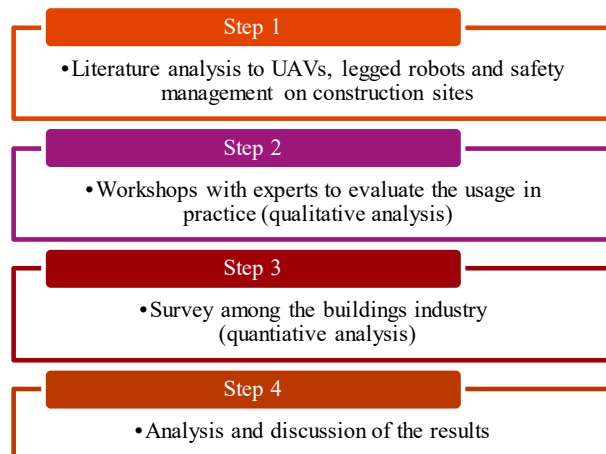


Figure 3: Methodology of this paper

The following sections explain the steps 2-4 more in detail: in section 4.1 the workshops and the results of the workshops (step 2) are described, in section 4.2 the results of the survey (step 3) are shown. The discussion (step 4) is shown in section 5.

4 Integration of robotics in safety management

The following sections describe the results of the workshops and the survey, that aimed to find out possible applications and added values of the integration of robotics in safety construction. The workshops, as a qualitative analysis, were based on the literature reviews. Based on that, the survey was conducted.

4.1 Workshop results

The 9 workshops with up to 3 experts were conducted between 01/08/2023 and 22/09/2023, to evaluate various perspectives on safety management on construction sites, the relevant stakeholders, such as site managers, health and safety coordinators, the German employer's liability insurance association (BG Bau), clients and planners were interviewed in the workshops, whereas respectively 3 participants from BG Bau and safety and health coordinators as well as 5 site managers took part in the workshops.

The workshops showed that the participants have different perceptions of the topic of robotics in safety management. This is because all those involved have a different focus in their work with the construction site and a different level of knowledge about the integration of robotics in safety management. The workshops revealed that none of the interviewees had already fully utilised robotics for safety management on construction sites. However, some workshop participants already had initial ideas on how robotics could be used. However, those who had already dealt intensively with the use of robotics on construction sites in general were also increasingly sceptical about the use of robotics for safety management in particular. This is particularly because of various details, such as driving licences for drone flights or the supervision of robotics, still represent decisive safety aspects for the successful use of robotics on construction sites.

On the other hand, the workshops showed, that the participants could imagine being supported their work by robots. For example, the robot could help with measuring, target/actual comparisons, and construction documentation. However, it was also clear that robots could not replace a visit to the construction site, especially due to

1. lack of human interaction, that is – from the perspective of the workshop participants - a key success factor in ensuring order and safety on the construction site,
2. data protection aspects, such as taking pictures that go beyond the construction site and could capture people on the construction site and
3. the need for a proper documentation of safety risks on the construction site.

All in all, the participants stated, that a robot can complement the work of safety management but cannot completely replace it.

4.2 Survey results

Based on these workshops results, a quantitative survey was conducted. The survey focused on German Safety and Health protection coordinator. The results of the survey are presented in the following three sections.

4.2.1 Situation analysis

The workshops have shown that there is potential to save time, particular by eliminating the need to travel between the office and the construction site. With this in mind, the first question asked was how much time the participants spent travelling between the office and the construction site each day. 40% of respondents indicated that they spend up to 2 hours travelling to and from the office and site in a working day. A further 25% of the respondents stated, that they spend up to 3 hours travelling. The results are shown in Figure 4.

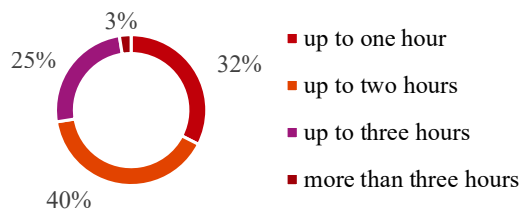


Figure 4: How much time do you spend travelling time between office and site each working day (n=40)

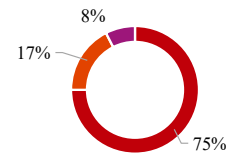
Furthermore, the participants were asked, what work they conduct on the construction site and how they document potential errors. Approx 98% of respondents said that they take photographs of conspicuous features using smartphones or cameras. In contrast, only 15% said that anomalies were recorded using an app on a tablet or smartphone.

These results show, that robots could save time for the construction safety management, by enabling remote working, which saves time by avoiding long journeys.

4.2.2 Robotics in general

Based on that, the participants were asked, if they believe that robots can assist them in their practical work on site in future, so that the time savings could be realised.

Figure 5 shows the results of this question. 75% of the respondents believe, that robots can support them in their daily work on the construction site. 17% of the respondents believe, that robots even could take over part of their daily work on the construction site. Only 8% believe, that robots have no value to their work on the site.



- ... can support me in my work on the construction site.
- ... be able to take over part of my work on the construction site independently.
- ... do not add any value to my work on site.

Figure 5: Answers to the question, if a robot could support the work on the construction site (n=40)

In addition to that, the participants were asked, what kind of robot could support their daily work. 62.5% of respondents believe that the use of a UAV can assist them in documenting the construction process.

Although the robots are seen as a possibility for documentation, the interviewees do not assume that this will lead to a reduction in on-site times at the construction site.

The minority (37.5% of all respondents) agrees with the statement that robots can reduce site attendance. On the other hand, 45% partially disagree with this statement and a further 17.5% disagree. This result is shown in Figure 6.

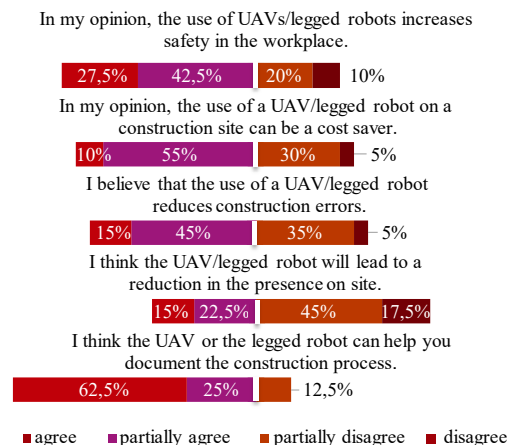


Figure 6: Answers to the question, whether the participants agree or disagree to the following expectations of the robot (n = 40)

On the other hand, Figure 6 also shows, the majority of respondents believe that robotics can improve safety on the construction site in the future. 27.5% of respondents agreed that the use of robots increases work safety on the construction site. A further 42.5% of respondents partially agree with this statement. One advantage is especially seen in the documentation of the construction process and safety aspects (88% agree or partially agree to that statement).

4.2.3 Use of robots - coordination of health and safety protection

According to the respondents, the following characteristics of the robot are a prerequisite for its use in health and safety coordination.

Respondents were divided on the need for the robot to be able to speak independently or interact with other people. Just under 50% see this as a requirement. 43.6% of respondents expect the robot to be able to process data independently. 47.5% expect the robot to be connected to a control centre. 35% expect this to some extent. 52.5% of respondents expect the robot to be able to move around the construction site independently. 72.5% of respondents said that the robot should be able to detect dangerous situations on its own. These results are shown in Figure 7.

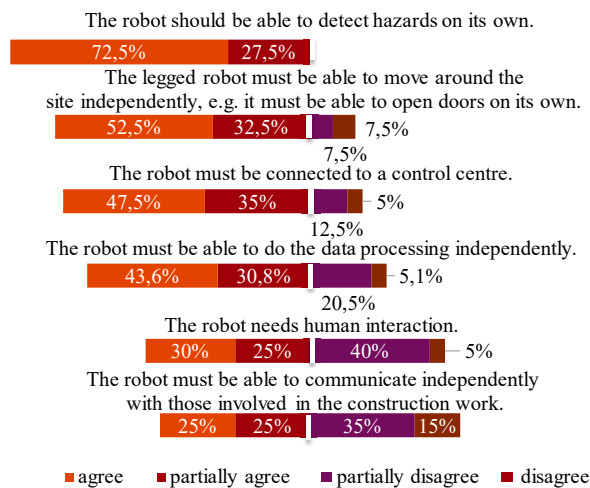


Figure 7: Answers to the question, what characteristic and necessary properties of the robot need to be implemented to support safety management (n=40)

In addition to the requirements on safety robots, the participants were also asked for occurring challenges by using robots for safety management.

Figure 8 shows the results. Especially the acceptance of construction stakeholders is seen as the key challenge (75% of the respondents). This corresponds with the statements from the workshops, which also see a lack of human interaction as a challenge.

In addition to that, the respondents see particular challenges in the cost of acquisition (63%). Furthermore, unclear legal requirements (50%) and privacy policies are seen as challenges. The acceptance of robots as figures of authority is only mentioned by 23% of the participants.

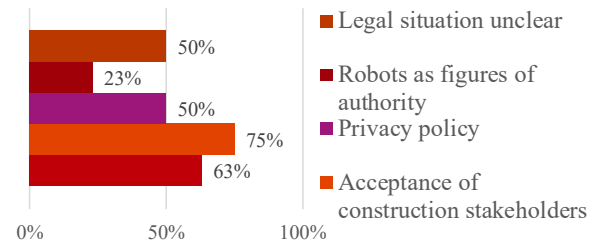


Figure 8: Answers to the question: What challenges do you see in using robots as support for safety management on construction sites (n=40)

Further possible information on the use of robotics in safety management could be entered via a free field at the end of the questionnaire. The participants suggested in that survey, that the robot should provide site security, replacing security guards or camera surveillance. It was also stated that an external service provider should use robots to record data on the construction site and make it available to all those involved in the construction. By doing so, the challenges of costs and flying licences could be avoided.

In addition, one respondent pointed out that the different types of construction sites would also pose different challenges. The use of robots would therefore not be suitable for all construction sites. Lastly, it was also pointed out, that health and safety protection is very difficult. Even as a human health and safety coordinator, it is difficult to instil common sense in the people on a construction site. This would be an impossible task for a robot.

5 Discussion

The results of the study conducted among health and safety coordinators show that the respondents consider the general use of robots on the construction sites to be an added value of supporting their work.

On the one hand, manual safety inspection processes lead to inconsistent, time-consuming and error-prone data collection, what is also supported by ref. [11]. On the other hand, many defects and safety issues can be prevented by real-time monitoring of construction progress and quality control (see also [23]). Robotic data collection can reduce the amount of labour and time required to collect data while improving quality (see also ref. [11]).

However, a closer look at the range of applications for robots, such as coordinating health and safety on the construction site, also reveals that 75% of respondents consider the acceptance of robots by construction workers to be the biggest challenge. The reasons for this became clear in the workshops. A health and safety coordinator needs to be able to deal with conflict situations and have a high level of social skills to mediate between those

involved in the construction work. These are qualities that a robot cannot currently demonstrate. However, respondents were divided on whether a robot should be able to speak independently or interact with other people. In addition, a robot equipped with a camera may create a feeling of surveillance among those involved in the construction work. Data protection should therefore not be overlooked in this context.

In addition to data protection, other legal and employment regulations must be observed. National and European laws must be observed when using UAVs. Prior permissions must be obtained from the aviation authorities to fly over buildings of more than 120 metres in height. The location of buildings, e.g. near airports or hospitals, also restricts the use of UAVs and requires additional permits. It can be concluded that the benefits of UAVs should outweigh the effort required to obtain permission and therefore they are not suitable for every construction site. This is also stated by [18] and is reflected in the survey result. For example, 50% of those respondents said that the (unclear) regulatory framework was one of the biggest challenges to the use of UAVs and legged robots.

Some of the requirements that respondents have for robots on construction sites are not currently feasible from a legal perspective. 52.5% of respondents consider it necessary for the robot to be able to carry out its work autonomously on site. UAVs that are automatically programmed and fly without a pilot cannot currently be licensed under European and international aviation law [11].

In contrast to the legal requirements for UAVs, there are currently no legal requirements for the use of legged robots. To date, there are only a few studies looking at the requirements for the safe use of legged robots on construction sites (e.g. [12]). However, human-robot collaboration on construction sites poses a risk. People should therefore keep a distance of 2 metres when operating a legged robot in order to avoid collisions [11].

Another prerequisite for the use of robots in health and safety on the construction sites is the ability to recognise hazards independently. For example, hazard could be caused by steps or pipes lying on the ground. These hazards cannot currently be reliably detected by a legged robot. Recent studies show that the legged robot has difficulty detecting objects less than 30 cm in height. It also fails to detect objects smaller than 3 cm, such as pipes lying on the floor. [11]

Literature and the interviewees also stated that it is difficult to monitor health and safety on the construction sites for robots, because construction sites differ according to the type of work being carried out. For example, there are building sites, civil engineering sites and pipeline construction sites. Each of these types of construction site has different health and safety challenges that the robot has to identify independently.

The cost of the robots is also seen as a challenge by 63% of respondents. Currently, a legged robot costs approx. € 130.000 and a UAV costs approx. €15.000 [24,25]. In the long term, however, costs can be saved through synergy effects. In addition to using the robot purely to coordinate health and safety, the survey showed that it is possible to compare target and actual values with the building model, create daily construction reports with photos and take measurements of trades. There is potential here for a new market segment on the construction site through a company offering centralised visual recording of the construction site using UAVs and legged robots. A company would be contracted to capture site data and make it available to contractors, clients, site managers, architects and health and safety coordinators for downstream processes.

6 Summary

The results show that construction workers generally see the use of robots on construction sites as adding value to their work. However, the use of robots to coordinate health and safety on construction sites is viewed critically. This is due to the need for social skills that robots cannot currently provide. The desired autonomous use is also currently feasible. The autonomous use of UAVs on construction sites is currently not possible under current aviation law. On the other hand, there are no studies on the risks associated with the use of legged robots on construction sites.

The high acquisition costs are also a major challenge, according to the respondents. However, the use of robots on construction sites offers the opportunity to increase efficiency and save time and resources on the construction site. One way of doing this is by bundling synergies. In addition to coordinating occupational safety, the robots will carry out comparisons with the building model, create daily construction reports with photos and take measurements of the trades. Robots could play an important role in the construction industry in the future by supporting those involved in construction in their work or even taking it over completely. Further developments are needed for their successful use. For example, the legal basis for the use of robots on construction sites needs to be reviewed. The risks of human-robot interaction need to be researched and the autonomy of robots needs to be further developed.

References

- [1] BG Bau. BG BAU stellt Zahlen für 2022 vor. Online: <https://www.bgbau.de/mitteilung/bg-bau-zahlen-2022>, Accessed: 06/12/2023.

- [2] Radtke R. Anzahl der meldepflichtigen Arbeitsunfälle in Deutschland nach ausgewählten Wirtschaftszweigen im Zeitraum 2012 bis 2022. On-line: <https://de.statista.com/statistik/daten/studie/209520/umfrage/gemeldete-arbeitsunfaelle-nach-wirtschaftszweigen/>, Accessed: 06/12/2023.
- [3] Kring F., Follmann J., Meyer G. and Dudek T. *Praxis-Handbuch SiGeKo*, Rudolf Müller, Cologne, 2020.
- [4] Lennartz K.M. Chancen agiler Methoden für das Bauprojektmanagement. In *Beiträge zum 29. BBB-Assistententreffen vom 06. bis 08. Juni 2018*, page 193, Brunswick, Germany, 2018.
- [5] Gartner H. *Praxisleitfaden Baustelle und Recht. Rechtsfragen erkennen und lösen*, Linde Verlag, Vienna, 2023.
- [6] Muro D. and Rahlmeyer N. *Sicherheit und Gesundheitsschutz auf der Baustelle*, Richard Boorberg, Stuttgart, 2023.
- [7] Bundesministerium der Justiz. BaustellV. On-line: <https://www.gesetze-im-internet.de/baustellv/BaustellV.pdf>, Accessed: 06/12/2023.
- [8] Rusch L.-P. *Basics Bauleitung*, De Gruyter, Basle, 2017.
- [9] Sohn P. *Architektenhaftung. Grundstrukturen in Haftpflicht und Deckung*, VVW GmbH, Karlsruhe, 2018.
- [10] Abbas M., Mneymneh B. E., Khoury H. Use of unmanned aerial vehicles and computer vision in construction safety inspections, In *Proceedings of International Structural Engineering and Construction*, pages 1-6, Kuching, Sarawak, Malaysia, 2016.
- [11] Afsari K., Halder S., Ensafi M., Devito S. and Serdakowski J. Fundamentals and Prospects of Four-Legged Robot Application in Construction Progress Monitoring. *Built Environment*, 2: 274-283, 2021.
- [12] Halder S., Afsari K. Robots in Inspection and Monitoring of Buildings and Infrastructure: A Systematic Review, *Applied Sciences*, 13(4): 2304, 2023.
- [13] Casini M. *Construction 4.0*, Woodhead Publishing, Cambridge, 2022.
- [14] Boston Dynamics. Spot Userguide. On-line: <https://www.generationrobots.com/media/spot-boston-dynamics/spot-user-guide-r2.0-va.pdf>, Accessed: 06/12/2023.
- [15] Halder S., Afsari K., Serdakowski J., DeVito S., Ensafi M. and Thabet W. Real-Time and Remote Construction Progress Monitoring with a Quadruped Robot Using Augmented Reality, *Buildings* 12(11):2027, 2022.
- [16] Alsamarraie M., Ghazali F., Hatem Z. M. and Flaih A.Y. A Review on the Benefits, Barriers of the Drone Employment in the Construction Site, *Jurnal Teknologi*, 84:121–131, 2022.
- [17] Landrock H. and Baumgärtel A. *Die Industriedrohne – der fliegende Roboter*, Springer Fachmedien, Wiesbaden, 2018.
- [18] Bundesministerium für Digitales und Verkehr. Rechtsgrundlagen für den Drohnenbetrieb. On-line: <https://www.dipul.de/homepage/de/informationen/allgemeines/rechtsgrundlagen/>, Accessed: 06/12/2023.
- [19] Bundesministerium für Digitales und Verkehr. Drohnen – Freiheit und Sicherheit für die unbemannte Luftfahrt. On-line: https://www.lba.de/SharedDocs/Downloads/DE/B/B5_UAS/Drohnenflyer_BMVI.pdf;jsessionid=9A2E46680AE1E78A9918F17BCE72B7DA.live11311?__blob=publicationFile&v=4, Accessed: 06/12/2023.
- [20] Hamledari H., Davari S., Sajedi S. O., Zangeneh P., McCabe B. and Fischer M. UAV Mission Planning Using Swarm Intelligence and 4D BIMs in Support of Vision-Based Construction Progress Monitoring and As-Built Modeling, in *Construction Research Congress 2018*, pages 43–53, New Orleans, Louisiana, 2018.
- [21] Freimuth H., Müller J. and König M. Simulating and Executing UAV-Assisted Inspections on Construction Sites, In *Proceedings of the 34th International Symposium on Automation and Robotics in Construction*, pages 647–654, Taipei, Taiwan, 2017.
- [22] Lobbyregister Deutscher Bundestag. VSGK Verband der Sicherheits- und Gesundheitsschutzkoordinatoren Deutschlands e.V. On-line: <https://www.lobbyregister.bundestag.de/suche/R005309/19424>, Accessed: 15/12/2023.
- [23] Abbasa R., Westlinga F.A., Skinnera C., Hanus-Smitha M., Harrisa A. and Kirchnerb N. BuiltView: Integrating LiDAR and BIM for Real-Time Quality Control of Construction Projects. In *Proceedings of the 37th International Symposium on Automation and Robotics in Construction*, pages 233–239, Kitakyushu, Japan, 2020.
- [24] RobotShop Europe. Boston Dynamics SPOT-Roboter. On-line: <https://eu.robotshop.com/de/products/boston-dynamics-spot-robot>, Accessed: 15/12/2023.
- [25] DJI. DJI Inspire 3. On-line: <https://store.dji.com/de/product/dji-inspire-3?from=store-nav&vid=136551>, Accessed: 15/12/2023.

The Influence of an Immersive Multisensory Virtual Reality System with Integrated Thermal and Scent Devices on Individuals' Emotional Responses in an Evacuation Experiment

Xiaolu Xia^a, Nan Li^{a*}, Jin Zhang^b

^a Department of Construction Management, Tsinghua University, Beijing 100084, China

^b State Key Laboratory of Hydro Science and Engineering, Tsinghua University, Beijing 100084, China

* Corresponding author (nanli@tsinghua.edu.cn)

Abstract –

Virtual reality (VR) technology has proven to be a valuable tool for exploring human behaviors during construction disasters. Enhancing VR systems with multisensory devices can amplify participants' immersive experiences and emotional response to emergency events. The potential of thermal and scent feedback to intensify negative emotions and immersion in virtual disaster scenarios remains unconfirmed. Addressing this gap, this study developed an immersive VR system equipped with thermal feedback and scent feedback mechanisms, embedded within a multisensory human-computer interface, to investigate their influence on participants' negative emotional response and sense of presence. Findings indicated that thermal feedback devices by allowing participants to detect notable changes in ambient temperature, markedly augmented negative emotional states, cardiac responses, and skin conductance reactions. Meanwhile, although participants could identify noxious smells emitted by the scent player, and while the scent player did evoke negative emotions, this impact was not significantly stronger than the absence of scent feedback. This research underscores the importance and efficacy of incorporating thermal feedback in virtual disaster environments, and advances the utility of VR as a more potent investigatory tool in the realm of individual behavior during building emergencies.

Keywords –

Virtual reality (VR); Evacuation behavior; Fire emergency; Thermal device; Scent player

1 Introduction

Virtual reality (VR) technology is increasingly being utilized in the field of construction, particularly for addressing challenges involving hazardous environments that cannot be easily experimented with in the real world. Acquiring accurate behavioral data during such disasters poses a unique challenge due to the difficulty in replicating real-world evacuation scenarios. VR technology offers a solution [1,2], with its high ecological validity, the ability for precise experimental design and control over key variables, ease of behavioral data collection, assurance of participant safety, and cost-effectiveness [3]. As an illustration, Bourhim and Cherkaoui [4] employed the Unity3D game engine and the HTC Vive head-mounted display (HMD) to create a virtual representation of a fire evacuation scene in a high-rise building. By comparing the results from this VR simulation to actual fire incident data, the efficacy of VR as a research tool for fire emergencies was underscored. To further enhance the validity and impact of VR in the construction domain, scholars are devising VR systems with multi-sensory interaction models.

Immersive multi-sensory VR devices have the potential to significantly influence users' sense of presence and emotional responses. Compared to conventional desktop display devices [5], these immersive VR systems have been shown to elicit presence and emotional responses more akin to real-world experiences [6]. The depth of sensory engagement in a virtual environment (VE) directly influences how realistically human behavior is replicated in simulations used for construction industry training and safety drills [7]. Essentially, the more deeply and extensively an individual's senses are immersed in the VE, the more realistic their experiences become [5]. A literature review

on multisensory interaction reveals that a staggering 84.8% of studies reported positive outcomes from multi-sensory VR systems [8]. For instance, spatial audio, when compared to non-spatial versions, has a pronounced effect on emotional responses within immersive virtual environments (IVE) [9]. Both visual and auditory immersion have been found to influence the sense of presence, and in some contexts, the emotional responses as well [10]. In the context of architectural fires, flames, smoke, and sounds such as fire alarms were identified as the primary factors influencing how people perceive fire emergency scene, and these elements can be effectively modeled using current VR technologies [11]. How the visual feedback of flames and smoke affects human emotional responses has been widely explored [11,12]. However, the tactile feedback of flames and the olfactory feedback of smoke still need to be further studied.

Multisensory stimuli, particularly tactile feedback, can significantly enhance the VR experience when accurately delivered [13]. Thermal devices, a subtype of tactile devices, produce sensory impressions by generating airflows that interact with human skin [14]. A VR system was developed in [15] that incorporated a thermal source device, linking temperature sensors to the user's HMD and proximity sensors to a heating element. This system increases the heater's temperature as the user approaches, detected by the proximity sensor, and reduces it when the user moves away. The effect of this device on individual responses has not been investigated [15]. While introducing a thermal source in VR scenarios can amplify the participants' VR experience and boost their sense of presence, there were instances where participants reported not sensing the thermal cues [2]. Studies on the efficiency of such training tools indicate that multisensory VR systems might not hold a clear edge over traditional VR platforms that merge visual and auditory feedback [16]. In the context of architectural fires, the fire's temperature plays a pivotal role in shaping individual evacuation behaviors. Yet, the influence of immersive virtual reality systems integrated with thermal feedback within a multisensory human-computer interface on emotional responses of people evacuating from virtual building fires is still to be explored.

In addition, the effect of smell VR devices on participants in building fire simulations has not received sufficient attention. In certain standard settings, exposure to agreeable odors can influence individuals' feelings of presence and emotional reactions [17]. For example, altering ambient temperature and introducing odors at specific moments in a 360° immersive film within a virtual space can amplify viewers' immersion. Conversely, the absence of these supplemental stimuli can diminish the authenticity of their experience [18]. When olfactory cues are introduced in VE, participants

tend to show an increased desire for the presented virtual food, and their ability to interact with this virtual sustenance intensifies this effect [19]. Furthermore, the integration of olfactory tools can improve participants' task performance in VR settings [17]. Nonetheless, within the framework of architectural disaster simulations, especially real-life building fires, the dominant smells are often off-putting. The ramifications of immersive VR systems that combine off-putting smell devices within a multisensory human-computer interplay remain uncharted.

This study delves into the effects of immersive virtual reality (IVR) systems with thermal feedback and smell devices on individuals' emotional responses. Our goal is to ascertain whether these multisensory devices influence an individual's sense of presence and emotional reactions in simulated building fire scenarios. Subsequent sections of this article will detail our research methodology, present the results, and provide a discussion and conclusion based on our findings.

2 Methods

2.1 Experimental Design

To explore the influence of an IVR system with thermal feedback and scent player on human emotional response and sense of presence, a two-factor mixed experiment is designed. Thermal feedback serves as a factor for repeated measurement, while scent feedback is a factor for non-repeated measurement. Participants' sense of presence is assessed using the Presence Questionnaire (PQ) [20] and their emotional response is measured by the Positive Affect and Negative Affect Scale (PANAS) [21]. Additionally, the electrodermal activity (EDA) [22] signals and electrocardiogram (ECG) signals [23], commonly used as markers of autonomic arousal during mental stress, are employed as objective measures of emotional response.

2.2 Experimental Environment

The experimental virtual environment was modeled based on a subway station in Beijing, including both its underground and ground floors. By strategically placing obstacles within the environment, we crafted two evacuation scenarios. As shown in Figure 1, both scenarios share architectural similarities and feature evacuation routes of identical lengths. The fire originates from a train approaching the station. Upon receiving a station-wide announcement, participants waiting for the train are instructed to evacuate immediately to designated exits within the station. Simultaneously, devices simulating temperature and smell effects dynamically adjust their intensity based on participants' proximity to

the fire source. The starting point for participant evacuation is situated on the ground floor. Each scenario includes only one safe exit: participants in Scenario 1 must navigate to Exit B1, while those in Scenario 2 must proceed to Exit B2. Successfully reaching any of these exits from their respective starting positions is considered a successful evacuation.

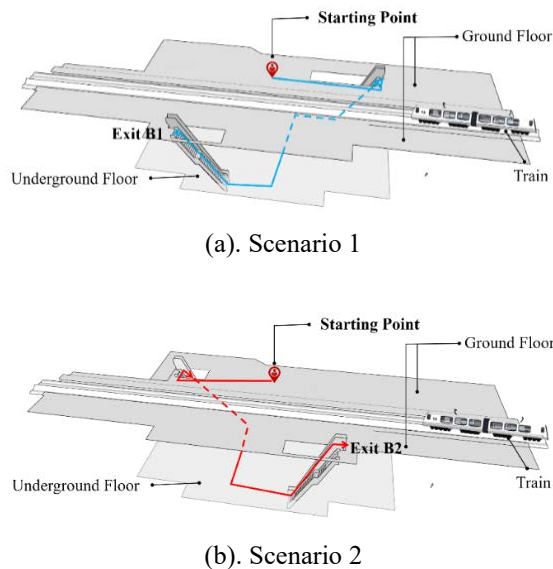


Figure 1. Virtual experimental scenes.

2.3 Experimental Equipment

A VR fire evacuation behavior experimental system based on multisensory human-computer interaction has been developed using the Unity3d engine. This VR system facilitates a multisensory interaction experience, incorporating visual, auditory, tactile, thermal, and olfactory sensations within an engaging IVE. The initial construction and rendering of the scene, including the firework models, were conducted using Revit software to create the fundamental structures. Subsequently, 3ds Max software was utilized to enhance the models through rendering before importing them into Unity3D software for the assembly of the scene prototype. To simulate the properties of virtual flames, such as their temperature and rate of spread, we used PyroSim and FDS, applying the resultant data directly to the flame models within Unity3D. In addition to the visual elements, the IVE was enriched with a diverse array of sound effects, ranging from evacuation broadcasts to the sounds of flames crackling. The use of VR HMD culminates in a deeply immersive experience for the participants, enhancing the realism and impact of the fire evacuation simulation.

To enhance the realism of the evacuation experiment, the system incorporates interactive devices that simulate the high temperatures and smells produced at a fire scene.

The purpose of the scent player is to mimic the irritating smells emitted at a fire scene. This is achieved through the use of a scent player (Figure 2) manufactured by the ScentrealmM Company [24]. The player is filled with non-toxic, side-effect-free scented materials available for selection, including smells such as burning plastic and burning paper. Depending on the participant's distance from the fire source in the virtual scene, the player can adjust the frequency of the scent released to simulate the irritating smell at a fire scene. The customized thermal interaction device consists of a heater and a lifting and rotating platform (Figure 3). The device controls the temperature changes in the participant's surrounding environment by adjusting the heater's settings and the deflection angle directly facing the participant. The lifting device can be adjusted according to the participant's height to optimize their experience. Real-time data transmission and feedback between the device and the scene are possible. In the virtual scene, as the participant approaches the fire source, the temperature interaction device automatically adjusts its settings and deflection angle based on the distance between the participant and the fire source, allowing the participant to feel varying degrees of warmth. As the distance between the participant and the fire source increases, the sensation of warmth noticeably decreases until it disappears. Before the experiment commenced, participants were positioned facing the thermal device at a distance of 0.9 meters.



Figure 2. Off-the-shelf scent player.



Figure 3. Customized thermal device.

2.4 Participants

A total of 36 participants, all of whom were Tsinghua University students, were recruited for this study. They had an average age of 21.34 years (S.D. = 3.531) and an equal gender distribution, with a 1:1 ratio of males to females. The participants were allocated into two distinct groups: both groups wore the device, and one utilized the scent player consistently throughout the entire session, while the others wore the scent player without it being activated. Every participant was subjected to both thermal-interactive and non-thermal-interactive evacuation scenarios, and whether they first engaged in the thermal-interactive or the non-thermal-interactive scenario was fully randomized. All individuals involved possessed normal visual capabilities and were free from any medical conditions that might deem them unfit for participation in VR-based experiments. Upon completion of the experiment, each participant was rewarded with a compensation of 60 RMB.

2.5 Experiment Process

This experiment proceeded in seven succinct steps: 1) Participants arrived at the lab, settled at a computer, and after a brief break, received and signed an informed consent form outlining the study's aims and data collection methods, along with a reassurance of their right to withdraw at any point without compensation. 2) Participants were briefed on the experiment, equipped with EDA and ECG BIOPAC devices, instructed on their use, and advised to minimize movement to reduce data artifacts. A one-minute pause allowed for the collection of baseline physiological data. 3) A pre-test questionnaire was conducted to gather demographic data, emotional states via the PANAS scale, simulator sickness via the simulator sickness questionnaire (SSQ) [25], wayfinding anxiety via Lawton's spatial anxiety scale (LSAS) [26], and sense of direction via the Santa Barbara sense of direction scale (SBSOD) [27]. 4) Participants proceeded to a training task, wearing VR helmets and practicing movement in a virtual empty room until they felt fully familiar with the controls. 5) Post-training, participants took a break, reviewed the instructions of the main experimental task detailing a virtual metro station scenario, and were informed that no questions would be answered once the main task began. They then re-entered the IVE to perform the main task, concluding this phase upon successful task completion. 6) Helmets off, participants completed a post-test questionnaire including the PANAS, SSQ, PQ, and questions to the utility of the thermal and smell devices. 7) Participants completed a second task by repeating step 5, followed by once again filling out the comprehensive post-test questionnaire from step 6.

3 Results

Data analysis was carried out using SPSS [28], with significance levels set at 0.1 for marginal, 0.05 for standard, and 0.01 for high significance. A normality assessment preceded all tests. The independent sample t-test evaluated mean differences between two groups for data following a normal distribution [29]. Conversely, for skewed data distributions, the Wilcoxon signed-rank test and the Mann-Whitney U test were employed for paired and independent samples, respectively [30]. No significant differences were observed between the two groups in terms of age, gender, educational background, wayfinding anxiety, or sense of direction, with all p-values exceeding 0.1.

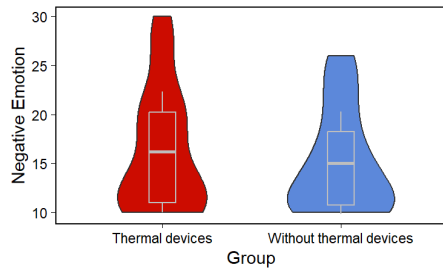
3.1 The Effect of Thermal Device

The results showed that thermal feedback did not significantly influence participants' sense of presence ($Z = 0.291$, $p > 0.1$). Regarding emotional response, no notable difference was observed in participants' positive emotions ($Z = 0.26$, $p > 0.1$). However, participants exposed to VR scenes with thermal feedback reported higher levels of negative emotion (Mean = 16.19, SD = 6.150) compared to those without the feedback (Mean = 15.06, SD = 5.226), with this increase being statistically significant ($Z = -2.948$, $p = 0.003 < 0.01$), as depicted in Figure 4(a).

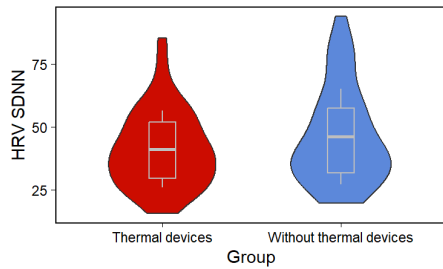
For physiological indicators, heart rate variability (HRV), specifically the standard deviation of NN intervals (SDNN), served as an ECG measure of emotional response, where a higher HRV SDNN indicates a more relaxed state [23]. Participants without thermal feedback exhibited a higher HRV SDNN (Mean = 41.13, SD = 15.275) compared to those with thermal feedback (Mean = 46.05, SD = 18.966), approaching marginal significance ($Z = -1.665$, $p = 0.096 < 0.1$), as illustrated in Figure 4(b).

Electrodermal activity (EDA), measured by the Tonic Standard Deviation (Tonic SD) [22], reflects emotional arousal, with lower EDA Tonic SD indicating a calmer state. Participants receiving thermal feedback exhibited a higher EDA Tonic SD (Mean = 41.13, SD = 15.275) compared to those without feedback (Mean = 46.05, SD = 18.966), with the difference being statistically significant ($Z = -1.665$, $p = 0.025 < 0.05$), as illustrated in Figure 4(c).

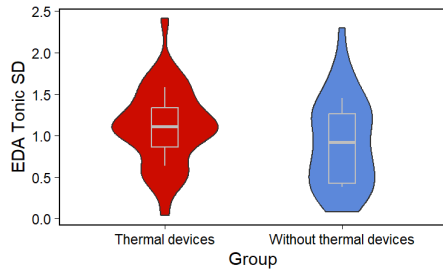
Regarding the perceived effectiveness of the thermal device (as shown in Table 1), significant differences were noted between the two groups regarding the sensation of temperature increase (measured by question 1 (Q1): $Z = -4.576$, $p = 0.000 < 0.01$) and the ability of the thermal device to alter temperature (measured by question 2 (Q2): $Z = -4.628$, $p = 0.000 < 0.01$).



(a). The effect of thermal feedback on participants' negative emotion



(b). The effect of thermal feedback on participants' HRV SDNN



(c). The effect of thermal feedback on participants' EDA Tonic SD

Figure 4. The effect of thermal devices on human performance.

Table 1. The perceived effectiveness of the thermal device.

Group	Mean	SD	Z	Sig.
Q1 (Thermal devices)	4.00	1.31	-4.576	0.000
Q1 (Without thermal devices)	2.06	1.12		
Q2 (Thermal devices)	3.25	1.23	-4.628	0.000
Q2 (Without thermal devices)	1.94	1.04		

3.2 The Effect of Scent Player

No substantial differences were observed between the two groups (with versus without a scent player) in terms of participants' sense of presence, negative and positive emotions, HRV SDNN, and EDA Tonic SD (all $p > 0.1$). However, a notable difference was found in the SSQ scores ($Z = -1.840$, $p = 0.066 < 0.1$), as depicted in Figure 5.

In terms of the effectiveness of the scent player (as shown in Table 1), the two groups exhibited significant difference in the sense of smell of paper burning (measured by question 3 (Q3): $Z = -2.349$, $p = 0.019 < 0.05$) and whether the device could produce the smell (measured by question 4 (Q4): $Z = -4.628$, $p = 0.018 < 0.05$).

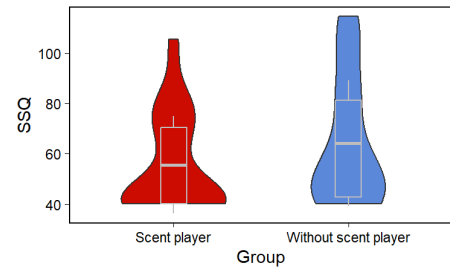


Figure 5. The effect of scent player on SSQ.

Table 2. The perceived effectiveness of the scent player.

Group	Mean	SD	Z	Sig.
Q3 (Scent player)	2.61	1.59	-2.349	0.019
Q3 (Without scent player)	1.78	1.02		
Q4 (Scent player)	2.53	1.75	-2.368	0.018
Q4 (Without scent player)	1.44	1.00		

4 Discussion and Conclusion

This study investigates the effects that a VR fire scenario, furnished with both a thermal device and a scent player, has on participants' sense of presence, emotional reactions, and perceptions of usefulness. The findings demonstrate that the thermal device significantly affected participants' negative emotional states, ECG, EDA, and directly influences their perceived effectiveness assessment of thermal devices. This highlights the ability of enhanced thermal sensations in a VR fire emergency scenario to significantly amplify participants' emotional responses. This finding is in line with previous research that showed a positive effect of thermal sensations on

human emotional responses in standard virtual built environments [15]. Being the largest perceptive organ in the human body, the skin's sensitivity to temperature variations plays a crucial role in shaping an individual's environmental perceptions [31]. Incorporating a heat source device within the virtual realm thus proves to be a potent trigger for negative emotional reactions.

Conversely, despite participants' ability to detect the release of potent scents from the scent disperser, this detection did not lead to changes in their sense of immersion, emotional responses, nor affected their perceived effectiveness of the scent player. Similar findings have also been echoed in several previous studies, highlighting that smell cues do not exert a substantial influence on the sense of presence [32], given that olfactory congruence with tactile and visual inputs does not sway the participants' immersive experience [33]. Both our study and preceding research grapple with the challenge of striking a balance between the efficacy of stimuli and ethical considerations. An excessively strong scent may elicit intense physiological responses, whereas a milder scent may fall short in bolstering the VR system's ecological validity. Identifying an optimal scent concentration presents itself as a potential avenue for future inquiry. Nevertheless, this research substantiates the efficacy and imperative of integrating temperature feedback within a virtual disaster setting, thereby contributing another layer of validation to the utility of virtual reality technology as a tool for studying human behavior in the context of architectural calamities. Lastly, it is also noteworthy that participants' past experiences in real-life fire emergencies may affect how they would react to thermal and scent stimuli in virtual experiments, and this factor could be further explored in future research.

Acknowledgments

The authors are grateful for the support of the National Natural Science Foundation of China (NSFC) under Grant Nos. 72374120 and 71603145. Any opinions, findings, conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agency.

References

- [1] Xia X. and Li N. and González V.A. Exploring the influence of emergency broadcasts on human evacuation behavior during building emergencies using virtual reality technology. *Journal of Computing in Civil Engineering*, 35(2):04020065, 2021.
- [2] Blomander A., Investigating the use of radiative heat panels to enhance perceived threat of fire in vr, LUTVDG/TVBB, 2020.
- [3] Li N. and Du J. and González V.A. and Chen J. Methodology for extended reality-enabled experimental research in construction engineering and management. *Journal of Construction Engineering and Management*, 148(10):04022106, 2022.
- [4] Bourhim E.L.M. and Cherkaoui A. Efficacy of virtual reality for studying people's pre-evacuation behavior under fire. *International Journal of Human-Computer Studies*, 142:102484, 2020.
- [5] Alam T.A. and Dibben N., A comparison of presence and emotion between immersive virtual reality and desktop displays for musical multimedia. In *Future Directions of Music Cognition 2021 Virtual Conference Proceedings*, Ohio State University Libraries, Ohio State, Amercia, 2021.
- [6] Chirico A. and Gaggioli A. When virtual feels real: Comparing emotional responses and presence in virtual and natural environments. *Cyberpsychology, Behavior and Social Networking*, 22(3):220-226, 2019.
- [7] Shaw E. and Roper T. and Nilsson T. and Lawson G. and Cobb S.V.G. and Miller D., The heat is on. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1-13, 2019.
- [8] Melo M. and Goncalves G. and Monteiro P. and Coelho H. and Vasconcelos-Raposo J. and Bessa M. Do multisensory stimuli benefit the virtual reality experience? A systematic review. *IEEE Transactions on Visualization and Computer Graphics*, 28(2):1428-1442, 2022.
- [9] Warp R. and Zhu M. and Kiprijanovska I. and Wiesler J. and Stafford S. and Mavridou I., Moved by sound: How head-tracked spatial audio affects autonomic emotional state and immersion-driven auditory orienting response in vr environments. In *Audio Engineering Society Convention 152*, Audio Engineering Society, 2022.
- [10] Pietra A. and Vazquez Rull M. and Etzi R. and Gallace A. and Scurati G.W. and Ferrise F. and Bordegoni M. Promoting eco-driving behavior through multisensory stimulation: A preliminary study on the use of visual and haptic feedback in a virtual reality driving simulator. *Virtual Reality*, 25(4):945-959, 2021.
- [11] Zou H. and Li N. and Cao L. Emotional response-based approach for assessing the sense of presence of subjects in virtual building evacuation studies. *Journal of Computing in Civil Engineering*, 31(5):04017028, 2017.

- [12] Perroud B. and Régnier S. and Kemeny A. and Mérienne F. Model of realism score for immersive vr systems. *Transportation Research Part F: Traffic Psychology and Behaviour*, 61:238-251, 2019.
- [13] Nilsson N.C. and Serafin S. and Steinicke F. and Nordahl R. Natural walking in virtual reality. *Computers in Entertainment*, 16(2):1-22, 2018.
- [14] Da Silveira A.C. and Rodrigues E.C. and Saleme E.B. and Covaci A. and Ghinea G. and Santos C.A.S. Thermal and wind devices for multisensory human-computer interaction: An overview. *Multimedia Tools and Applications*, 82(22):34485-34512, 2023.
- [15] Valente L. and Feijó B. and Ribeiro A. and Clua E. Pervasive virtuality in digital entertainment applications and its quality requirements. *Entertainment Computing*, 26:139-152, 2018.
- [16] Lawson G. and Shaw E. and Roper T. and Nilsson T. and Bajorunaite L. and Batool A. Immersive virtual worlds: Multi-sensory virtual environments for health and safety training. *arXiv preprint arXiv:1910.04697* 2019.
- [17] Andonova V. and Reinoso-Carvalho F. and Jimenez Ramirez M.A. and Carrasquilla D. Does multisensory stimulation with virtual reality (vr) and smell improve learning? An educational experience in recall and creativity. *Frontiers in Psychology*, 14:1176697, 2023.
- [18] Jones S. and Dawkins S., The sensorama revisited: Evaluating the application of multi-sensory input on the sense of presence in 360-degree immersive film in virtual reality, *Augmented reality and virtual reality*, 183-197, 2018.
- [19] Tuanquin N.M.B. and Hoermann S. and Petersen C.J. and Lindeman R.W., The effects of olfactory stimulation and active participation on food cravings in virtual reality. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 709-710, IEEE, 2018.
- [20] Witmer B.G. and Jerome C.J. and Singer M.J. The factor structure of the presence questionnaire. *Presence: Teleoperators & Virtual Environments*, 14(3):289-312, 2005.
- [21] Qiu L. and Zheng X. and Wang Y. Revision of the positive affect and negative affect scale. *Chinese Journal of Applied Psychology*, 14(3):249-254, 2008.
- [22] Mansi S.A. and Pigliautile I. and Arnesano M. and Pisello A.L. A novel methodology for human thermal comfort decoding via physiological signals measurement and analysis. *Building and Environment*, 222:109385, 2022.
- [23] Pereira T. and Almeida P.R. and Cunha J.P.S. and Aguiar A. Heart rate variability metrics for fine-grained stress level assessment. *Computer Methods and Programs in Biomedicine*, 148:71-80, 2017.
- [24] Scentrealm. On-line: <https://www.qiweiwangguo.com/>, Accessed: 17/08/2023.
- [25] Kennedy R.S. and Lane N.E. and Berbaum K.S. and Lilienthal M.G. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, 3(3):203-220, 1993.
- [26] Lawton C.A. and Kallai J. Gender differences in wayfinding strategies and anxiety about wayfinding: A cross-cultural comparison. *Sex roles*, 30(11-12):765-779, 1994.
- [27] Hegarty M. and Richardson A.E. and Montello D.R. and Lovelace K. and Subbiah I. Development of a self-report measure of environmental spatial ability. *Intelligence*, 30(5):425-447, 2002.
- [28] IBM, The ibm spss® software. On-line: <https://www.ibm.com/analytics/spss-statistics>, Accessed: 17/08/2023.
- [29] Winter J.C.F.d. Using the students t-test with extremely small sample sizes. *Practical Assessment, Research, and Evaluation*, 18(10):1-12, 2013.
- [30] Hoeffding W. A non-parametric test of independence. *The Collected Works of Wassily Hoeffding*, XIX(4):214-226, 1994.
- [31] Zimmerman A. and Bai L. and Ginty D.D. The gentle touch receptors of mammalian skin. *Science*, 346(6212):950-954, 2014.
- [32] Murray N. and Lee B. and Qiao Y. and Muntean G.-M. Olfaction-enhanced multimedia: A survey of application domains, displays, and research challenges. *ACM Computing Surveys (CSUR)*, 48:1-34, 2016.
- [33] Baus O. and Bouchard S. and Nolet K. Exposure to a pleasant odour may increase the sense of reality, but not the sense of presence or realism. *Behaviour & Information Technology*, 38(12):1369-1378, 2019.

Envisioning Extraterrestrial Construction and the Future Construction Workforce: A Collective Perspective

Amirhosein Jafari¹, Yimin Zhu¹, Suniti Karunatilake², Jennifer Qian³, Shinhee Jeong⁴, Ali Kazemian¹, and Andrew Webb⁵

¹ Bert S. Turner Department of Construction Management, Louisiana State University, USA

² Department of Geology & Geophysics, Louisiana State University, USA

³ Lutrill & Pearl Payne School of Education, Louisiana State University, USA

⁴ Department of Counseling, Leadership, Adult Education and School Psychology, Texas State University, USA

⁵ Division of Computer Science and Engineering, Louisiana State University, USA

ajafari1@lsu.edu, yiminzhu@lsu.edu, sunitiw@lsu.edu, jqian@lsu.edu, shinheejeong@txstate.edu,
kazemian1@lsu.edu, andrewwebb@lsu.edu

Abstract

Humanity's evolution toward an interplanetary species poses a new frontier for the construction industry: extraterrestrial construction. With fast technological advancements in manufacturing and robotics, it is a matter of when, not if, before humans make perpetual habitats on nearby planetary bodies. We envision the emerging frontier of the construction industry as extraterrestrial construction, where the role of workers and their required skills will change dramatically. Due to the extreme and hazardous outer-space conditions, the future of extraterrestrial construction will be technology-intensive, from using onsite machine and robot operations to using cyber-physical systems for managing logistical operations. Accordingly, the role and skillsets of construction workers in future extraterrestrial construction projects will contrast with the current practices on Earth. We aim to present a collective perspective on the nature of future extraterrestrial construction and the hierarchy of the skills required by future workers, as well as emerging technologies that can be used for developing a future-ready workforce. To accomplish this, we convened a national interdisciplinary workshop, engaging diverse stakeholders in the United States, including researchers, educators, and professionals, from academia, industry, and government. This paper summarizes the outcomes of our workshop, structured around three core themes: future work (extraterrestrial construction), future workers (extraterrestrial workforce), and future technology (emerging technologies for workforce training). The detailed exploration within these three themes marks an initial endeavor to chart a course for training in extraterrestrial construction, particularly with NASA's Moon to Mars Program in mind.

Keywords – Extraterrestrial Construction; Workforce Training; Construction Technology; Robotics.

1 Introduction

The construction industry is evolving at an unprecedented speed through advancements in materials, manufacturing, and technologies. In addition, in recognition of the 50th anniversary of the first manned lunar landing, the National Aeronautics and Space Administration (NASA), together with the European Space Agency (ESA), revealed plans to resume diversely crewed exploration missions and to establish permanent human habitats on the Moon and Mars by 2040 [1]. The vision of becoming an interplanetary species presents an unprecedented challenge to the construction industry: extraterrestrial construction. Extraterrestrial construction, constructing structures on a planetary body outside of the Earth and its atmosphere, will be a significant part of the future market for the construction industry [2].

We consider extraterrestrial construction as the theory and practice of building anthropogenic environments in outer space to enable temporary or perpetual human presence or activities [3]. Conventional construction processes, which require massive human-labor intervention, may be less suited for extraterrestrial construction due to near-vacuum conditions, meteors, radiation, and other extremes. Building a human habitat in uninhabitable environments on other planetary bodies will require automated construction processes optimized for that environment [4]. To promote feasible and sustainable space exploration, these habitats and other structures are envisioned to require in-situ construction [5], shaping a new generation of the construction

workforce. Regarding the construction process, NASA researchers believe robotics and additive manufacturing (AM) technologies have enormous potential to become a game-changer for space exploration. However, we will need appropriate construction technology to advance and simplify the complexity of the construction process in hostile and extreme extraterrestrial environments [6, 7]. To achieve such advancement and simplicity, NASA believes robotics and AM technologies require advancement beyond their current achievements [8]. Robotic and AM technologies can also transform future construction work on Earth [9]. However, the non-repetitive nature of construction tasks may result in semi-automated construction robots, requiring humans and robots to collaborate [10]. In this scenario, human workers must learn to adaptively perform construction work (i.e., improvise) based on evolving site conditions, supervisor input, the constraints of the co-worker robots, and accrued experience [11]. Therefore, advancement in construction technologies (both in near-future terrestrial construction and far-future extraterrestrial construction) may require new roles for human workers. Accordingly, the role of construction workers in future extraterrestrial construction projects will be significantly different from the current practice on Earth.

On the other hand, the current construction industry is facing enormous technological and institutional transformations with their resultant difficulties and challenges [12]. The lack of adequately trained personnel has been identified as a major hindrance to adopting and implementing emerging construction technologies [13]. This issue is becoming more significant due to the unprecedented speed of technological advancements in this industry [14]. Therefore, construction workers must deal with a rapid pace of technological change, a highly interconnected world, and complex problems that require multidisciplinary solutions. While it is obvious that training programs must maintain some likeness to the current practices of the construction industry, keeping the construction workforce development in line with future needs is an important challenge, given that the construction industry is poised for rapid transformation. With the advancement of construction technologies and the changing role of construction workers, a new set of skills is required to enable the future workforce, which necessitates closer scrutiny of the procedures for workforce development. The need for a future-ready workforce that can be adapted to meet the constantly evolving workplace has turned attention to the development and integration of transferable skills into training programs [15].

Industry 4.0 emerging technologies rooted in digitalization and artificial intelligence can enable this future construction workforce. In the wake of the fourth industrial revolution, highly advanced technologies have

made a paradigm shift in the construction workforce training environment. Traditional training methods (e.g., classroom lectures) insufficiently reflect the dynamism of the construction site and replicate the physical construction scenarios, making them ineffective and unengaging [16]. With the advent of computational technologies, such as artificial intelligence and mixed reality, some future learning concepts have been highlighted, overcoming the limitations of traditional approaches. For example, immersive learning, supported by digital media simulations, involves creating artificial environments where learners can experience and interact with the learning contents and lessons [17]. The immersive nature enables learners to grasp abstract concepts and gain hands-on experience in low-risk/secure virtual environments while feeling as though they are actually there (e.g., the construction sites), ideal for extraterrestrial construction workforce training [18]. Another example is adaptive intelligent learning, which uses automated machine learning technologies (e.g., artificial intelligence) to analyze and adapt to a learner's needs and progress; as a result, providing a personalized learning experience [19]. Additionally, this adaptiveness is ideal for an extraterrestrial workforce training environment where workers from diverse backgrounds are involved.

The goal of this study is to converge on the nature of future extraterrestrial construction and build a hierarchy of skills required by future workers, as well as identify emerging technologies that can help integrate transferable skills for developing a future-ready workforce from a diverse set of perspectives. We organized an interdisciplinary perspectives workshop that brought together the expertise and representatives of the different disciplines that will be involved in the envisioned extraterrestrial construction workforce development, including engineering, construction science, architecture, computer science, robotics, human-computer interaction, planetary geoscience, workforce development, learning sciences, and social science. The outcome of this workshop is presented in this paper, which can provide a collective perspective of extraterrestrial construction and the future construction workforce in terms of technological needs and challenges.

2 Methodology

One of the powerful tools for brainstorming about a subject is conducting workshops where varied perspectives can interact with each other to generate new knowledge [20-22]. In addition, Nielsen [23] highlighted that to enhance the effectiveness of brainstorming interaction in workshops, not only do group discussions need to be created, but also group members must be encouraged to cooperate with each other. In such an

environment, participants are assisted in revealing their thoughts and engaging in a social process of clarifying, developing, and refining ideas, which will increase the effectiveness of the workshop. Nevertheless, a successful brainstorming workshop should make the participants active learners, allowing them to be engaged instead of passive participants [23].

To create a perspective of visionary extraterrestrial construction and the future construction workforce, we used the brainstorming workshop technique as suggested by the literature [20-23]. A brainstorming workshop was used to identify the needs, challenges, and solutions for creating such a collective perspective from diverse viewpoints. The goal of this workshop was to converge on the nature of future extraterrestrial construction and build the hierarchy of skills required by future workers, as well as emerging technologies that can be utilized for developing a future-ready workforce from a diverse set of perspectives. The workshop was structured as a series of charrettes in which small and diverse groups brainstormed and designed to answer questions associated with each theme. The intended workshop activities include invited plenary talks, panel discussions, and breakout sessions on the three main themes. In addition, instructions for the moderators of these sessions were prepared to facilitate discussion and data collection.

3 Workshop Information

A two-day virtual interdisciplinary workshop, entitled “Roadmap to Future-ready Extraterrestrial Construction Workforce Development: Technological Visions and Needs,” was organized on February 20-21, 2023. The experts were meticulously chosen from a pool of distinguished scientists and researchers, meeting three criteria: 1) having authored at least one published paper pertinent to the themes of the workshop, 2) possessing a minimum of five years of experience in the relevant field, and 3) submitting a comprehensive vision statement outlining their perspective on the workshop’s topic. In conjunction with the organizational team, a cohort of 26 experts, each bringing unique backgrounds and expertise to the table, actively participated in the event. Table 1 shows the demographic information of the participants.

Table 1 Demographic information of the participants

Category	Level	No (Percent)
Organization	NASA-Affiliated	6 (23.1%)
	Technological/Educational	3 (11.5%)
	R1 Research Institution	14 (53.8%)
	R2 Research Institution	3 (11.5%)
Expertise	Extraterrestrial Construction	9 (34.6%)
	Human-Robot Collaboration	6 (23.1%)
	Human-Computer Interaction	4 (15.4%)
	Technology in Education	7 (26.9%)

The brainstorming workshop was structured around three core themes, as illustrated in Figure 1:

- Theme 1: Future work (Extraterrestrial Construction) to answer questions such as: How does the construction site environment on the Moon and Mars differ from Earth? What are the main characteristics of construction technologies used in extraterrestrial construction on the Moon and Mars? Is there any robot function in extraterrestrial construction that can’t/shouldn’t be fully automated?
- Theme 2: Future Workers (Extraterrestrial Construction Workforce) to answer questions such as What are the main anthropic tasks, roles, knowledge, and competencies/skills needed by the future civil workforce in extraterrestrial construction? To collaborate with robots effectively, what skills/knowledge are required for the future workforce? How the identified skills can be framed, categorized, or prioritized?
- Theme 3: Future Technology (Emerging Technologies for Workforce Training) to answer questions such as: What are the strategies and considerations in developing training scenarios in an AI-XR-based training environment for future workforce development? Among the skills and competencies identified for the extraterrestrial construction workforce, which ones are more suitable to be trained using an AI-XR-based training environment? How can the trained skills and competencies using an AI-XR-based training environment be evaluated in the context of extraterrestrial construction?

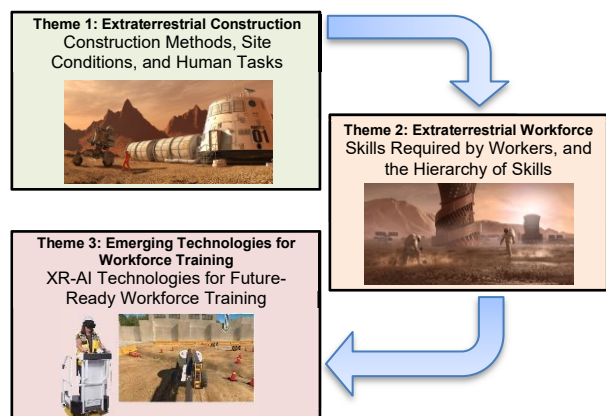


Figure 1. The workshop framework

The answers to these questions, as an outcome of the organized workshop and supported by the literature review, helped to understand how future technologies can help address the needs of future extraterrestrial construction workers in the context of their future work.

4 Summary of Outcomes

4.1 Theme 1: Extraterrestrial Construction

Extraterrestrial construction is an emergent area in which workers have no prior planetary experience. In this situation, planning and testing any idea for extraterrestrial construction is important to transform it into a process of exploration, analysis, discovery, design, implementation, and feedback. Although there are many uncertainties about future extraterrestrial construction, certainties include:

1. The differences between terrestrial and extraterrestrial site conditions should be clearly determined. For example, the intensity and frequency of atmospheric disturbances and quakes will differ across planets. Also, the harsh environment is a major factor that needs to be considered (e.g., radiation, gravity differences, vacuum, diurnal temperature range, abrasive Lunar regolith, variations in soil-regolith composition, and electrostatically adhering dust).
2. The differences between terrestrial and extraterrestrial construction protocols should be clearly identified, particularly in the design phase. In other words, we need to understand how conventional construction tasks change in planetary conditions. For example, in extraterrestrial construction, underground installations would be rare, and even small excavation forces may dislodge massive (i.e., high momentum) objects due to the lower gravity on Mars and the Moon. In addition, different expectations across the building lifecycle, such as construction, operations, and maintenance, may play an essential role during the design phase.
3. For extraterrestrial construction, In-situ Resource Utilization (ISRU) is a necessity, given the prohibitive costs of delivery from Earth. It is important to understand the available materials on different planets and be able to identify and predict those materials and resources from orbit. Also, the resiliency and longevity of construction materials and structures become vital factors in extraterrestrial construction, which are still poorly known and need to be researched.
4. Extraterrestrial construction will be complex in nature, with the involvement of numerous parties (experts in different disciplines from different countries). Therefore, it is important to develop a technical taxonomy to ensure consistent use of technological terms, derived SI units, and even labeling, which may otherwise lead to disasters – epitomized by the unit conversion error that destroyed the Mars Climate Orbiter mission; and the drogue parachute's failure to deploy in the OSIRIS-REx mission.
5. Although the specifications of robots for extraterrestrial construction may not be clear, it is evident that extraterrestrial construction will be robot-intensive. Although some level of autonomy is required for specific tasks (e.g., establishing landing pads for future landers to a base camp), the role of human workers cannot be overlooked in extraterrestrial construction. The human-in-the-loop approach is required due to:
 - a. Due to the general inability of machines to repair themselves or other equipment (even though the machines will be designed for easier maintenance or easier replacement of components), the role of human workers will still be essential.
 - b. There is greater efficiency when human workers guide or intervene throughout the construction process, i.e., the human workers can ensure that occasional malfunctions (defects and errors) of equipment do not become catastrophes. Human supervision would always be required even if the monitoring systems provided automated reports.
 - c. The machines' inability to replace human instinct or creativity - even automating knowledge exchange via generative large language models (e.g., ChatGPT, Bard) has limited applicability, massive energy demands, and cannot fully replace human knowledge oversight. For example, humans can use subtle visual changes in planetary landscape topography as immediate early warnings of structural weaknesses in construction settings.
6. Human-robot collaboration (HRC) will be a salient feature of extraterrestrial construction. Robots will continue to depend on human workers' ability to anticipate or preempt changes in response to dynamic construction conditions. Such collaborations will likely change in phases, mirroring how construction as a field evolved on Earth. HRC will be in all aspects of extraterrestrial construction, from automatically built landing platforms to enclosed structures using teleoperations to corporeal collaborations onsite for complex habitats.
7. To effectively integrate human-in-the-loop HRC in extraterrestrial construction, workers need the ability to remotely operate and troubleshoot equipment or processes. In such situations, teleoperation will be the desired and safe solution. However, teleoperation might have several technological limitations in extraterrestrial construction (e.g., latency, typically exceeding 20 min between Earth and Mars), which can significantly impact human behaviors, abilities, and work performance.

4.2 Theme 2: Extraterrestrial Construction Workforce

The future workforce in extraterrestrial construction will need to possess a range of knowledge, skills, and competencies to be successful in the new work environment to complete the challenging tasks. Extraterrestrial construction is a process of exploration with numerous uncertainties. Therefore, the future extraterrestrial construction workforce must be ready to expect the unexpected. Given the low population density anticipated for early habitats, workers may face responsibilities typically distributed across professions (e.g., a construction worker at one moment may perform surgical triaging the next). To enable such a workforce, the training should focus on what is certain to foster skills and cognition to handle uncertainties. This requires:

1. Situational Awareness, dating to NASA's first astronaut training, is crucial to make workers focus on their immediate tasks despite constantly monitoring for unrecognized site hazards.
2. Knowing the dramatic changes in geology (including atmospheres) and space environments to construct on the Moon and Mars safely and successfully, including but not limited to
 - a. Extreme thermal gradients and low external (confining) fluid pressure (e.g., thin Martian atmosphere and minimal atmosphere on the Moon) and how they affect the degree of mobility in wearing protective suits.
 - b. Mission operation times in relation to the planetary body's rotation period (e.g., local day) and how it can affect the workforce.
3. Understanding the associated microenvironment of individual workers in their suits. Given the human body's operational differences under microgravity conditions confined to a spacesuit, training in simulated conditions can provide dexterity.
4. Operating equipment remotely (teleoperation). Previous experience in the transition to teleoperation (e.g., Military transition from fighter pilots to drone operators) can be used for training such workforces. They will have to learn how to deal with several constraints in extraterrestrial construction, such as handling time delay (latency), working on Global Navigation Satellite System (GNSS)-unavailable areas, preventing robotic collisions or breakage, dealing with environmental factors (e.g., ambient temperature, dust, wind), and maintaining robots.
5. Training in Safety. Planetary differences may not necessarily cause collectively more hazards. However, isolated settings with effectively non-existing medical support will dramatically increase the risk of realized hazards.
6. Trusting in technology. The level of trust in technology systems depends on the characteristics of

the user, the technological system, and the environment. The dimensions of trust include explainability and interpretability, performance and robustness, reliability and safety, and privacy and security. Preventing credulity is critical, as made obvious in recent fatalities of drivers trusting Tesla's Level 2 autonomy (Society of Automotive Engineers scale) as Level 5.

7. Familiarity with the differences between hybrid collaboration (e.g., human-robot, human-robot-robot) versus human collaboration. Humans and robots may need to build structures as complementary partners, perhaps to the point of co-creating concepts. Therefore, the future workforce needs to understand the HRC and how to use those skills in new circumstances.
8. Compartmentalizing tasks to handle large amounts of information in stressful situations, especially in physically isolated situations.
9. IT Awareness to understand how to run and program robots during teleoperation.
10. Cross-disciplinary collaboration. While the idea of cross-disciplinary collaboration may seem like a cliché, it is still highly relevant and may become more important and prevalent in future work environments.
11. Understanding human-related factors, particularly in new and challenging work environments. Examples of human-related factors include fatigue, collaboration, boredom, isolation, and adaptability to the new work environment.
12. Cross-training at the intersection of mechatronics, electrical engineering, and programming. Skills will go outside the basic skills of a single domain. They will need to understand safety procedures, blueprint reading, programming (main requirement), processes and timing of moving robotic arms, and maintenance operations of such arms. They will need a basic understanding of how AI models are trained and how to employ machine learning algorithms.

4.3 Theme 3: Emerging Technologies for Workforce Training

The future extraterrestrial construction workforce should be trained using hands-on and real-world experiences. Practice in real-world environments is essential in training the future workforce; however, currently, extraterrestrial construction experiences are limited to orbital space stations. In this situation, simulations can provide the key immersive training required to address the differences in tasks (e.g., manipulating objects under low gravity, the limited set of available tools, and different lighting conditions) for extraterrestrial construction settings compared to Earth

counterparts. In this situation, AI-XR can provide immersive, interactive, and intelligent training environments that other educational technologies cannot match. Additionally, digital twin systems are an effective tool in such simulations. The simulation should also expose workers to the inevitable sensory deprivation when trapped in a space suit on a remote planetary site. Enabling the refinement of the framing with a feedback loop can ensure advancements from failures and lessons learned. Sensory degradation (such as visual ambiguity and limited perception of orientation extending to proprioception) and physiological impacts should be considered in such simulations. Such simulations should have the following features:

- **Realism.** The training environment should be as realistic as possible, representing extraterrestrial construction scenarios. But this does not suggest all training scenarios or their objects should be in their highest fidelity to reality. Use cases and technologies (e.g., HMD) define the level of fidelity that is needed and possible. The level of fidelity required depends on the goal of the simulation. However, the simulation of real extraterrestrial environments may be very complex, making the creation of learning scenarios using AI-XR technology computationally intensive and, as a result, expensive.
- **Interaction.** The training environment would be more effective if it were interactive. Interactive training fosters the development of essential skills by providing opportunities for individuals to collaborate, share ideas, and tackle challenges.
- **AI-enabled.** The use of AI to generate unforeseen scenarios since much of the planning and training focuses on what we expect to happen. AI-enabled training can bring in the worst-case and adaptable scenarios for training safety and situational awareness.
- **VR.** VR-based training enables more unique scenarios than other approaches. In a simulation, there is a scenario, and there is a consequence. Current training scenarios are very repeatable. Different variables of a simulation can be used to create a unique scenario for different trainees. This is where VR-based training becomes effective.
- **Physical and Emotional Responses.** Simulations need to immerse other senses beyond visual while creating both physical and emotional responses. Learning outcomes will be improved if trainees have a experience that evokes not only physical reactions but emotional ones. Effective use of visual, auditory, and haptic feedback will be necessary.
- **Decision-making protocol.** It is imperative to train the workforce to follow guidelines or procedures used to make decisions in extraterrestrial construction environments.

In addition, it is important to develop an internally consistent knowledge taxonomy for training purposes, which may help prevent unexpected skill failures. Furthermore, considering the complexity of extraterrestrial construction work, the needed skills and knowledge should be prioritized/classified by their importance, frequency of use, and practicality (hands-on experience) and broken down into small, digestible pieces for trainees.

An AI-XR training environment can be used to train several skills for enabling the future extraterrestrial construction workforce, including:

1. *Technical skills*
 - a. Construction procedures/protocols specific to extraterrestrial construction
 - b. Human-in-the-loop teleoperation for troubleshooting and repairing the robots.
 - c. HRC skills such as trust-related issues, collaboration/teamwork, and safety
 - d. Safety training
2. *Social skills*
 - a. The capacity to resolve interpersonal conflicts between workers without risking human lives. Such conflicts may disrupt the team's operations together and, in some cases, even cause the entire project to fail.
 - b. Cross-disciplinary collaboration in teleoperation to work with diverse people with different cultures.

To ensure successful training, two distinct layers of evaluation should be considered (1) evaluation of the action and (2) evaluation of the learning. They affect each other but are not the same learning objectives; thus, two different kinds of evaluations using different criteria are required. There are several evaluation tools that can be used to evaluate the action:

- **Performance metrics analysis.** A quantitative performance assessment to track progress over time and identify areas of improvement.
- **Data collection/analysis with AI.** Using AI to collect performance data and provide immediate and in-depth performance evaluation.
- **NASA task load index (TLX).** The NASA task load index is a tool for measuring and conducting a subjective mental workload assessment.

Also, the evaluation of the learning can be measured by assessing skill acquisition, skill retention, and performance satisfaction (i.e., usability and user experience metrics as well as skill-based performance metrics).

5 Conclusions and Takeaways

We found the following strategic knowledge gaps in each area:

Future extraterrestrial construction and the role of humans: Building a human habitat in extreme outer-space environments on other planetary bodies will require robotic construction processes optimized for that environment. Regarding the construction process, NASA researchers believe robotics and additive manufacturing technologies have enormous potential to enable space colonization. However, suitable and versatile construction technologies are needed to advance and simplify the complexity of the construction process in hostile and extreme outer-space environments in extraterrestrial construction. While the future of extraterrestrial construction remains uncertain, certain aspects are well-established, including (1) There are significant differences between terrestrial and extraterrestrial site conditions, environment, and climate, which must be determined clearly; (2) It is evident that extraterrestrial construction will be robot intensive. Although some level of autonomy is required for specific tasks, the role of human workers cannot be overlooked. Robots will continue to depend on human workers' ability to anticipate changes in response to dynamic construction conditions, making HRC a salient feature of extraterrestrial construction; and (3) To effectively integrate human-in-the-loop robot operations in extraterrestrial construction, future workers will need the ability to remotely operate and troubleshoot equipment or processes. In such situations, teleoperation is the desired and safe solution. However, teleoperation presently has several technological limitations in extraterrestrial construction.

Future construction workforce and competencies required: The lack of adequately trained personnel has been identified as a major hindrance to adopting and implementing emerging construction technologies. This issue is becoming more significant due to the unprecedented speed of technological advancements in this industry. In addition, the non-repetitive nature of construction tasks may result in semi-automated construction robots, which require humans and robots to work jointly. In this situation, human workers must learn how to perform construction work adaptively (i.e., improvise) based on evolving site conditions, supervisor input, and accrued experience. The future workforce in extraterrestrial construction will need to possess a range of knowledge, skills, and competencies to succeed in the new work environment to complete challenging tasks. Extraterrestrial construction is a process of exploration with numerous uncertainties. Therefore, the future extraterrestrial construction workforce must be ready to handle the unexpected with an acceptable level of success. To enable such a workforce, the training should focus on what is certain to foster skills and abilities to handle uncertainties. This requires (1) situational awareness as a necessity for the future extraterrestrial construction

workforce; (2) operating equipment remotely (teleoperation) while dealing with several constraints in extraterrestrial construction; (3) trust in technology, which will play an important role in the future extraterrestrial construction workforce. The level of trust in technology systems depends on the characteristics of the user, the technological system, and the environment; (4) familiarity with the differences between robots and humans, building structures as complementary partners, and how to use those competencies in new circumstances.

Future training environment for future-ready workforce development: While it is obvious that training programs must maintain some likeness to the current practices of the construction industry, keeping the construction workforce development in line with future needs is an important challenge, given that the construction industry is poised for rapid transformation. Traditional training methods (e.g., classrooms) cannot reflect the real dynamics of the construction site or replicate the actual construction scenarios. It is emphasized that the future extraterrestrial construction workforce should be trained using hands-on and real-world experiences. It requires a training transition from traditional behaviorist and cognitivist learning to learner-centered constructivist learning. Constructivist learning, as a contemporary learning paradigm, emphasizes the importance of an immersive and intelligent learning experience. Immersive learning, supported by digital media simulations, will create artificial environments where trainees experience and interact with real-world complexities. Its immersive nature will enable trainees to grasp abstract concepts, processes, and thinking, and gain hands-on experience in low-risk/secure virtual environments while feeling like they are actually there (e.g., the construction sites), ideal for harsh environments (e.g., extraterrestrial construction) workforce training. Adaptive intelligent learning uses AI technologies to analyze and adjust to learners' needs and progress, thus providing a personalized learning experience. It also will be an ideal feature for an extraterrestrial workforce training environment where workers from diverse backgrounds are involved.

6 Acknowledgment

We extend our sincere appreciation to all participants for their invaluable contributions during the brainstorming workshop, which greatly enriched our collaborative efforts. Their insights and feedback have been instrumental in shaping the success of this endeavor.

This research was supported by NSF Award No. 2222890. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation (NSF).

References

- [1] Naser, M.Z. (2019), "Extraterrestrial construction materials," *Progress in Materials Science*, Volume 105, August 2019, 100577.
- [2] King, C. (2019), "The Future of Extraterrestrial Construction," *Construction Executive (CE)* <https://constructionexec.com/article/the-future-of-extraterrestrial-construction>
- [3] Osburg, J., Adams, C.M., Sherwood, B. (2003), "A Mission Statement for Space Architecture," 33rd International Conference on Environmental Systems (ICES), Vancouver, British Columbia, Canada, 7–10 July 2003.
- [4] Lim, S. Prabhu, V.L., Anand, M., and Taylor, L. (2017), "Corrigendum to "Extraterrestrial construction processes – Advancements, opportunities and challenges," *Advances in Space Research*, 60 (2017) 1413–1429.
- [5] Khoshnevis, B., Carlson, A., Leach, N., and Thangavelu, M. (2012), "Contour Crafting Simulation Plan for Lunar Settlement Infrastructure Build-Up," NIAC Phase-I Final Project Report: https://www.nasa.gov/pdf/716069main_Khoshnevis_2011_PhI_Contour_Crafting.pdf
- [6] Isachenkov, M., Chugunov, S. Akhatov, I., and Shishkovsky, I. (2021), "Regolith-based additive manufacturing for sustainable development of lunar infrastructure – An overview," *Acta Astronautica*, Volume 180, 2021, Pages 650–678.
- [7] Giwa, I., Moore, D., Fiske, M., and Kazemian, A. (2022), "Planetary Construction 3D Printing Using Lunar and Martian In Situ Materials," the proceeding of Earth and Space (ASCE) 2022.
- [8] Metzger, T., Muscatello, A., Mueller, P., Mantovani, J. (2013), "Affordable, rapid bootstrapping of the space industry and solar system civilization," *Journal of Aerospace Engineering*, 26 (1), 18–29.
- [9] Delgado, J., Oyedele, L., Ajayi, A., Akanbi, L., Akinade, O., Bilal, M., and Owolabi, H. (2019), "Robotics and automated systems in construction: Understanding industry-specific challenges for adoption," *Journal of Building Engineering*, Volume 26, 2019, 100868.
- [10] Liang, C., Wang, X., Kamat, V., and Menassa, C. (2016), "Human-Robot Collaboration in Construction: Classification and Research Trends," *Journal of Construction Engineering and Management*, Volume 147, Issue 10 - October 2021.
- [11] Manyika, J., Chui, M., Miremadi, M., Bughin, J., George, K., Willmott, P., and Dewhurst, M. (2017), "A future that works: automation, employment, and productivity," McKinsey Global Institute, San Francisco,
- [12] Becerik-Gerber, B., Gerber, D., and Ku, K. (2011), "The pace of technological innovation in architecture, engineering, and construction education: integrating recent trends into the curricula," *Journal of Information Technology in Construction*, Vol. 16, pg. 411-432.
- [13] Sacks, R., and Barak R. (2010), "Teaching Building Information Modeling as an Integral Part of Freshman Year Civil Engineering Education," *Journal of Professional Issues in Engineering Education and Practice*, Volume 136, Issue 1.
- [14] Dainty, A., Cheng, M., and Moore, D. (2005), "Competency-Based Model for Predicting Construction Project Managers' Performance," *Journal of Management in Engineering*, Volume 21, Issue 1 - January 2005.
- [15] Muhamad, S. (2012), "Graduate Employability and Transferable Skills: A Review," *Advances in Natural and Applied Sciences*, 6(6): 882-885, 2012.
- [16] Jeelani, I., Han, K., and Albert, A. (2017), "Development of Immersive Personalized Training Environment for Construction Workers," *ASCE International Workshop on Computing in Civil Engineering 2017*, June 25–27, 2017 | Seattle, WA.
- [17] Tan, B., Cheng, G., Yu, R., and Cai, S. (2021), "Online Course Construction and Multi-Dimensional Mixed Teaching: Research and Practice," 16th International Conference on Computer Science & Education (ICCSE), (pp. 920-925), IEEE.
- [18] Palombi, T., Galli, F., Giancamilli, F. et al. (2023) The role of sense of presence in expressing cognitive abilities in a virtual reality task: an initial validation study. *Scientific Report* 13, 13396.
- [19] Samhoury, M. S., Idwan, S., and Mukattash, A. (2011), "A system for adaptive intelligent learning (SAIL)," *International Review on Modeling & Simulations*, 4(5), 2651-2656.
- [20] Woodson, T. S., Harsh, M., Bernstein, M. J., Cozzens, S., Wetmore, J., & Castillo, R. (2019). Teaching Community Engagement to Engineers via a Workshop Approach. *Journal of Professional Issues in Engineering Education and Practice*, 145(4).
- [21] Ozkaynak, M., Sircar, C. M., Frye, O., & Valdez, R. S. (2021). A systematic review of design workshops for health information technologies. *Informatics*, 8(2).
- [22] Parija, C., & Adkoli, B. V. (2020). Communication skills for organizing workshops. In *Effective Medical Communication: The A, B, C, D, E of it*.
- [23] Nielsen, M. F. (2012). Using artifacts in brainstorming sessions to secure participation and decouple sequentiality. *Discourse Studies*, 14(1).
Burger, M. (2007). Designing successful workshops for digital tools. *Reference Services Review*, 35(4).

Trends, Challenges, and Opportunities in Assistive and Robotic Kitchen Technologies for Aging Society: A Scoping Review

Rongbo Hu¹, Ami Ogawa², Borja García de Soto³ and Thomas Bock⁴

¹Kajima Technical Research Institute Singapore, Kajima Corporation, Singapore

²Department of System Design Engineering, Keio University, Japan

³S.M.A.R.T. Construction Research Group, Division of Engineering, New York University Abu Dhabi (NYUAD), Experimental Research Building, Saadiyat Island, P.O. Box 129188, Abu Dhabi, United Arab Emirates

⁴Chair of Building Realization and Robotics, Technical University of Munich, Germany (www.rod.de)
r.hu@kajima.com.sg, ami_ogawa@keio.jp, garcia.de.soto@nyu.edu, bockrobotics@web.de

Abstract –

The world's population is rapidly aging. According to the United Nations, nearly 10% of the world's population is over 65 years old. The aging population crisis is arguably one of the most pressing challenges to the future well-being of humanity. The kitchen is an essential place for human activities, and cooking holds significant importance for older adults, serving as a key measure of their capacity for independent living. Moreover, it plays a crucial role in ensuring that older individuals receive the necessary nutrition to maintain good health. Naturally, the investigation of assistive kitchen technologies emerges as a compelling and significant field of research. However, there are no existing reviews that specifically address assistive kitchen and cooking technologies designed for vulnerable populations such as older adults or individuals with disabilities. Therefore, this paper provides a scoping review of the trends, challenges, and opportunities in emerging assistive kitchen and cooking technologies for the past decade. The Preferred Reporting Items for Systematic Reviews and Meta-Analysis Extension for Scoping Reviews (PRISMA-ScR) methodology was adopted for the review. As a result, 100 peer-reviewed publications from 29 countries were included in this review. These studies cover six major research topics on assistive and robotic kitchens. As a result, the trends, challenges, and opportunities of the current technologies are analyzed, and a modular and adaptable smart kitchen is proposed. The study will fill the scientific gap and lay the groundwork for future development of assistive kitchen technology.

Keywords –

Activity Recognition; Artificial Intelligence; Disabilities; Cooking; Gerontechnology; Population Aging; PRISMA; Robotics; Sensing; Smart Home

1 Introduction

The world's population is aging at an unprecedented pace. According to the United Nations, as of 2022, there are nearly 8 billion individuals living on this planet, of which 9.6% (771 million) are over 65 years old. Globally, the population aged 65 years or over is the fastest-growing age group, whose proportion increased from 6.9% in 2000 to 9.3% (i.e., “aging society”) in 2020 and is projected to reach 15.9% (i.e., “aged society”) by 2050 and 22.6% (i.e., “super-aged society”) by 2100 [1, 2]. The population aging crisis is not only a severe crisis for the developed world, but also an imminent threat to emerging economies. It is arguably one of the most imminent challenges facing the future prosperity of humankind.

Kitchens are one of the places where humans spend the most time while awake. A new poll commissioned by Bosch Home Appliances suggests that the average American spends over an hour per day in the kitchen, adding to more than 400 hours a year [3]. Although the time that the current generation spends in the kitchen is much less than the older generation, potential pandemics like COVID-19 will only likely increase people's time in the kitchen again [4]. More importantly, cooking is a significant activity for older adults, not only because it is a crucial indicator of older adults' ability to maintain independent living, but also because appropriate cooking ensures the nutrition intake of older adults to maintain good health. In fact, previous studies even suggested that frequent cooking may improve the survival rates in older adults [5]. Over time, kitchens have evolved to be more convenient and technological. However, average

kitchens are not designed to be inclusive and adaptive, especially for the characteristics and needs of the growing aging population. As a result, the investigation of assistive kitchen technologies becomes a compelling and significant research area.

There have been a number of review studies on the impact of assistive technology on older people. For example, Shishehgar et al. examined how various robots can help older adults [6]. Zhu et al. analyzed the global trends in the study of smart healthcare systems for older adults with a special focus on artificial intelligence (AI) solutions [7]. Ghafurian et al. reviewed smart home devices for supporting older adults [8], while Nthubu reviewed sensors, design, and healthcare technologies in smart homes [9]. In addition, Facchinetti et al. analyzed how smart home technologies can help older adults manage their chronic conditions [10], and Ohneberg et al. identified existing assistive robotic applications in nursing settings in a scoping review [11]. Furthermore, Fasoli et al. made an in-depth online search on emerging technologies in aged care, which went beyond only the literature review [12]. It is worth noting that Singh et al. reviewed automated cooking machines and food service robots [13], but the study focused more on robotic solutions for the food service industry rather than for private homes and individuals.

The studies mentioned above provide insightful knowledge in regard to the impact of emerging technologies on older adults. However, reviews focusing on assistive kitchen and cooking technologies for vulnerable groups such as older adults and disabled individuals are non-existent. Therefore, in order to fill the scientific gap and lay the groundwork for the future development of assistive kitchen technology, this paper provides a scoping review of the trends, challenges, and opportunities in emerging assistive kitchen and cooking technologies for the aging population [14].

2 Research questions

The review aims to examine the following research questions (RQs):

- RQ1: What smart assistive kitchen and cooking technologies have been developed for older adults or individuals with disabilities in the past decade?
- RQ2: What are the focuses and trends of the reviewed studies on developing smart kitchen devices for older adults?
- RQ3: What are the challenges in the reviewed studies that need to be taken into account as potential considerations for developing smart kitchen technologies for supporting older adults and individuals with disabilities in the future?

3 Methodology

In order to conduct a thorough scoping review, the Preferred Reporting Items for Systematic Reviews and Meta-Analysis Extension for Scoping Reviews (PRISMA-ScR) methodology was adopted [14]. The details of the application of the PRISMA-ScR methodology, including the selection criteria, publication search strategy, and data screening process, are reported as follows.

3.1 Study selection and eligibility criteria

Peer-reviewed articles published in journals, books, and conference proceedings were included. When a conference proceeding and a journal article are based on the same research conducted by the same team, only the journal articles are kept.

The main reason for the defined time period for selecting the publications (i.e., articles that were published in the past 10 years) is that earlier studies may contain technologies that are already outdated or obsolete. In fact, even some early studies included in this research already have this problem. For example, the use of Adobe Flash is noticed in Blasco et al. [15], which has been an obsolete technology for many years [16].

Sociological articles that do not involve any emerging assistive technologies or novel designs are not included in this study since they are outside the scope of this study.

Articles that are written in languages other than English are excluded altogether due to the limited language skills of the authors as well as the pursuit of impartiality.

Articles that are not relevant to the RQs are also excluded. This could be papers featuring technologies that are meant to be applied outside of the kitchen or cooking context, such as the living room or bedroom.

Articles that are not yet fully peer-reviewed, such as preprints that are under review at the time of the analysis, are also excluded to achieve high reliability.

Articles with target groups that do not include older adults or individuals with disabilities are not included. For instance, several identified studies only focus on adolescents or children whose physiological and psychological characteristics are vastly different from older adults.

Table 1 summarizes the inclusion and exclusion criteria of this review.

Table 1 Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
<ul style="list-style-type: none"> Peer-reviewed articles published in journal and conference proceedings Articles published in the past 10 years (from 2013 to 2023) 	<ul style="list-style-type: none"> Articles with no emerging technologies or novel designs involved Articles that are written in languages other than English Articles that are essentially different versions of the same research Articles that are not relevant to the RQs Articles that are not yet peer-reviewed Articles' target users do not include older adults or individuals with disabilities

3.2 Databases for publication search

Popular research databases for science and engineering, including Scopus, Web of Science, and IEEE Xplore were utilized, which should cover the vast majority of the relevant literature. Google Scholar was not considered due to its limitations, such as limited functionality, opacity, and inaccuracy [17]. Keio University Library and the Technical University of Munich Library synergistically provided a full-access database for the literature items.

The search protocol was defined based on the RQs as well as the inclusion and exclusion criteria. The following search query was adapted and applied to each database:

(TITLE-ABS-KEY (kitchen OR cooking) AND TITLE-ABS-KEY (sensor OR sensing OR smart OR intelligent OR assistive OR gerontechnology) AND TITLE-ABS-KEY (old OR elderly OR older OR aging OR ageing OR senior OR disabled OR disability OR impairment)) AND PUBYEAR > 2012 AND PUBYEAR < 2024.

The final search was performed on August 22, 2023, covering all publications that have been included in these databases until that date.

3.3 Data screening process

In total, 794 items were identified in the initial data search. After removing duplicate records, 576 unique records were kept for abstract screening. In this step, a total of 433 records were excluded due to their irrelevance. Next, the records were sought for retrieval,

and the full text of 10 reports could not be retrieved. Finally, the full texts of the remaining 133 records were fully assessed, and 33 publications were further excluded based on the exclusion criteria. This resulted in 100 publications eventually being included in the scoping review.

Figure 1 shows the identification and screening processes of this study based on the PRISMA flowchart.

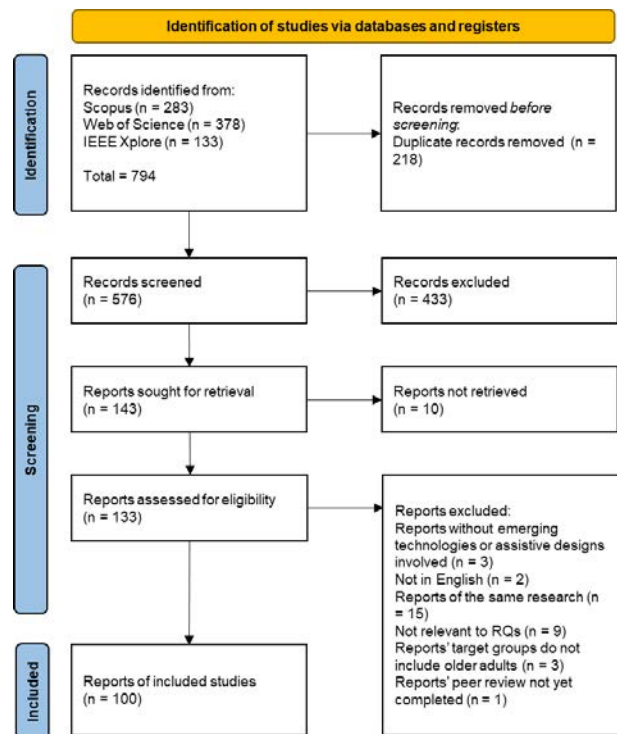


Figure 1. The PRISMA flowchart for this study

4 Results

In this section, the collective characteristics of the included studies are analyzed and visualized as follows.

4.1 Sources of the publications

The study finally included 100 publications for the scoping review, including 26 scholarly journal articles and 74 book chapters or conference proceedings, which suggests a good timeliness of the included studies (Figure 2).

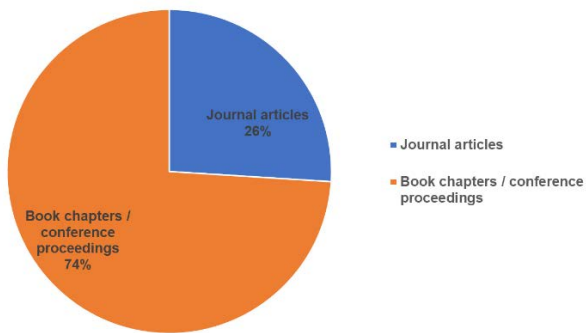


Figure 2. Sources of the publications

4.2 High-frequency keywords

A word cloud based on the titles of the included studies was generated using Free Word Cloud Generator [18] (Figure 3). The larger the word size, the more frequently these words appear in the titles. The word cloud largely goes in line with the data search protocol with a few exceptions such as activity recognition, which turns out to be one of the major topics in the assistive smart kitchen research but was not expected in the data search process.



Figure 3. Word cloud representing the frequency of the words used in the titles of the analyzed publications

4.3 Countries or regions of authors

The countries or regions of the authors were counted. Authors who contributed to multiple articles were counted multiple times. It can be observed that Canada is the strongest country in terms of smart kitchen research, followed by the UK, Italy, Germany, Japan, USA, Greater China, India, Greece, and France. The top 10 countries on the list indicate a combination of advanced countries in gerontechnology and countries with great culinary cultures (Figure 4).

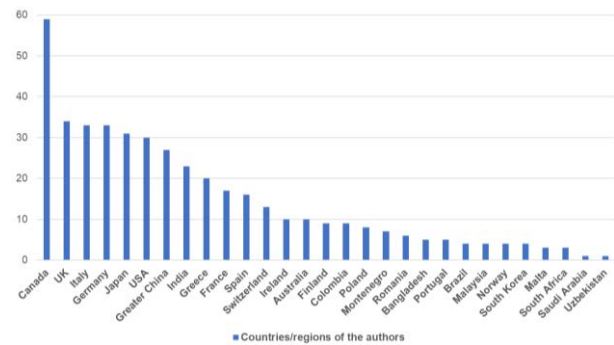


Figure 4. Countries or regions of the authors

4.4 Years of publications

Figure 5 shows the number of analyzed publications per year in the past 10 years. The year 2023 was not included since only publications until August 2023 were available when performing the data search. There is a general upward trend in the number of publications each year in the past 10 years, which indicates that the topic of smart and assistive kitchen technology is gaining popularity.

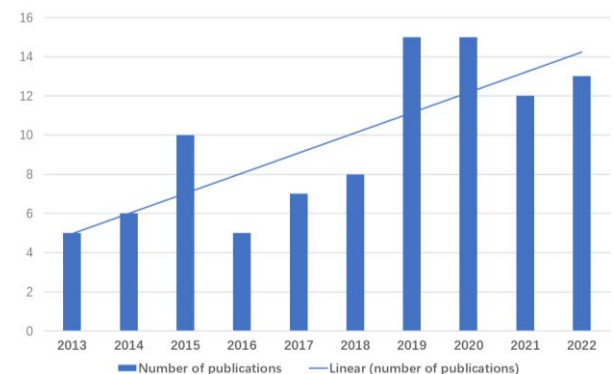


Figure 5. Numbers of publications per year from 2013 to 2022

4.5 Categories of study purposes

Regarding the functionality of the studies, the plurality of analyzed studies ($n = 28$) focus on human activity recognition in the home environment, including the kitchen. Such studies are often not about assistive kitchen technologies per se, but these endeavors provide unique insights into older adults' activities in the kitchen, which may serve as a guide for future research and development of assistive kitchen technology. The second largest category of studies is kitchen assistance, followed by safety monitoring, cooking instruction, and inclusive design. Only one study is about improving comfort level (i.e., room temperature) in the kitchen environment (Figure 6).

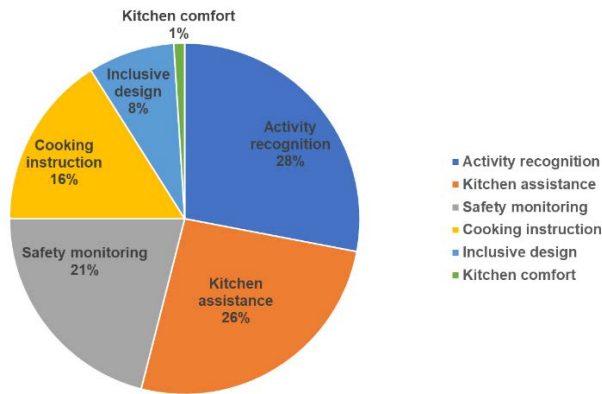


Figure 6. Categories of study purposes

4.6 Technologies applied

Based on the comprehensive analysis of the selected studies, the following categories of technologies were utilized:

- 1) Stationary / ambulant sensors
 - Cameras (e.g., RGB cameras, depth cameras, infrared cameras, etc.)
 - Audio sensors
 - Temperature sensors
 - Gas and smoke sensors
 - Vibration sensors
 - Load sensors
 - Wireless sensor nodes
- 2) Wearable sensors
 - Biometric sensors
 - Motion sensors
 - Smart glasses
- 3) Robotics
 - Mobile service robots
 - Robotic arms
 - Humanoid robots
- 4) Extended reality
 - Augmented reality
 - Mixed reality
- 5) Interactive user interface
 - Mobile apps
 - Websites
 - Physical buttons
 - Optical scanners
 - Ambient lighting
 - Smart speakers
- 6) Ergonomics
 - Inclusive design
 - Mechatronics

Many studies utilized more than one category of technologies. On top of these technologies, a majority of the selected studies were powered by AI solutions, including machine learning and neural networks, to realize the target smart functions (Figure 7).

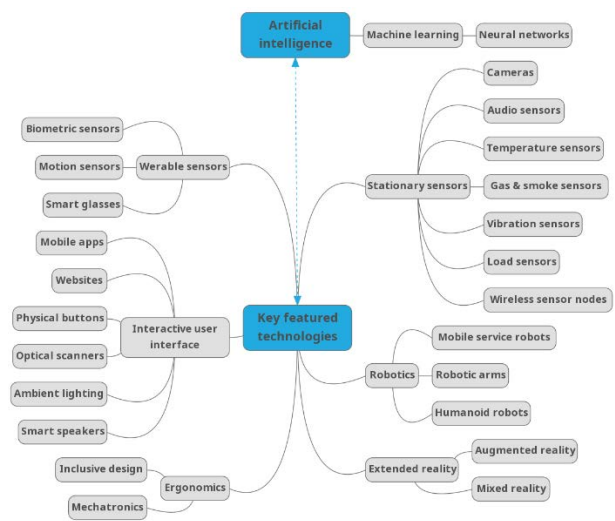


Figure 7. Mapping the key featured technologies applied in the studies (created by MindMup 2)

4.7 Study design

In the total examined studies, 64 employed at least one study design with voluntary test participants, while 36 proposed systems were only tested by the authors and researchers with no test participants as end users (Figure 8). This is relevant because one of the main aims of this paper is to reveal research trends in smart assistive kitchen technologies rather than the soundness of these systems.

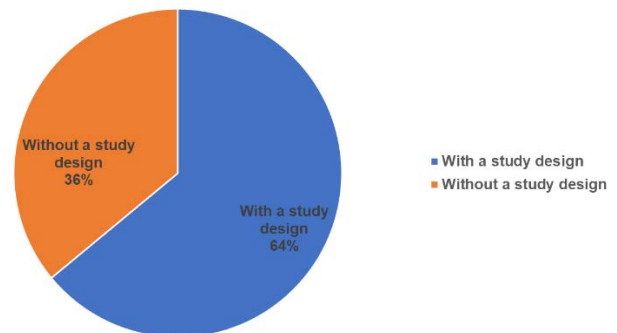


Figure 8. Percentage of publications with or without a study design

5 Discussion

In this section, the trends, challenges, and opportunities in Assistive and Robotic Kitchen Technologies for older adults and people with disabilities will be discussed, which answers the RQs of this literature review.

5.1 Trends

As defined in the inclusion and exclusion criteria, the target users of all selected studies included older adults, with the largest number of studies targeting older adults in general. Other noticeable target users include older adults who live alone, older adults with cognitive decline (e.g., dementia), and older adults with physical disabilities. In short, the vast majority of systems attempt to improve the quality of life for a wide variety of seniors by allowing them to use the kitchen and cook more independently.

A kitchen seems to be an ideal environment to create a dataset for machine learning of human activity recognition due to the complexity and uniqueness of human activities that occur in that space. A large number of studies focus on human activity recognition in kitchens, which may not directly assist older adults in cooking but will be helpful in understanding the behavior patterns of older adults in the kitchen and may serve as prerequisites to develop assistive products to help older adults cook.

Many studies focus on using external add-on technologies to assist existing kitchen facilities rather than redesigning and redeveloping kitchens to assist older adults, with a few exceptions.

Stationary sensors, which in some cases can be regarded as ambulant sensors (e.g., sensors embedded in wheelchairs and kitchen utensils), are the most commonly used technologies in the analyzed studies. They were chosen due to various reasons such as low cost and non-intrusiveness.

On the contrary, wearable sensors are only employed by a few studies ($n = 3$) due to their limitations, such as intrusiveness and limited battery life.

Robotics is another important technology that many studies employed ($n = 13$). Robotics has many advantages, such as good flexibility, high efficiency, and high mobility, but it also has obvious disadvantages, such as high complexity and high cost, which leads to a low level of willingness to adopt among older adults. In this study, those robotic solutions were only applied by authors from developed countries. However, it is foreseeable that as robotics becomes more widespread and the price of robots becomes more affordable, more assistive smart kitchens will integrate this technology.

Extended reality is also a promising category of assistive kitchen technologies ($n = 4$). However, it seems that only augmented reality and mixed reality are useful in the kitchen context. Virtual reality was not employed by any of the selected studies as its immersive experience isolates its users from the real environment.

Interactive user interfaces (e.g., graphical user interfaces, physical buttons, optical scanners, ambient lighting, etc.) are another noteworthy area of research because they are typically inexpensive, tangible, intuitive, highly versatile, and can be seamlessly embedded into

inhabited environments, thus creating a technological environment that is appropriate for the characteristics of older adults.

The technologies proposed, especially in developing countries, are primarily focusing on practical solutions with local characteristics. Because cooking is a highly regional activity in different countries, many technologies developed may not be widely applicable or demanded in other countries, such as smart LPG cylinders for India and smart rice cookers for China.

5.2 Challenges

Reliability and usefulness: As mentioned in Section 4.7, less than two-thirds of the studies employed a study design with test participants. As a result, a major issue is that many proposed technologies were tested only by the authors or engineers, or in some cases by students or healthy volunteers, not by real target users of these studies who are older adults or people with disabilities. Therefore, the reliability and usefulness of these developed technologies for their actual target groups were largely unknown; thus, their acceptance could not be guaranteed.

Accuracy: The accuracy (i.e., the “intelligence”) of many AI-powered monitoring systems needs to be improved. For example, the fire detection system proposed by Mukhiddinov et al. can sometimes falsely identify sunsets, sunshine, lighting, and electric lamps as fires [19].

Compatibility: In some cases, robotic solutions become highly complex and difficult to achieve due to their incompatibility with the environment and target objects. For example, in the study conducted by Odabasi et al. [20], the detection rate of the water bottles is low due to their transparency, which indicates that the concept of robot-oriented design (ROD) should be implemented already in the process of configuring the compliant environment to reduce the difficulties in robot development [21].

Usability: Ease of use is one of the most important attributes of these kitchen technologies for older adults. In many studies, older adults complained that the proposed technologies were too difficult to use. The simpler the operability of the proposed systems, the easier they will be accepted by older adults.

Affordability: Affordability is a main concern not only for emerging economies but also for developed countries as well. Therefore, improving the affordability of the technologies is another key to increasing user acceptance.

5.3 Opportunities

According to the findings of this literature review, some of the studies proposed adaptable designs, but they

were centered around cabinets rather than the cooking area. When people age, their physical characteristics will also undergo a series of changes, especially their height and posture. Some people may even rely on wheelchairs. Meanwhile, older adults and young people with different physical characteristics may share a kitchen. However, the ergonomics of these existing assistive kitchen designs optimized for the physical characteristics of older adults and individuals with disabilities and scenarios of kitchens shared by people with drastically different physical characteristics were seldom considered.

Therefore, modular smart kitchens that can be adapted to the physical characteristics of older adults and individuals with disabilities (e.g., people using wheelchairs) may be a promising future research direction. As a follow-up of this study, the authors aim to propose a modular adaptable kitchen to improve the quality of life for older adults and people with physical impairment.

The modular adaptable smart kitchen features key functions such as adaptable modular lifting cupboards, an interactive media center, ambient cueing signs, an Azure Kinect motion capture sensor, barrier-free lifting operating platform for cooking. A preliminary design of a modular and adaptable kitchen system is shown in Figure 9.

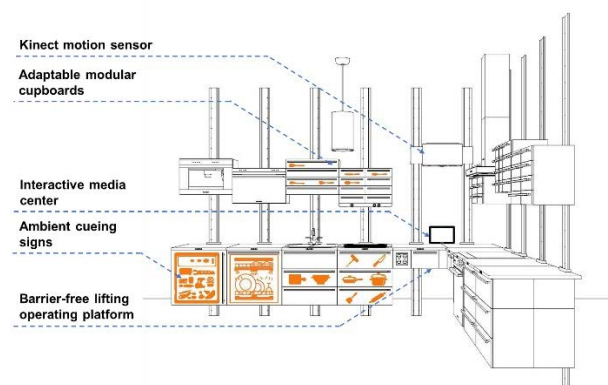


Figure 9. Preliminary design of the modular adaptable smart kitchen proposed by the authors' team

Currently, a low-fidelity prototype of the proposed kitchen is being built in the Architecture-Human Interaction System Lab at Keio University. Furthermore, approximately 10 student volunteers will be recruited from Keio University to carry out an initial usability test as a proof of concept. After the initial test, high-fidelity prototypes will be built and tested by more older adults in a real-world setting as a key step toward the marketization of the proposed modular and adaptable smart kitchen system.

5.4 Limitations of the study

There are several limitations in this study. The analyzed items are only in English, which is potentially biased against many studies in non-English-speaking countries. In the next step, research items in native languages in non-English speaking gerontechnological stronghold countries such as Italy, Germany, and Japan need to be further examined.

In this study, only research articles, rather than products on the market, are analyzed. Many research articles are experimental, and are far from marketization. Patents were not considered, either.

Also, the quality of the included studies is uneven because the sources are inclusive of journals, book chapters, and conferences, which, on the other hand indicates high timeliness and new trends.

In addition, due to page limitations, each reference to the included studies, as well as many details, cannot be fully revealed in this paper. However, an extended report of this research will soon be revealed in detail in an upcoming publication in a scholarly journal.

5.5 Conclusion

In this paper, a scoping review on assistive kitchen and cooking technologies for aging society in the past decade was performed following the PRISMA guideline. By analyzing the included publications, insights regarding the trends, challenges, and opportunities of these technologies were discussed. This work fills the scientific gap and lays the groundwork for future development of assistive kitchen technology. The paper will be a trusted reference for students, researchers, and practicing engineers who are interested in developing assistive kitchen and cooking technologies for vulnerable groups and beyond.

Acknowledgment

The study is funded by the Japan Society for the Promotion of Science (Grant ID No. PE23703). The authors would like to thank Mr. Yushun Zhuo from the Technical University of Munich for his support in the data retrieving process. The paper was written during the time when the co-author Rongbo Hu was at Department of System Design Engineering, Keio University, Japan.

References

- [1] Kim, K. W., & Kim, O. S. Super Aging in South Korea Unstoppable but Mitigatable: A Sub-National Scale Population Projection for Best Policy Planning. *Spatial Demography*, 8(2), 155–173, 2020. <https://doi.org/10.1007/s40980-020-00061-8>

- [2] United Nations. World Population Prospects, 2022. On-line: <https://population.un.org/wpp/Publications/>, Accessed: 21/11/2023.
- [3] Kitanovska, S. Americans Feel Most at Home in Their Kitchens, Survey Reveals. Newsweek. On-line: <https://www.newsweek.com/americans-feel-most-home-their-kitchens-survey-reveals-1730629>, Accessed: 21/11/2023.
- [4] Francis, G. Average person spends “half as much time” cooking as parents’ generation, poll claims. Independent. On-line: <https://www.independent.co.uk/life-style/home-cooking-meal-time-kitchen-microwave-parents-a9361236.html>, Accessed: 21/11/2023.
- [5] Chen, R. C.-Y., Lee, M.-S., Chang, Y.-H., & Wahlqvist, M. L. Cooking frequency may enhance survival in Taiwanese elderly. *Public Health Nutrition*, 15(7), 1142–1149, 2012. <https://doi.org/10.1017/S136898001200136X>
- [6] Shishehgar, M., Kerr, D., & Blake, J. A systematic review of research into how robotic technology can help older people. *Smart Health*, 7–8, 1–18, 2018. <https://doi.org/https://doi.org/10.1016/j.smhl.2018.03.002>
- [7] Zhu, Y., Yang, Q., & Mao, X. Global Trends in the Study of Smart Healthcare Systems for the Elderly: Artificial Intelligence Solutions. *International Journal of Computational Intelligence Systems*, 16(1), 105, 2023. <https://doi.org/10.1007/s44196-023-00283-w>
- [8] Ghafurian, M., Wang, K., Dhode, I., Kapoor, M., Morita, P. P., & Dautenhahn, K. Smart Home Devices for Supporting Older Adults: A Systematic Review. *IEEE Access*, 11, 47137–47158, 2023. <https://doi.org/10.1109/ACCESS.2023.3266647>
- [9] Nthubu, B. An Overview of Sensors, Design and Healthcare Challenges in Smart Homes: Future Design Questions. *Healthcare*, Vol. 9, Issue 10, 2021. <https://doi.org/10.3390/healthcare9101329>
- [10] Facchinetti, G., Petrucci, G., Albanesi, B., De Marinis, M. G., & Piredda, M. Can Smart Home Technologies Help Older Adults Manage Their Chronic Condition? A Systematic Literature Review. *International Journal of Environmental Research and Public Health*, Vol. 20, Issue 2, 2023. <https://doi.org/10.3390/ijerph20021205>
- [11] Ohneberg, C., Stöbich, N., Warmbein, A., Rathgeber, I., Mehler-Klamt, A. C., Fischer, U., & Eberl, I. Assistive robotic systems in nursing care: a scoping review. *BMC Nursing*, 22(1), 72, 2023. <https://doi.org/10.1186/s12912-023-01230-y>
- [12] Fasoli, A., Beretta, G., Pravettoni, G., & Sanchini, V. Mapping emerging technologies in aged care: results from an in-depth online research. *BMC Health Services Research*, 23(1), 528, 2023. <https://doi.org/10.1186/s12913-023-09513-5>
- [13] Singh, A., Chavan, A., Kariwall, V., & Sharma, C. A systematic review of automated cooking machines and foodservice robots. In *Proceedings of 2021 International Conference on Communication Information and Computing Technology (ICCICT)*, 1–6, 2021. <https://doi.org/10.1109/ICCICT50803.2021.9510121>
- [14] Peters, M., Godfrey, C., McInerney, P., Khalil, H., Larsen, P., Marnie, C., Pollock, D., Tricco, A. C., & Munn, Z. Best practice guidance and reporting items for the development of scoping review protocols. *JBIC Evidence Synthesis*, 20(4), 2022. <https://doi.org/10.11124/JBIES-21-00242>
- [15] Blasco, R., Marco, Á., Casas, R., Cirujano, D., & Picking, R. A Smart Kitchen for Ambient Assisted Living. *Sensors*, Vol. 14, Issue 1, pp. 1629–1653, 2014. <https://doi.org/10.3390/s140101629>
- [16] Adobe. Adobe Flash Player EOL Enterprise Information Page. On-line: <https://www.adobe.com/products/flashplayer/enterprise-end-of-life.html>, Accessed: 21/11/2023.
- [17] Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & Delgado López-Cózar, E. Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *Journal of Informetrics*, 12(4), 1160–1177, 2018. <https://doi.org/https://doi.org/10.1016/j.joi.2018.09.002>
- [18] Free World Cloud Generator. On-line: <https://www.freewordcloudgenerator.com/>, Accessed: 21/11/2023.
- [19] Mukhiddinov, M., Abdusalomov, A. B., & Cho, J. Automatic Fire Detection and Notification System Based on Improved YOLOv4 for the Blind and Visually Impaired. *Sensors*, Vol. 22, Issue 9, 2022. <https://doi.org/10.3390/s22093307>
- [20] Odabasi, C., Graf, F., Lindermayr, J., Patel, M., Baumgarten, S. D., & Graf, B. Refilling Water Bottles in Elderly Care Homes With the Help of a Safe Service Robot. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 101–110, 2022. <https://doi.org/10.1109/HRI53351.2022.9889391>
- [21] Bock, T. Robot-Oriented Design. In R. Ishikawa (Ed.), In *Proceedings of the 5th International Symposium on Automation and Robotics in Construction (ISARC)*, pp. 135–144, Tokyo, Japan, 1988. <https://doi.org/10.22260/ISARC1988/0019>

Guiding Visual Attention in Remote Operation: Meaning, Task and Object in Post-Disaster Scenarios

Xiaoxuan Zhou¹ and Xing Su¹

¹College of Civil Engineering and Architecture, Zhejiang University, Hangzhou, Zhejiang, 310000
xiaoxuan@zju.edu.cn, xsu@zju.edu.cn

Abstract –

In post-disaster scenarios where on-site operations are unfeasible, remote operation of robots or drones by human operators presents an effective and promising solution for survey and search-and-rescue (SAR) missions. These critical missions require human operators to rapidly process extensive visual data and allocate attention within dynamic, complex, and hazardous environments. While previous research has largely focused on the influence of salience and meaning in routine environments, our study shifts this focus towards the unique challenges faced by human operators in emergency settings. Utilizing eye-tracking technology and constructing feature maps of the environment, this research quantitatively assesses the roles of salience, meaning, task demands, and object relevance from the perspective of human operators in post-disaster environments, exploring their interrelationships. Our findings reveal that task demands and object relevance significantly affect how human operators allocate their visual attention, with this influence being modulated by factors such as salience and meaning, while meaning continues to play a predominant role in guiding attention. This study advances our understanding of the visual attention dynamics of human operators in critical SAR missions, providing essential insights for the design of more effective remote operation systems for emergency response.

Keywords –

Visual Attention; Remote Operation; Disaster Management; Eye Tracking

1 Introduction

The increase in severe natural disasters calls for innovative approaches in emergency management[1,2]. Remote search-and-rescue (SAR) has been propelled to the forefront of effective emergency response with the advances in robotics and unmanned aerial vehicles (UAVs), offering the ability to access areas otherwise

unreachable or dangerous by human responders[3]. Key players in this process are operators who pilot drones or robots in disaster zones. These operators manage equipment and process extensive visual data from the field to guide SAR efforts. The challenge encompasses more than technology operation; it involves analyzing and prioritizing received continuous stimulus. Their effectiveness depends on efficiently allocating attention and identifying crucial information to the visual feed, requiring complex attention skills.

Research into operators' visual attention can elucidate patterns of attention allocation, enabling predictions about focus distribution and, based on these insights, provide tailored support to enhance their operational performance. The study of attention has acknowledged that attention is a limited resource and is influenced by multiple factors[4,5]. Visual attention, as a specific subset of attention research, has been extensively explored, offering valuable frameworks for understanding how individuals process visual stimuli[6]. The framework is divided into two main directions, i.e., visual attention in a scene is driven by bottom-up, low-level image features, such as color, luminance, object feature, and edge orientation, which are combined to form a saliency map[7,8]; and attention is guided by top-down, high-level cognitions, such as knowledge, semantics, and scene context[9,10]. However, because cognition is difficult to represent and compute directly, many current studies on attention guidance still focus primarily on image saliency[11].

However, while these models and theories provide a comprehensive understanding of visual attention in general, they fall short of addressing the specific challenges faced by remote operators in disaster scenarios. The existing literature, while rich, often focuses on controlled environments or tasks that are less complex than those encountered in real-world disaster response. Moreover, the specific aspect of visual attention allocation in interpreting and analyzing real-time imagery, as required in remote operation, has not been extensively studied. Considering the distinctive challenges inherent in disaster environments—including swiftly changing scenarios, profound emotional impacts,

and the necessity for prompt decision-making—the discrepancy between extant research findings and practical needs is starkly highlighted. Research has substantiated that variables such as fear emotions[12], visual stimuli[13], and cognitive burdens critically affect both information reception and the distribution of visual attention.

Therefore, this study focuses on post-disaster high-risk task scenarios, aiming to investigate the attention mechanism in the specific environment of remote operations after natural disasters and explore the influencing factors of attention under dynamic and high-pressure tasks. It is expected to provide more efficient and usable solutions for emergency response and disaster management while enriching the theoretical framework of attention allocation modeling.

2 Related works

In the realm of post-disaster remote rescue operations, understanding the visual attention of the driver is critical for effective management and response. The dynamics of visual attention, especially in high-stress and complex environments such as post-disaster scenarios, are multifaceted and are influenced by a multitude of factors ranging from scenario characteristics to individual cognitive processes.

Visual attention is an important component of human visual perception. When confronted with a complex visual scene, human beings will efficiently localize the parts of interest and analyze the scene by selectively processing some regions of the visual input, a process that is also known as prioritized allocation of attention in the presence of limited attention. In order to understand the mechanism of human visual attention, there have been many scholars who have extensively explored the underlying theories of visual attention. Feature-Integration Theory[4], saliency-based visual attention model[7] and graph-based visual saliency model[8] provide essential insights into how visual features are processed and prioritized. These theories highlight the significance of bottom-up stimuli characteristics in directing attention.

With the development of computational vision, many studies have begun to emphasize the effects of top-down high-level features on visual attention. Researches demonstrate that local scene semantics guide attention during natural visual search in scenes[10,11]. Additionally, the work of Vö [14] delves into the deeper layers of scene structure and meaning, further emphasizing the role of high-level cognitive factors in visual attention. This is especially important in disaster scenarios, where quickly understanding the structure of the scene and recognizing meaningful elements in a scene full of debris can be life-saving.

Nowadays, most of the research on visual attention is based on the living environment and simple visual tasks, but the tasks and environments in remote rescue operations are a new dimension and challenge for visual attention research. Fan, Li, and Su[15] discuss the construction of human visual attention maps in teleoperation, pertinent to remote rescue scenarios where operators navigate through rubble remotely. The augmentation of reality, as explored by Eyraud, Zibetti, and Baccino[16], demonstrates how technological enhancements can alter the allocation of visual attention, a factor critical in designing remote rescue operation interfaces.

Understanding visual attention in the context of post-disaster remote operations has practical implications. Driewer, Schilling, and Baier[17] discuss human-computer interaction in rescue systems, highlighting the need to tailor these systems to accommodate the visual attention patterns of operators. Additionally, Rea et al.[18] emphasize the importance of attention-grabbing techniques in robot teleoperation, a key component in remote rescue scenarios.

In this study, we synthesize multiple factors such as salience, meaning, scene, and object semantics to explain and predict patterns of human attention in non-routine scenarios. The study delves into the factors influencing operators' attention in post-disaster remote search and rescue, with a special focus on the particular mechanisms of attention in high-risk tasks and continuously changing scenarios.

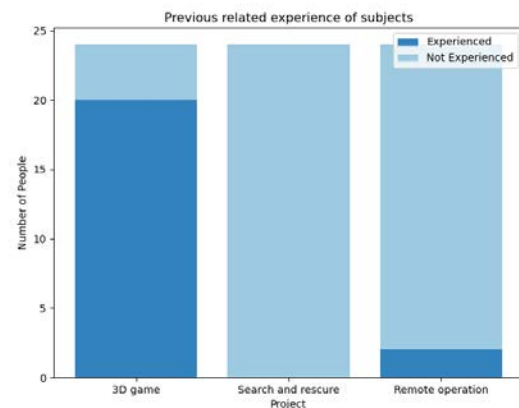


Figure 1. Previous related experience of participants

3 Method

3.1 Eye movements: an experiment in post-disaster remote SAR

We employed a rigorously designed experimental framework utilizing VR technology. A virtual 3D scene of a two-story building after an earthquake was built in

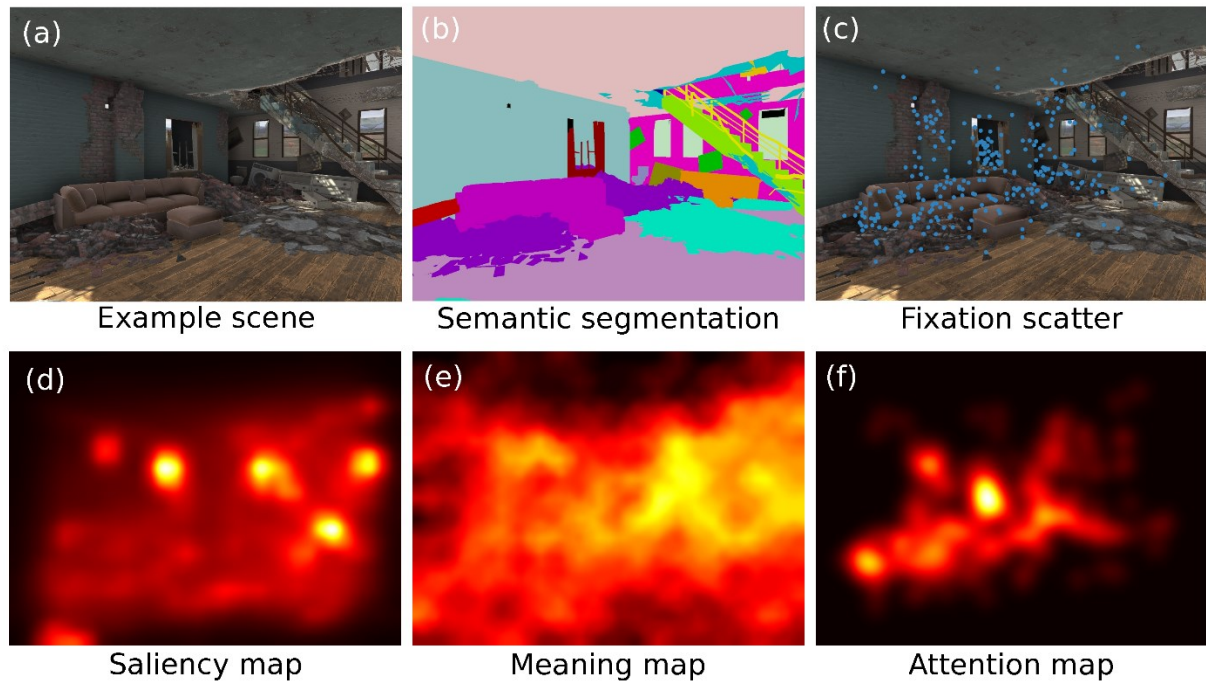


Figure 2. Maps for an example scene. (a) Scene (b) Semantic segmentation mask (c) Fixations scatter plot (d) Saliency map (e) Meaning map (f) Attention map

Unity as a stimulus for the experiment. Users were able to roam around the model following a specified route while completing tasks related to search and rescue. Recognizing the critical role of stress in such high-stakes environments, we integrated time constraints and multitasking requirements to induce realistic stress levels. A pilot test with two participants was conducted before the formal experiment to verify the appropriateness of the experimental design and ensure that the system functions properly.

3.1.1 Participants

A total of 24 (12 males and 12 females) college students volunteered to participate in this experiment. The average age was 23.1 years old. As Figure 1, none of the participants had experience in remote or on-site search and rescue, ensuring a baseline level of expertise consistent across the sample. However, most of the participants had experience with 3D gaming and teleoperation, reflecting a proficiency with virtual environments. This participant profile was intentionally selected to elucidate the intrinsic characteristics of visual attention in SAR scenarios, absent of specialized training biases. Such a selection criteria facilitate the extrapolation of our findings to a broader audience, potentially enhancing the inclusivity and effectiveness of remote SAR training programs.

Table 1 Categories of objects

Object type	Object class
Danger	broken wall, concrete floor, cooking bench, droplight, fallen wall, light, picture, rebar
Information	bed, broken bricks, broken rubble, broken door, closet, door, stairs, toilet, window
Environment	background, bedside table, book, cabinet, scattered ground object, chair, cupboard, kitchen hood, curtain, ventilation, grid, fridge, tableware, shelf, sofa, table, wall, washbasin, washing machine, wooden floor

3.1.2 Scene design

The stimulation scene of the experiment is the indoor environment of a two-story building after an earthquake developed in Unity modeled after a real scene (Figure 2(a)). The model was meticulously designed to closely mimic post-disaster scenarios, facilitating immersive task execution by the participants. Considering the post-earthquake SAR mission, the building's attributes were set as a residence, and the interior design was arranged according to common home life scenes, including living

room, kitchen, bathroom, bedroom, and other spaces. In the scene, various kinds of rich objects were arranged to enhance the realism of the scene, and based on this, after removing some object classes that were small in size and could be recognized as having no impact on the study, the points of interest were formed based on the object classes as a unit. As shown in Table 1, the objects were categorized into three categories (danger, information, and environment) according to the task.

3.1.3 Apparatus

Eye movements were recorded using a Tobii Eye Tracker 4C at a sampling frequency of 90Hz. The tracking accuracy of the eye tracker was 38° and 29° horizontally and vertically, respectively. Participants were seated 85 cm from the 21-inch screen, giving the scene a viewing angle of approximately $26.5^\circ \times 20^\circ$ at 1024×768 pixels.

3.1.4 Experimental procedure

The experiment was structured into four distinct phases. Initially, participants were provided with a standardized guide to familiarize themselves with the experimental scenario, roles, and tasks. Main tasks entailed identifying lifeforms, ensuring UAV and future personnel safety, and learning about the indoor environment for later tasks. Operators viewed a video of the urban road scenario for task understanding, with 5 minutes allocated for mastery.

Next, participants underwent a 5-point calibration with an eye-tracker for accuracy. In the second stage, they practiced the task for 1 minute. The third stage, the main test, lasted 15 minutes, featuring an automated scene change along a fixed route with static intervals for detailed observation. Throughout the route, there were 30 designated pause points where participants were required to identify and click on objects necessitating attention following task comprehension, all within an 8-second window at each location.

The final stage involved completing a questionnaire about the experiment and gathering essential information.

3.2 Analysis

3.2.1 Data processing

For the eye movement data, gaze and sweep were distinguished by an Initial-Velocity Threshold (I-VT) filter, with the threshold set at $30^\circ/\text{s}$. The I-VT filter was based on the assumption that if the eye movement data exceeded the threshold, then the eye movement was most likely intentional, otherwise it could be recognized as noise or other non-intentional eye movements. In addition, any gaze shorter than 60 ms and longer than 1500 ms was excluded as an outlier (Figure 2(c)).

For the experimental scenarios, each type of object was replaced in unity with a specific color material without illumination to form a 3D scene with a semantic segmentation mask (Figure 2 (b)).

3.2.2 Attention maps

Attention maps were generated as described in Henderson and Hayes[9]. Briefly, a fixation frequency matrix based on the locations (x, y coordinates) of all fixations was generated across participants for each scene. a Gaussian low-pass filter with a circular boundary and a cutoff frequency of - 6 dB was applied to each matrix, to account for foveal acuity and eye-tracker error. The spatial extent of the low-pass filter was 152 pixels in diameter (Figure 2(f)).

In addition, the significance and saliency maps were normalized to a common scale using image histogram matching using the gaze density maps of each scene as a reference image before statistical analysis.

3.2.3 Visual saliency maps

The saliency maps (Figure 2(d)) for the 30 stops in the scene routes were computed using the Graph-Based Visual Saliency (GBVS) toolbox with default settings[8]. GBVS uses information about the global structure of the image to improve the efficiency of saliency computation. is a well-known saliency model that performs well in complex scenes.

3.2.4 Meaning Maps

Meaning maps of 30 stops in a scene route were created according to the method proposed by Henderson and Hayes.

The stimuli of the session consisted of images of the 30 stopping points. Each image was segmented into circular patches on two scales: fine (300 patches/image) and coarse (108 patches/image.) There were ultimately 9000 unique fine patches and 3240 unique coarse patches.

A total of 132 participants rated the patches on the web application we built. All participants were university students from Zhejiang University, and each of them was allowed to participate in this experiment only once. Each participant would rate 300 randomized patches after reading the experiment content and the two patch examples of low meaning and high meaning. Participants were asked to rate each patch on a 6-point likert scale (very low, low, somewhat low, somewhat high, high, and very high) based on their understanding of the degree of meaning of each patch. Each unique patch was rated by at least 3 independent raters. Finally, the meaning map of each stopping point image was obtained by averaging, smoothing, and combining the ratings of both segmentation scales (Figure 2(e)).

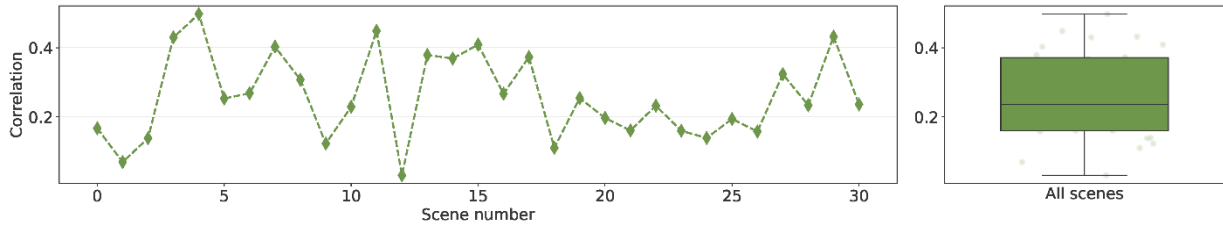


Figure 3. Correlation between meaning and saliency maps

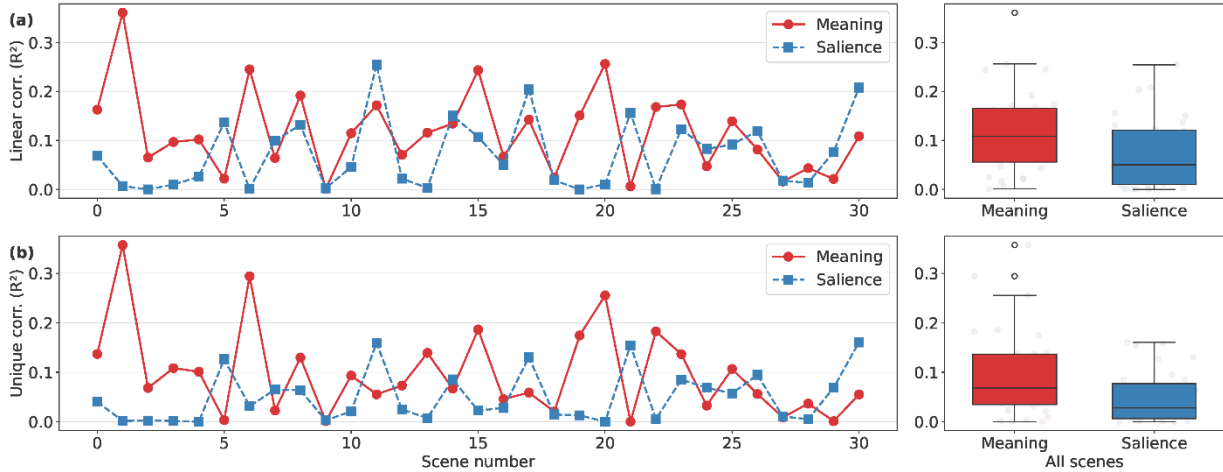


Figure 4. Squared linear correlation and semi-partial correlation by scene

4 Results

4.1 Meaning and visual saliency

Following Henderson and Hayes[9], we used squared linear and semi-partial correlation to analyze meaning maps and saliency maps across the whole scene.

It is suggested that meaning and visual saliency are highly correlated across scenes and that the guidance of attention by visual saliency may come from meaning. This argument has been validated to some extent in everyday scenarios after the significance map approach was proposed. **Error! Reference source not found.** shows the correlation between meaning and saliency for each scene in a complex environment after a disaster. On average, the correlation is 0.26 (s.d. = 0.12) across the 30 scenarios. A one-sample t-test confirmed that the correlation was significantly greater than zero, $t(29) = 11.7$, $p < 0.0001$, 95% confidence interval (CI) [0.21, 0.30]. This demonstrates that meaning and visual saliency are correlated even in non-daily scenarios. It is important to consider the relationship between these two when investigating their role in visual attention modeling.

Figure 4(a) presents the correlation of meaning and saliency on attention for all scenes. For the mean squared linear correlation across the 30 scenes, meaning

explained 12% of the change in attention ($M = 0.12$, s.d. = 0.08), while saliency explained 7% of the change in attention ($M = 0.07$, s.d. = 0.07). A two-tailed t-test showed that this difference was statistically significant, $t(58) = 5.63$, $p < 0.0001$, 95% confidence interval [0.004, 0.084].

The ability of meaning and saliency to independently explain attention was further explored by computing squared semi-partial correlations, controlling for shared variance in attention. Figure 4(b) illustrates the unique variance of meaning and saliency on attention across all scenarios. On average, meaning independently explained 10% of the variance in attention ($M = 0.10$, s.d. = 0.09), double the ability of saliency to explain it ($M = 0.05$, s.d. = 0.05). This suggests that meaning, controlling for saliency, produces 10% additional variance in the attention graph; whereas saliency only produces 5%. This difference remained significant by a two-tailed t-test ($t(58) = 2.56$, $p < 0.01$, 95% CI [0.01, 0.08]). This suggests that meaning relative to visual saliency continues to dominate directing attention in cluttered post-disaster scenes.

4.2 Object value

Some studies[19,20] have demonstrated that objects predict the allocation of attention points in non-search

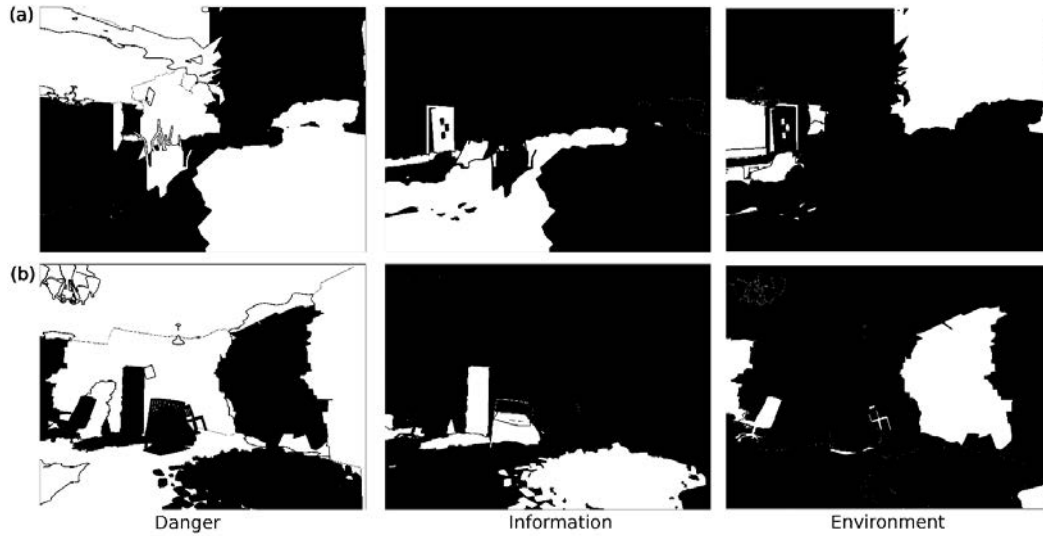


Figure 5. Object semantic category maps of two example scenes

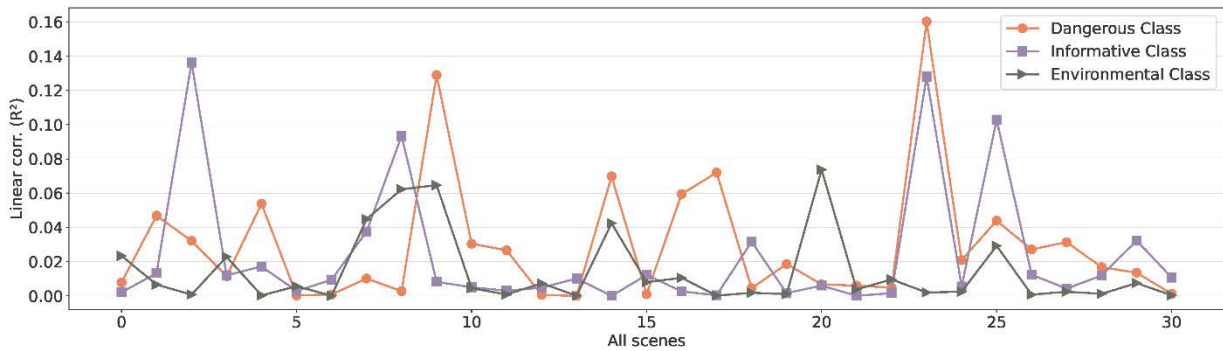


Figure 6. Squared linear correlation between object semantics and attention

tasks. That is, when performing a non-search task, people's attention tends to be drawn to specific objects in an image.

In the experiments, task and object semantics were highly correlated. As in Figure 5, segmenting the object semantics and distinguishing labels according to different categories (belonging to the category and not belonging to the category) forms an attribute map of the object semantics.

In visual search, attention may be guided by the target or by distractors. Figure 6 shows the squared linear correlations between different semantic categories of objects and attention in all scenes. The danger category explained 2.9% of the variation in attention ($M = 0.029$, $s.d. = 0.04$), with a one-sample t-test significantly greater than 0 ($t(29) = 4.36$, $p < 0.001$, 95% CI [0.016, 0.043]); the information category explained 2.3% of the variation in attention ($M = 0.023$, $s.d. = 0.04$), with a one-sample t-test significantly greater than 0 ($t(29) = 3.43$, $p < 0.01$, 95% CI [0.009, 0.037]); and the environment category explained 1.4% of the attentional change ($M = 0.014$, $s.d.$

$= 0.02$), with a one-sample t-test significantly greater than 0 ($t(29) = 3.73$, $p < 0.001$, 95% CI [0.006, 0.022]). This result demonstrates that object semantics guided visual attention in the task and that the target objects (hazard class, information class) were able to play more of a role relative to distractors (environment class), achieving a more task-appropriate guidance.

Figure 7 illustrates the results of the linear correlation analysis of semantic attribute maps with meaning. For all scenarios on average, the correlation coefficient between the danger category and the meaning map was -0.14 ($M = -0.14$, $s.d. = 0.20$), with a significant one-sample t-test ($t(29) = -3.80$, $p < 0.001$, 95% CI [-0.217, -0.065]); the correlation coefficient between the information category and the meaning map was 0.076 ($M = 0.076$, $s.d. = 0.15$), with a significant one-sample t-test ($t(29) = 2.76$, $p < 0.01$, 95% CI [0.020, 0.132]); the correlation coefficient of the environmental category with the significance map was -0.018 ($M = -0.018$, $s.d. = 0.146$) and the one-sample t-test was non-significant ($t = -0.6712$, $p = 0.507 > 0.05$). This suggests that the target objects (danger category,

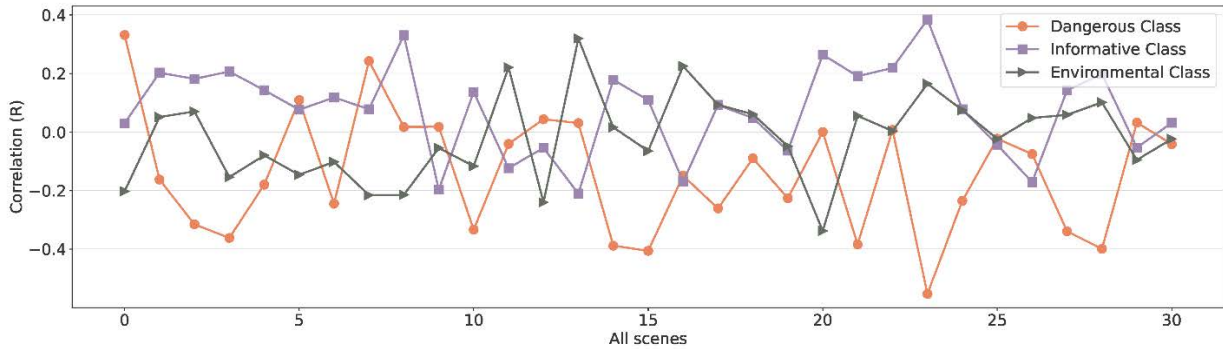


Figure 7. Linear correlation between object semantics and meaning

information category) are correlated with the meaning map and cannot be considered independent of their guidance of visual attention with meaning. Among them, the danger category is negatively correlated with the meaning map. Dangerous objects are more difficult to recognize than everyday scene objects because of the characteristics of breakage, loss of original order, confusion, and mismatch with inherent cognition, making semantics slightly negatively correlated in the representation of meaning. The distractors (environment class), on the other hand, are uncorrelated with the meaning map, which further suggests that distractors may be difficult to provide valuable recognition information in the task.

Since object semantics and meaning are partially correlated, the independent bootstrapping of salience, meaning, and semantics on attention was further investigated by multiple linear regression. Table 2 shows the independent influence of either factor on attention, controlling for the other factors. In particular, meaning independently explained 9.4% of the variation in attention ($M = 0.094$, $s.d. = 0.092$), whereas object semantics remained relevant to attention independently of significance, with each category explaining about 2% of the variation in attention. Relative to other factors, the guidance of attention by meaning remained dominant.

Table 2. Multiple linear regression result

Category	$M(R^2)$	S.D.	p-value
Meaning	0.094	0.092	5.3e-19
Saliency	0.050	0.055	5.9e-48
Danger	0.025	0.028	2.5e-4
Information	0.013	0.019	2.2e-12
Environment	0.018	0.024	2.5e-2

5 Conclusion and future works

Previous work has suggested numerous factors that will have an impact on visual attention, which will be directly expressed in human behavior and task efficiency and accuracy. Most of the past researches base their

experiments on daily environment, using some real world or simplified feature expressions that people often come into contact with as stimuli, but in fact, high-stress and high-risk environments form stronger stimuli and cognitive loads, which make the human attention pattern more complex. At the same time, the analysis about scenes and semantics has mainly focused on the quantitative expression of a certain semantic feature, while failing to completely describe the post-disaster environment and the work faced by search and rescue personnel.

In this study, a post-earthquake cluttered environment was designed as an experimental stimulus, and routes and tasks were designed according to the remote search and rescue approach. Our methodological approach, employing eye-tracking technology alongside the development of scenario-specific feature maps, enabled an in-depth analysis of visual attention in high-stress SAR conditions. The interactions between the guiding factors of salience, meaning, task, and object were investigated, and it was found that the effect of any factor on attention was not independent and that meaning remained dominant. This suggests that the construction of an attention model needs to take multiple factors into account, and that a more comprehensive model will lead to a better explanation of attention. The inclusion of participants without SAR background revealed their capability to allocate attention to targets as per task demands, providing empirical support for mobilizing a broader workforce in post-disaster rescue efforts. This study also used semantic attribute maps to represent scene features and understand how objects affect visual attention. These findings provide a more comprehensive perspective for understanding human visual attention mechanisms.

In summary, the main findings of this study are:

- Meaning and visual salience are significantly correlated in post-disaster scenes. They both predicted the distribution of attention, but after controlling for the relationship between significance and salience, only significance

contributed to the unique variance of the attention distribution.

- Objects do not direct attention independently of meaning in semantically related tasks, although the correlation between the two is weak.
- Object relative value is also able to direct visual attention, with guidance from the target object being more significant relative to guidance from distractors.

Although this study provides a robust analysis of the mechanisms of visual attention for teleoperated drivers in natural disaster scenarios, it does have some limitations. In the study of dynamic scenes, time and cognitive updating formed by repetitive images were not added to the analysis of attention to form a model of attention with a temporal sequence. As a result, the study did not provide a complete picture of the possible effects of successive visual images. Subsequent research could build on the foundation of this study and try to establish a more complete theory of visual attention and construct a more accurate predictive model of attention. Further work could also assist the operator to locate the target through multimodal modeling, give real-time positive guidance to attention, and improve the efficiency and success rate of search and rescue.

References

- [1] S. Mandyam, S. Priya, S. Suresh, and K. Srinivasan, A Correlation Analysis and Visualization of Climate Change using Post-Disaster Heterogeneous Datasets, , 2022.
- [2] United Nations Office for Disaster Risk Reduction Centre for Research on the Epidemiology of Disasters, The human cost of disasters: an overview of the last 20 years, , 2020.
- [3] B. Mishra, D. Garg, P. Narang, and V. Mishra, Drone-surveillance for search and rescue in natural disaster, *Computer Communications*. 156:1–10, 2020.
- [4] A.M. Treisman, The effect of irrelevant material on the efficiency of selective listening, *The American Journal of Psychology*. 77:533–546, 1964.
- [5] D.E. Broadbent, Perception and communication, Elsevier, 2013.
- [6] J.M. Wolfe, and T.S. Horowitz, Five factors that guide attention in visual search, *Nat Hum Behav*. 1:1–8, 2017.
- [7] L. Itti, C. Koch, and E. Niebur, A model of saliency-based visual attention for rapid scene analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 20:1254–1259, 1998.
- [8] J. Harel, C. Koch, and P. Perona, Graph-Based Visual Saliency, in: *Advances in Neural Information Processing Systems*, MIT Press, 2006.
- [9] J.M. Henderson, and T.R. Hayes, Meaning-based guidance of attention in scenes as revealed by meaning maps, *Nat Hum Behav*. 1:743–747, 2017.
- [10] J.M. Henderson, T.R. Hayes, G. Rehrig, and F. Ferreira, Meaning Guides Attention during Real-World Scene Description, *Sci Rep*. 8:13504, 2018.
- [11] C.E. Peacock, T.R. Hayes, and J.M. Henderson, Meaning guides attention during scene viewing, even when it is irrelevant, *Atten Percept Psychophys*. 81:20–34, 2019.
- [12] E.B. Foa, and M.J. Kozak, Emotional processing of fear: Exposure to corrective information, *Psychological Bulletin*. 99:20–35, 1986.
- [13] J.D. Fraustino, J.Y. Lee, S.Y. Lee, and H. Ahn, Effects of 360° video on attitudes toward disaster communication: Mediating and moderating roles of spatial presence and prior disaster media involvement, *Public Relations Review*. 44:331–341, 2018.
- [14] M.L.-H. Võ, The meaning and structure of scenes, *Vision Research*. 181:10–20, 2021.
- [15] J. Fan, X. Li, and X. Su, Building Human Visual Attention Map for Construction Equipment Teleoperation, *Front Neurosci*. 16:895126, 2022.
- [16] R. Eyraud, E. Zibetti, and T. Baccino, Allocation of visual attention while driving with simulated augmented reality, *Transportation Research Part F: Traffic Psychology and Behaviour*. 32:46–55, 2015.
- [17] F. Driewer, K. Schilling, and H. Baier, Human-Computer Interaction in the PeLoTe rescue system, in: *IEEE International Safety, Security and Rescue Robotics, Workshop, 2005.*, IEEE, 2005: pp. 224–229.
- [18] D.J. Rea, S.H. Seo, N. Bruce, and J.E. Young, Movers, Shakers, and Those Who Stand Still: Visual Attention-grabbing Techniques in Robot Teleoperation, in: *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, Association for Computing Machinery, New York, NY, USA, 2017*: pp. 398–407.
- [19] M.S. Castelhana, M.L. Mack, and J.M. Henderson, Viewing task influences eye movement control during active scene perception, *J. Vision*. 9:6, 2009.
- [20] Y. Chen, and G.J. Zelinsky, Is There a Shape to the Attention Spotlight? Computing Saliency Over Proto-Objects Predicts Fixations During Scene Viewing, *J. Exp. Psychol.-Hum. Percept. Perform*. 45:139–154, 2019.

Conceptualizing Digital Twins in Construction Projects as Socio-Technical Systems

Mohammed Abdelmegid¹, Algan Tezel², Carlos Osorio-Sandoval², Zigeng Fang², and William Collinge¹

¹School of Civil Engineering, University of Leeds, UK

²Department of Civil Engineering, University of Nottingham, UK

m.abdelmegid@leeds.ac.uk, algan.tezel@nottingham.ac.uk, carlos.osorio@nottingham.ac.uk,
zigeng.fang@nottingham.ac.uk, w.h.collinge@leeds.ac.uk

Abstract –

Digital Twin (DT) has been proposed in the construction sector as the evolution of static Building Information Modeling (BIM) by enabling seamless data flow between physical and digital environments. This paper presents an integrated framework that conceptualizes DT in construction projects as a socio-technical system. This framework aims to address the social dynamics inherent in construction projects when implementing digital technologies. The proposed framework is developed as an amalgamation of components from existing DT and socio-technical systems to enable simultaneous focus on technical requirements and social readiness. It represents different DT maturity stages spanning from static digital models to fully functional DT with bi-directional data flow. These stages are mapped against the dimensions of a socio-technical system including goals, people, technology, processes, infrastructure, and culture. The findings of this study can serve as a research roadmap for creating a comprehensive approach for DT implementation in the construction industry. This approach can ultimately evolve into a DT execution plan (DTEP), which can be requested by clients and as an addendum to contracts in the future. Future work involves validating the framework and benchmarking it against established frameworks in other industries. Additionally, the framework can be extended to incorporate external factors such as policy, legal, and commercial factors.

Keywords –

Digital twins; Socio-technical systems; implementation frameworks

1 Introduction

Digital Twin (DT) is often described as a system of three components: a physical entity, a virtual replica, and

information flow between both. Despite the multiple definitions of DTs in the literature, common characteristics can be identified such as the intelligence and agency of DTs, enabling them to sense and change the environment [1]. DTs find different applications in construction, spanning from health and safety (e.g. worker/plant interface, worker posture, crowd management etc), sustainability (e.g. near real time carbon estimates over site operations), logistics planning, and structural reliability monitoring of temporary and newly built elements. Several benefits have been reported to the use of DTs in construction. They can significantly boost productivity, reduce lifecycle costs, improve environmental performance, and advance safety standards [1, 2]. In decision-making processes, DTs provide valuable transparency into the impacts of decisions on both human factors and the natural environment, as stressed by Council [3].

DT technology has evolved significantly in recent years, emerging from its application in several industries such as aerospace and manufacturing [1]. However, there has been slow progress in embracing the technology in the construction industry with small pockets of adoption throughout a building's lifecycle [4]. Several studies highlighted the confusion of DT with Building Information Modeling (BIM) in the industry, highlighting that DT can be considered the natural progression from the more static and design-focus applications of BIM to a more dynamic and holistic approach, enabling real-time monitoring, analysis, and simulation of physical asset throughout their lifecycle [2, 4, 5].

One of the key misconceptions about the adoption of new technologies in the construction industry is that they will inherently improve current practices [6]. Such misconceptions can be observed in both academia and industry, assuming that organizations can organically embrace new technologies, neglecting the need for clearly defined implementation frameworks, execution plans, and the right skill sets [7]. With regard to DT,

several studies emphasize the need for specialized research that addresses not only technological challenges but also the organizational, managerial, and behavioral aspects for successful implementation of DT in construction [6, 8]. Agrawal, et al. [1] stressed the need to investigate human-DT interactions in construction projects to ensure appropriate role allocation. Boje, et al. [2] highlighted the complex social systems around built environments, emphasizing the challenge of developing intelligent systems that can adapt to human dynamics and provide insights for decision-making. Council [3] underlined the need to extend DT research beyond technical solutions, calling attention to the necessity of addressing human and organizational factors, including ethics, management, and social considerations.

The aim of this study is to conceptualize a framework for the implementation of DT in construction projects, emphasizing the integration of social and technical aspects into construction projects through socio-technical systems. The following objectives are set to facilitate this aim:

- Investigate current frameworks for DT implementation in construction to identify key elements and common practices.
- Investigate methods of modelling digital technologies as socio-technical systems in the context of construction projects.
- Integrate the identified elements from DT frameworks with a socio-technical model, aiming to create a cohesive and adaptable framework that addresses the nature of construction projects.

2 Background

2.1 Digital Twin Frameworks

Several frameworks have been developed to facilitate DT implementation in construction. Some were initiated to extend BIM applications to include the use of Internet of Things (IoT) and Data Mining algorithms to capture live data, incorporate it to a BIM model, and have a deeper understanding of how data can be translated to support decision making [9]. Other frameworks focused on data flow in construction DT. For example, Pregolato, et al. [10] proposed a workflow consisting of five steps (data and need acquisition, digital modelling, dynamic data transmission, data/model integration, operation) interlinked with four components (real, link, virtual, experience). Zhao, et al. [11] introduced a bottom-up framework with six layers: preparation, data acquisition, data processing, data transmission, model logic, application. Similarly, Honghong, et al. [12] developed a framework for lifecycle digital transformation for bridge engineering consisting of 8 steps: data collection, data

preprocessing, data transfer and storage, DT model building, model update, fusion, feedback, and human-computer interaction. Several other studies presented generic higher-level frameworks that link the physical, the digital, and the application layers of construction projects [2, 5, 8, 13, 14]. Among these studies, few indicated the social aspects of construction projects as key components of DT implementation. Xie, et al. [14] stressed the need to develop a framework that enables humans to become active elements of any DT environment. Agrawal, et al. [8] presented a digitalization framework to balance technology push (data/model and performance) with business pull (value and transformation). More specifically, Boje, et al. [2] developed a DT framework for conventional Life Cycle Sustainability Assessment. This framework evaluates social life cycle from six perspectives: working conditions, human rights, health and safety, cultural heritage, socio-economic, and governance. In a broader sense, Lu, et al. [13] introduced a road map for building and city level DT with three layers: trust, function, and purpose. In the trust layer, a sustainable plan is established, including society, economy, and environment. Although several frameworks have been presented in the literature, there are limitations in relation to providing a holistic view of DT implementation that addresses the key barriers in the construction industry from both technical and non-technical perspectives.

2.2 Socio-technical systems and Digital Twins

The socio-technical theory is based on the principle that understanding and enhancing the design and performance of organizational systems necessitate the simultaneous consideration of both 'social' and 'technical' aspects, treating them as interdependent components within a complex system [15]. This theory views an organization as a group of co-dependent subsystems engaged in dynamic interactions [16]. The historical development of socio-technical systems can be traced to Trist and Bamforth [17], which conceptualized the complex relationship between technological and social dimensions in heavy industries [18]. The application of socio-technical systems has evolved to encompass diverse industries [19]. Therefore, socio-technical systems form a holistic framework that has become fundamental to understand and improve the complex dynamics of organizational structures [20].

The adoption of a socio-technical perspective emerges as a critical imperative for the successful and sustainable implementation of new technologies. Münch, et al. [22] reported that the failure of new systems to meet user requirements often results from a biased focus on technological needs, neglecting the equally important social needs. In contrast to the traditional approach of developing technology first and then fitting people to it,

socio-technical systems place the same importance on addressing technological and social needs concurrently [23]. In addition to the interdependent role humans play in system performance, a socio-technical system should also consider how humans are influenced by the system [16]. The emphasis of socio-technical systems on joint optimization, adaptability, and human-centric design establishes their key role in navigating the complexities of modern organizational challenges.

Socio-technical systems have been applied effectively within the construction context. Li [20] reported several studies on conceptualizing BIM implementation as a socio-technical system such as Sackey, et al. [24], which highlighted that success in BIM implementation necessitates prioritizing people and processes over technology and information. In addition to BIM, socio-technical approaches have been applied to Distributed Ledger Technologies (DLT) and Smart Contracts (SCs) by Li, et al. [18]. Ang, et al. [15] investigated the application of socio-technical systems to model an intelligent robot technology project. Their study examined stakeholder interactions, from ideation to successful prototyping, illustrating the efficacy of socio-technical systems in driving innovation.

A recent research theme emerges in DT literature related to the conceptualization of DTs as sociotechnical systems. According to Barn [26], DT encapsulates a socio-technical system, where social and technical elements are seamlessly integrated to achieve goal-directed behavior. The complexity of DT extends into a socio-technical dilemma as it [27]necessitates real-time adaptation to users and responding effectively to daily changes [27]. This dynamic interplay between technical functionalities and human interactions characterizes the intricate nature of DTs [2].

According to Lei, et al. [28], DT should not only be technology-driven but should also encourage public participation and be understandable to a wider audience. Collaboration, both within and between organizations, is identified as a critical foundation for implementing DT, emphasizing the need to consider social, legal, and commercial/business perspectives for a more comprehensive understanding. Rebentisch, et al. [29] confirm that implementing DT within a socio-technical context not only enhances organizational goals, such as business performance and product lifecycle management, but also contributes to broader societal benefits, including social and environmental sustainability.

In the context of the built environment, DT represents a clear example of socio-technical systems. Jiang, et al. [30] emphasize the integration of various information and communication technologies in construction DTs for collaborative management and operation. This application reinforces the notion that effective DT in construction requires a harmonious relationship between

technological advancements and social processes.

In summary, the incorporation of socio-technical systems in DT research extends the focus on technology to acknowledge the critical role played by social processes, collaboration, and comprehensive system reasoning in the implementation of DTs.

2.3 Relevant socio-technical frameworks

Several studies highlighted the importance of applying frameworks to guide the design of socio-technical systems to ensure that the whole ecosystem is considered when implementing new technologies [31]. Yu, et al. [21] presented a framework based on socio-technical systems theory to facilitate integrating BIM with Blockchain comprising three components: process, technic, and context. Similarly, Li, et al. [18] adopted a framework for DLT and SCs covering the technical, process, social, and policy dimensions. They demonstrated the role of socio-technical systems in addressing the barriers of technology implementation in the construction industry. The socio-technical hexagon by Davis, et al. [19] illustrates the interdependent components of dynamic socio-technical systems including goals, people, buildings/infrastructure, technology, culture, and processes/procedures. Ivanov [32] presented a DT framework based on socio-technical systems encapsulating key aspects of DT for supply-chain and operations management. This framework consists of seven elements: people, organization, modeling, task, scope, technology, and management. This study emphasizes that DTs are complex socio-technical phenomena incorporating human-artificial intelligence interactions.

3 A Conceptual Framework for Construction DT based on Socio-technical systems

As described in the previous section, several frameworks have been developed for construction DTs and socio-technical systems separately. We propose an integrated approach to conceptualize DT as a socio-technical system in construction projects by developing a matrix that illustrates how the key components of the socio-technical hexagon capture the requirements of DT at different implementation levels. We adopt the three levels of DT by Kritzing, et al. [33], which differentiates between a static digital representation of the physical system (Digital Model), a partially integrated system with one-way data flow (Digital Shadow), and a fully integrated two-way data flow between the physical and digital system (Digital Twin). The Matrix in Table 1 outlines our proposed framework.

Table 1: A matrix illustrating the conceptual framework

Elements of the socio-technical system	Digital Model	Digital Shadow	Digital Twin
Goals and Metrics	<ul style="list-style-type: none"> • Utilize digital models for project visualization and planning. • Set objectives for accurate representation of physical objects in the digital space. 	<ul style="list-style-type: none"> • Utilize the digital shadow for real-time monitoring and basic data exchange. • Set objectives for improved responsiveness and monitoring capabilities. 	<ul style="list-style-type: none"> • Optimize project goals based on continuous two-way data exchange with the digital twin. • Set objectives for enhanced collaboration and decision-making through the digital twin.
People/Human Factors	<ul style="list-style-type: none"> • Train personnel on BIM tools for creating and updating digital models. 	<ul style="list-style-type: none"> • Train personnel on utilizing real-time data from the digital shadow. 	<ul style="list-style-type: none"> • Create work interfaces where construction teams can interact with digital twins. • Provide training on advanced tools and features of the digital twin.
Technology/Tools	<ul style="list-style-type: none"> • Implement BIM software for creating accurate digital representations. • Ensure compatibility with common industry tools for data exchange. 	<ul style="list-style-type: none"> • Integrate sensors and automated data capture devices for real-time data. • Implement tools for monitoring and basic analysis of digital shadow data. 	<ul style="list-style-type: none"> • Utilize advanced IoT devices and communication systems for seamless data exchange. • Implement AI algorithms for advanced real-time data analytics for decision support. • Integrate digital twin and its spin off technologies with legacy IT systems.
Processes/Practices	<ul style="list-style-type: none"> • Establish workflows for creating and updating digital models throughout the project life cycle. • Integrate digital models into planning and design processes. • Incorporate risks related to the digital model implementation (e.g., data accuracy, interoperability, etc) into the project risk management plan. 	<ul style="list-style-type: none"> • Integrate digital shadow data into existing monitoring and reporting processes. • Establish protocols for responding to changes identified in the digital shadow. • Incorporate risks related to the digital shadow implementation (e.g., sensor malfunctions, data overload, etc) into the project risk management plan. 	<ul style="list-style-type: none"> • Adapt construction processes for real-time collaboration and decision-making with the digital twin. • Establish workflows that leverage bi-directional data exchange for optimization. • Incorporate risks related to the digital twin implementation (e.g., data synchronization, ethical implications, etc) into the project risk management plan.
Physical Infrastructure	<ul style="list-style-type: none"> • Ensure hardware supports BIM and visualization. • Provide access to digital models across relevant project teams. 	<ul style="list-style-type: none"> • Upgrade infrastructure to support real-time data flow for the digital shadow. • Ensure connectivity and 	<ul style="list-style-type: none"> • Implement IoT infrastructure for bi-directional data exchange in the digital twin. • Ensure robust

Elements of the socio-technical system	Digital Model	Digital Shadow	Digital Twin
		regular maintenance of remote sensing devices. • Provide database infrastructure for recorded data.	cybersecurity measures for protecting bi-directional data flow. • Ensure connectivity and regular maintenance of remote sensing and actuator devices. • Provide database infrastructure for bi-directional recorded data
Culture	<ul style="list-style-type: none"> • Foster a culture of using digital models for better understanding, project planning, and decision support. • Encourage knowledge sharing and avenues for learning from experience such as communities of practices and mentoring. 	<ul style="list-style-type: none"> • Encourage teams to rely on the digital shadow for timely decision support. • Establish means to enable sharing knowledge through the digital shadow. 	<ul style="list-style-type: none"> • Foster a collaborative environment through the digital twin by building trust in the knowledge created and shared through the digital twin. • Ensure alignment between organizational culture and digital twin practices.

We demonstrate the application of the framework in an example of production planning and control of piling operations in a construction site. For this example, we assume a large construction site with different soil and water profiles to create a reasonable level of complexity suitable for DT. There are different types of equipment required to perform this operation including piling rigs, cranes, water pumps, concrete trucks, haul trucks, and excavators. Also, there is a need to manage inventory and

supply of materials at the site (e.g., concrete and reinforcement steel) and move excavated soil to a dumping site. The aim of using DT is to support production planning and control of the piling operations by minimizing idle time of machinery, optimizing inventory on site, and improving schedule reliability. Table 2 demonstrates the application of the framework to implement a DT for the piling operations.

Table 2: Illustrative example of piling operations

Elements of the socio-technical system	Digital Twin
Goals and Metrics	<ul style="list-style-type: none"> • Develop a dynamic model that automatically adjusts production plans and authorizes work orders to all production teams.
People/Human Factors	<ul style="list-style-type: none"> • Ensure clear role allocation is established between DT and decision-makers.
Technology/Tools	<ul style="list-style-type: none"> • Ensure the facilitation of different stakeholders' communication with DT central hub. • Embed AI tools (e.g., machine-learning, NLP, expert systems, etc) into the model to evaluate and predict project performance. • Utilize advanced IoT devices and communication systems for seamless data exchange between the model and construction equipment.
Processes/Practices	<ul style="list-style-type: none"> • Automatically update production plans to reflect optimum solutions.
Physical Infrastructure	<ul style="list-style-type: none"> • Enable the digital twin to authorize work orders to production teams and suppliers. • Deploy interconnected sensors and actuators to enable two-way information flow. • Utilize reliable cybersecurity tools and services to maintain project safety and security.
Culture	<ul style="list-style-type: none"> • Establish the digital twin as the central hub for knowledge sharing between academic, industry, and public organizations.

4 Discussion

The proposed conceptual framework offers an initial methodology to support DT implementation in construction projects by harmonizing components from DT and socio-technical systems. A key strength of the proposed framework is the emphasis on balancing social and technical factors across different DT implementation stages. By mapping these stages to socio-technical dimensions, construction teams can have a mutual understanding of the requirements to successfully implement DT. The matrix representation allows an incremental implementation of DT while ensuring appropriate social adaptation. Moreover, incorporating change management practices can mitigate organizational challenges of deploying new technologies [3]. In this section, we summarize the components of the framework in relation to the 'Digital Twin' stage in Table 1 to highlight the key requirements and challenges of DT implementation compared to the other lower levels.

4.1 Goals and Metrics

The framework establishes a goal to enable data-driven decision-making leveraging the ability of DTs to synchronize virtual and physical environments. Key performance indicators, such as model accuracy and optimization effectiveness, can be employed to ensure that the DT aligns with real-world situations. The DT should aim to generate insight for enhanced outcomes not only for the current project but for future ones. The framework provides a trajectory for DT to support cross-learning and continuous improvement throughout the project lifecycle.

4.2 People/Human Factors

The framework highlights DT alignment with organizational structure as a key factor for integrating construction teams with DT workflow. This includes clear allocation of roles and responsibilities for staff interacting with the DT to alleviate the adverse consequences of incorrect role allocation such as cost increases, unrealistic expectations, and strategic misalignment [1]. Training programs should aim at capability building and establishing trust in DT outcomes, thus, enhancing organizational readiness for widespread adoption of DT for optimized project performance.

4.3 Technology/Tools

This component of the framework emphasizes the need for robust bi-directional data flow through advanced integration of interconnected sensors and automation technologies. However, one major concern is the integration of DTs and their spin-off technologies with legacy IT systems through appropriate Application

Programming Interfaces (APIs) while ensuring cybersecurity and data protection in organizations across the supply chain. For example, large-scale infrastructure projects (e.g., highways, pipelines, water channels, railways, etc) rely heavily on Geographic Information Systems (GIS) for spatial coordination. In such projects, geospatial data should be connected with DTs in a similar way to current integrated GIS/BIM approaches.

4.4 Processes/Practices

The framework advocates for a paradigm shift necessitating the need for process engineering to accommodate the multi-disciplinary effort for DT implementation. As DT technology is rapidly evolving and increasing in complexity, overwhelming the internal capabilities of construction organizations, a new business sector might emerge to provide specialized consultancies and non-profit organizations to guide construction companies in establishing processes and practices for DT implementation.

4.5 Physical Infrastructure

The framework points out to the need for interconnected sensor and actuator systems to enable bi-directional data flow. However, decisions related to infrastructure requirements should carefully consider cost-benefit analysis. The cost of investment in DT infrastructure and maintenance can be necessary in large complex infrastructure projects but difficult to justify in small construction projects that run on tight profit margins, which might lead to 'pseudo-DT' applications much like some examples in the BIM domain. Reusability of DT infrastructure across different projects can be one of the key factors for cost-benefit analysis. Leasing such equipment from specialized subcontractors can be another business decision in future construction projects. In addition to economic challenges, practical and legal challenges require attention when deploying DT infrastructure such as licensing for drones and security of site equipment.

4.6 Culture

DT implementation requires strong commitment across all industry levels to foster buy-in at individual, team, department, organization, supply chain, and the overall sector. This buy-in should ideally be intrinsic and not merely forced through government mandates, contractual requirements, or secondary reasons, e.g., to create an innovative company persona or image. Achieving such level of intrinsically motivated adoption is a key challenge for construction projects as can be observed even with less advanced technologies such as BIM. Hence, building trust in DT capabilities should be

in the core of any implementation plans. Pilot demonstrations, participatory decision-making, contractual incentives, and awareness programs can be among the promoting techniques for DT cultural transition. However, it is crucial to address ethical risks and underlying biases in a transparent manner in any effort to influence cultural transition.

5 Conclusion

This paper introduces an integrated framework to conceptualize construction DT as a socio-technical system. It aimed at bridging a recognized gap in construction DT research regarding the lack of managerial, human, and social considerations in existing construction DT approaches, which primarily focus on the technical aspects of DT such as data integration and analytics. The framework addresses this gap by balancing the focus between technical and social requirements in DT implementation in construction projects. Therefore, this study suggests a more holistic view of DT across critical areas such as performance metrics, training, process re-engineering, and cultural readiness.

The outcomes of this study can form a research roadmap to develop a holistic approach for systematic DT implementation in the construction industry. This approach will eventually evolve into a DT execution plan (DTEP) that can be used for contractual and governance requirements by construction companies to successfully deploy DT technologies and realize their value to business performance as well as social outcomes. Due to significant investment requirement and data richness of DT systems, such DTEP document will be more critical than current BIM execution plans. Future research involves testing and enhancing the framework in industry settings and benchmarking against frameworks in other sectors. Such testing should incorporate expert validation to gain better insight into the enablers, challenges, and practical implications of implementing the proposed framework. In addition, examining the dynamic interrelations within the socio-technical system when implementing DT is vital to ensure alignment between all components of the framework. Modeling the framework development in a simulation environment that can capture long-term business and social outcomes (e.g., Systems Dynamics) is a possible approach for testing and validation. Finally, we recognize the need to incorporate other external aspects to the framework such as policy, legal, and commercial factors that influence the diffusion of digital technologies in the industry.

References

- [1] A. Agrawal, R. Thiel, P. Jain, V. Singh, and M. Fischer, "Digital Twin: Where do humans fit in?," *Automation in Construction*, vol. 148, 2023, doi: 10.1016/j.autcon.2023.104749.
- [2] C. Boje *et al.*, "A framework using BIM and digital twins in facilitating LCSA for buildings," *Journal of Building Engineering*, vol. 76, 2023, doi: 10.1016/j.jobbe.2023.107232.
- [3] G. Council, & Lamb, K., "Gemini Papers: Why connected digital twins," 2022.
- [4] F. Jiang, L. Ma, T. Broyd, and K. Chen, "Digital twin and its implementations in the civil engineering sector," *Automation in Construction*, vol. 130, 2021, doi: 10.1016/j.autcon.2021.103838.
- [5] J. M. Davila Delgado and L. Oyedele, "Digital Twins for the built environment: learning from conceptual and process models in manufacturing," *Advanced Engineering Informatics*, vol. 49, 2021, doi: 10.1016/j.aei.2021.101332.
- [6] M. Lövgren Moazzami and M. Brandt, "Delivering the undefined: The value potential of digital twins: A qualitative study on digital twins in the Swedish AEC/FM industry," 2023.
- [7] O. Vigren, A. Kadefors, and K. Eriksson, "Digitalization, innovation capabilities and absorptive capacity in the Swedish real estate ecosystem," *Facilities*, vol. 40, no. 15/16, 2022, doi: 10.1108/F-07-2020-0083.
- [8] A. Agrawal, M. Fischer, and V. Singh, "Digital Twin: From Concept to Practice," *Journal of Management in Engineering*, vol. 38, no. 3, 2022, doi: 10.1061/(ASCE)ME.1943-5479.0001034.
- [9] Y. Pan and L. Zhang, "A BIM-data mining integrated digital twin framework for advanced project management," *Automation in Construction*, vol. 124, 2021, doi: 10.1016/j.autcon.2021.103564.
- [10] M. Pregnolato *et al.*, "Towards Civil Engineering 4.0: Concept, workflow and application of Digital Twins for existing infrastructure," *Automation in Construction*, vol. 141, 2022, doi: 10.1016/j.autcon.2022.104421.
- [11] J. Zhao, H. Feng, Q. Chen, and B. Garcia de Soto, "Developing a conceptual framework for the application of digital twin technologies to revamp building operation and maintenance processes," *Journal of Building Engineering*, vol. 49, 2022, doi: 10.1016/j.jobbe.2022.104028.
- [12] S. Honghong, Y. Gang, L. Haijiang, Z. Tian, and J. Annan, "Digital twin enhanced BIM to shape full life cycle digital transformation for bridge engineering," *Automation in Construction*, vol. 147, 2023, doi: 10.1016/j.autcon.2022.104736.
- [13] Q. Lu *et al.*, "Developing a Digital Twin at Building and City Levels: Case Study of West Cambridge Campus," *Journal of Management in Engineering*, vol. 36, no. 3, 2020, doi: 10.1061/(ASCE)ME.1943-5479.0000763.

- [14] H. Xie, M. Xin, C. Lu, and J. Xu, "Knowledge map and forecast of digital twin in the construction industry: State-of-the-art review using scientometric analysis," *Journal of Cleaner Production*, vol. 383, 2023, doi: 10.1016/j.jclepro.2022.135231.
- [15] K. C. S. Ang, S. Sankaran, D. Liu, and P. Shrestha, "From Lab to Field: A Sociotechnical Systems View of the Organizational Cycle for Intelligent Robotics Technologies Ideation, Planning, Design, Development and Deployment," in *Proceedings of the International Symposium on Automation and Robotics in Construction*, 2023, pp. 294-301, doi: 10.22260/ISARC2023/0041.
- [16] C. Xia and Y. Hu, "A Review of a Socio-Technical System Approach for Interdependent Infrastructure Systems Resilience Analysis: Present Status and Future Trends," in *Construction Research Congress 2022: Infrastructure Sustainability and Resilience - Selected Papers from Construction Research Congress 2022*, 2022, vol. 1-A, pp. 416-426, doi: 10.1061/9780784483954.043.
- [17] E. L. Trist and K. W. Bamforth, "Some Social and Psychological Consequences of the Longwall Method of Coal-Getting: An Examination of the Psychological Situation and Defences of a Work Group in Relation to the Social Structure and Technological Content of the Work System," *Human Relations*, vol. 4, no. 1, 1951, doi: 10.1177/001872675100400101.
- [18] J. Li, D. Greenwood, and M. Kassem, "Blockchain in the built environment and construction industry: A systematic review, conceptual models and practical use cases," *Automation in Construction*, Article vol. 102, 2019, doi: 10.1016/j.autcon.2019.02.005.
- [19] M. C. Davis, R. Challenger, D. N. W. Jayewardene, and C. W. Clegg, "Advancing socio-technical systems thinking: A call for bravery," *Applied Ergonomics*, vol. 45, no. 2, Part A, 2014, doi: 10.1016/j.apergo.2013.02.009.
- [20] J. J. Li, "A socio-technical framework to guide implementation and value realisation of distributed ledger technologies (DLT) in the construction sector," Northumbria University, 2023.
- [21] J. Yu, H. Zhong, and M. Bolpagni, "Integrating blockchain with building information modelling (BIM): a systematic review based on a sociotechnical system perspective," *Construction Innovation*, 2023, doi: 10.1108/CI-04-2023-0082.
- [22] C. Münch, E. Marx, L. Benz, E. Hartmann, and M. Matzner, "Capabilities of digital servitization: Evidence from the socio-technical systems theory," *Technological Forecasting and Social Change*, vol. 176, 2022, doi: 10.1016/j.techfore.2021.121361.
- [23] S. H. Appelbaum, "Socio - technical systems theory: an intervention strategy for organizational development," *Management decision*, vol. 35, no. 6, 1997, doi: 10.1108/00251749710173823.
- [24] E. Sackey, M. Tuuli, and A. Dainty, "Sociotechnical systems approach to BIM implementation in a multidisciplinary construction context," *Journal of management in engineering*, vol. 31, no. 1, 2015.
- [25] G. Hagan, D. Bower, and N. Smith, "Managing complex projects in multi-project environments," in *Association of Researchers in Construction Management, ARCOM 2011 - Proceedings of the 27th Annual Conference*, 2011, vol. 2, pp. 787-796.
- [26] B. S. Barn, "The Sociotechnical Digital Twin: On the Gap between Social and Technical Feasibility," in *Proceedings - 2022 IEEE 24th Conference on Business Informatics, CBI 2022*, 2022, vol. 1, pp. 11-20, doi: 10.1109/CBI54897.2022.00009.
- [27] C. Boje, A. Guerriero, S. Kubicki, and Y. Rezgui, "Towards a semantic Construction Digital Twin: Directions for future research," *Automation in Construction*, vol. 114, 2020, doi: 10.1016/j.autcon.2020.103179.
- [28] B. Lei, P. Janssen, J. Stoter, and F. Biljecki, "Challenges of urban digital twins: A systematic review and a Delphi expert survey," *Automation in Construction*, vol. 147, 2023, doi: 10.1016/j.autcon.2022.104716.
- [29] E. Rebentisch, D. H. Rhodes, A. L. Soares, R. Zimmermann, and S. Tavares, "The Digital Twin as an Enabler of Digital Transformation: A Sociotechnical Perspective," in *IEEE International Conference on Industrial Informatics (INDIN)*, 2021, vol. 2021-July, doi: 10.1109/INDIN45523.2021.9557455.
- [30] Y. Jiang, M. Li, D. Guo, W. Wu, R. Y. Zhong, and G. Q. Huang, "Digital twin-enabled smart modular integrated construction system for on-site assembly," *Computers in Industry*, vol. 136, 2022, doi: 10.1016/j.compind.2021.103594.
- [31] J. Li and M. Kassem, "Applications of distributed ledger technology (DLT) and Blockchain-enabled smart contracts in construction," *Automation in Construction*, vol. 132, 2021, doi: 10.1016/j.autcon.2021.103955.
- [32] D. Ivanov, "Conceptualisation of a 7-element digital twin framework in supply chain and operations management," *International Journal of Production Research*, 2023, doi: 10.1080/00207543.2023.2217291.
- [33] W. Kritzinger, M. Karner, G. Traar, J. Henjes, and W. Sihn, "Digital Twin in manufacturing: A categorical literature review and classification," *IFAC-PapersOnLine*, vol. 51, no. 11, 2018, doi: 10.1016/j.ifacol.2018.08.474.

Coupled Risk Assessment of Construction Workers' Unsafe Behaviors in Human-Robot Interactions

Jing Lin¹, Qinyuan Li¹, Runhe Zhu², Yuhang Cai¹, Longhui You¹, Lewen Zou¹ and Siyan Liu¹

¹ Department of Construction Management, Dalian University of Technology, China

² Moss School of Construction, Infrastructure and Sustainability, Florida International University, United States

linjing@dlut.edu.cn, liqinyuan@mail.dlut.edu.cn, rzhu@fiu.edu, caiyuhang@mail.dlut.edu.cn,
2777988504@qq.com, 2975286849@mail.dlut.edu.cn, 2793817353@qq.com

Abstract –

Human-robot interaction (HRI) is expected to play an important role in the construction industry in the coming decades. However, construction workers face new safety challenges that stem from both external environmental risks and individual internal risks in the HRI environment. Therefore, this study aims to identify safety risk factors related to the HRI environment and develop a coupled risk assessment method for construction workers' unsafe behaviors in the HRI environment.

The research methodology involves literature review, questionnaire surveys, hierarchical analysis, and fuzzy evaluation. As a result, this study identified 59 risk factors (37 external and 22 internal) related to the HRI environment and established a three-level coupled degree indicator system. There are 35 experts invited to rate the influence and occurrence possibility of each indicator. Hierarchical analysis was employed to assign weights to the indicators, taking into account the experts' opinions and using the entropy weight method to improve the accuracy. The composite indicators of the second-level and first-level indicators were calculated to evaluate the coupled risk of unsafe behaviors in the HRI environment.

The findings revealed that the interaction between construction equipment and site condition factors has a strong effect on construction workers' unsafe behavior in the HRI environment, indicating a need for strengthened control measures such as developing construction guidelines. The present study provides a scientific basis for evaluating the safety of the HRI environment in building construction and designing safety management systems and measures for intelligent construction.

Keywords –

Human-robot interaction; Unsafe behavior; Coupled risk; Risk assessment; Construction safety

1 Introduction

The construction industry faces challenges due to an aging population, high casualty rate and shortage of workers, which hinder its healthy growth. Collaborative research dedicated towards the effective use of robotics and automation is suggested to solve the challenges in the construction industry [1]. The Human-robot interaction (HRI) environment refers to the workspace where both construction workers and robots are working at the same time, aiming at accomplishing a different or the same task.

Compared to traditional construction equipment, construction robots possess higher levels of automation and intelligence. However, the increased safety level of construction robots may lead to workers' compensatory risk behavior (e.g., overly trusting the intelligent system or unintentionally entering the interaction area), which may result in more unsafe behaviors and eventually accidents [2]. Therefore, understanding causes of workers' unsafe behaviors in the HRI environment is crucial for construction safety.

In the HRI environment, the causes of unsafe behavior include both robot-related risks (e.g., robot malfunction and mishandling) and traditional construction risks (e.g., workers' safety knowledge and experience) [3]. These two types of risks act in tandem with each other, hereafter referred to as HRI-related risks collectively, and lead to the emergence of unsafe behaviors in the HRI environment. Therefore, understanding interaction risks is critical to understanding construction workers' unsafe behaviors in the HRI environment.

This paper aims to analyze the coupled mechanisms of safety risks affecting construction workers' unsafe behaviors in the HRI environment. Drawing on the findings obtained through the questionnaire-based approach, this study designates influences characterized by high coupled effects as high-impact safety risks. The outcomes pertaining to high-impact safety risks, along with the coupled safety risks, were deliberated upon to augment construction safety.

2 Background

2.1 Safety Risks of the HRI Environment in Construction

Previous research has identified various safety risks associated with the HRI environment. For example, Chung et al. [4] identified seven risk categories (i.e., human, control, unauthorized Access or Operational Situation Awareness, mechanical concerns, environmental sources, power systems and improper installation) related to the HRI environment. Physical safety risks, attentional cognitive safety risks, and physiological response safety risks have been identified as the three main types of safety challenges in Unmanned Aerial Vehicles (UAVs) construction [5]. A variety of factors such as system malfunction, operator error, and worker stress have also been shown in the literature to trigger safety risks during construction workers' interaction with the machine [6]. In addition to this, existing studies have developed assessment tools for risks in the HRI environment, including 8 categories totaling 40 security risks [7]. Meanwhile, to address these safety risks, researchers proposed solutions and evaluated them through the Hierarchy of Control (HoC) method to verify the effectiveness of these solutions [6]. However, existing research has not yet considered the coupled relationship of internal and external risk factors in the HRI environment, which hinders effective support for elucidating safety management in such environments.

2.2 Assessment of Coupled Safety Risks

In the field of safety risk assessment, researchers have increasingly focused on the assessment of coupled safety risks. Various research methods have been employed, including Bayesian networks and N-K model [8]. Zhi et al. [9] analyzed the tunnel construction risk based on N-K and coupling degree models, and calculated the

coupling degree of each component of single-factor and two-factor risk coupling models respectively. Wang et al. [10] established a risk network model based on the complex network theory, analyzed the topological characteristics and key risk characteristics of the tower crane safety network, and then revealed the evolution law and coupling relationship of the safety risks in the whole process of the tower crane, and realized the quantification of key risk characteristics.

While existing studies have focused on safety risks in construction, research on coupled safety risks related to unsafe behaviors in the HRI environment is rare. Therefore, this study aims to address this gap by employing coupling evaluation method to construct a coupled risk assessment method for unsafe behaviors in the HRI environment.

3 Research Method

This study conducted a comprehensive literature review to establish a three-level indicator system for assessing the coupled degree. Specifically, four first-level indicators of HRI safety risks, namely, organization, robot, environment and equipment, and workers, were identified based on reference to traditional construction safety classifications. Then, a review of relevant literature was conducted based on the four first-level indicators, which led to the identification of second-level and third-level indicators. This paper constructs a systematic framework by summarizing the factors of three high-quality review papers, then enriches the second-level indicators and third-level indicators based on literature review of empirical studies. To determine the contribution value of each factor in the indicator system, the analytic hierarchy method was employed. Expert assessment method was chosen because the collection of objective data is challenging. The research methodology is shown in Figure 1

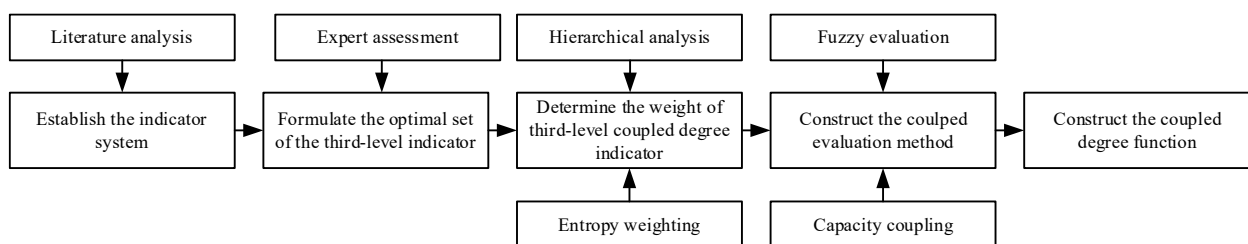


Figure 1. Flow chart of coupled risk assessment for the HRI environment

3.1 Establish the Indicator System

To identify safety risks in the HRI environment, a literature review of publications in databases, including Web of Science and Google Scholar, was conducted. Search keywords include “construct*”, “coupling risk

OR coupled risk” and “evaluat* OR assess*”. The retrieved literature serves as the basis for the indicator system of the risk coupled degree in the HRI environment. The indicator system consists of 4 first-level indicators, 13 second-level indicators and 59 third-level indicators, as shown in Table 1. The three levels of the indicator

system were firstly constructed based on three high-quality review studies [4], [7], [11]. Then, the indicators at each level were enriched and refined by reviewing

related empirical research, such as those that consider physiology factors and fatigue levels [7], [13].

Table 1. Indicator system of coupled degree of HRI risk system

First-level indicators	Second-level indicators	Third-level indicators
Organization	Safety management system	Safety leadership, safety supervision, and safety regulation
	Safety climate	Safety education and training, incentives, colleague safety behavior, safety investment and cost, and safety culture
	Construction technical plan	Schedule, project characteristics
	Enterprise organization	Enterprise revenue, enterprise size, enterprise reputation, and financial condition
Robot	Human error	Improper robot control, improper robot assembly and installation, unauthorized access, and inspection and maintenance
	Robot problem	Potential component failures, quality control errors, mechanical failures, and robot aging, wear, and renewal
	Environmental induced robot malfunction	Electromagnetic and radiation interference, electrical failures and overload, and dust
Environment and equipment	Site condition	Vision, auditory, dirt, site layout, overhead load environment, obstructions and congested sites, ground condition, climate conditions, cross operation, blind area, and power system
	Construction equipment	Equipment characteristics-ergonomics, inspection and maintenance of equipment, and aging, wear, and renewal of equipment
Construction workers	Demographic factor	Age, education level, personality, job position, income level, family, occupation type, and qualification
	Work skills	Safety knowledge, working knowledge, working experience, accident experience, and working ability
	Physiology factor	Fatigue level, health, and lifestyle habits
	Psychology factor	Psychological burden, safety awareness, mental state, trust in equipment, and perceived control

Note: Original questionnaire could be provided on request via email.

3.2 Formulate the Optimal Set of the Third-Level Indicator

In this research, the coupling degree is calculated based on the coupling degree model, which has the advantages of low sample demand and efficient calculation [14].

3.2.1 Determine the Optimal Indicator Vector

In this paper, an improved grey multi-level theory was employed to calculate the composite indicator of each coupled subsystem of safety risks in the HRI environment. The HRI risk coupled system is divided into three levels, the first-level (organization- C_1 , robot- C_2 , equipment and conditions- C_3 and workers- C_4), the second-level ($F_i, i=1, 2, \dots, m$), and the third-level ($F_{ij}, j=1, 2, \dots, n$).

To gather data for the coupled degree indicators, an

expert assessment method was adopted. Experts were invited to participate in the study and assess the third-level indicators. The requirement for experts includes a minimum of three years of working experience in the relevant field and a leader of three or more people, or at least a master's degree and publications of peer-reviewed articles in areas such as construction safety, unsafe behavior, or HRI.

The questionnaire administered to the experts is used to assess the degree of influence of each third-level indicator, as well as the likelihood of its occurrence. Both the degree of influence and the likelihood of occurrence are assessed by a 5-point Likert Scale. Thus, the value of each parameter is calculated by the product of the degree of influence and the likelihood of occurrence with a maximum value of 25 points. The average product of influence degree score and possibility score is taken as the importance degree of each third-level indicator. This approach enabled the researchers to quantitatively determine the significance of each indicator within the system.

d_{ij} is the original value of the third-level coupled degree indicator corresponding to the second-level indicator F_{ij} , expressed by the matrix D_i as:

$$D_i = (d_{i1}, d_{i2}, \dots, d_{ij}, \dots, d_{in})^T \quad (1)$$

where D_i represents the initial vector of the i th second-level indicator; d_{ij} represents the original value of the j th third-level indicator under the i th second-level indicator; n represents the number of third-level indicators included in second-level indicator.

Subsequently, the matrix D_i was standardized to obtain the standard matrix for each evaluation indicator. Suppose d_{ij}^{max} is the optimal value of the tertiary indicator F_i corresponding to the secondary indicator F_{ij} , then:

$$D_i^{max} = (d_{i1}^{max}, d_{i2}^{max}, \dots, d_{ij}^{max}, \dots, d_{in}^{max})^T \quad (2)$$

where D_i^{max} represents the optimal value vector of the second-level coupled degree indicator.

3.2.2 Determine evaluation vector

Take D_i as the vector to be compared and take D_i^{max} as the reference vector, then the correlation degree between the third-level coupled degree indicator and the optimal indicator value can be calculated as:

$$\lambda_{ij} = \frac{\min|d_{ik} - d_{ik}^{max}| + \rho * \max|d_{ik} - d_{ik}^{max}|}{|d_{ij} - d_{ij}^{max}| + \rho * \max|d_{ik} - d_{ik}^{max}|} \quad (3)$$

where λ_{ij} represents the correlation coefficient between the actual value and the optimal value of the third-level coupled degree indicator F_{ij} ; k represents the ordinal number of third-level coupled degree indicators, $k=1, 2, \dots, n$; ρ represents resolution coefficient, in general $\rho \leq 0.5$.

Formula 3.3 is used to calculate the correlation degree between the measured value and the optimal value of each indicator. The ρ of 0.5 is used to calculate the correlation degree. Then the evaluation vector of the third-level indicators is:

$$P_i = (\lambda_{i1}, \lambda_{i2}, \dots, \lambda_{ij}, \dots, \lambda_{in})^T \quad (4)$$

where P_i represents the coupling evaluation matrix of third-level indicator.

3.2.3 Determine the Weight of Third-level Coupled Degree Indicator

Since this study is reasoning and analyzing about construction workers' unsafe behaviors and risks from a qualitative point of view, and thus obtaining the indicators used for the analysis. Therefore, it was not possible to assign separate weights to each indicator for analysis from a quantitative perspective. For example, the safety risk of equipment, unlike storms with the

objective property of risk frequency, is difficult to measure directly. Therefore, hierarchical analysis is used in this study to assign the indicator weight of coupled degree. At the same time, in order to avoid the influence of expert method on the reliability and accuracy of data, entropy weight method was used to modify the analytic hierarchy process, so as to make full use of the available information to improve the accuracy of weight value.

In this study, questionnaire data were used as the measured value of the importance degree of each third-level coupled indicator, and the pairwise division was obtained according to the interval division to obtain the importance degree comparison score a_{ij} . The importance degree comparison score is separate from the influence degree score and possibility score above. According to the 5-point system, 1 point represents that the two were equally important, and 5 points represent that the former was extremely important than the latter. In order to avoid numerical errors in the above methods, this study invited 5 experts in related research fields to score the comparison of importance between second-level indicator and third-level indicator as backup data for comparative analysis.

The pair comparison matrix A_{ij} can be obtained by comparison and transformation, each factor in the matrix is not only positive, but also has integer characteristics, that is, the judgment matrix should meet $a_{ij} > 0$, $a_{ji}=1/a_{ij}$, and when $i=j=1, 2, \dots, m$, $a_{ij}=1$.

$$A_{kl} = \begin{bmatrix} 1 & a_{12} & \dots & a_{1j} & \dots & a_{1m} \\ 1/a_{12} & 1 & \dots & a_{2j} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 1/a_{1j} & 1/a_{2j} & \dots & 1 & \dots & a_{jm} \\ \vdots & \vdots & \dots & \vdots & 1 & \vdots \\ 1/a_{1m} & 1/a_{2m} & \dots & 1/a_{jm} & \dots & 1 \end{bmatrix} \quad (5)$$

where A_{kl} represents the paired comparison matrix of the l th second-level indicator of the k th first-level indicator; k represents the number of first-level indicator ($k=1, 2, 3, 4$); l represents the number of second-level indicator among corresponding first-level indicators; and a_{ij} represents the importance ratio between the i th and j th third-level indicator.

The root method is adopted to solve the weight vector,

$$\omega_i^0 = (\prod_{j=1}^n a_{ij})^{\frac{1}{n}} \quad (6)$$

where ω_i^0 represents the root of the judgment matrix. Normalization is then performed,

$$w_i' = \frac{\omega_i^0}{\sum_i \omega_i^0} \quad (7)$$

where $w_i'=[w_i', w_i', \dots, w_i']^T$ represents the eigenvector.

By importing the ω_i^0 obtained from the questionnaire into software yaanp and conducting normalization

processing, the eigenvector of the judgment matrix W'_i and the maximum eigenvalue of the matrix λ_{max} can be obtained. The consistency test is carried out, and the entropy weight method is used to modify it. Then, the final weight w_i and the desired feature vector $W_i = [w_1, w_2, \dots, w_n]^T$ can be obtained.

3.2.4 Construct the coupled evaluation method

Calculate the composite indicator of the second-level indicator. The formula for calculating the comprehensive indicator of the second-level indicator is as follows:

$$X_i = W_i^T P_i \quad (8)$$

where X_i represents the composite indicator of the second-level indicator.

Calculate the composite indicator of the first-level indicator. For the comprehensive evaluation of second-level indicators, $X_i = W_i^T P_i (i=1, 2, \dots, n)$, calculate the evaluation value vector of second-level indicators in each subsystem, which can be composed of $R_{C_i} = [X_1, X_2, \dots, X_n]^T$. Then, the vector composed of the optimal set of corresponding indicators is found out, and the correlation degree is calculated by Formula 3.3 to obtain the evaluation matrix Q_{C_i} of the second-level indicators. Finally, $X_{C_i} = Q_{C_i}^T W_{C_i}$ was used to calculate the composite indicator of each subsystem.

(3) Construction function

Suppose the variable $u_i (i=1, 2, \dots, n)$ is the ordinal parameter of HRI risk system, and u_{ij} is the j th indicator of the i th order parameter, whose value is $X_{ij} (i, j=1, 2, \dots, n)$. α_{ij} and β_{ij} respectively represent the upper and lower limits of the j th indicator of the i th order parameter when the system is stable. In this study, the upper and lower limits that can be reached by the evaluation indicator under ideal conditions are adopted for calculation, $\alpha_{ij}=1, \beta_{ij}=0$. Then, for the HRI risk system, the efficacy function of the system coupled indicator can be expressed as:

$$u_{ij} = \begin{cases} X_{ij} - \beta_{ij}/\alpha_{ij} - \beta_{ij}, & (X_{ij} \text{ has positive effect}) \\ \alpha_{ij} - X_{ij}/\alpha_{ij} - \beta_{ij}, & (X_{ij} \text{ has negative effect}) \end{cases} \quad (9)$$

where $u_{ij} \in [0, 1]$ represents the effect of indicators on the order degree of the system, that is, the contribution of indicators to the functional realization of HRI risk system.

3.2.5 Construct the coupled degree function

Using the concept of capacity coupling and the model of coupled coefficient in the field of physics, the coupled degree calculation model of any two or more first-order indicator systems can be determined respectively. The coupled degree formula of any two first-level indicator systems is as follows:

$$C_{ij} = [U_i \cdot U_j / (U_i + U_j)(U_j + U_i)]^{\frac{1}{2}} \quad (10)$$

where C_{ij} represents the coupled degree of the i th first-level indicator system and the j th first-level indicator system.

The coupled degree formula of multiple first-level indicator systems is as follows:

$$C = \sqrt[m]{\frac{U_1 \cdot U_2 \cdot \dots \cdot U_m}{\prod_{j=1}^m [\prod_{i=1, i \neq j}^m (U_i + U_j)]}} \quad (11)$$

where C represents the coupled degree of the entire system.

Value range of coupled degree between systems is $C \in [0, 1]$. According to the classification of coupled states in physics, $C \in [0, 0.3)$ represents weak interaction, $C \in [0.3, 0.7)$ represents medium interaction, and when $C \in [0.7, 1)$ represents strong interaction. According to the above classification, the results of coupled degree between systems can be discussed.

4 Data Analysis and Discussion

4.1 Coupled Risk Assessment

In this study, a total of 33 questionnaires were received, out of which 25 were valid and 8 were invalid. The invalid questionnaires were primarily due to identical responses for each question, resulting in low data reliability. The correlation and weight values were calculated based on the valid questionnaire data collected (Table 2). The results indicate that construction workers exhibit the highest correlation and weight values among the first-level indicators, highlighting their high impacts on HRI risk systems. Moreover, the influence of construction workers shows the strongest correlation with the other three influencing factors. Regarding internal organizational factors, the safety management system holds the utmost importance, while the enterprise organization exhibits the strongest correlation with the other three factors. Noteworthy third-level indicators with high correlation and weight values include safety supervision (r (correlation)=0.843; w (weight)= 0.339), project characteristics ($r=0.810$; $w=0.380$), robot control system errors ($r=0.957$; $w=0.261$), aging wear and update of construction equipment ($r=0.966$; $w=0.348$), poor health condition ($r=0.836$; $w=0.340$), and so on.

The utility values of the first-level and second-level indicators are shown in Table 3. Among the first-level indicators, construction workers exhibit the highest utility value, consistent with the results of correlation and weights. The results indicate that construction workers are more susceptible to engaging in unsafe behavior in the HRI environment. Concerning the second-level

indicators, robot failures resulting from enterprise organization, environmental factors, and construction equipment significantly impact their respective first-level indicators. Due to space limitations, the detailed calculation results for all third-level indicators were not explicitly provided in this article.

The results of this study demonstrate that the coupled degree C of unsafe behavior risk in the HRI environment is 0.230. The results indicate a low level of interaction within the entire system, suggesting the presence of an coupled relationship between various subsystems that influences unsafe behavior in the HRI environment. The coupled degree of the first-level indicators is as follows: Organization (C_o) = 0.147, Robot (C_r) = 0.263, Construction equipment and site conditions (C_e) = 0.498, and Construction workers (C_w) = 0.158.

The coupled state (R) of organizations, robots, and

construction workers within their respective subsystems is low, implying a weak correlation between the second-level indicators within each subsystem. The coupled degrees between subsystems are as follows: Organizations-Environment (C_{or}) = 0.497, Organizations-Construction equipment (C_{oe}) = 0.499, Organizations-Construction workers (C_{ow}) = 0.494, Robots-Environment (C_{re}) = 0.499, Robots-Construction workers (C_{rw}) = 0.500, and Environment-Construction equipment (C_{ew}) = 0.498.

The coupled degree between on-site conditions and construction equipment is relatively high, indicating a strong interaction relationship between these factors. The coupled effect resulting from this interaction increases the risk of unsafe behavior, thereby posing a greater danger to the overall system.

Table 2. Results of correlation degree and weight

First-level indicators	Correlation	Weight	Second-level indicators	Correlation	Weight
Organization	0.609	0.249	Safety management system	0.439	0.323
			Safety climate	0.615	0.261
			Construction technical plan	0.640	0.258
			Enterprise organization	1.000	0.158
Robot	0.834	0.218	Human error	0.709	0.600
			Robot problem	0.776	0.217
			Environmental induced robot malfunction	1.000	0.183
Construction equipment and site conditions	0.706	0.225	Site condition	0.490	0.583
			Construction equipment	1.000	0.417
Construction Workers	1.000	0.309	Demographic factor	0.748	0.214
			Work skills	0.865	0.337
			Physiology factor	0.766	0.237
			Psychology factor	1.000	0.212

Table 3. Results of utility

First-level indicators	Utility value u_{ij}	Second-level indicators	Utility value u_{ij}
Organization	0.626	Safety management system	0.864
		Safety climate	0.922
		Construction technical plan	0.928
		Enterprise organization	0.978
Robot	0.777	Human error	0.940
		Robot problem	0.947
		Environmental induced robot malfunction	0.966
Construction equipment and site conditions	0.702	Site condition	0.812
		Construction equipment	0.956
Construction workers	0.845	Demographic factor	0.869
		Work skills	0.895
		Physiology factor	0.873
		Psychology factor	0.918

4.2 Comparison with the Traditional Construction Environment

First, coupled safety risks of unsafe behavior in the traditional environment assessed in the literature [15] are higher than the coupled safety risks of unsafe behavior in the HRI environment assessed in the present study ($0.6997 > 0.230$). A possible explanation is that the application of robots reduces the interaction between traditional safety risks and the intelligence of robots improves the level of risk management.

Second, robot-related safety risks have coupled effects on unsafe behavior with traditional safety risks. As (upgraded) substitutes of workers and machines, robots conduct construction tasks with workers or independently. The nature of construction tasks and environment causes coupled safety risks between robots and other safety risks, as validated in the present study. Thus, safety management of the HRI environment should consider the new coupled safety risks found in the present study.

Third, workers are the most important indicator of unsafe behavioral safety risks for the both traditional environment and HRI environment. The cognitive process of workers' unsafe behaviors has become one of the core issues in construction site safety research and many scholars have analyzed workers' unsafe behaviors from a cognitive perspective [16]. Thus, all safety risks

may interact with workers and then cause unsafe behavior. Considering the intelligence application of robots in construction industry, training and education of construction workers on robot-related knowledge and skills should be well designed and strengthened for safety.

4.3 Reliability and Validity

Due to time constraints and limited availability of experts, this study employed a relatively small sample size. Two data collection methods were utilized to determine the weights, including direct scoring of importance comparison by 5 experts and 30 experts. The coupled analysis results are shown in Table 4. Although there is a difference in the way the questionnaires were filled out by the 5 experts and 30 experts, the results obtained are similar and therefore the results enhance the reliability and validity of the study's findings. This study was somewhat unaffected by the bias inherent in the expert assessment method, which was eliminated as much as possible by screening the qualifications of the experts, actively collecting future field data, designing standardized research steps, and repeating the means of validation.

Moreover, the design of research steps, selection of research methods and the consistent results among two groups of experts ensure the validity of the methodology of this study.

Table 4. Comparison of coupled results

Indicators	Results of 5 experts	Results of 30 experts
Organization	0.155	0.147
Robot	0.263	0.262
Construction equipment and site conditions	0.498	0.498
Construction workers	0.159	0.158
Total (HRI system)	0.233	0.230

5 Conclusion

This study built a model to quantitatively assess the coupled risk of unsafe behavior among construction workers in the HRI environment. The research findings indicate the presence of coupled safety risks in this environment, particularly the interaction between construction equipment and on-site conditions, as well as a high correlation between construction workers and other factors. Therefore, it is recommended to enhance safety management in the HRI environment through regular inspection and maintenance of construction equipment, as well as by providing efficient safety education for construction workers. Strengthening safety management measures for construction equipment, on-site conditions, and construction workers is crucial.

Although the coupled degree of robot-related factors may not be the highest, it is essential to consider their coupled effects in the design and implementation of safety management measures for robots, as the existing measures are still being explored. This study introduces the concept of coupled effects into the safety risk assessment of the HRI environment for construction operations, presenting a three-level coupled indicator system and evaluation method for analyzing the unsafe behavior of construction workers. This approach expands the knowledge in the field of the HRI environment safety for construction by considering coupled safety risks.

The HRI risk system constructed in this study could provide a theoretical basis for designing construction safety management documents, guiding safety training and inspection, etc.

Due to the limited number of construction projects in the HRI environment, this study used questionnaire-based survey as a method to obtain data. The application and promotion of the proposed method in practice should be studied in the future. Future research could utilize the questionnaire and coupled calculation methods proposed in this study with larger sample sizes, which will enhance the reliability and validity of the research results.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (NSFC) under Grant No. 72301053, and the Institute for Guo Qiang, Tsinghua University under Grant No. 2020GQG0007. Any opinions, findings, conclusions, or recommendations expressed in the paper are those of the authors and do not necessarily reflect the views of the funding agency.

References

- [1] Lin J., Cai Y. and Li Q. Development of Safety Training in Construction: Literature Review, Scientometric Analysis, and Meta-Analysis. *Journal of Management in Engineering*, 39(6): 03123002, 2023.
- [2] Hasanzadeh S., Garza J. and Geller E. S. Latent Effect of Safety Interventions. *Journal of Construction Engineering and Management*, 146(5): 4020033, 2020.
- [3] Yang J., Ye G., Xiang Q., Kim M., Liu Q. and Yue H. Insights into the mechanism of construction workers' unsafe behaviors from an individual perspective. *Safety Science*, 133(174): 105004, 2021.
- [4] Chung F. and Ashuri B. Investigating Hazards and Safety Risks Inherent in Human-Robot Interactions. *Construction Research Congress 2022: Project Management and Delivery, Controls, and Design and Materials - Selected Papers from Construction Research Congress 2022*, pages 631–640, Arlington, Virginia, 2022.
- [5] Jeelani I. and Gheisari M. Safety challenges of UAV integration in construction: Conceptual analysis and future research roadmap. *Safety Science*, 144: 105473, 2021.
- [6] Xu Y., Turkan Y., Karakhan A. A. and Liu D. Exploratory Study of Potential Negative Safety Outcomes Associated with UAV-Assisted Construction Management. *Construction Research Congress 2020*, pages 809–818, Tempe, Arizona, 2020.
- [7] Okpala I., Nnaji C. and Gambatese J. Assessment Tool for Human–Robot Interaction Safety Risks during Construction Operations. *Journal of Construction Engineering and Management*, 149(1): 04022145, 2023.
- [8] Jiang J., Liu G. and Ou X. Risk Coupling Analysis of Deep Foundation Pits Adjacent to Existing Underpass Tunnels Based on Dynamic Bayesian Network and N–K Model. *Applied Sciences*, 12(20): 10467, 2022.
- [9] Shan Z., Qiu L., Chen H. and Zhou J. Coupled Analysis of Safety Risks in Bridge Construction Based on N-K Model and SNA. *Buildings*, 13(9): 2178, 2023.
- [10] Wang L., Chen W., Liu J. and Li Z. Research on Key Risks of Tower Crane Based on Complex Network. *International Conference on Construction and Real Estate Management*, pages 398–405, Beijing, China, 2021.
- [11] Muñoz-La R. F., Mora-Serrano J. and Oñate E. Factors Influencing Safety on Construction Projects (Fscps): Types and Categories. *International Journal of Environmental Research and Public Health*, 18(20): 10884, 2021.
- [12] Okpala I., Nnaji C. and Gambatese J. Assessment Tool for Human–Robot Interaction Safety Risks during Construction Operations. *Journal of Construction Engineering and Management*, 149(1): 04022145, 2023.
- [13] Zacharaki A., Kostavelis I., Gasteratos A. and Dokas I. Safety Bounds in Human Robot Interaction: A Survey. *Safety Science*, 127: 104667, 2020.
- [14] Zhang H. and Sun Q. Risk Assessment of Shunting Derailment Based on Coupling. *Symmetry*, 11:1359, 2019.
- [15] Pan H., Gou J., Wan Z., Ren C., Chen M., T. Gou and Luo Z. Research on Coupling Degree Model of Safety Risk System for Tunnel Construction in Subway Shield Zone. *Mathematical Problems in Engineering*, 2019.
- [16] Ye G., Yue H., Yang J., Li H., Xiang Q., Fu Y. and Cui C. Understanding the Sociocognitive Process of Construction Workers' Unsafe Behaviors: An Agent-Based Modeling Approach. *International Journal of Environmental Research and Public Health*, 17(5):1588, 2020.

Enhancing Public Engagement in Sustainable Systems through Augmented Reality

Yu-Chen Lee, S.M.ASCE ¹, Abdullah Alsuhaibani ², Chih-Shen Cheng, Ph.D., M.ASCE ³, Fernanda Leite, Ph.D., P.E., F.ASCE ⁴

¹ Ph.D. candidate, Sustainable Systems Program, Maseeh Dept. of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, Austin, TX, USA.

² Ph.D. candidate, Civil Engineering and Project Management Program, Maseeh Dept. of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, Austin, TX, USA.

³ Postdoctoral Fellow, Civil Engineering and Project Management Program, Maseeh Dept. of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, Austin, TX, USA.

⁴ Associate Dean for Research, Cockrell School of Engineering, Joe J. King Professor in the Maseeh Dept. of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, Austin, TX, USA.

yuchenlee@utexas.edu, alsuhaibani@utexas.edu, chihshen.cheng@utexas.edu, fernanda.leite@utexas.edu

Abstract –

The sustainability of building infrastructure has been a growing critical global concern. However, stakeholders often overlook the sustainability aspect of systems in their everyday lives. This paper presents an exploration of the potential of augmented reality (AR) technology as a tool for enhancing public engagement around sustainable systems in buildings. We introduce an innovative approach to address the challenge of gap between sustainable infrastructure and user awareness. Through data collection such as access counts, usage duration, and user feedback, we can evaluate the effectiveness of AR applications in improving awareness and engagement in sustainability. This study illustrates how AR technology can facilitate a deeper connection between the public and sustainable systems, thus paving the way for a more environmentally friendly society.

Keywords – Building Sustainability, Augmented Reality, Public Engagement

1 Introduction

According to the Intergovernmental Panel on Climate Change fifth assessment report, the global average temperature has significantly increased over the past century, with carbon dioxide concentrations identified as one of the most prominent drivers. Greenhouse gas emissions have risen globally by 70% from 1970 to 2004 [1]. Consequently, the care for the environment and sustaining the earth's resources have multiplied in recent years [2].

Sustainability, focused on using and conserving

natural resources [3], has become a prevalent term in recent years. However, it is often seen as a complex concept with a lack of shared understanding among stakeholders about its effects, causes, roles, and strategies [4]. The lack of public awareness is commonly identified as a major barrier to implementing sustainability practices [5], [6], [7], [8]. This issue has led organizations such as UNESCO to advocate for sustainability education for several decades. For instance, in a UNESCO conference in 1999, the goal of environmental education was outlined as to “clearly show the economic, political and ecological interdependence of the modern world, in which decisions and actions by the different countries can have international repercussions. Environment should, in this regard, help to develop a sense of responsibility and solidarity among countries and regions” [9].

Furthermore, aging infrastructure and frequent climate-related disasters present an evident challenge for city-makers and communities across the U.S. Yet, the inaction of institutions, policymakers, and the public hinders significant progress toward climate-resilient infrastructure. We must convey the critical need for resilient infrastructure and allow these stakeholders to be more engaged in the process. In addition, in the built environment domain, with the increased focus on sustainable developments, recent buildings and infrastructures often contain sustainable systems, which are the integrated sustainable components in buildings or as Cabezas [10] defined them as systems that balance needs of human and the environment. However, many occupants are not aware of those sustainable systems, which may make some of them less efficient [11]. Moreover, it is worth mentioning that awareness, which

is knowledge about the sustainable system, is the first step to have the public engage with sustainable systems, which is the activities, decisions, and behaviours that promote sustainability.

Given the advancement of computing power, technologies such as augmented reality (AR), can make invisible, hard-to-imagine concepts visible to the public in meaningful ways. The use of AR has shown a positive impact on education, enhancing traditional learning methods and creating immersive and engaging

experiences [12] Also, AR was found to attract those who are less interested in the topic [13]. To this end, this paper explored the potential of AR applications to help the public visualize, adapt, and engage in a conversation about the sustainable systems our community depends on to build resilience. Connecting scientific literacy, visceral experiences, and compelling storytelling provide a powerful pathway to motivate policymakers and the public to take the crucial next steps needed to prepare for the climate crisis.

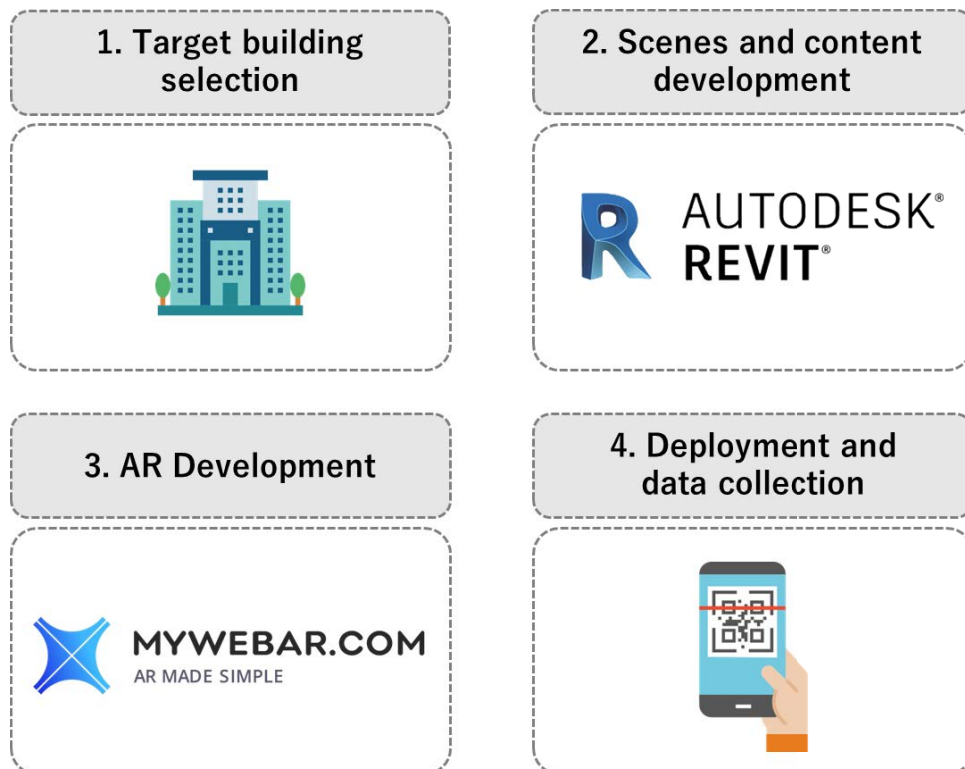


Figure 1. The proposed methodology

Additionally, to address the aforementioned challenges and better improve the awareness among building occupants regarding sustainable systems, this paper proposes a framework leveraging mobile AR to promote sustainability through increasing public awareness about sustainable systems in our daily lives. By scanning QR codes at specific locations within a building, community members and visitors can interact with the developed AR experience to learn about the building itself and sustainability principles.

2 Background

Despite an extensive body of research on sustainability, a notable gap exists in public awareness of sustainability

concepts. Many studies that investigated challenges with implementing sustainability found low public awareness among the top barriers [14]. For instance, AlSanad [15] conducted a study to evaluate the obstacles to sustainable construction in Kuwait and found that lack of awareness was ranked as the top barrier, followed by cost. While many of the studies that investigated the barriers were focused locally on different countries such as Chile [16] and Ghana [17], the lack of awareness or similar terms is almost always among the top barriers [5].

To tackle the issue of limited sustainability awareness, various researchers have investigated the use of AR technologies. For example, Bekaroo et al. [18] examined AR for educating about sustainable electronic device usage by showing energy consumption. Moreover,

Strada et al. [19] introduced a collaborative AR serious game focusing on urban development sustainability. Similarly, Wang et al. [20] developed an AR game to educate about climate change. In the circular economy context, Katika et al. [21] tested an AR engagement tool in two Greek case studies, focusing on enhancing public understanding. Huh et al. [13], who used AR to educate students about fine dust, mentioned that the use of AR triggers the students' interest and increases participation leading to improved learning activity. These examples demonstrated the positive impact of AR technologies in increasing sustainability awareness. This supports the findings of Alsaati et al. [22], who indicated that using AR in education can lead to environmental and economic benefits. Recently, Boncu et al. [23] thoroughly reviewed the use of mobile applications, including mobile AR, to improve public awareness. Their findings generally suggested that integrating these mobile technologies can be a highly effective approach for enhancing sustainability education.

While several studies leveraged AR to promote sustainability, existing literature focusing on sustainable building systems is still limited. Additionally, many of the solutions suggested in prior studies are not easily accessible to the public, as they often require specialized hardware or software that may not be readily available to everyone. Therefore, drawing inspiration from the success demonstrated by previous research that employed AR technologies, this paper proposed a framework for developing a web-based AR experience to enhance public engagement of sustainable systems.

3 Methodology

The framework for crafting a mobile AR experience focused on engaging the public in sustainability encompasses four steps: (1) Target building selection, (2) Scenes and content development, (3) AR development, and (4) Deployment and data collection. Figure 1 illustrates the proposed methodology of this study. Each step unfolds as follows:

1. Target building selection

The initial step involves targeting a building that incorporates sustainable building systems. For enhanced impact, it is recommended to select a building with high daily foot traffic. For example, campus buildings or public libraries tend to have high traffic, thus are appropriate for engaging the public given their high traffic. Upon selecting a facility, identifying sustainable design features within the building can be achieved by reviewing its Leadership in Energy and Environmental Design (LEED) documents or contacting the general contractor. It is crucial to opt for sustainable building

systems that are easily accessible and situated in or near high-traffic areas of a building, such as hallways or gates. Additionally, selecting a wide variety of building systems can help the public acquire a broader spectrum of knowledge. In this study, the authors selected several building systems (such as daylighting and water systems) in a campus building, which received LEED Silver green building certification. A detailed example use case is presented subsequently.

2. Scenes and content development

With a list of selected systems and locations to be displayed, scenes and contents can be crafted based on this list. For each system, the scene should feature multimedia content such as audio, video, text, or a combination thereof to introduce that system. Incorporating 3D models or interactions can enhance the AR experience, making it more engaging. For example, Autodesk® Revit® can be employed to create 3D models of that building or specific components. In one of our use cases, the authors leveraged Autodesk® Revit® to create 3D models and simulation of daylighting, providing the user with precise insights into natural light distribution and sustainable design features within the spaces. In this phase, developers should also strategize for the deployment of this AR experience. Within the framework, the authors created QR codes for each system selected and placed them in designated locations within the selected facility. By scanning the QR codes, users can be directed to the AR experience, elucidating the sustainable feature present in that location, and facilitating knowledge acquisition.

3. AR development

In this research, the authors utilize MyWebAR (<https://mywebar.com/>), a web-based AR development platform, to construct an AR experience accessible effortlessly through scanning QR codes, eliminating the need app downloads. This way, we can ensure that all users can access the experience through any type of mobile device. Following the preparation of all requisite resources, the files are imported into MyWebAR for mobile AR experience development. The AR experience undergoes multiple pilot tests to confirm accessibility across all devices and proper display of all content.

4. Deployment and data collection

After finalizing the AR experience, the QR codes are printed and placed in designated locations. The authors suggest gathering information for future analysis, such as the kind of device a user uses, how much time they spend on each system, and how many people scan QR codes. With this data, developers can attain a better understanding of the AR experience. This insight allows for iterative refinements to enhance user engagement.

4 Use cases

This study develops four AR experiences in lighting, daylighting, water refilling, and heating, ventilation, and air-conditioning (HVAC) systems. Let's consider one of our AR experiences, where a water refilling machine is an example of sustainable infrastructure within a campus building. Typically, building occupants (e.g., faculty, students, staff, and visitors) encounter this machine, yet

often without recognizing its important role as a sustainable feature designed to reduce the consumption of disposable plastic bottles. Moreover, users may not fully understand the machine's eco-conscious purpose, highlighting the need for our effective communications through AR to convey the sustainability message and engage users. To address the challenge, an AR experience of the water-refilling machine was developed, as visualized in Figure 2.

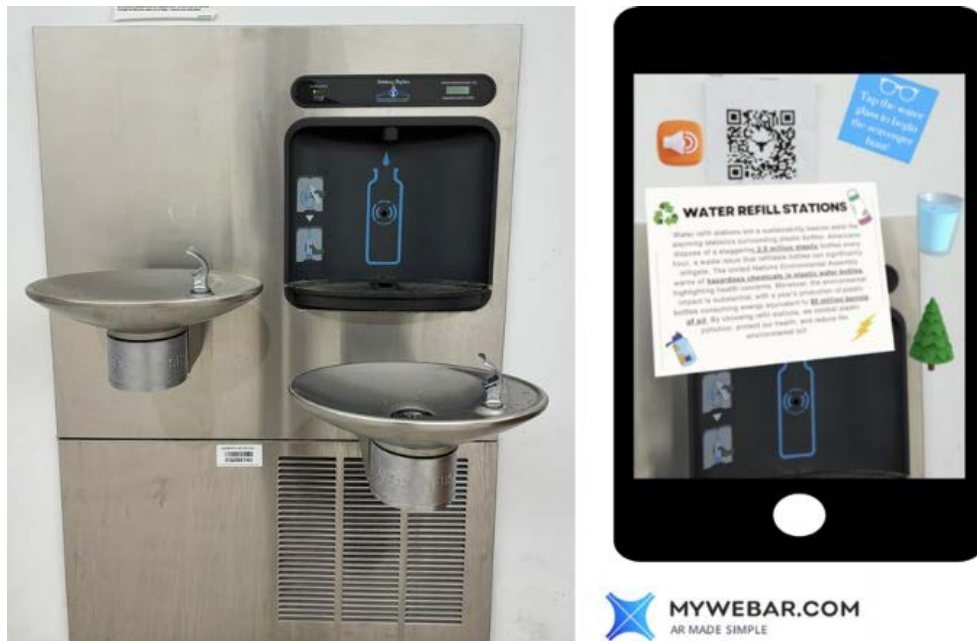


Figure 2. AR implementation: Introducing a water refilling station in a campus building.

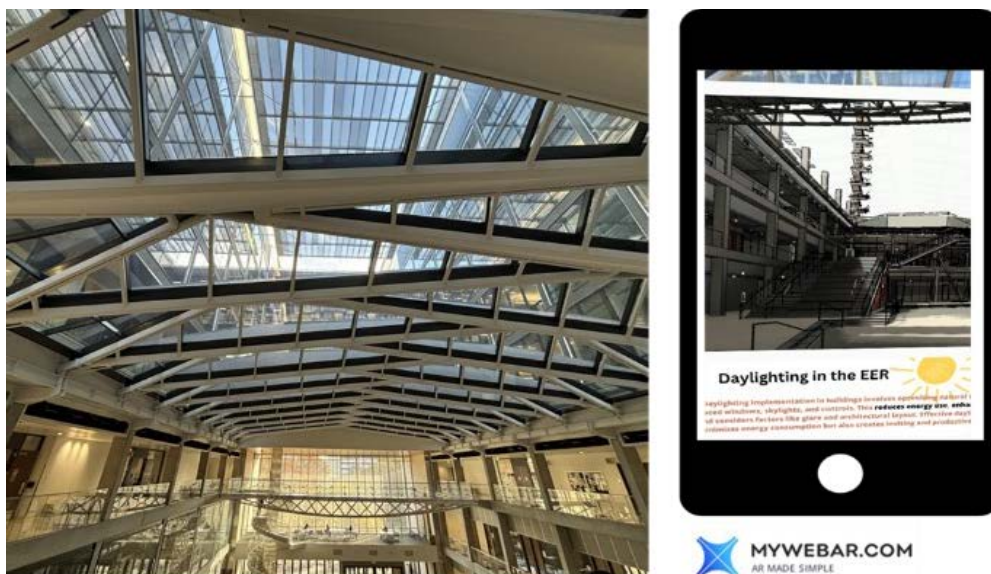


Figure 3. AR implementation: Introducing a daylighting system in a campus building.

The daylighting system is another example of a selected sustainable system in the same building, as shown in Figure 3. The roof structure design is an architectural feature that incorporates transparent materials to allow natural sunlight to penetrate interior spaces. These systems can provide multiple benefits, including energy savings through reduced reliance on artificial lighting, enhanced aesthetics, and improved occupant well-being by developing visually appealing indoor environments. By leveraging AR, users can immerse themselves into the developed AR environment and interactively gain a deeper insight into the sustainable benefits of daylighting systems. Overall, this AR-based approach enables users to actively engage with the concept of daylighting and its energy-saving potential.

In our future work, the authors plan to collect and analyze users' data, including tracking their usage frequency, accessed AR content (e.g., water refilling machine information), and engagement duration. Qualitative feedback will also be gathered through an online survey. Particularly, three types of data will be collected, including access counts, usage duration, and general feedback from users.

We will then use the data to evaluate the impact on public engagement in sustainable systems. Combining the quantitative or qualitative insights from the data, we will be able to evaluate the feasibility of AR in raising awareness and engagement regarding sustainability within the building. For example, access frequency and usage duration can be used to gauge user engagement levels while the user feedback provides qualitative insights into their experiences. Additionally, the general feedback from users provides qualitative insights into their experiences, offering strategies for improving the applications and identifying which elements are successful in enhancing public engagement in sustainability. Note that the metrics related to evaluating public awareness and engagement are crucial for measuring the effectiveness of our initiatives. The metrics will serve as important indicators for demonstrating effectiveness to engage the participants and to improve the awareness, which is also a potential direction of our future research.

5 Conclusions

This research explored the use of AR with an aim to improve the public engagement in sustainable systems. Existing literature has highlighted the need for raising awareness about sustainability as well as the challenge of communicating sustainability to the public despite the abundance of related technical publications. To address this issue, this manuscript proposed an AR-based framework to facilitate educating building occupants about sustainable systems. In this study, four sustainable

systems were identified within an engineering building on campus, which included lighting, daylighting, and HVAC systems. Subsequently, corresponding AR experiences were developed for these systems.

The research team faced some challenges when building this framework that are present in almost all existing AR applications, including the compatibility between software systems used, and the accuracy of augmentation. Due to the inherent limitations of paper-based presentations, our use case demos were confined to 2D imagery. It is worth noting that our use cases also include the visualization of 3D objects in motion and auditory instructions. In our future work, we will collect user data and analyze these data to assess user engagement levels and gather feedback. The collected data is beneficial for us to refine the AR experiences, identifying successful elements and strategies for enhancing public engagement in sustainability. Beyond this, as we continuously enhance the AR experience and expand its coverage to include a broader range of sustainable systems (e.g., sustainable construction materials), this visualization technology could contribute towards a more sustainability-aware society.

Acknowledgments

This work was supported by Planet Texas 2050, a research grand challenge at the University of Texas at Austin, and the University of Texas at Austin Green Fund. Planet Texas 2050's and the Green Fund's support are gratefully acknowledged. Any opinions, findings and conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of Planet Texas 2050.

References

- [1] T. F. Stocker *et al.*, "Climate Change 2013: The Physical Science Basis.," Intergovernmental Panel on Climate Change, Cambridge, United Kingdom and New York, NY, USA, 2013. Accessed: Nov. 08, 2023. [Online]. Available: <https://www.ipcc.ch/report/ar5/wg1/>
- [2] J. A. Du Pisani, "Sustainable development – historical roots of the concept," *Environ. Sci.*, vol. 3, no. 2, pp. 83–96, Jun. 2006, doi: 10.1080/15693430600688831.
- [3] K. E. Portney, *Sustainability*. in The MIT Press Essential Knowledge series. The MIT Press, 2015. Accessed: Nov. 08, 2023. [Online]. Available: <https://mitpress.mit.edu/9780262528504/sustainability/>
- [4] T. Mezher, D. Noamani, A. Abdul-Malak, and B. Maddah, "Analyzing sustainability knowledge in

- the Arab world,” *Sustain. Dev.*, vol. 19, no. 6, pp. 402–416, 2011, doi: 10.1002/sd.465.
- [5] A. Darko and A. P. C. Chan, “Review of Barriers to Green Building Adoption,” *Sustain. Dev.*, vol. 25, no. 3, pp. 167–179, 2017, doi: 10.1002/sd.1651.
- [6] Y.-C. Lee and F. Leite, “Enhancing Engineering Education on Circular Economy in Water Systems Leveraging Mixed Reality [Manuscript submitted for publication],” *Maseeh Dep. Civ. Archit. Environ. Eng. Univ. Tex. Austin*, 2024.
- [7] Y.-C. Lee and F. Leite, “Mixed Reality Promoting Circular Economy in Urban Water Systems,” in *Computing in Civil Engineering 2023*, Corvallis, Oregon, U.S.A: American Society of Civil Engineers, Jan. 2024, pp. 102–109. doi: 10.1061/9780784485248.013.
- [8] Y.-C. Lee, F. Leite, and K. Lieberknecht, “Prioritizing selection criteria of distributed circular water systems: A fuzzy based multi-criteria decision-making approach,” *J. Clean. Prod.*, vol. 417, p. 138073, Sep. 2023, doi: 10.1016/j.jclepro.2023.138073.
- [9] UNESCO, “Educating for a sustainable future: a transdisciplinary vision for concerted action,” Brazil, 1999. Accessed: Nov. 09, 2023. [Online]. Available: <https://unesdoc.unesco.org/ark:/48223/pf0000110686>
- [10] H. Cabezas, C. W. Pawlowski, A. L. Mayer, and N. T. Hoagland, “Sustainability: Ecological, Social, Economic, Technological, and Systems Perspectives,” in *Technological Choices for Sustainability*, S. K. Sikdar, P. Glavič, and R. Jain, Eds., Berlin, Heidelberg: Springer, 2004, pp. 37–64. doi: 10.1007/978-3-662-10270-1_3.
- [11] D. Grierson, “Towards a sustainable built environment,” *CIC Start Online Innov. Rev.*, no. 1, pp. 70–78, Dec. 2009.
- [12] M. Kesim and Y. Ozarslan, “Augmented Reality in Education: Current Technologies and the Potential for Education,” *Procedia - Soc. Behav. Sci.*, vol. 47, pp. 297–302, Jan. 2012, doi: 10.1016/j.sbspro.2012.06.654.
- [13] J. R. Huh, I.-J. Park, Y. Sunwoo, H. J. Choi, and K. J. Bhang, “Augmented Reality (AR)-Based Intervention to Enhance Awareness of Fine Dust in Sustainable Environments,” *Sustainability*, vol. 12, no. 23, Art. no. 23, Jan. 2020, doi: 10.3390/su12239874.
- [14] J. Á. Jaramillo, J. W. Z. Sossa, and G. L. O. Mendoza, “Barriers to sustainability for small and medium enterprises in the framework of sustainable development—Literature review,” *Bus. Strategy Environ.*, vol. 28, no. 4, pp. 512–524, 2019.
- [15] S. AlSanad, “Awareness, Drivers, Actions, and Barriers of Sustainable Construction in Kuwait,” *Procedia Eng.*, vol. 118, pp. 969–983, Jan. 2015, doi: 10.1016/j.proeng.2015.08.538.
- [16] A. Serpell, J. Kort, and S. Vera, “Awareness, actions, drivers and barriers of sustainable construction in Chile,” *Technol. Econ. Dev. Econ.*, vol. 19, no. 2, pp. 272–288, Jun. 2013, doi: 10.3846/20294913.2013.798597.
- [17] D.-G. J. Opoku, J. Ayarkwa, and K. Agyekum, “Barriers to environmental sustainability of construction projects,” *Smart Sustain. Built Environ.*, vol. 8, no. 4, pp. 292–306, Jan. 2019, doi: 10.1108/SASBE-08-2018-0040.
- [18] G. Bekaroo, R. Sungkur, P. Ramsamy, A. Okolo, and W. Moedeem, “Enhancing awareness on green consumption of electronic devices: The application of Augmented Reality,” *Sustain. Energy Technol. Assess.*, vol. 30, pp. 279–291, Dec. 2018, doi: 10.1016/j.seta.2018.10.016.
- [19] F. Strada *et al.*, “Leveraging a collaborative augmented reality serious game to promote sustainability awareness, commitment and adaptive problem-management,” *Int. J. Hum.-Comput. Stud.*, vol. 172, p. 102984, Apr. 2023, doi: 10.1016/j.ijhcs.2022.102984.
- [20] K. Wang, Z. D. Tekler, L. Cheah, D. Herremans, and L. Blessing, “Evaluating the Effectiveness of an Augmented Reality Game Promoting Environmental Action,” *Sustainability*, vol. 13, no. 24, p. 13912, 2021.
- [21] T. Katika, I. Karaseitanidis, D. Tsiakou, C. Makropoulos, and A. Amditis, “Augmented Reality (AR) Supporting Citizen Engagement in Circular Economy,” *Circ. Econ. Sustain.*, vol. 2, no. 3, pp. 1077–1104, Sep. 2022, doi: 10.1007/s43615-021-00137-7.
- [22] T. Alsaati, S. El-Nakla, and D. El-Nakla, “Level of Sustainability Awareness among University Students in the Eastern Province of Saudi Arabia,” *Sustainability*, vol. 12, no. 8, Art. no. 8, Jan. 2020, doi: 10.3390/su12083159.
- [23] Ștefan Boncu, O.-S. Candel, and N. L. Popa, “Gameful Green: A Systematic Review on the Use of Serious Computer Games and Gamified Mobile Apps to Foster Pro-Environmental Information, Attitudes and Behaviors,” *Sustainability*, vol. 14, no. 16, Art. no. 16, Jan. 2022, doi: 10.3390/su141610400.

Virtual Construction Safety Training System: How Does it Relate to and Affect Users' Emotions?

Fan Yang¹ and Jiansong Zhang, Ph.D.^{1*}

¹School of Construction Management Technology, Purdue University, United States of America

* Corresponding Author

yang2352@purdue.edu, zhan3062@purdue.edu

Abstract -

The construction industry is witnessing an increasing adoption of virtual reality (VR) technology for training and education purposes. Given this trend, it becomes essential to critically investigate the impact it has on learners, especially when compared to traditional paper-based learning method. In this paper, the authors developed a close-to-reality virtual system using the Unity3D game engine. Participants engage in learning safety protocols, operating a virtual crane, and assembling a steel structure within this environment. Corresponding paper-based instructional materials were also developed for comparison. The study involved 16 participants who were randomly assigned to either the VR training or the traditional paper-based training, their brainwaves data were recorded through electroencephalography (EEG) headset during the training progress to assess their emotions. Results show that an individual is most likely to experience exciting emotions when they are training in the VR system compared with the traditional training method. The correlation with actual safety performance, however, remains unclear and requires further investigation.

Keywords -

Virtual reality; Construction training; EEG; Emotion analysis

1 Introduction

Construction is a complex, high-risk industry due to the diverse set of skilled workers' involvement [1]. Consequently, the identification of hazards is of paramount importance within the construction industry. Extensive analysis of construction accidents reveals that unsafe site conditions, lack of proper training, unsafe worker behaviour, and unsafe method and task sequencing are the main causes of accidents [2]. Despite the construction industry's consistent emphasis on on-the-job training, security risk assessments using two-dimensional drawings make it difficult to identify many potential hazards [3].

The use of Virtual Reality (VR) technology in the context of construction training has attracted considerable interest in recent research publications. Several studies have

demonstrated how VR applications may help with critical issues that the construction industry faces, especially those related to skill development, safety training, and operational efficiency.

To address these challenges, the authors proposed a VR-based construction training system. This system aims to furnish users with comprehensive training before engaging in on-site activities. This highly realistic three-dimensional (3D) environment was developed using Unity 3D game engine, allowing the system to simulate detailed safety protocol training, crane operation and steel structure assembly procedures, thereby delivering virtual on-the-job training. Although there have been many studies using VR for construction training, a significant research gap exists in comparing the emotional impacts on users between virtual reality method in construction training and traditional paper-based training method. Emotions are key indicators of an individual's thoughts, underlying psychological conditions, and resultant behaviors [4]. Thus, discerning users' emotional states is vital for evaluating the efficacy of varying training approaches. Electroencephalography (EEG) signals can reflect people's emotional state in real-time objectively [5]. To ascertain the participant's emotional state during the training, this study employed EEG to quantitatively assess brain activity, providing a standardized gauge of user experience. In our controlled experiment, 16 participants were assigned to two groups for comparing VR and traditional learning methods, analyzing arousal and valence dimensions from EEG data.

2 Background

Despite attention to construction site injuries, accident rates in this high-risk industry double the industrial average [6], often due to limited safety knowledge and awareness among workers and engineers. Traditional education methods fail to provide adequate safety training for on-site performance [7].

2.1 Game simulation in construction education

The scope of traditional learning methods, such as paper-based materials, DVD videos and discussions has

been questioned as ineffective for the construction industry [8]. According to [9], active project-based approaches in construction education are more effective in fulfilling educational goals than traditional lecture-based methods. Especially when Generation Y students, known as the "Internet Generation," have transitioned into the workforce. They are more interested in game-based learning with an entertainment twist than traditional learning methods [10]. The use of game simulations for educational training dates back to 1991 with Microsoft Flight Simulator [11]. Despite not utilizing VR technology and instead relying on a combination of mouse, keyboard, and joysticks, the experiments have demonstrated a notable positive impact on learners and the method was subsequently widely used in the aircraft industry. With the technology advancement, game technology and game engines are now being increasingly utilized in the construction domain. For example, [12] proposed the integration of BIM and gaming, demonstrating a virtual, interactive framework's potential in design education; [13] described an innovative safety assessment approach using game technology, which involves immersing workers in 4D virtual risk scenarios specific to their projects, coupled with a range of potential action strategies, thereby enhancing their knowledge and skill levels in real-world situations; [14] introduced a BIM-based simulation method combining game technology and robotics to enhance automation in modular wood construction. And later, the same authors implemented this method in the game engine-based environment, where a workflow encompassing extraction of BIM information, development of the wood frame assembly process, and validation through game engine simulations was conducted, laying the groundwork for future productivity assessment with physical robotic integration [15].

2.2 Virtual reality in construction

An immersive, interactive digital environment opens new opportunities for enhancing educational experiences [16]. For example, [17] presented a virtual training program, in which workers actively engage in realistic scenarios, making decisions based on their knowledge or intuition. This interactive approach significantly enhances workers' safety awareness and skills; [18] conducted comparative training experiments, revealing that virtual reality-based construction safety training significantly enhances learning retention and attention span; [7] reviewed the need for effective safety education in construction, where traditional methods fall short in providing practical experience. It proposes an innovative online social VR system with role-playing and interactive learning modules, proven effective in enhancing construction safety education; [19] presented an immersive safety training environment for construction workers, addressing

poor hazard recognition in dynamic work environments. It emphasizes personalized, realistic training using 360° panoramic images and videos, offering a more immersive training environment with a high degree of presence. As VR gains prominence in construction, researchers have advocated for its inclusion as a standard component within the construction curriculum system [20], and also beginning to pay more attention to evaluating the effectiveness of VR in their studies. For example, [21] conducted a comparison between VR training and in-person training for the control task of a demolition robot. Experts evaluated the participants' behaviour along three dimensions: knowledge acquisition, operational skills, and safety behaviour, by measuring against specific targeted operational skills and safety behaviours in each learning module; [22] introduced a VR training module tailored for the precast/prestressed concrete industry, concentrating on PPE, strand tensioning, and suspended loads training. The module's efficacy compared with traditional video-based training methods was evaluated using questionnaires. The advent and evolution of VR headsets have marked a shift in research from screen-based VR to immersive headset experiences in recent years. While this change enhances immersion, it also brings issues like VR sickness in many users [23]. Consequently, it is evident that there is a growing emphasis among researchers on users' psychological and physiological responses in construction training.

2.3 Psychophysiological studies in construction

Construction safety is directly related to the physical and emotional states of construction workers which have been studied using different methods. For example, [24] used heart rates as a reference to measure user response to hazardous construction scenarios; [25] used physiological measurements of energy expenditure, including oxygen consumption and heart rate data, to assess construction workers' physiological thresholds. Recently, Heart rate, heart rate variability and skin temperature were used as the physiological impact of the construction VR experience study [26]. "EEG is a noninvasive measurement of the brain's electrical activity, which is generated by firing neurons in the brain" [27]. EEG data is commonly used to examine two fundamental properties of affective experience: valence and arousal. Arousal is associated with the user's level of excitement/calmness and valence is associated with the user's level of positivity/negativity [28]. There has been a century-long discussion on their relationship. According to Kron et al., [29] valence and arousal can be separated in emotional experiences. Based on the famous active learning theory [30], creating excitement in the classroom can motivate students to be active learners. Recent years have seen a surge in the use of EEG signals in construction studies, primarily because wire-

less wearable EEG devices have expanded its functionality non-intrusively, moreover, EEG offers rich information about an individual's mental status [27]. [31] compared three different light design alternatives by using VR simulation and the EEG of the participants was recorded to assess their emotions. [32] collected EEG data from users when they were using VR for construction safety training, through statistical modelling and machine learning-driven EEG analysis, 6 workers with high-risk health conditions were identified successfully. [33] provided a method for continuous monitoring of mental fatigue in construction workers, utilizing EEG data and advanced quantitative methods to prevent fall hazards.

3 Approach

This study was accomplished through the following four steps: in the first step, an immersive VR construction laboratory environment was developed in the Unity 3D game engine, version 2022.3.8f1. Unity is a versatile and widely used game engine, providing a comprehensive set of software solutions that facilitate the creation, execution, and monetization of interactive 3D content. Secondly, training elements were selected and integrated to enhance the ability of VR to interact with users, using the Oculus Quest as the wearable device and the OVR SDK in the Unity platform as the VR interaction toolkit, working in concert to form a prototype of the training system. Thirdly, paper-based training materials were derived for this study from lecture documents of the CM150 course at Purdue University. Finally, Comparative experiments were conducted and EEG data were collected through user experiments to determine the user's emotional state. This study employed the famous James Lange's theory to categorize emotions into two basic dimensions of emotional arousal and emotional valence, respectively [34]. Emotional arousal is the extent to which an individual's emotions are activated in reaction to a stimulus, ranging from passive to active. Conversely, emotional valence denotes the degree of pleasantness or unpleasantness of an experience, spanning from negative to positive.

4 Virtual Reality Training System

This training system is developed using the Unity 3D and Visual Studio, using primarily programming language C#. To experience the developed virtual world, the participant will need to wear the Oculus headset. Participants can navigate in the module using touch controllers. By using the left controller, the participant can move forward and backward, while the right controller helps to rotate 90 degrees. Each controller is represented as a virtual hand. These two virtual hands can be used to perform tasks in the virtual world. The audio and visual instructions are

provided whenever necessary to show the VR module of the developed hardware facility.

Before creating the VR environment for this study, guided by VR development theory [35], the authors addressed three critical questions:

1. *Q*: "Who is the user?" *A*: The user is defined as an individual with novice-level or limited experience in the construction domain.
2. *Q*: "What do the users need?" *A*: The users require a comprehensive grasp of crucial safety protocols, crane operations, and steel structure assembly before real-life construction site involvement.
3. *Q*: "Can VR meet these needs?" *A*: Some of the needs that surpass the capabilities of traditional paper-based learning methods can be addressed through the advantages offered by VR.

While fieldwork in steel assembly encompasses extensive detailed knowledge and hands-on techniques, replicating every aspect in a VR environment is impractical. Guided by VR design theories [35], we have selectively chosen aspects of this knowledge that are most amenable to VR representation, such as visual elements, 3D spatial interaction, physical manipulation, and procedural learning. Based on this, we developed the virtual reality system including three major parts: safety protocol training, crane operation, and steel structure assembly. The below sections present these three parts.

4.1 Safety protocol training

The authors sourced Personal Protective Equipment (PPE) models in .fbx format from open-source websites and imported them into Unity assets, including high visibility undershirts, safety goggles, protective hats, and hard-toe boots. The training guide's audio, corresponding to each PPE, was created using a text-to-speech artificial intelligence platform. Users will complete this part of training by grabbing the correct PPE and properly outfitting the worker model, while also listening to corresponding instructional audio throughout the process, as shown in Figure 1.

4.2 Crane operation

In the second phase of the training system, the authors imported a fixed-post jib crane model into Unity. The virtual crane replicated its real-world three-degree-of-freedom movement by enabling rotation, lateral movement, and vertical motion across three distinct parts of the crane model. Concurrently, a virtual operator panel mirroring a real-world counterpart was developed, featuring six buttons (i.e. Lift up, Lift down, Move Left, Move



Figure 1. Safety protocol training part in VR

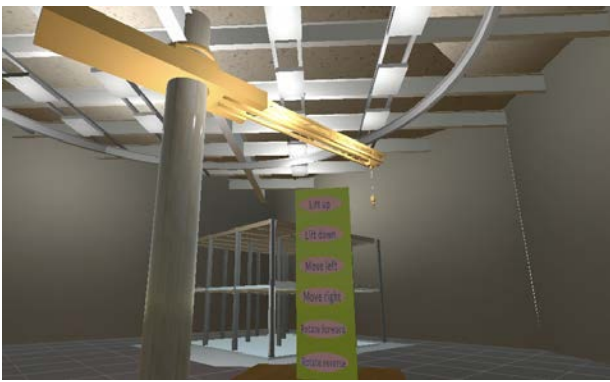


Figure 2. Crane operation part in VR

Right, Rotate Forward, Rotate Reverse) that allow users to control the crane's movement in six directions within the virtual environment, as shown in Figure 2.

4.3 Steel structure assembly

In the third part of the training, the users undertake the assembly of a classic steel structure, which is composed of columns, beams, floors and roof. The users will familiarize themselves with the steps involved in steel assembly starting by getting familiar with the steel components. Two steel columns and a steel beam are placed on a table, and the user needs to grab them in VR. When the user grabs a column or beam, a highlighting function will be triggered at the target position in the steel structure assembly space to guide the user there, and when the user places the component, there will be an alignment function to accurately align the component to the target position and angle to reduce the user's workload. Considering the numerous components in the steel structure, manually installing each one would be overly burdensome. To streamline this process, users are enabled to click buttons for batch installation, simplifying and accelerating the assembly, as shown in Figure 3.

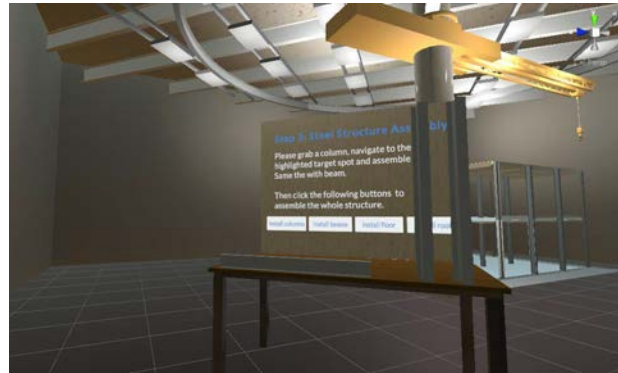


Figure 3. Steel structure assembly part in VR

4.4 Interaction with users

VR interface metaphors are the conceptual tools that translate the experiences of the real world into the digital environment to help users understand and interact with VR systems more intuitively. During the development of this study, key metaphors were implemented to facilitate user interaction within the VR environment:

1. Direct Manipulation: Enables users to reach out and grab objects directly by the virtual hands in the virtual environment.
2. Ray Casting: Allows for the selection of objects by casting a ray from the user's head or hand towards the target.
3. Movement: Empowers users to navigate and move freely throughout the VR environment.

Affordances design in VR refers to the practice of designing objects and environments that intuitively communicate their purpose and usage to the user. A good cognitive model requires that the behaviour of the mapped object meets expectations. Adhering to this design principle, the VR environment in this study features interactive elements such as clickable virtual buttons, objects that can be picked up, descriptive virtual billboards, and a control panel for operating the virtual crane.

5 Experiment

This study has recruited 16 participants including eight females and eight males. They reported no history of neurological diseases, mental/physical disabilities, cognitive impairment, or heart diseases, and were randomly divided into two groups: one for VR training and the other for traditional paper-based training. Following an initial demonstration and overview of the experimental procedures, participants in both groups engaged in self-directed learning. Details on each step will be described in the following sub-sections.

5.1 Device description

This study involves the use of two devices. Oculus Quest is a VR HMD with a headset and two hand controllers, allowing the users to view and operate the VR system. Emotiv EPOC+ EEG headset is a brain-computer interface for EEG raw data collection, it has 14 electrodes (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4) and two reference channels. In this study, it captures EEG data at a sampling rate of 128 Hz.

5.2 Experimental setup and procedures

Before commencing each experiment, participants were thoroughly briefed on the experimental procedure and the specific training objectives, this included detailed instructions on the functionalities of the equipment and its proper usage. For the VR-based training group, strict sanitary measures were implemented: the equipment was disinfected with alcohol wipes before each session. After that, participants received a comprehensive demonstration of the hand controller's fundamental operations, encompassing navigation, viewpoint alteration, virtual ray casting, and object manipulation within the VR environment. Then participants wear the VR device and the EEG device at the same time, immerse themselves in the VR environment detailed in the preceding section, following the provided clues and procedures within the VR to complete the training.

For the traditional training group, instructional documents were collected from the lecture slides and reading materials of CM150 course at Purdue University, a fundamental course for undergraduate construction management technology students. Relevant content covers: 1) Safety Training. Instructions on properly wearing helmets, boots, reflective vest, goggles, etc. before entering the construction site. 2) Crane Operation. The concept of crane degrees of freedom, corresponding control panel, and hazardous areas on construction sites with cranes. 3) Steel Structure Erection. Types of structural steel components, common parameters, sequence of construction, etc. These were printed out and used as materials for traditional learning. Figure 4 shows the scene of the experiment.

The EEG device's felt pads were fully hydrated with saline before starting each experiment, and after mounting them on the user's head, the contact quality and EEG signal quality of each channel were confirmed to be green using the Emotive Pro software. EEG data recording started after the commissioning of the EEG equipment was completed, and stopped after the participants signaled completion of VR training or paper-based training. Problems encountered during the data recording phase include:

1. The incongruity between wearing VR and EEG device at the same time led to mild discomfort for users.

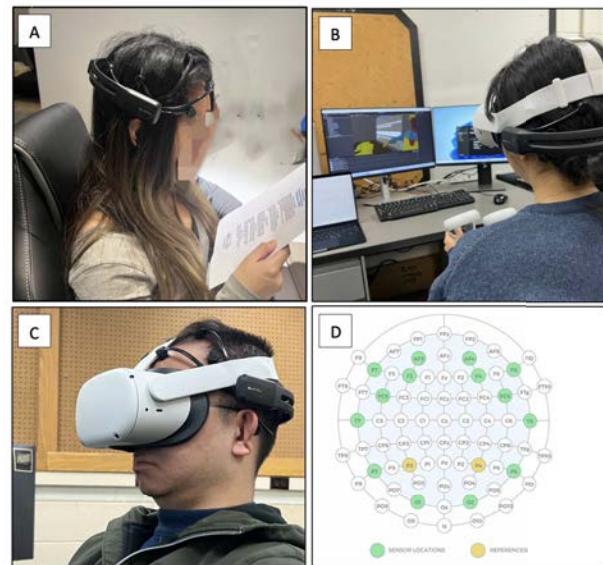


Figure 4. (a) Participant in paper-based training; (b) Participant in VR-based training; (c) Details of how a participant wears two devices; (d) Channels and references of EEG device.

2. Dense hair affects the quality of the electrode-scalp connection.
3. The EEG device that has been tuned for connectivity and signal quality was interfered with while wearing the VR device.
4. The bodily and head movements of users had impact on EEG signal quality, occasionally leading to transient signal disruptions.

5.3 Data acquisition and EEG data preprocessing

EEG raw data were acquired at 125 Hz sampling rate from users by using Emotiv EPOC+ device and Emotiv Pro software. EEG signals can be categorized by frequency into Delta wave, Theta wave, Alpha wave, Beta wave, and Gamma wave. Among them, alpha waves are in the range of 8-14 Hz, linked to activities involving mental coordination, calmness, alertness, and learning. An increase in alpha wave activity can indicate a state of relaxation or meditation. Beta waves are in the range of 14-30 HZ, present during states of alertness, problem-solving, decision-making, and focused mental activity. Beta waves are indicative of a highly engaged mind and are often encountered during tasks that require active thinking, focus, and concentration. The famous James Lange's theory divides emotions into two basic dimensions that are emotional arousal and emotional valence [34]. This study uses data from alpha and beta waves for arousal and valence

analysis was based on the calculation method proposed by [36], the arousal level of the mind can be determined by computing the ratio between the powers of beta and alpha waves (i.e., Beta/Alpha). The formula below was used to calculate the valence level using alpha and beta waves in channels F3 and F4:

$$Valence = Alpha_{F4}/Beta_{F4} - Alpha_{F3}/Beta_{F3}$$

Before further analysis, preprocessing the captured EEG signals is crucial. A high signal-to-noise ratio can significantly impact the data. In this study, the EEG data were preprocessed in EEGLAB (MATLAB toolbox) as the following steps:

1. The data were high-pass filtered at 1 Hz and low-pass filtered at 50 Hz.
2. Independent Component Analysis (ICA) was used to decompose the signals into 14 independent components. Then, these components were separated from heart, line noise, eye, and muscle components.
3. The data underwent a manual review to meticulously eliminate any residual artifacts, including eye blinks and other potential disturbances.

6 Results and Analysis

Figure 5 summarizes the emotional state analysis results for the 16 participants, the valence and arousal level for the two groups was shown as spots in the figure. We adopt Whissel's arousal-valence emotion space [37] as our taxonomy. Arousal (how exciting/calming) and valence (how positive/negative) are the two basic emotion dimensions found to be most important and universal. Based on the results, it was found that: (1) EEG data shows VR-based learning elicits higher arousal levels, showing that users felt significantly more excited about VR-based learning; (2) At the valence level, both paper-based and VR-based learning modalities exhibit comparable means and distributions, indicating that users did not show a significant difference in the degree of positivity or negativity towards VR-based learning compared to traditional paper-based learning methods; (3) The distribution of arousal levels in paper-based learning demonstrates greater centralization, whereas virtual VR learning exhibits a more dispersed pattern, implying that there are significant individual differences in participants' experience of how exciting/calming VR is compared to paper-based learning. Based on active learning theory [30], the excitement brought by a high arousal level can lead to a more active learning experience for users.

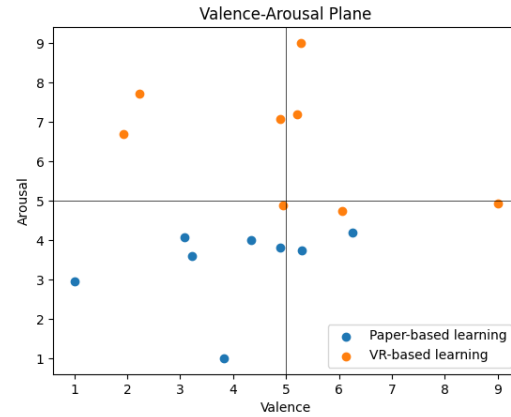


Figure 5. Arousal-valence emotion plane

7 Conclusion and Limitations

This research introduces a VR-based training system, designed in alignment with VR design theory, to virtualize the knowledge and skills involved in steel construction into three components: safety learning, crane operation, and steel assembly. A case study with 16 participants, divided into VR and traditional paper-based learning groups, was conducted to assess the impact of this VR system on users' emotional states. Analysis of EEG data revealed that VR-based learning induced a high arousal level, indicating a more exciting emotional state, yet both VR and traditional learning methods elicited similar valence (positive/negative) responses. This finding suggests that VR can effectively make users more excited, serving as a valuable adjunct to traditional paper-based methods. This finding suggests that VR can effectively stimulate user excitement by amplifying arousal levels, which leads to more active learning. However, the users who are in a negative valence space (such as aversive motivation or unpleasant emotions) should exercise caution with VR use. It has the potential to shift their emotional state from low-arousal negative valence (such as sadness) to high-arousal negative valence (such as anger) according to Whissell space [37].

Future work will explore how this data will be related to user experience and learning efficiency. The following limitations are acknowledged: 1) Despite efforts to align the content of VR and paper-based learning, the distinct nature of these mediums may lead to variations in learning stressors and experiences, thereby affecting EEG outcomes. 2) The participants' general unfamiliarity with VR technology could have influenced their emotional responses as they adapted to the VR skills, a factor absent in paper-based learning. Therefore, emotional changes related to device usage should be isolated in future studies. 3) With the growing use of machine learning algorithms

for EEG data analysis, further research comparing these methods' results is merited to gain deeper insights. 4) This study presents the initial phase of establishing a virtual construction lab, and therefore the focus limits to basic functionality and elements from learning materials. Future efforts will incorporate AI-assisted VR learning and scenario-based demonstrations, enriching the educational realism of the construction training experience.

8 Acknowledgement

The authors would like to thank the National Science Foundation (NSF). This material is based on work supported by the NSF under Grant No.1827733. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

References

- [1] Quang Tuan Le, Do Yeop Lee, and Chan Sik Park. A social network system for sharing construction safety and health knowledge. *Automation in Construction*, 46:30–37, 2014.
- [2] T. Michael Toole. Construction site safety roles. *Journal of Construction Engineering and Management*, 128(3):203–210, 2002.
- [3] Heng Li, Greg Chan, and Martin Skitmore. Visualizing safety assessment by integrating the use of game technology. *Automation in construction*, 22: 498–505, 2012.
- [4] Celia Shahnaz, SM Shafiul Hasan, et al. Emotion recognition based on wavelet analysis of empirical mode decomposed eeg signals responsive to music videos. In *2016 IEEE Region 10 Conference (TENCON)*, pages 424–427. IEEE, 2016.
- [5] Chaofei Yu and Mei Wang. Survey of emotion recognition methods using eeg information. *Cognitive Robotics*, 2:132–146, 2022.
- [6] Steve Rowlinson. *Construction safety management systems*. Routledge, 2004.
- [7] Quang Tuan Le, Akeem Pedro, and Chan Sik Park. A social virtual reality based construction safety education system for experiential learning. *Journal of Intelligent & Robotic Systems*, 79:487–506, 2015.
- [8] Shunsuke Someya, Kazuya Shide, Hiroaki Kani-sawa, Zi Yi Tan, and Kazuki Otsu. Research on the relationship between awareness and heart rate changes in the experience of safety education materials using vr technology. In *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*, volume 40, pages 333–340, 2023.
- [9] M Betts, SJ Rickard Liow, and RW Pollock. Different perceptions of importance of educational objectives. *Journal of Professional Issues in Engineering Education and Practice*, 119(3):317–327, 1993.
- [10] James Goedert, Yong Cho, Mahadevan Subramaniam, Haifeng Guo, and Ling Xiao. A framework for virtual interactive construction education (vice). *Automation in Construction*, 20(1):76–87, 2011.
- [11] William F Moroney and Brian W Moroney. Utilizing a microcomputer based flight simulation in teaching human factors in aviation. In *Proceedings of the Human Factors Society Annual Meeting*, volume 35, pages 523–527. SAGE Publications Sage CA: Los Angeles, CA, 1991.
- [12] Wei Yan, Charles Culp, and Robert Graf. Integrating bim and gaming for real-time interactive architectural visualization. *Automation in Construction*, 20(4): 446–458, 2011.
- [13] Heng Li, Greg Chan, and Martin Skitmore. Visualizing safety assessment by integrating the use of game technology. *Automation in Construction*, 22: 498–505, 2012. doi:10.1016/j.autcon.2011.11.009.
- [14] Oscar Wong Chong and Jiansong Zhang. Game Simulation to Support Construction Automation in Modular Construction Using BIM and Robotics Technology—Stage I. In *Computing in Civil Engineering 2019*, Atlanta, Georgia, 2019. American Society of Civil Engineers. doi:10.1061/9780784482438.048.
- [15] Oscar Wong Chong, Jiansong Zhang, Richard M. Voyles, and Byung-Cheol Min. BIM-based simulation of construction robotics in the assembly process of wood frames. *Automation in Construction*, 137: 104194, 2022. doi:10.1016/j.autcon.2022.104194.
- [16] James Paul Gee. What video games have to teach us about learning and literacy. *Computers in entertainment (CIE)*, 1(1):20–20, 2003.
- [17] Dong Zhao and Yincheng Ye. Using virtual environments simulation to improve construction safety: An application of 3d online-game based training. In *Future Control and Automation: Proceedings of the 2nd International Conference on Future Control and Automation (ICFCA 2012)-Volume 1*, pages 269–277. Springer, 2012.

- [18] Amotz Perlman Rafael Sacks and Ronen Barak. Construction safety training using immersive virtual reality. *Construction Management and Economics*, 31(9):1005–1017, 2013. doi:10.1080/01446193.2013.828844.
- [19] Idris Jeelani, Kevin Han, and Alex Albert. *Development of Immersive Personalized Training Environment for Construction Workers*. 2017. doi:10.1061/9780784480830.050.
- [20] Jeffrey Kim and Tom Leathem. Virtual Reality as a Standard in the Construction Management Curriculum. 2018.
- [21] Pooya Adami, Patrick B. Rodrigues, Peter J. Woods, Burcin Becerik-Gerber, Lucio Soibelman, Yasemin Copur-Gencturk, and Gale Lucas. Effectiveness of VR-based training on improving construction workers' knowledge, skills, and safety behavior in robotic teleoperation. *Advanced Engineering Informatics*, 50:101431, 2021. doi:10.1016/j.aei.2021.101431.
- [22] Sayali Joshi, Michael Hamilton, Robert Warren, Danny Faucett, Wenmeng Tian, Yu Wang, and Junfeng Ma. Implementing Virtual Reality technology for safety training in the precast/prestressed concrete industry. *Applied Ergonomics*, 90:103286, 2021. doi:10.1016/j.apergo.2020.103286.
- [23] Eunhee Chang, Hyun Taek Kim, and Byounghyun Yoo. Virtual Reality Sickness: A Review of Causes and Measurements. *International Journal of Human-Computer Interaction*, 36(17):1658–1682, 2020. doi:10.1080/10447318.2020.1778351.
- [24] Shunsuke Someya, Kazuya Shide, Hiroaki Kani-sawa, Zi Yi Tan, and Kazuki Otsu. Research on the Relationship between Awareness and Heart Rate Changes in the Experience of Safety Education Materials Using VR Technology. In *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*, volume 40, pages 333–340, Waterloo, Canada, 2023.
- [25] Tariq S Abdelhamid and John G Everett. Physiological demands during construction work. *Journal of construction engineering and management*, 128(5): 427–437, 2002.
- [26] Gilles Albeaino, Patrick Brophy, Idris Jeelani, Masoud Gheisari, and Raja R. A. Issa. Psychophysiological Impacts of Working at Different Distances from Drones on Construction Sites. *Journal of Computing in Civil Engineering*, 37(5):04023026, 2023. doi:10.1061/JCCEE5.CPENG-5225.
- [27] Houtan Jebelli, Sungjoo Hwang, and SangHyun Lee. Eeg-based workers' stress recognition at construction sites. *Automation in Construction*, 93:315–324, 2018.
- [28] Richard D Lane, Phyllis ML Chua, and Raymond J Dolan. Common effects of emotional valence, arousal and attention on neural activation during visual processing of pictures. *Neuropsychologia*, 37(9):989–997, 1999.
- [29] Assaf Kron, Maryna Pilkiw, Jasmin Banaei, Ariel Goldstein, and Adam Keith Anderson. Are valence and arousal separable in emotional experience? *Emotion*, 15(1):35, 2015.
- [30] Charles C Bonwell and James A Eison. *Active learning: Creating excitement in the classroom*. ASHE-ERIC higher education reports. ERIC, 1991.
- [31] Nidia Bucarelli, Jiansong Zhang, and Chao Wang. Maintainability Assessment of Light Design Using Game Simulation, Virtual Reality, and Brain Sensing Technologies. pages 378–387, 2018. doi:10.1061/9780784481264.037.
- [32] Dongjin Huang, Xianglong Wang, Jinhua Liu, Jinyao Li, and Wen Tang. Virtual reality safety training using deep EEG-net and physiology data. *The Visual Computer*, 38(4):1195–1207, 2022. doi:10.1007/s00371-021-02140-3.
- [33] Behnam M. Tehrani, Jun Wang, and Dennis Truax. Assessment of mental fatigue using electroencephalography (EEG) and virtual reality (VR) for construction fall hazard prevention. *Engineering, Construction and Architectural Management*, 29(9): 3593–3616, 2021. doi:10.1108/ECAM-01-2021-0017.
- [34] Peter J Lang. The emotion probe: Studies of motivation and attention. *American psychologist*, 50(5): 372, 1995.
- [35] Sri Kurniawan. Interaction design: Beyond human-computer interaction by preece, sharp and rogers. *Universal Access in the Information Society*, 3:289–289, 2004.
- [36] Rafael Ramirez and Zacharias Vamvakousis. Detecting emotion from eeg signals using the emotive epoc device. In *International Conference on Brain Informatics*, pages 175–184. Springer, 2012.
- [37] Cynthia M Whissell. The dictionary of affect in language. In *The measurement of emotions*, pages 113–131. Elsevier, 1989.

REBAPose -A Computer vision based Musculoskeletal Disorder Risk Assessment Framework

Sivakumar K S¹, Nikhil Bugalia² and Benny Raphael³

^{1,2,3}Department of Civil Engineering, Indian Institute of Technology Madras, Chennai, India
ce22s018@smail.iitm.ac.in , nikhilbugalia@gmail.com , benny@civil.iitm.ac.in

Abstract –

Work-related Musculoskeletal Disorders (WMSDs) are one of the prominent challenges facing the construction industry, and their effective management requires a risk-quantification strategy. A comprehensive method to assess WMSDs for the construction sector is the Rapid Entire Body Assessment (REBA). The conventional risk-quantification frameworks relying on manual analysis of postures and images are resource-intensive and not scalable. Hence, automated approaches using Computer Vision (CV) are gaining attention. Current CV-based methods to automatically estimate REBA scores do not track all the key points necessary for accurate computation; they rely on heuristics. The present study is part of a wider research effort to develop a CV-based, fully automatic REBA risk calculator. An important task in this effort is to develop and validate a key-point annotation strategy for accurately estimating various body parts and joints necessary for REBA score calculation. This study re-annotated 149,813 human images present in the open-access COCO dataset. About 0.5 million key points were added for three body parts: the midpoint of head, neck, and center hip. Then, this data was used to train the state-of-the-art CV algorithms, MMPose and Alphapose. A comparison of the ML models developed on the newly annotated data with the pre-existing heuristics-based approaches demonstrates significant performance gains in precision and accuracy. The proposed model has the potential to be widely adopted for the precise and quick estimation of the REBA WMSD risk-assessment framework in different industries.

Keywords –

Ergonomics; Computer vision; Pose analytics; REBA Score; COCO; Construction.

1 Introduction

The construction industry worldwide continues to perform poorly regarding Occupational Health and

Safety (OHS) issues [1][2]. Among these, Work-related Musculoskeletal Disorders (WMSD), preventable disorders affecting nerves, tendons, muscles, and supporting structures, are receiving increasing academic and practitioner attention due to their significant impact on the construction industry. For example, 41% of registered workers in Hong Kong reported facing WMSD [3]. In the USA, workday loss due to WMSD has increased from 8 days in 1992 to 13 days in 2014 [4]. Hence, developing an appropriate WMSD risk-assessment framework is an important task.

Observation-based methods have been prominently used for WMSD risk quantification and management in construction. In these methods, the worker's postures are observed through work-sampling approaches and classified into risk categories. Several risk-assessment frameworks, such as OVAKO, REBA, and RULA, have been extensively used on construction sites for WMSD risk assessment [5][6]. An important step in this is to classify postures in images and videos into risk categories. However, these conventional observation-based approaches rely on extensive manual inputs. Hence, they are not scalable to the needs of the rapidly increasing construction sector. To overcome this problem, Computer Vision (CV) has been used to identify postures in images and quickly analyze large quantities of images and videos. A highly accurate CV-based risk assessment framework for OVAKO exists [7]. However, to the authors' knowledge, an accurate REBA score estimator has not yet been developed, even though few previous studies have attempted heuristic approaches [8]. Compared to OVAKO, REBA is a comprehensive and accurate risk-assessment framework that is more widely applicable for construction. Therefore, a CV-based approach for accurate REBA estimation may significantly benefit many practitioners.

The present study is part of a wider research effort to develop a CV-based, fully automatic REBA risk calculator called REBAPose. An important task in this effort is to accurately locate various body parts and joints in the images. This work is focused on the task mentioned above and aims to develop and validate a strategy for key-point annotation. The performance of the proposed

strategy is compared with the currently adopted heuristics approach to estimate the key points necessary for REBA score estimation. The scope of the present study is limited to developing an accurate training dataset by adding 3 more annotation points to the existing COCO dataset of 149813 images and evaluating the performance of ML algorithms on detecting these newly annotated key points. Such an assessment is expected to make the REBA scoring assessment more accurate.

The study is structured as follows. Section 2 provides an overview of the literature and identifies the essential gaps. Section 3 describes the essentials of CV-based WMSD assessment and the analytical methodology adopted in the current study. Results have been summarized in section 4, followed by discussions in section 5. Conclusions have been outlined in section 6.

2 Literature Review

2.1 Overview of WMSD assessment frameworks

Several methods of WMSD risk assessment are used in construction and other industries. Popular direct observation methods are the OVAKO working posture assessment system (OWAS) [9], Rapid Entire Body Assessment (REBA) [10], and Rapid Upper Limb Assessment (RULA) [11]. REBA is a common assessment method with high accuracy. Recent extensive literature on the topic suggests that REBA is likely to be a comprehensive and accurate analysis method for the construction industry [6]. However, a generic challenge with all observation-based methods is their extensive reliance on manual intervention for analyzing human postures (identifying body parts joints and angles between them) from large quantities of images collected at the site. Hence, a CV-based approach for automatic REBA risk estimation is a great tool of practical relevance that could aid the work of safety practitioners.

2.2 Computer vision frameworks for posture analysis

In recent years, extensive research has been carried out in automatic human pose estimation using images for various applications. Several advanced Machine Learning (ML) algorithms have been developed, such as MMPose [12], and Alphapose [13]. These have been reported to have achieved more than 90% accuracy in identifying various joints and body parts on large-scale standard datasets such as the open-access COCO validation dataset. The current study leverages these standardized algorithms to develop REBAPose.

2.3 Available Posture Datasets

A fundamental need for any CV-based algorithm is an annotated dataset for training the ML algorithms. Over the years, several large-scale standardized publicly available human-annotated datasets have been created, allowing researchers from several areas to quickly develop and validate their own ML approaches. Popular datasets include MS COCO [14], MPII [15], Human36M [16], and AIC[17]. Using pre-existing datasets has several advantages compared to developing a new pose estimation model from scratch. Manual efforts can be reduced, and comparative assessments become possible using published results.

For a reliable estimate of the REBA score, 18 key points need to be estimated. However, none of the existing datasets can track all these 18 key points necessary for REBA score estimation (see Table 1).

Table 1. Body parts and Datasets

Body Part	COCO	BlazePose	MPII	AIC	REBA
Midpoint of Head	X	X	X	X	Yes
Nose	✓	✓	X	X	Yes
Neck	X	X	✓	✓	Yes
Right Shoulder	✓	✓	✓	✓	Yes
Left shoulder	✓	✓	✓	✓	Yes
Right Elbow	✓	✓	✓	✓	Yes
Left Elbow	✓	✓	✓	✓	Yes
Right Wrist	✓	✓	✓	✓	Yes
Left Wrist	✓	✓	✓	✓	Yes
R Index Finger	X*	✓	X	X	Yes
L Index Finger	X*	✓	X	X	Yes
Right Hip	✓	✓	✓	✓	Yes
Left Hip	✓	✓	✓	✓	Yes
Center Hip	X	X	X	X	Yes
R Knee	✓	✓	✓	✓	Yes
L Knee	✓	✓	✓	✓	Yes
R Ankle	✓	✓	✓	✓	Yes
L Ankle	✓	✓	✓	✓	Yes
Total KPs	13	15	13	13	18

X* - Not available in the original COCO dataset but is available in COCO Whole body dataset

Due to the limitations of the dataset, many previous studies on automatic REBA score estimators use heuristics to estimate the 3 key points necessary for REBA score calculation. For example, previous studies have relied on the point “Nose,” available in the COCO dataset, instead of the midpoint of head for angle calculations. Section 3.4 of the current study presents more examples of such heuristics.

Overall, it is evident that none of the existing datasets are sufficient to be readily used for accurate REBA posture analytics, and annotation on additional key points may be necessary for training and testing purposes of the CV model. The current study adopts to improve the existing annotations of the COCO dataset, as this dataset covers huge variations in the type of images collected and has been extensively used in previous studies, allowing for easy benchmarking.

3 Methodology

3.1 Methodology Overview

A three-step process is adopted in the current study and is summarized in Figure 1. The steps include annotation, training, and testing. The detail of each step is presented in the subsequent sections. As indicated in Table 1, key points at the midpoint of HEAD, NECK, and Center HIP are required in addition to the information already available in the pre-annotated COCO dataset. These three points were annotated in the current study to prepare an extensive new dataset for accurate REBA score estimation. It is important to note that the information on the two index fingers is also required for REBA score estimation. While such information was unavailable in the original COCO dataset, it was available from the COCO Whole Body dataset implemented in another study [18]. Hence, these readily available key points have been assimilated in the current study without the need to re-annotate them.

3.2 Annotation Framework

Crowdsourcing is extensively used to annotate large quantities of datasets. However, the quality of crowdsourced work is difficult to assess and control. Hence, we deployed 3rd party vendor to annotate the images for 3 new pose points. Employees of this organization are specialized in annotation services. The current study sought the help of these paid professional data annotators with extensive efforts to quality control. For this purpose, a toolset compatible with the JSON files in the COCO dataset was developed specifically to annotate pose key points. The tool can use the existing annotated key points from datasets such as COCO and build on top of it to annotate additional points.

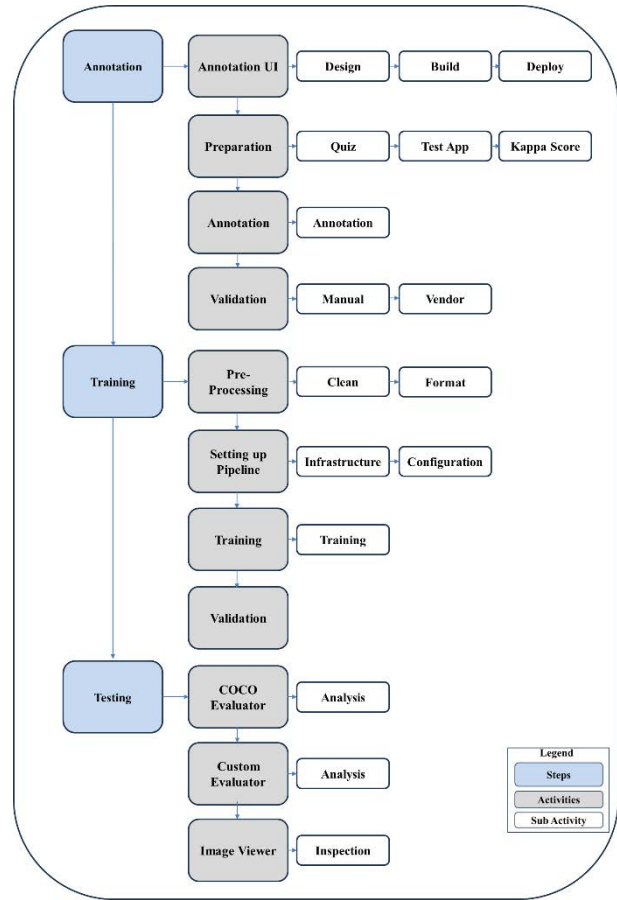


Figure 1. Methodology overview

3.2.1 Annotation UI details

Figure 2 shows the annotation toolset's intuitive user interface. The central area showcases the bounding box of the human being annotated. The pixel coordinates of each key point, along with their human interpretable names, are shown on the left side. Various annotation support functions are on the right side. To help in faster annotation, the framework enabled quick shortcut keys, such as Zoom functionality for the image and changing the color palette for better visibility. The tool was web-based so that multiple annotators could work simultaneously.



Figure 2. Annotation toolset UI Screen

3.2.2 Methods to minimize the bias in annotation

The current study devised a rigorous approach to ensure the annotation process's quality across the 20 different professional annotators. Before the annotation process, each annotator had to complete a ten-question quiz. The questions provided specific directions for certain tasks and could be answered based on information available only for successful completion. Annotators who scored above 90% are made eligible to carry out an annotation. Inter-rater reliability of all the annotators through kappa statistics [19] was also assessed on 100 images. Kappa statistics compares the agreement between two or more annotators by calculating the degree of agreement without considering the probability of consistency due to chance. The coefficient ranges between 0 (likely to be coincidence chance) and 1 (perfect agreement). A kappa value above 0.6 represents an adequate annotator agreement. The current study obtained a Kappa score of 0.67 for 20 annotators.

Once the annotation task was complete, the third-party annotation company was asked to re-check the quality of annotations using different personnel. The research team also randomly checked the annotation for 1000 persons and ensured the quality of the work done. Through this extensive process, an entire training set of 149813 persons and a validation set of 6352 persons were annotated for each of the three new key points. This annotated dataset is used for training and evaluation during testing and can be made accessible to all upon a reasonable request.

3.3 Training of ML Models

Training is a process where machine learning models learn from labeled data. The key steps in this process are preprocessing, setting up ML architecture, training, and validation.

3.3.1 Pre-processing of data

In this pre-processing step, data is cleaned and arranged so that models can learn efficiently. A single image in the COCO dataset may contain multiple persons; hence, a top-down approach is used to train the ML algorithms. The bounding boxes are first generated for every person in the image. Each bounding box is then taken as a single JSON file and shown to annotators for ease of key point annotation. Hence, multiple JSON files could contain annotation information for a single image in the COCO dataset. However, for training purposes, the algorithms require all information in a single file; hence, all individual annotated files are clubbed into a single file using the bounding box information and image ID.

3.3.2 Setting up the ML Model Architecture

For retraining purposes using the modified key points,

the current study leverages the ML algorithm architectures of the two state-of-the-art ML models known to have good accuracy on the COCO dataset. These two ML algorithms are MMPose and Alphapose. Both these architectures adopt the most accurate HRNet V2 [20] as the backbone for training.

3.3.3 Training

The two ML models were initialized based on the optimal hyperparameters reported in the previous literature [12][13]. The training validation is calculated at every 5th epoch, and a decision is made whether to stop the training or continue. Percentage Correct Key Point (PCK) is widely used as an accuracy parameter in pose analytics [22]. It is defined as the percentage of key points detected correctly compared to the total number of key points available in the annotated dataset of ground truth. The key points are predicted correctly if the prediction lies within a circle of threshold distance radius, i.e., 20% of the diagonal length of the bounding box [22]. In the current study, the training was stopped at the 100th epoch as PCK at 20% threshold value reached an accuracy of 94.34% and loss (Mean Squared Error - MSE) saturated at a low value of 0.0012 (See Figure 3). Overall, training took around 3 days for the total 149813 images with 100 epochs.

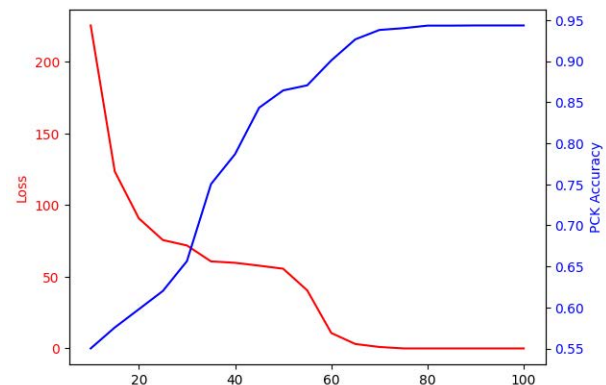


Figure 3. Training accuracy and Loss values

3.4 Testing

To confirm the quality of the newly trained ML model with 18 key points, evaluation is carried out by comparing the key points predicted by the ML algorithm to the validation data annotated by professional humans. Consistent with the state-of-the-art literature on the topic, several metrics were used to evaluate the performance of the ML model, such as a.) The difference in pixel distances between ML predicted and annotated key points b.) commonly used metrics such as Average precision (AP) and Average Recall (AR) based on Object Key point Similarity (OKS) [21] criteria, and c.) Computational performance in inference time, i.e., time taken to detect the key points from the given image.

Further, multi-stage experiments have been designed to evaluate the proposed REBAPose's comparative performance compared to several ML algorithms. In the first stage, the performance of Alphapose and MMPose algorithms is compared with other commonly adopted ML algorithms, such as the Detectron, Mediapipe, YOLO, Movenet, and PoseNet, for the original COCO dataset. Such a step is necessary to evaluate if state-of-the-art algorithms such as Alphapose and MMPose are the best choice, even for the added number of key points.

In the second stage, the predictions of the newly trained models and heuristic methods are compared. One heuristics method [8] computes the five key points as follows: The neck point is estimated as the midpoint of the ML algorithm's right and left shoulder prediction. The hip point is estimated by dividing the distance between the right and left hip points by 2. The midpoint of the head is estimated as the midpoint of the right and left eye points. Index finger values are derived by adding 0.02 times the diagonal length of the bounding box to the wrist point.

Testing is performed on unseen images that are not used for training the algorithm. The COCO dataset contains a set of images for evaluation and a default evaluator algorithm. However, a custom evaluation set and code were necessary for the current study for several reasons. First, the COCO validation dataset image contains many attributes, such as object detection, segmentation, and key points about human joints. Hence, not all images in the validation dataset can be evaluated; only images with humans become relevant. Out of 5000 images, only 2346 with 6352 persons could be used.

The default COCO validation dataset does not contain information about all the persons and the corresponding bounding boxes and key points. This is due to low-precision annotation done by a few annotators in the earlier annotation work. Few images in the COCO dataset have multiple humans. However, the COCO validation dataset has only a few persons annotated with key points and not all. Many modern algorithms can detect more humans than what is annotated in the default COCO validation dataset. The number of persons detected varies depending on the ML algorithm. For example, the *detectron* algorithm could detect about 11,000 persons, whereas the media pipe algorithm could detect only 2000 persons in the same dataset. Hence, only the images of 245 persons commonly detected by all the algorithms are used for comparative evaluation in the current study.

4 Results

4.1 Visual Comparison

Figure 4 visually compares the results obtained from

the newly trained REBAPose models and the heuristic approaches. The three essential key points for REBA, i.e., the neck, midpoint of head, and center hip, as estimated by the heuristic approaches, are shown using the blue dots in Figure 4. The same key points estimated based on the newly annotated data are shown using the yellow dots.



Figure 4. Visual representation of study results

4.2 Pixel Distance Comparison

A comprehensive heatmap of differences in pixel distances between ML predictions and annotated data for various algorithms for each of the body joints is shown in Figure 5. The data in Figure 5 shows the average results obtained for 245 persons (148 single-person images, 97 persons from images containing more than 1 person) from the COCO validation dataset. In Figure 5, except for the results of the two REBAPose architectures, i.e., REBAPose (Alphapose) and REBAPose (MMPose), the results for body parts (midpoint of head, neck, and center hip) from the remaining algorithms are estimated based on heuristics approaches, as described in section 3.4. For REBAPose (Alphapose) and REBAPose (MMPose), the results are obtained using retraining of the algorithm based on the newly annotated key points. Several important inferences can be made from these results.

First, state-of-the-art methods such as MMPose and Alphapose perform better than commonly adopted frameworks, even when only heuristics-based methods are considered. Such results reaffirm the technical superiority of these ML architectures.

The results for the heuristics method for the REBA essential key points, such as neck, midpoint of head, and center hip, further demonstrate that the heuristics-based methods for estimating the three-body parts result in large errors compared to the annotated ground truth, irrespective of the ML methodology applied.

Finally, the results demonstrate that Alphapose and MMPose architectures, when applied to the newly annotated data, greatly improved the prediction accuracy for all three key points that are newly annotated and re-

trained, while there is a slight improvement on two index finger points which are retrained and derived COCO whole

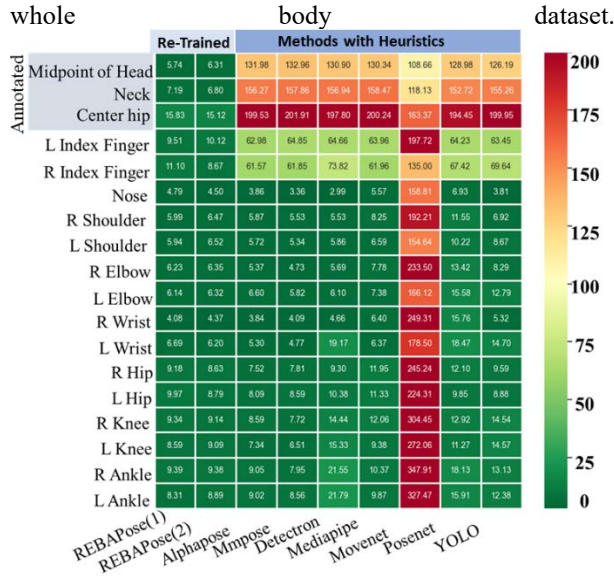


Figure 5. Pixel variation of detection and ground truth

4.3 OKS Analysis

OKS-based criteria are often used to assess how far or close the ML predictions are compared to the ground truth to estimate the performance of CV algorithms for pose estimations. OKS considers both the body part area and the image scale and is defined as a threshold value like 0.5 or 0.75. If the CV model predicted key point pixel is within the threshold distance governed by OKS criteria of the annotated pixel, the prediction is deemed *True Positive*. Similarly, information about false positives, false negatives, etc., can be estimated, and such matrices then help evaluate the necessary metrics, such as the Average Precision (AP) and Average Recall (AR), to understand the efficiency of pose estimation.

A higher OKS threshold is less permissive and requires key points to be more precisely aligned to be considered correct. The results from accuracy parameters at OKS thresholds 0.5 (loose metric), and 0.75 (strict metric), along with their Mean AP (Average of values at 0.5 and 0.95) are given in Tables 2 and 3.

Table 2. Average Precision Results

Methodology	mAP	AP @0.5	AP @0.75
REBAPose(A)	0.588	0.833	0.677
REBAPose(M)	0.530	0.800	0.584
Alphapose	0.224	0.733	0.015
MMPose	0.240	0.767	0.019

REBAPose (A) is based on Alphapose architecture,

and REBAPose (M) is based on MMPose architecture.

Table 3. Average Recall Results

Methodology	mAR	AR@0.5	AR @0.75
REBAPose(A)	0.686	0.877	0.785
REBAPose(M)	0.642	0.858	0.723
Alphapose	0.316	0.817	0.094
MMPose	0.329	0.834	0.108

Results indicate that the extensive annotation and retraining on 5 out of 18 REBA key points done in the current study (for REBAPose (A) and (M) algorithms) has resulted in multi-fold accuracy improvement, which is expected for the critical WMSD assessment framework.

4.4 Computational Performance

For CV-based WMSD assessment, the inference time is important, as real-time estimation is often desirable. The quicker models can potentially be used even on mobile phones. Table 4 compares the speed of models.

Table 4. Inference Results

Methodology	Inference time in seconds
REBAPose(M)	0.2025
REBAPose(A)	0.0388
Detectron	0.1138
YOLO	0.0898
Mediapipe	0.0252
Movenet	0.0161
Posenet	0.0169

5 Discussion

5.1.1 Technical advantages of REBAPose

Accurate estimation of the key points (or the body parts) is one of the first steps toward an accurate REBA score calculation. If the body parts are not correctly predicted, the angles between them will have errors directly affecting the REBA score estimations. The REBAPose estimator developed in this study has significantly higher accuracy compared to conventionally adopted heuristics-based methods (see results in Figure 5, Tables 2 and 3). Due to annotation and retraining, there was improvement towards precise detection in pixel difference between detected and annotated ground truth for the neck, center hip, and midpoint of the head. For example, for the midpoint of head, REBAPose variation from ground truth was just about 5 pixels (Figure 5), while for other heuristic approaches such a variation was a few orders of magnitude apart at about 130 pixels (see Figure 5). Similar improvements can also be seen for the two other

newly trained points as compared to heuristics approaches (Figure 5). Due to such a significant improvement in the accuracies of each of the three newly annotated points, the overall accuracy of REBAPose is also improved by about 100 % (see Tables 2 and 3).

Another important parameter to measure the performance of ML models is inference time. Results in Table 4, indicate the inference time of all models. Results indicate that the REBAPose (Alphapose), despite the high accuracy does not suffer from longer inference time and is the best model regarding both improved accuracy and inference time.

5.2 Key academic contribution

One of the essential academic contributions of the current study is the creation of a unified and comprehensive dataset that can be readily used for WMSD risk-assessment tasks, especially using the REBA method. Such a task has been achieved by developing an integrated annotated dataset comprising 13 key points from the original COCO dataset (Table 1), 2 key points annotated from the COCO whole body dataset (Left and Right Index figures), and 3 newly annotated key points (neck, midpoint of head and center hip) for about 1,50,000 images available in the COCO dataset. Previous studies have not carried out such an extensive annotation. Previous studies have retrained the datasets with only a limited number of images.

REBAPose essentially removes the heuristic approach on three body parts neck, midpoint of head, and center hip, and hence provides accurate angle measurements forming the basis for REBA scoring. Because of annotation and retraining, the accuracy of all points improved.

The extensively annotated dataset can be used for training and testing future CV-based algorithms for higher technical performance. Moreover, the REBAPose estimator developed as part of the current study can be used as a pre-trained model for accurately estimating the human postures for REBA scores in various fields where there is a concern about the OHS performance of their workers.

The newly annotated information about the center hip and the neck can enable a more accurate estimation of the spine, which is necessary for correcting yoga postures or for medical researchers treating spondylitis.

5.3 Limitations and Future Work

While the currently developed REBAPose model is industry agnostic, its accuracies can be further evaluated for specific trades. For example, future studies could test the applicability of the REBAPose models for construction-specific datasets such as the MOCS datasets [23]. The present study focussed only on 2D pose

estimation and did not address 3D pose improvements which may be covered in future work. As the REBA score depends on angles between body parts, 3D pose estimation is expected to get more precise angles than 2D points.

In the current study, human-annotated data was considered as the ground truth to estimate the accuracy of the work. However, the accuracy of these human annotations must be validated. The data from actual safety practitioners, such as industrial hygienists, could also be collected and compared against the predictions made by REBAPose. REBA scoring mechanism involves computing three scores A, B and C and these scores rely on angles formed between body parts. Since the present study can accurately detect body parts (key joints), it is likely to find precise angles that are formed between body parts. So present study is likely to have better REBA score calculations to reflect real ground conditions. However, the impact of an improved algorithm for enhanced key point detection performance on the enhancement of the accuracy of the REBA score needs to be validated in future studies.

The human pose estimation is just one of the components in REBA score calculation. Several other factors cannot be determined using a CV-based pose estimation technique. Examples include the weight of the objects held by personnel and the coupling factors. In future studies, this aspect will be considered.

6 Conclusions

Existing approaches for WMSD risk assessment using REBA rely on readily available pre-annotated datasets. While such approaches using heuristics for estimating REBA scores are easy and quick, they lack precision. This study aims to address this gap by developing new REBAPose estimators. More than 150,000 images selected from the COCO dataset were re-annotated to add 3 essential key points for calculating the REBA score. The REBAPose estimators in this study use state-of-the-art ML architectures, including Alphapose and MMPose. The proposed approach results in a 200% improvement in mAP and mAR values compared to predictions made using the heuristics approaches (see Tables 2 and 3). The inference time of the REBAPose estimators is also comparable, making them a fast and accurate method for pose estimations in future studies.

Acknowledgments

We acknowledge the contributions of Hridhhi Mondal, Antony Samuel, and Likhith Kumara of Tata Consulting Services (TCS) for building and deploying annotation UI. We are also thankful for the exploratory research grant provided by IIT Madras (Project Number

RF22230544CERFER008976).

Data Availability

The dataset used in the current study can be made available for academic use upon written request to authors.

References

- [1] K. Bhattacharjee, N. Bugalia, and A. Mahalingam, "An analysis of safety practices for small, medium, and large construction projects: A resilience engineering perspective," *Saf Sci*, vol. 169, Jan. 2024, doi: 10.1016/j.ssci.2023.106330.
- [2] N. Bugalia, V. Tarani, J. Kedia, and H. Gadekar, "Machine Learning-Based Automated Classification Of Worker-Reported Safety Reports In Construction," *Journal of Information Technology in Construction*, vol. 27, pp. 926–950, 2022, doi: 10.36680/j.itcon.2022.045.
- [3] W. Yi and A. Chan, "Health profile of construction workers in Hong Kong," *Int J Environ Res Public Health*, vol. 13, no. 12, Dec. 2016, doi: 10.3390/ijerph13121232.
- [4] X. Wang, X. S. Dong, S. D. Choi, and J. Dement, "Work-related musculoskeletal disorders among construction workers in the United States from 1992 to 2014," *Occup Environ Med*, vol. 74, no. 5, pp. 374–380, May 2017, doi: 10.1136/oemed-2016-103943.
- [5] D. Kee, "Systematic Comparison of OWAS, RULA, and REBA Based on a Literature Review," *International Journal of Environmental Research and Public Health*, vol. 19, no. 1, MDPI, Jan. 01, 2022, doi: 10.3390/ijerph19010595.
- [6] T. Jones and S. Kumar, "Comparison of ergonomic risk assessment output in four sawmill jobs," *International Journal of Occupational Safety and Ergonomics*, vol. 16, no. 1, pp. 105–111, 2010, doi: 10.1080/10803548.2010.11076834.
- [7] S. Pinzke and L. Kopp, "Marker-less systems for tracking working postures: Results from two experiments," 2001.
- [8] S. O. Jeong and J. Kook, "CREBAS: Computer-Based REBA Evaluation System for Wood Manufacturers Using MediaPipe," *Applied Sciences (Switzerland)*, vol. 13, no. 2, Jan. 2023, doi: 10.3390/app13020938.
- [9] O. Karhu, P. Kansil, and I. Kuorinka, "Correcting working postures in industry: A practical method for analysis," 1977.
- [10] S. Hignett and L. M. Ergonomist, "Rapid Entire Body Assessment (REBA)," 2000.
- [11] L. Mcatamney and E. N. Corlett, "RULA: a survey method for the investigation of work-related upper limb disorders," 1993.
- [12] T. Jiang et al., "RTMPose: Real-Time Multi-Person Pose Estimation based on MMPose," Mar. 2023, [Online]. Available: <http://arxiv.org/abs/2303.07399>
- [13] H.-S. Fang et al., "AlphaPose: Whole-Body Regional Multi-Person Pose Estimation and Tracking in Real-Time," Nov. 2022, [Online]. Available: <http://arxiv.org/abs/2211.03375>
- [14] T.-Y. Lin et al., "Microsoft COCO: Common Objects in Context," May 2014, [Online]. Available: <http://arxiv.org/abs/1405.0312>
- [15] M. Andriluka, L. Pishchulin, P. Gehler, and B. Schiele, "2D human pose estimation: New benchmark and state of the art analysis," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, Sep. 2014, pp. 3686–3693. doi: 10.1109/CVPR.2014.471.
- [16] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu, "Human3.6M: Large scale datasets and predictive methods for 3D human sensing in natural environments," *IEEE Trans Pattern Anal Mach Intell*, vol. 36, no. 7, pp. 1325–1339, 2014, doi: 10.1109/TPAMI.2013.248.
- [17] J. Wu et al., "AI Challenger: A Large-scale Dataset for Going Deeper in Image Understanding," Nov. 2017, doi: 10.1109/ICME.2019.00256.
- [18] S. Jin et al., "Whole-Body Human Pose Estimation in the Wild," Jul. 2020, [Online]. Available: <http://arxiv.org/abs/2007.11858>
- [19] J. Cohen, "A coefficient of agreement for nominal scales," *Educ Psychol Meas*, vol. 20, no. 1, pp. 37–46, 1960.
- [20] Z. Zhao, A. Song, S. Zheng, Q. Xiong, and J. Guo, "DSC-HRNet: a lightweight teaching pose estimation model with depthwise separable convolution and deep high-resolution representation learning in computer-aided education," *International Journal of Information Technology (Singapore)*, vol. 15, no. 5, pp. 2373–2385, Jun. 2023, doi: 10.1007/s41870-023-01297-5.
- [21] "COCO evaluation." Accessed: Dec. 14, 2023. [Online]. Available: <https://cocodataset.org/#keypoints-eval>
- [22] Y. Yang and D. Ramanan, "Articulated human detection with flexible mixtures of parts," *IEEE Trans Pattern Anal Mach Intell*, vol. 35, no. 12, pp. 2878–2890, 2013, doi: 10.1109/TPAMI.2012.261.
- [23] A. Xuehui, Z. Li, L. Zuguang, W. Chengzhi, L. Pengfei, and L. Zhiwei, "Dataset and benchmark for detecting moving objects in construction sites," *Autom Constr*, vol. 122, Feb. 2021, doi: 10.1016/j.autcon.2020.103482.

Developing Joint-level Scoring Models Tailored to Whole-body Ergonomic Assessment of Construction Workers

Zirui Li¹, Yantao Yu², and Qiming Li¹

¹Department of Construction and Real Estate, Southeast University

²Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology
seulzr@126.com, ceyantao@ust.hk, njlqming@163.com

Abstract –

Construction workers frequently engage in manual operations at workplaces, increasing their ergonomic risks of developing Work-Related Musculoskeletal Disorders (WMSDs). To proactively assess and prevent such risks, ergonomic scales have been widely employed, incorporated cutting-edge technologies to achieve advancements in automation. However, these scales have demonstrated limited accuracy in risk identification, mainly attributed to unreliable joint-level assessment rules based on discrete boundaries and binary rules. Although previous attempts have incorporated fuzzy logic to improve accuracy, the involved subjective determination of function shapes and threshold settings remains a persistent hindrance. To address this limitation, the present study aims to develop data-driven joint-level scoring models for replacing these conventional rules. This process leverages pose data from the Construction Motion Data Library (CML) dataset and employs a robust and heuristic data-driven approach named Heuristics Gaussian Cloud Transformation (H-GCT). The results, with all Confusion degrees below the threshold value of 0.64, demonstrate the significant independence of the developed scoring models, ensuring accurate identification of ergonomic risk. Furthermore, a comparison is conducted with previous studies that employed fuzzy logic to improve REBA. This process highlights the superiority of the data-driven H-GCT in developing scoring models. This study contributes to the existing body of knowledge by providing joint-level scoring models to improve the applicability of ergonomic assessment in construction. Future studies can further enhance this work by expanding pose data, enriching assessment modules, and refining the data-driven approach.

Keywords –

Postural ergonomic risk assessment; Data-driven approach; WMSDs prevention; Occupational health and safety

1 Introduction

Construction workers frequently engage in physically demanding manual operations at workplaces [1]. As a result, they face remarkable ergonomic risks of developing Work-Related Musculoskeletal Disorders (WMSDs). In Hong Kong, statistics from the Pilot Medical Examination Scheme (PMES) reveal that around 41% of registered construction workers suffer from WMSDs-related injuries [2]. Moreover, amidst the challenges posed by an aging labor force, escalating labor wage, and manpower shortages in construction, ergonomic risks may give rise to increasingly serious repercussions. Therefore, it is crucial to assess and prevent ergonomic risks of workers' operations.

To proactively prevent ergonomic risks for construction workers, various systematic observation methods have been employed. These methods typically involve the assessment by experienced experts and incorporate observational ergonomic scales such as Rapid Entire Body Assessment (REBA) [3], Rapid Upper Limb Assessment (RULA) [4], and Ovako Working Postures Assessment System (OWAS) [5]. Serving as assessment tools, these scales define the rules for coding posture-related data (e.g., joint angles) and subsequently rate ergonomic risks for construction workers based on the data [6]. For instance, REBA codes the postures of the trunk, neck, legs, upper arms, lower arms, and wrists, while assigning whole-body postures with ergonomic risk scores ranging from 1 to 15.

However, attributed to manual implementation, these practical methods are subject to notable constraints in objectivity and cost-effectiveness. To address these challenges, recent advancements have introduced cutting-edge data collection technologies, such as computer vision [7] and wearable sensors [8], to enable automated detection of ergonomic risks. In line with these initiatives, several studies have adopted machine learning and deep learning algorithms to classify postural ergonomic risks. For example, Zhang, Yan [9] compared the accuracy of multiple classification algorithms (BP-

ANN, DT, SVM, KNN, and EC) and developed optimal posture classifiers for the arms, back, and legs. Moreover, from the perspective of whole-body posture, Antwi-Afari, Qarout [8] trained a gated recurrent units (GRU) network to accurately classify five types of awkward working postures. However, an evident limitation emerges as the insufficient interpretability regarding the relationship between inputs and outputs in these black-box classification methods, particularly in comparison to risk rating based on well-established rules. In response to this challenge, a significant body of research transfers 2D [10] or 3D [7] joint coordinates extracted by pose estimation technologies into joint parameters. These joint parameters further serve as inputs for ergonomic scales such as REBA.

Despite significant progress in automated implementation of ergonomic scales, these scales have demonstrated limited accuracy in assessing ergonomic risk for construction workers [11]. This limitation is primarily attributed to the unreliable rules for joint-level assessment. Specifically, these scales rely on discrete boundaries and binary rules (specified angle ranges and positions) to determine risk categories for assessed modules [3, 4]. However, the formulation of these ranges for classifying joint-level risks lacks reliable support by statistical data on joint angles obtained from construction workers [12]. Furthermore, the oversimplified binary rules are susceptible to subjectivity resulting from manual observations and errors introduced by the pose estimation tools. Despite attempts to improve the joint-level rules by fuzzy logic [11, 13], the subjective determinations of function shapes and transition ranges continue to hinder the accuracy. Therefore, the development of joint-level scoring models, utilizing extensive 3D pose data of construction workers and a robust data-driven algorithm, holds potential for improvement. This approach allows for proactive mining of implicit knowledge from voluminous data and

effectively captures the characteristics of variations in joint-level risks [14, 15], thereby enhancing the credibility of scoring models.

To address the limitation in joint-level assessment rules of ergonomic scales, this study attempts to develop data-driven scoring models for replacing the unreliable discrete boundaries and binary rules. Given the objectivity of the REBA scale, which scores ergonomic risks based on joint angle values [7], and its emphasis on whole-body assessment covering critical body segments such as the trunk, neck, legs, upper arms, lower arms, and wrists, the joint-level modules of REBA serve as the foundational framework in this study. To achieve the goal, this study implements Heuristic Gaussian Cloud Transformation (H-GCT) [16], which consists of two phases: (1) Heuristic Gaussian Transformation (H-GT) for data clustering and (2) Forward Cloud Generation (FCG) for enhancing uncertainty representation at the boundaries between adjacent scoring models. Prior to implementing this robust data-driven approach, statistical data on joint angles are extracted from the Construction Motion Data Library (CML) dataset. Subsequently, these developed scoring models are evaluated using Confusion degree (CD) values. This indicator effectively reflects the independence of scoring models and their accuracy in risk identification. Ultimately, a comparison is implemented with previous studies that adopted fuzzy logic to improve REBA, aiming to highlight the superiority of the data-driven H-GCT in developing scoring models.

2 Methodology

Figure 1 illustrates the workflow of the scoring model development process, which involves extracting statistical data from CML dataset and developing scoring models through H-GCT.

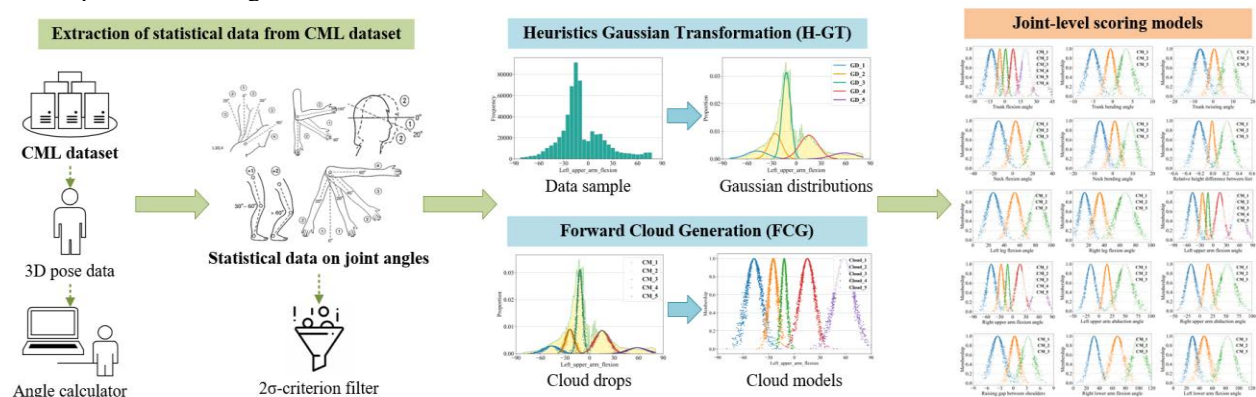


Figure 1. Workflow for the development of joint-level scoring models using H-GCT

2.1 Extracting statistical data from CML dataset

To obtain statistical data on joint angles of construction workers, the Construction Motion Data Library (CML) developed by Tian, Li [17] is selected. This dataset contains 4,333 samples of construction-related activities specifically curated for ergonomics analysis, covering five major types observed on construction sites, namely production activities, unsafe activities, awkward activities, common activities, and other activities. The dataset provides useful 3D pose data (20-joint system) for analysis.

In this study, based on REBA [3], 15 joint-level assessment modules are selected to cover five critical body segments: trunk, neck, legs, upper arms, and lower arms. Table 1 provides an overview of these modules and corresponding score distributions. Taking the Trunk Flexion module as an example, its assessment rules in REBA are as follows: an upright position (i.e., 0° flexion) is assigned a score of 1, a flexion or extension angle of 0-20° is assigned a score of 2, a flexion angle of 20°-60° or an extension angle greater than 20° is assigned a score of 3, and a flexion angle greater than 60° is assigned a score of 4. Following the progression of trunk flexion angles from negative to positive (with extension angles defined as negative), the corresponding scores change sequentially as 3, 2, 1, 2, 3, and 4, respectively (as illustrated in Table 1). Using the extracted 3D pose data, a vector-based calculator is utilized to obtain statistical data on joint angles for the 15 modules. Prior to developing the scoring models, the 2σ -criterion is applied, which sets the threshold at two standard deviations from the mean [18]. This criterion effectively eliminates potential outliers while preserving underlying data distributions.

2.2 Developing scoring models through H-GCT

Building upon the filtered statistical data on joint angles, the Heuristic Gaussian Cloud Transformation (H-GCT) is subsequently utilized to generate scoring models. Compared to other data-driven heuristic and clustering algorithms, such as Gaussian Mixture Model (GMM) [19] and K-Means [20], H-GCT effectively leverages prior knowledge from REBA to predefine the number of generated models while showcasing the capability to synthetically describe uncertainty [16]. The implementation of H-GCT follows two steps [21, 22]: Firstly, the H-GT generates a predefined number of clusters that align with joint-level assessment modules of REBA (i.e., the Gaussian distributions), based on the input data samples. Subsequently, the Forward Cloud Generation (FCG) develops cloud model to enhance

uncertainty representation at boundaries between adjacent scoring models.

2.2.1 Data clustering by H-GT

Based on the score distributions of the assessment modules, the Heuristic Gaussian Transformation (H-GT) is selected to generate a predefined number of Gaussian Distributions (GDs) [19]. The H-GT algorithm leverages prior knowledge from REBA to determine the number of clusters, denoted as M , within each specific module. This allows a set of data samples to be partitioned into a superposition of M GDs. For a random variable x in the problem domain, which represents joint angles for each module, the frequency distribution function $p(x)$ can be constructed. After the H-GT process, the mathematical expression for $p(x)$ is given by formula (1) [16].

$$p(x) \rightarrow \sum_{i=1}^M (w_i G) \quad (1)$$

In formula (1):

$$G = \frac{1}{\sqrt{(2\pi)^d |Cov_i|}} e^{-\frac{1}{2}(x-\mu_i)^T \Sigma_i^{-1}(x-\mu_i)} \quad (2)$$

Among these parameters, w_i , μ_i , and Cov_i represent the amplitude, expectation, and covariance matrix of the i -th GD after transformation, respectively, with the w_i satisfying the condition $\sum_{i=1}^M w_i = 1$. Furthermore, d represents the dimension of the data sample while M is the predefined number of clusters.

Table 1 An overview of considered assessment modules

Joint-level assessment module	Score distribution
Trunk Flexion	3, 2, 1, 2, 3, 4
Trunk Bending	1, 0, 1
Trunk Twisting	1, 0, 1
Neck Flexion	2, 1, 2
Neck Adjustment	1, 0, 1
Legs Support	2, 1, 2
Legs Flexion (left)	0, 1, 2
Legs Flexion (right)	0, 1, 2
Upper Arms Flexion (left)	2, 1, 2, 3, 4
Upper Arms Flexion (right)	2, 1, 2, 3, 4
Upper Arms Abduction (left)	1, 0, 1
Upper Arms Abduction (right)	1, 0, 1
Gap between Shoulders	1, 0, 1
Lower Arms Flexion (left)	2, 1, 2
Lower Arms Flexion (right)	2, 1, 2

During the H-GT process, the Expectation Maximization (EM) algorithm is employed as an iterative optimization technique for parameter estimation in GDs with hidden variables [23]. It combines with Maximum Likelihood Estimation (MLE) and involves two main steps: the E-step for posterior probability calculation and the M-step for reassessment and optimization.

For the implementation, the EM algorithm is configured to iterate 100 times by default. The number of components, which corresponds to the generated GDs, is determined for each module based on the respective score distribution. Powered by the scikit-learn library and Python (version 3.9.12), the statistical data on joint angles for the 15 considered modules undergo transformation into 15 sets of GDs.

2.2.2 Cloud models generation by FCG

The H-GT process facilitates the transformation of continuous data samples into a superposition of multiple GDs. Each element within the generated GDs is associated with a membership degree determined by the probability values on each relevant GD (one-to-one correspondence). However, the absence of interpreting probable one-to-many correspondence at the boundaries between adjacent models introduces uncertainty. To address this, the Cloud Model (CM), functioning as a cognitive model capable of synthetically describing the uncertainty, is employed [24]. Consequently, it optimizes the robustness of assessment.

During the FCG process, three numerical characteristics, namely Ex (expectation), En (entropy), and He (hyper entropy) are routinely employed to provide a comprehensive representation of a CM. Specifically, Ex represents the most representative data sample within a cluster, En quantifies the granularity scale of the cluster, and He portrays the uncertainty of the cluster granularity. As a result, a set of CMs can serve as the specific scoring models for a considered assessment module [25].

To determine the above Ex , En , and He before generation of CMs, the following steps are involved:

(i) For the k -th GD after H-GT, its mean and standard deviation are μ_k and σ_k , respectively.

(ii) The current standard deviation is considered as the maximum granularity parameter of the concept, while keeping the expectation constant to achieve equal scaling. Subsequently, the scaling ratios α_1 and α_2 of the k -th GD are computed to avoid overlap between adjacent clusters. These scaling ratios are determined using the following formulas (3) and (4) [16]:

$$\mu_{k-1} + 3\alpha_1\sigma_{k-1} = \mu_k - 3\alpha_1\sigma_k \quad (3)$$

$$\mu_k + 3\alpha_2\sigma_k = \mu_{k+1} - 3\alpha_2\sigma_{k+1} \quad (4)$$

Then, the variation range of standard deviation caused by unclear conceptual partition in the k -th GD is $[\alpha \times \sigma_k, \sigma_k]$, where $\alpha = \min(\alpha_1, \alpha_2)$.

(iii) According to the theory of Gaussian Cloud, that is, the standard deviation follows a GD, En is the expectation of standard deviation, and He is the standard deviation of the standard deviation [16]. The parameters and Confusion degree (CD) of the k -th CM can be determined as formulas (5) to (8) [16]:

$$Ex_k = \mu_k \quad (5)$$

$$En_k = (1 + \alpha) \times \sigma_k / 2 \quad (6)$$

$$He_k = (1 - \alpha) \times \sigma_k / 6 \quad (7)$$

$$CD_k = 3 \times He_k / En_k = (1 - \alpha) / (1 + \alpha) \quad (8)$$

Following the FCG, the 15 sets of GDs corresponding to the considered modules are transformed into 15 sets of CMs. These CMs serve as the scoring models for the assessment. Each CM consists of numerous cloud drops, which allow for the representation of uncertainty between adjacent models.

Table 2 Confusion degrees of all generated CMs

Assessment module	Confusion degrees (CDs)
Trunk Flexion	0.56, 0.56, 0.51, 0.45, 0.38, 0.36
Trunk Bending	0.43, 0.49, 0.49
Trunk Twisting	0.53, 0.53, 0.51
Neck Flexion	0.46, 0.46, 0.36
Neck Adjustment	0.47, 0.47, 0.46
Legs Support	0.45, 0.45, 0.38
Legs Flexion (left)	0.40, 0.40, 0.39
Legs Flexion (right)	0.50, 0.50, 0.39
Upper Arms Flexion (left)	0.47, 0.51, 0.51, 0.29, 0.29
Upper Arms Flexion (right)	0.48, 0.48, 0.46, 0.30, 0.28
Upper Arms Abduction (left)	0.31, 0.40, 0.40
Upper Arms Abduction (right)	0.33, 0.40, 0.40
Gap between Shoulders	0.46, 0.55, 0.55
Lower Arms Flexion (left)	0.59, 0.59, 0.40
Lower Arms Flexion (right)	0.31, 0.42, 0.42

3 Results and discussions

3.1 Developed joint-level scoring models

Through the H-GCT process, 15 sets of scoring models are developed for the 15 considered joint-level assessment modules in Table 1. As shown in Figure 2, these scoring models consist of multiple CMs (cloud models) that intuitively represent the correspondences between input joint angles and membership degrees of each risk category. The CMs synthetically describe the one-to-many uncertainty at the edges of adjacent scoring models, thereby optimizing the robustness of joint-level ergonomic assessment.

During the assessment process, the joint parameters for a specific module (e.g., joint angles) are inputted into

the corresponding joint-level scoring model (Figure 2). The output of this process is a fuzzy set, which comprises a series of membership degrees. Each degree corresponds to an individual CM included in the scoring model. To calculate the joint-level ergonomic risk score, the Center of Area (CoA) method is employed for defuzzification, benefiting from its comprehensive consideration of membership information [11]. In the case of the Trunk Flexion module, assuming an input angle value of 30° , the resulting membership degrees from the six CMs are $2.81\text{e-}14$, $6.91\text{e-}75$, $5.82\text{e-}50$, $3.33\text{e-}17$, $1.48\text{e-}02$, and $9.85\text{e-}01$, respectively. According to Table 1, these CMs correspond to the score values of 3, 2, 1, 2, 3, and 4. By utilizing the CoA method for defuzzification, the risk

score for this module is determined to be 3.99.

To evaluate the performance of the H-GCT, Confusion degree (CD) is utilized as a measure of the independence between generated CMs. This indicator exclusively quantifies the degree of overlap between adjacent CMs. It is worth noting that a more distinct division between CMs, characterized by minimal overlap, offers advantages in achieving more accurate assessment performance [16]. According to the computation results in Table 2, all generated CMs exhibit a significant level of independence. with all 52 CDs are below the threshold value of 0.64. This suggests the minimal of overlap or confusion in the core area of CMs [16].

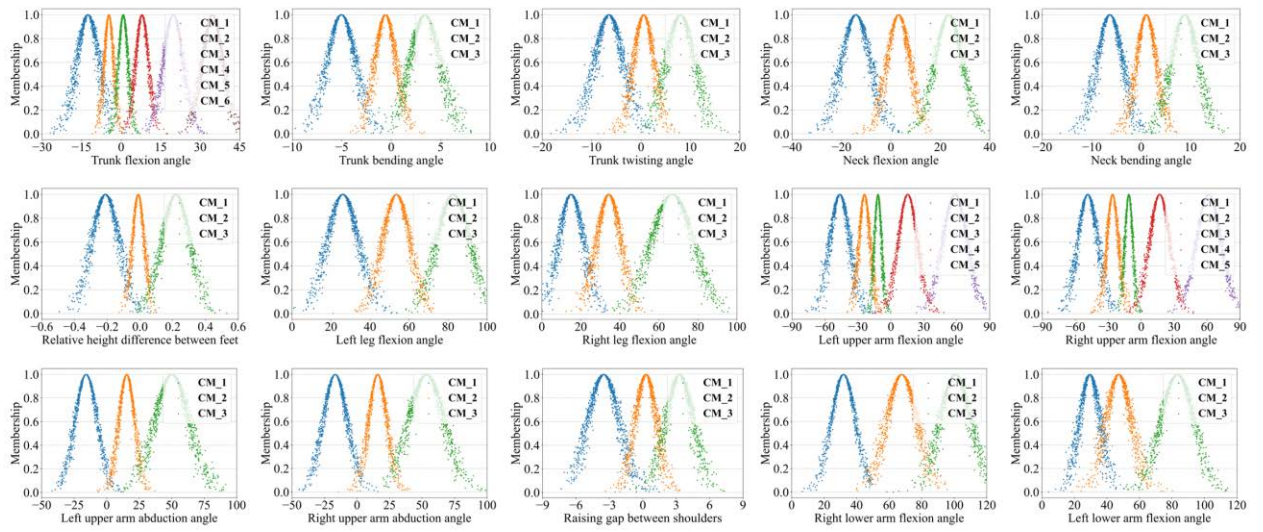


Figure 2. Developed 15 sets of joint-level scoring models

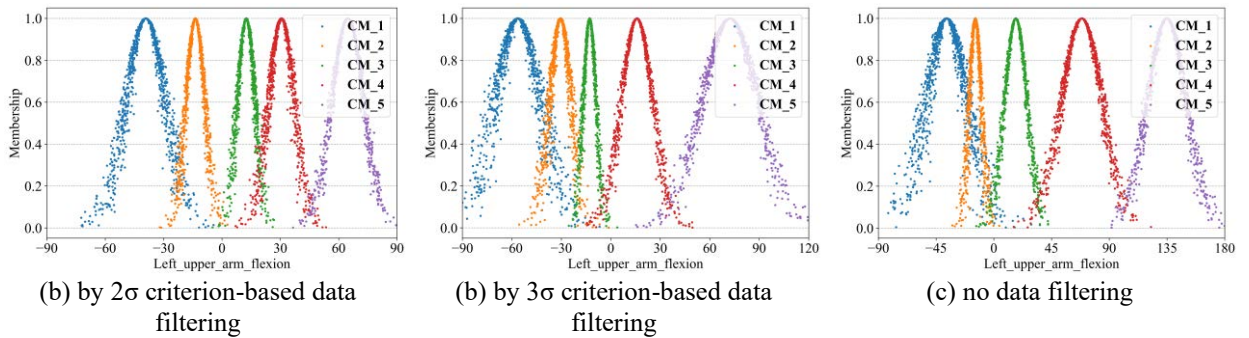


Fig.3 Comparison of generated CMs for Upper arm flexion (left) under different data filtering criteria

3.2 Performance of H-GCT

This section focuses on two crucial parameters that exert substantial influence on the performance of H-GCT. In specific, the threshold criterion for data filtering and the number of iterations during H-GT are investigated.

Firstly, the performance of the 2σ -criterion is compared with that of the 3σ -criterion (resulting in less data being deleted) and no data filtering. The Upper Arm

Flexion (left) assessment module's complexity is used as an example in Figure 3 to illustrate the generated CMs under these three criteria. It is evident that increasing the amount of deleted data leads to a reduction in overlaps between adjacent CMs, and the 2σ -criterion has generated CMs with the highest level of independence.

Secondly, the iteration process of H-GT is examined, specially focusing on the calculation of the likelihood estimation value to observe the EM (Expectation

Maximization) algorithm. Figure 4 reveals that H-GT for upper flexion angles data has undergone sufficient transformation as it approached the default 100th iteration and beyond. This is evident from the nearly highest estimated value and the reduction of fluctuations (notable fluctuations observed from the 10th to the 80th iterations). Furthermore, the generated GDs (Gaussian Distributions) at the 80th, 90th, 100th, 110th, and 120th iterations are plotted to provide a detailed representation of the iterative progression towards the optimal state.

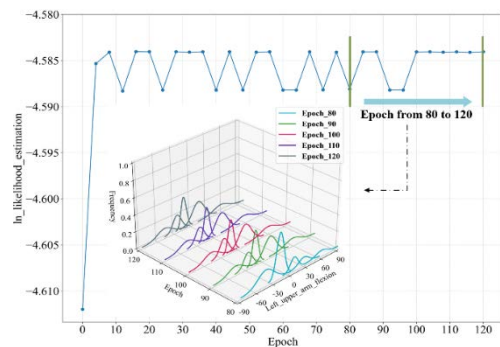


Figure 4. Iteration process of H-GT for statistical data on left upper arm flexion angles

3.3 Comparative analysis of scoring models with prior methods

This section conducts a comprehensive comparison of the generated scoring models with previous methods. Prior explorations have addressed the issue of sharp transitions between risk categories in REBA by employing fuzzy logic [1, 11, 13, 26]. These efforts aim to mitigate the impact of instrument errors [6] and inevitable human perceptual biases [4] on assessment accuracy. To facilitate the comparison, the Upper Arm Flexion (left) and Trunk Twisting modules are selected as representative examples. Figure 5 and Figure 6 are presented to illustrate the comparison of cloud models (abbreviated as CM) and membership functions (abbreviated as F) involved. These two modules represent typical joint-level assessment modules of REBA, characterized by discrete boundaries and binary rules, respectively.

Firstly, notable differences can be observed in the input domains of generated cloud models and corresponding membership functions in REBA. This disparity primarily arises from the fact that REBA is

developed based on the knowledge and experience of ergonomists [3], while the data-driven H-GCT generates scoring models based on extensive pose data from construction workers. Although the utilization of fuzzy logic by Wang, Han [11] effectively improve the traditional REBA, its focus remains limited to membership functions' shape and threshold settings, without fundamental altering the input domains of these functions.

Notably, the function shapes exhibit differences between the scoring models and other membership functions. Previous studies introduced fuzzy logic to improve REBA involving predetermined shapes such as trapezoidal and triangular functions [11, 26]. In contrast, in this study, the function shape is uniformly initialized as Gaussian distribution and iteratively optimize during the H-GCT process [16]. The incorporation of fuzzy logic effectively mitigates sharp transitions between risk categories, as evidenced by Figure 5(b). However, within each membership function improved by fuzzy logic, there are still notable instances where changes in angle do not correspond to changes in the assigned risk category, indicating certain unrealistic aspects.

Moreover, the threshold settings for fuzzy logic impact a direct influence the degree of overlap between adjacent functions, thereby affecting the distribution of membership functions. Previous studies typically relied on predetermined threshold values, such as 5° threshold adopted by Wang, Han [11]. On the contrary, the data-driven H-GCT approach employs an iterative optimization process to dynamically adjust the thresholds (i.e., overlap area) between adjacent cloud models. This dynamic adjustment process serves to minimize the Confusion degrees (CD) and, consequently, enhance the accuracy of assessment.

Ultimately, it is worth noting that previous studies failed to effectively improve the binary rule-based assessment modules of traditional REBA [1, 11, 13, 26]. These modules exhibit remarkable susceptibility to instrument errors and human perceptual biases, given their utilization of a limited and unrealistic input domain for certain membership functions. For instance, the practice of "assigning a risk score of 0 when the trunk twisting angle is 0°" fails to accurately reflect the true ergonomic risk. To overcome this limitation, the H-GCT generates scoring models that incorporate a reasonably expanded input domain for these functions, as illustrated in Figure 6(a).

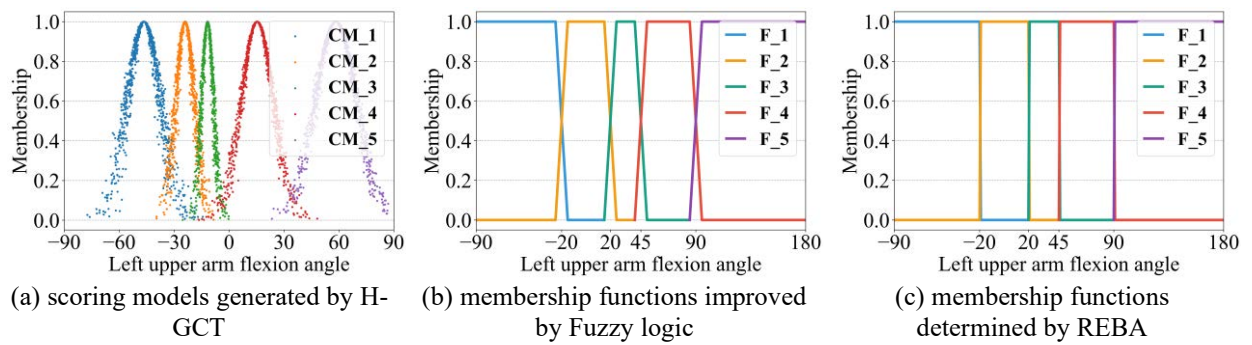


Figure 5. Comparison of scoring models and membership functions for Upper Arms Flexion (left) module

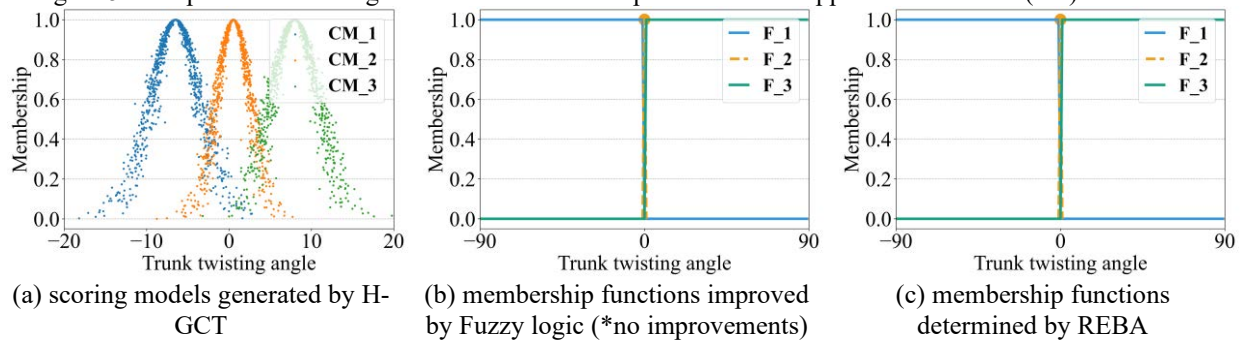


Figure 6. Comparison of scoring models and membership functions for Trunk Twisting module

4 Conclusion

This study primarily contributes to the existing body of knowledge concerning the occupational health and safety management on construction sites. It accomplishes this by developing joint-level scoring models tailored to whole-body ergonomic assessment of construction workers. Specifically, a robust and heuristic data-driven approach named H-GCT is employed, leveraging statistical data on joint angles extracted from a comprehensive CML dataset.

The results demonstrate that the CMs involved in generated scoring models exhibit a substantial level of independence, as evidenced by all Confusion degrees remaining below the threshold value of 0.64. This high level of independence contributes to the accurate identification of ergonomic risk. Furthermore, a comparison is conducted with previous studies that employed fuzzy logic to improve REBA. This process highlights the superiority of the data-driven H-GCT in developing scoring models.

In terms of practical use, the output of these scoring models can be collaborated with fuzzy inference to achieve both accurate and continuous whole-body risk scores, enhancing the applicability. Moreover, for future studies, it is recommended to improve the work by expanding pose data, enriching assessment modules, and refining the data-driven approach.

5 Acknowledgement

Financial support for this research was provided by the National Natural Science Foundation of China (grant numbers: 52378492, 72201226, and 51978164).

References

- [1] Wang, Li, Han and Al-Hussein. 3D standard motion time-based ergonomic risk analysis for workplace design in modular construction. *Automation in Construction*. 147: 104738, 2023.
- [2] Yi and Chan. Health Profile of Construction Workers in Hong Kong. *International Journal of Environmental Research and Public Health*. 13(12): 1232, 2016.
- [3] Hignett and McAtamney. Rapid Entire Body Assessment (REBA). *Applied Ergonomics*. 31(2): 201-205, 2000.
- [4] McAtamney and Nigel Corlett. RULA: a survey method for the investigation of work-related upper limb disorders. *Applied Ergonomics*. 24(2): 91-99, 1993.
- [5] Karhu, Kansu and Kuorinka. Correcting working postures in industry: A practical method for analysis. *Applied Ergonomics*. 8(4): 199-201, 1977.

- [6] Yu, Umer, Yang and Antwi-Afari. Posture-related data collection methods for construction workers: A review. *Automation in Construction*. 124: 103538, 2021.
- [7] Yu, Yang, Li, Luo, Guo and Fang. Joint-Level Vision-Based Ergonomic Assessment Tool for Construction Workers. *Journal of Construction Engineering and Management*. 145(5): 04019025, 2019.
- [8] Antwi-Afari, Qarout, Herzallah, Anwer, Umer, Zhang and Manu. Deep learning-based networks for automated recognition and classification of awkward working postures in construction using wearable insole sensor data. *Automation in Construction*. 136: 104181, 2022.
- [9] Zhang, Yan and Li. Ergonomic posture recognition using 3D view-invariant features from single ordinary camera. *Automation in Construction*. 94: 1-10, 2018.
- [10] Lee and Lee. SEE: A proactive strategy-centric and deep learning-based ergonomic risk assessment system for risky posture recognition. *Advanced Engineering Informatics*. 53: 101717, 2022.
- [11] Wang, Han and Li. 3D fuzzy ergonomic analysis for rapid workplace design and modification in construction. *Automation in Construction*. 123: 103521, 2021.
- [12] Buchholz, Paquet, Punnett, Lee and Moir. PATH: a work sampling-based approach to ergonomic job analysis for construction and other non-repetitive work. *Appl Ergon*. 27(3): 177-87, 1996.
- [13] Golabchi, Han, Fayek and AbouRizk. Stochastic Modeling for Assessment of Human Perception and Motion Sensing Errors in Ergonomic Analysis. *Journal of Computing in Civil Engineering*. 31(4): 04017010, 2017.
- [14] Habib, Ayankoso and Nagata. *Data-Driven Modeling: Concept, Techniques, Challenges and a Case Study*. in *2021 IEEE International Conference on Mechatronics and Automation (ICMA)*. 2021.
- [15] Niu, Fan and Ju. Critical review on data-driven approaches for learning from accidents: Comparative analysis and future research. *Safety Science*. 171: 106381, 2024.
- [16] Li and Du, *Artificial intelligence with uncertainty*, CRC press, Beijing, 2017.
- [17] Tian, Li, Cui and Chen. Construction motion data library: an integrated motion dataset for on-site activity recognition. *Scientific Data*. 9(1): 726, 2022.
- [18] Schilling, Gerum, Krauss, Metzner, Tziridis and Schulze. Objective Estimation of Sensory Thresholds Based on Neurophysiological Parameters. *Frontiers in Neuroscience*. 13, 2019.
- [19] Reynolds. Gaussian mixture models. *Encyclopedia of biometrics*. 741(659-663), 2009.
- [20] Sinaga and Yang. Unsupervised K-Means Clustering Algorithm. *IEEE Access*. 8: 80716-80727, 2020.
- [21] Xu and Wang. A novel cognitive transformation algorithm based on Gaussian cloud model and its application in image segmentation. *Numerical Algorithms*. 76(4): 1039-1070, 2017.
- [22] Deng, Wang and Zhang. A novel hybrid water quality time series prediction method based on cloud model and fuzzy forecasting. *Chemometrics and Intelligent Laboratory Systems*. 149: 39-49, 2015.
- [23] Bilmes. A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models. *International computer science institute*. 4(510): 126, 1998.
- [24] Li, Liu and Gan. A new cognitive model: Cloud model. *International Journal of Intelligent Systems*. 24(3): 357-375, 2009.
- [25] Wang, Xu and Li. Generic normal cloud model. *Information Sciences*. 280: 1-15, 2014.
- [26] Ghasemi and Mahdavi. A new scoring system for the Rapid Entire Body Assessment (REBA) based on fuzzy sets and Bayesian networks. *International Journal of Industrial Ergonomics*. 80: 103058, 2020.

Enhancing Human-Robot Teaming in Construction through the Integration of Virtual Reality-Based Training in Human-Robot Collaboration

Adetayo Onososen¹, Innocent Musonda² and Christopher Dzuwa³

^{1,2,3}Centre for Applied Research and Innovation in the Built Environment (CARINBE), Faculty of Engineering and the Built Environment (FEBE), University of Johannesburg, South Africa

221046022@student.uj.ac.za, imusonda@uj.ac.za, christopher.dzuwa@gmail.com

Abstract -

Human-robot teaming (HRT) has emerged as a pivotal research area in the construction industry, focusing on collaborative efforts between humans and robots for enhanced productivity and safety. This study delves into the dynamics of HRT within the immersive context of virtual reality (VR) learning environments. The research integrates established theories, including the Transtheoretical Model, the Cognitive-Affective Theory of Learning with Media (CATLM), and the Technology Acceptance Model (TAM), to construct a robust theoretical foundation. Key theoretical insights highlight the interplay of cognitive and emotional factors in influencing learning experiences and collaboration intentions. The study emphasizes the need for ergonomically compliant VR tools, addressing technical challenges to optimize user experiences. Findings underscore the significance of sustained exposure to VR features and the critical role of realism and user control in shaping the level of presence in VR environments. Practical implications emphasize the importance of usability in VR systems, encouraging educators and designers to prioritize user-friendly interfaces. The study suggests that multiple sessions are essential for VR features to substantially influence collaboration intentions. Realism and control over virtual elements are identified as key factors directly impacting the immersive experience. This research contributes to the theoretical depth of HRT in the construction industry, providing practical guidelines for developing and implementing VR systems. The insights garnered have implications for educators, designers, and practitioners, fostering a comprehensive understanding of the intricate interplay between human cognition, emotion, and technology in collaborative construction settings.

Keywords -

Construction Robot; Virtual Reality; Training; Learning; Collaboration, Human-Robot Teaming;

1 Introduction

In today's swiftly advancing technological terrain, the construction industry confronts various complex chal-

lenges that necessitate inventive resolutions [1]. Among these challenges, a crucial solution point emerges in the area of construction human-robot collaboration. While this collaboration holds the potential for heightened efficiency and productivity, it brings forth distinctive obstacles, predominantly centered on training and the potential for injuries [2, 3]. Despite relying on traditional construction methods for decades without substantial advancements in safety and productivity [4], it is imperative to embrace novel systems, innovations, and methodologies to address the ongoing challenges in the industry [2, 5].

The rise of advancements in robotics, artificial intelligence, the internet of things, and intelligent tools presents significant opportunities to propel the built sector beyond outdated traditional methods toward a more efficient and productivity-boosting industry. Among various emerging technologies, robotics has been emphasized as particularly crucial due to its potential to address targeted issues. It is poised to enhance safety by mitigating on-site risks faced by humans while improving productivity and fostering more sustainable labor costs [4].

As the utilization of robotics in construction continues to evolve, cautious consideration is given to the fact that, in numerous applications, robots still necessitate human presence within a collaborative system to accomplish project objectives [5, 6]. Due to the absence of a defined structure at construction worksites, it is essential for on-site workers to engage in collaborative efforts with robots before fully implementing autonomous robotics on construction sites [2, 3]. This collaborative approach allows for the identification of improvement areas, leading to more well-designed solutions. Moreover, it has the potential to mitigate concerns related to job displacement by robots. The imperative for human-robot collaboration (HRC) is underscored by the need to combine the strengths of robots, such as autonomy and the ability to perform repetitive tasks, with the skills possessed by human team members [5, 7]. The evolving role of robots as collaborators in the workplace suggests a complementary relationship with human labor, enhancing overall production efficiency rather than entirely replacing it [5].

Collaborating with robotics in a human-robot team poses challenges, especially considering the limited understanding of the effectiveness of such collaboration, leading to training difficulties [8, 9]. The situation is further complicated by the high costs associated with obtaining on-site training robots and the absence of hands-on experience in classroom instruction focused on human-robot collaboration [2, 3]. To overcome the constraints associated with traditional training methods, there is a proposal to implement immersive virtual reality (IVR) training for the construction workforce. This approach allows participants to immerse themselves in simulated scenarios that would otherwise present significant risks on actual construction sites. Beyond creating secure and cost-effective training environments, IVR technology provides workers with unique safety training opportunities. These opportunities include immediate guidance, empathy development, exposure to consequences, future scenario projection, feedback delivery, and emotional self-regulation, all of which serve as interventions for changing behavior [4, 10].

Moreover, the features of the virtual environment, encompassing its visual interface, immersive characteristics, interactive elements, and sense of presence, present diverse scenarios that have the potential to enhance workers' safety awareness in a physically secure setting [6]. The growing enthusiasm for utilizing virtual reality (VR) in education is well-supported by multiple meta-analyses consistently demonstrating positive educational outcomes when employing VR and simulations [3]. In some instances, these outcomes surpass those achieved through traditional classroom training. Numerous studies have affirmed that VR can impart knowledge and facilitate behavior change [9]. It is crucial to examine this knowledge gap comprehensively to establish a compelling rationale for why VR serves as an effective training method for improving behavior change outcomes in human-robot teaming.

This exploration should then drive the imperative to address all potential risks and hazards essential for ensuring a secure and reliable collaboration between humans and robots [1, 6]. However, it is equally vital to address other knowledge gaps, such as how VR can be employed in skills development, safety training, and hazard awareness in the context of human-robot team collaboration. The absence of research in this specific focus area could impede the effective adoption of robotics in construction work environments, hindering the realization of its full potential.

Hence, when individuals lack the necessary expertise to engage safely, efficiently, and effectively in interactions between machines and humans within these systems, the potential for safety risks and hazards significantly increases due to errors and incorrect assessments, resulting in serious consequences [8]. Additionally, flawed and risky dynamics in the interactions between machines and hu-

mans may lead to errors, as humans anticipate linear time behavior while machines exhibit nonlinear behavior [3, 7]. Grasping the intricacies of this relationship is crucial for establishing trust in human-robot teams and ensuring their successful integration. With robots' growing prevalence in personal spaces and workplaces, human-robot relationships are becoming increasingly widespread [6]. To tackle the challenges mentioned earlier, our research has developed a framework centered around Virtual Reality (VR) to enhance collaboration between humans and robots. This framework is designed to seamlessly integrate VR training applications into the realm of human-robot teamwork, with primary goals of improving the quality of collaboration and ensuring the safe integration of robots in built environments. The core aim of this framework is dual-fold: firstly, it harnesses the immersive capabilities of virtual reality to provide a comprehensive training experience, equipping individuals with the essential skills to collaborate effectively with robots. Secondly, it endeavors to enhance the overall standard of human-robot collaboration, promoting increased efficiency and safety at construction sites. The primary goal is to foster a harmonious partnership between individuals and robots in infrastructure delivery, ultimately benefiting both the industry and society on a broader scale.

Through the adoption of this VR-based approach, our research aims to transform the training of construction industry professionals and the integration of robots into construction processes. It not only tackles existing challenges but also sets the stage for a future in the construction sector that is more efficient, safer, and technologically advanced.

2 Theoretical Framework

In recent years, human-robot teaming (HRT) in the construction industry has emerged as an increasingly popular research topic because of its potential benefits. HRT, or Human-Robot Teaming, refers to the joint effort of humans and robots to carry out construction tasks with high productivity, efficiency, and safety [7]. The study adopted and incorporated technology-mediated learning theories, including the Transtheoretical theory, Cognitive theories (CVTAE), and the theory of planned behavior. These theoretical underpinnings not only provided a robust foundation for the research but also furnished a conceptual framework that guided the study. By integrating these theories, the study established a strong theoretical and conceptual foothold for the investigation, contributing to the theoretical depth and rigor in our field.

As stated by Mngadi [8], the role of theory is frequently used to elucidate specific circumstances and occurrences within society. It can be conceptualized as a collaboration of cognitive processes that have gradually aligned in consensus over time. This forms the basis for what is referred to as

a theoretical framework, emphasized as a "structure capable of upholding or bolstering a theory in research work," akin to a research guide or blueprint. The Technology Acceptance Model (TAM) is a conceptual framework that elucidates the way individuals perceive and embrace novel technologies [9]. The hypothesis of the theory suggests that users' behavior in accepting computers is significantly influenced by their perceptions of usefulness (U) and ease of use (EOU). Perceived usefulness (U) is characterized as the prospective user's subjective likelihood that using a particular application system will enhance their job performance within an organizational context. Perceived ease of use (EOU) pertains to the extent to which the prospective user anticipates that the target system will be effortless to use [10]. The implementation of a technology-driven learning approach has the capacity to offer valuable insights into the effectiveness of virtual reality (VR) technology in the field of education. Technology-mediated learning denotes an educational environment where learners interact with educational materials, peers, and/or instructors through the use of advanced information technologies. This research concentrates on the Immersive Virtual Environment (IVE) learning setting, where learners engage with educational content through the application of Virtual Reality (VR) technology. Thus, employing a technology-mediated learning framework is considered appropriate for this specific investigation. The theory of planned behavior (TPB) is a conceptual framework used to clarify and predict human behaviour. Developed by Icek Ajzen in the 1980s, this model has gained widespread application in the fields of psychology, sociology, and marketing [11].

In Ajzen's [11] theory of planned behavior, the main factor influencing behavioral change is identified as behavioral intention. This theoretical framework adopts a cognitive viewpoint to explain behavior based on an individual's attitudes and beliefs. The Theory of Planned Behaviour (TPB) posits that human behaviour is shaped by three primary determinants: attitudes, subjective norms, and perceived behavioural control. Attitudes encompass an individual's affective evaluations, either favourable or negative, towards a particular behaviour. Subjective norms pertain to the societal influences and pressures that encourage or discourage an individual's engagement in said behaviour. Perceived behavioural control pertains to an individual's self-assessment of their capability to execute the behaviour in question successfully.

While established theories in the cognitive domain provide a thorough framework for understanding the cognitive processes of learning, the interaction between emotional factors and cognitive elements in promoting learning is an understudied area that demands further exploration. The transtheoretical model (TTM) is a theoretical framework used to explain the process of behavior transfor-

mation, outlining five distinct stages. These stages encompass pre-contemplation, contemplation, preparation, action, and maintenance. In the pre-contemplation stage, individuals are unaware of the issue and don't actively consider changing their behavior. Moving to the contemplation stage, individuals become conscious of the problem and contemplate making a change in the near future. The preparation stage involves actively planning to initiate the desired behavior change. The action stage sees individuals actively engaging in the desired behavior, while the maintenance stage focuses on sustaining the behavior change over time. This model integrates behavior change processes and principles derived from an extensive examination of twenty-five prominent psychotherapy theories. Widely applied across disciplines such as health, psychology, sociology, and communication, the transtheoretical model, or stages of change model, proves valuable for its comprehensive conceptualization of a multi-stage process [12].

In their research, Lee et. al. [13] employed a structural equation modeling (SEM) approach to introduce a theoretical framework emphasizing the importance of the learning experience. This experience is assessed through individual psychological characteristics and its impact on learning outcomes within a virtual reality (VR) learning environment.

The study revealed that virtual reality (VR) features, such as representational fidelity, immediacy of control, and usability (encompassing quality and accessibility), act as predictors for various intermediate factors. These mediating factors include presence, motivation, cognitive benefits, control and active learning, and reflective thinking. Additionally, it was observed that these mediating factors have an impact on perceived learning, satisfaction, and performance achievement. Furthermore, the model suggests that virtual reality (VR) features indirectly influence these mediating factors by affecting usability. These mediating elements encompass both affective and cognitive factors, aligning with well-established affective frameworks like CATLM, INTERACT, and CVTAE.

A notable distinction between the aforementioned models and the proposed framework lies in its integration of significant factors from media literature, particularly virtual reality (VR) features and presence. These components are deemed indispensable for fostering effective learning in VR environments. Additionally, Lee et al.'s model offers a comprehensive portrayal of cause-and-effect relationships among key variables. Aligned with the Control-Value Theory of Achievement Emotions (CVTAE), established motivational theories like interest theory and the Cognitive-Affective Theory of Learning with Media (CATLM) shed light on the nexus between motivation and learning, providing invaluable insights for their applica-

tion in VR learning. Interest theory underscores the importance of grasping the motivational allure of e-learning technologies to facilitate the learning process, as situational interest can serve as an initial catalyst for learning promotion.

The CATLM serves as a theoretical framework aimed at elucidating the interconnections among cognitive, metacognitive, and motivational elements within technology-enhanced learning interventions. According to this model, heightened motivation is apt to yield improved learning outcomes. This is attributed to the belief that learners devoid of motivation are unlikely to engage in generative processing, even if they possess the cognitive capacity for it. Moreover, metacognitive factors play a pivotal role in regulating cognitive processing and influencing motivation.

On the other hand, in exploring employees' behavioral intentions concerning the utilization of robots, the Technology Acceptance Model (TAM) is employed to probe the relationship between the perceived ease of use of robots, perceived usefulness, and individuals' attitudes toward collaborating with robots. This inquiry is expanded by factoring in the perception of the risk of physical danger to elucidate how the sense of safety in working with robots impacts their adoption. While extant theories on behavior change and learning experiences adequately describe human behavior, they often fall short in integrating innovative systems such as actions within an immersive virtual environment (IVE). The discourse thus far underscores the significance of the transtheoretical model in scrutinizing behavior change in Human-Robot Collaboration (HRC). It's pertinent to note that this study, centered on the intention to collaborate in HRC through IVE learning, only considers the first two stages of the transtheoretical model.

The Transtheoretical Model encompasses five distinct stages of change: pre-contemplation, contemplation, preparation, action, and maintenance. In the pre-contemplation stage, individuals are oblivious to the problem and do not actively contemplate changing their behavior. Transitioning to the contemplation stage, individuals become cognizant of the problem and mull over potential changes in the near future. The preparation stage entails planning and gearing up for the desired behavior change, while the action stage witnesses individuals actively engaging in the desired behavior. Finally, the maintenance stage concentrates on sustaining the behavior change over time [12].

For the conceptual framework of this research, the initial two stages of behavior modification, specifically pre-contemplation and contemplation, are incorporated. In this framework, the features of the VR system indirectly influence collaboration through the mediation of the VR-HRC learning experience. This experience involves psy-

chological factors such as cognitive behaviors (Cognitive, Affective, and Active Learning), presence, and player experience. The learning knowledge stage (contemplation) signifies the initial phase of behavior change, wherein individuals begin to acknowledge a problem, contemplate the need for behavior change, and seek information on potential solutions and activities.

The research includes outcome variables related to behavior change, specifically examining changes elucidated through the human-robot interaction theory. This is evident in the behavioral intention to collaborate with robots, which is derived from both the intention to use robots and the intention to collaborate with them. Additionally, behavior change is observed in terms of satisfaction and intrinsic motivation.

Based on the amalgamation of concepts and theories discussed, specifically employing the Control-Value Theory of Achievement Emotions (CVTAE) and the transtheoretical model, this research anticipates that the evaluation of VR features can be conducted through the metrics of representational fidelity and co-presence. The incorporation of co-presence is additionally influenced by the human-robot interaction theory. Furthermore, the usability of Virtual Reality is assessed based on perceived ease of use, perceived usefulness, intention to use, intention to collaborate, and the perception of physical danger. The theoretical foundation for this assessment is grounded in the Technology Acceptance Model (TAM). A conceptual model has been formulated based on an input, process, and output framework, with a focus on the psychological aspects of learning. This model is designed to serve as a guide for the research design, specifically aimed at evaluating the impact of virtual reality (VR) on behavior change in Human-Robot Collaboration (HRC). The input factors that could influence the learning process, subsequently affecting learning outcomes, include the user's state of mind, VR features, and VR usability. The independent variables are the user's state of mind, measured by enthusiasm and a sense of motivation; VR features, measured by fidelity, co-presence; and VR usability, measured by perceived ease of use, perceived usefulness, perceived enjoyment, and risk of physical danger.

The VR learning experience functions as the mediating variable, while the behavior change outcome serves as the dependent variable. In terms of the learning process, internal psychological aspects of the learning experience, such as presence, cognitive behavior, and player experience, are assessed to understand the type of learning experience enhanced by VR and the significance of the learning experience in predicting behavior change outcomes. Additionally, this framework underscores the influence of VR features on behavior change outcomes through mediating factors within the learning experience.

3 Research Method

Various methods have been developed and applied to measure a range of phenomena within the construction robotics (CR) field. Questionnaire surveys and reviews of existing literature have been employed to define the study domains [14]. Utilizing virtual reality (VR) training with actual machinery in the construction industry as an initial training method shows substantial promise in comparison to the traditional paper-based or classroom approach. Empirical methodologies are commonly employed in robotics research and have recently integrated virtual reality into experimental pursuits [15].

The increasing popularity of VR can be attributed to its widespread availability and broad utilization in research studies. Nevertheless, its effective implementation requires the collaboration of diverse teams with expertise in both construction and VR/computer programming [16]. As a result, the number of tasks and scenarios incorporated in these studies has been limited. However, this study goes beyond that norm by integrating various tasks representative of a construction site.

In contrast to previous research, it was concluded that construction operators commonly employ manual techniques are the most suitable individuals to provide valuable insights into their experiences. A notable limitation in earlier experimental approaches is the exclusive reliance on student participants, who often lack practical familiarity with similar systems. Therefore, in this study, operators already using specific simulated construction machines were surveyed alongside students in a between-subjects experiment. The inclusion of input from operators was anticipated to enhance and diversify the quality of the responses.

3.1 Experiment in an Immersive Environment

To examine the hypothesis of this research, a between-subjects experiment was carried out, evaluating the evaluation capabilities of both operators and non-operators regarding human-robot collaboration. The study utilizes an immersive virtual environment to gauge participants' perceptions of their interactions with both robots and individuals in a simulated construction site. As part of the experiment, participants were involved in a human-robot collaboration task. A multitasking approach was implemented, mirroring the strategy used in earlier studies. The tasks involved human-drone collaboration tasks, human-forklift collaboration tasks, human-loaders collaboration tasks, and human-cranes collaboration tasks. These were simulated in the virtual environment with human team members responsible for instructing the robots and helping other human team members achieve the tasks. This was achieved through a multiplayer design in the immer-

sive virtual environment. The human-drone collaboration tasks were selected because drones are the most adopted form of robots currently used in the industry. Two, it is already commercially available and has huge potential for adoption. The Crane, forklift and loader tasks were selected as they represent a mix of transporting equipment and earth moving equipment currently largely used and with immense potential for adoption.

3.1.1 Experimental procedure

In order to guarantee safety during the operation, the immersive virtual world implemented a room-scale border, which the participants closely monitored. The implementation of this solution aimed to prioritize the safety of the participants when using the Head-Mounted Display (HMD). The Head-Mounted Display (HMD) limits the field of vision outside of the screen, requiring the presence of a clear and unobstructed area. Two individuals were present in the setting simultaneously alongside the robots, as depicted in Figure 1 (Showing the models and VR environment).

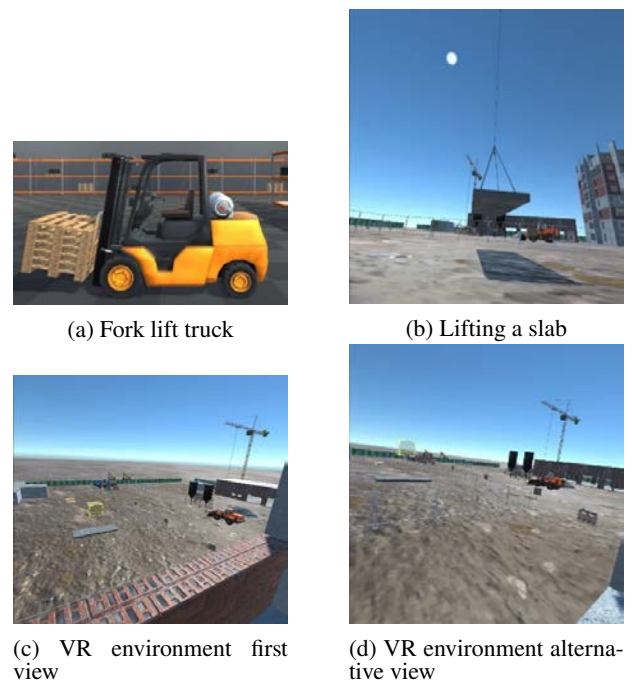


Figure 1. Captured scenes of the imported assets and assembled VR environment.

The experimental design comprised three main stages: an initial pre-questionnaire step, an experimental phase involving the robot in an immersive virtual environment (IVE), and a post-questionnaire phase. The initial phase was designated as the preparation step, which involved a 5-minute pre-experiment briefing. This entailed a com-

prehensive elucidation of experimental protocols and ethical principles. Next, the pre-experiment consent took place, during which participants confirmed their consent and provided their background information. This was determined through the pre-experiment virtual reality (VR) training, where participants familiarised themselves with the VR controls and safety requirements, completing the preparation stage in a total of 14 minutes.

4 Results

4.1 Participant's Characteristics

The analysis was conducted using empirical data obtained from pre-experiment and post-experiment responses of 42 participants. These participants included both postgraduate students and trade workers in the built environment. The tasks encompassed various scenarios introducing different models of robots, including the Forklift, Crane, Drone, Loader, Dumper, and Dump Truck. The scenarios involved tasks such as picking construction materials, delivering them to designated positions autonomously, and assigning goals and processes to the robots while the human team members observed and learned. The tests involved postgraduate students specializing in Engineering (7%), Construction/Project Management (20.9%), Quantity Surveying (25.6%), and Architecture (9.3%), all of whom were associated with the built environment. The experiment involved the participation of robotized construction equipment operators/trainees, specifically Drone Pilots/Trainees (9.3%), Crane Operators/Trainees (11.6%), Forklift Operators/Trainees (7%), and Dumper/Loader Operators/Trainees (9.3%).

Regarding the respondents' experience with robotic systems, 18.6% reported utilizing drones, 11.6% reported experience with Dump Trucks, 27.9% reported using Cranes, 18.6% reported using Forklifts, and 23.35% reported using Dumper loaders. Subsequent analysis found that these events were reported by operators/trainees of the systems rather than professionals.

The pre-experiment survey also assessed individuals' prior exposure to virtual reality; a mere 20.9% reported having engaged with virtual environments, whilst the majority of participants, 79.1%, had no prior contact with virtual reality.

4.2 Measurement Model

The measurement model is evaluated for construct reliability and validity at the start of Partial Least Squares Structural Equation Modelling (PLS-SEM) [17]. All constructs in the study were deemed acceptable as they met the threshold of Cronbach's alpha (α), > 0.70 and composite reliability scores (ρ_c), > 0.70 . Furthermore, it is worth noting that all of the study constructs successfully met the

Average Variance Extracted (AVE) test criteria. This indicates their satisfactory performance, as an AVE value greater than 0.5 is considered acceptable. The measuring model has convergent validity and internal consistency, according to the study [18]. The evaluation of the measurement model lays the groundwork for our subsequent analysis of the structural model.

4.3 Structural Model Assessment

Our research examines elements that may affect knowledge, motivation, and behaviour. This study focuses on the factors that can improve knowledge, motivation, and self-efficacy through virtual reality as seen in Figure 2. Thus, VR learning requires acknowledging its vulnerability to numerous elements. The variable inflation factor (VIF) established collinearity between the identified constructs. For the inner model, all the VIF values were less than 3.5, suggesting that these subdomains have an independent contribution to higher-order constructs. In addition, we employed bootstrapping to determine the significance of the path coefficients, demonstrating statistical importance at a level of 0.01. The study included five variables: Learner's State of Mind, VR Features, VR Usability, and VR-HRC Learning Experience. Behavioural intention to collaborate was the dependent variable. After testing the direct effect, we assessed the indirect effect through the mediating variables. Both outputs are shown individually. The mediating variable relationship was examined to determine why input-output learning occurred.

5 Discussion of Findings

The connection from the mindset of learners to virtual reality features and the learning experience was validated and affirmed as significant. These findings are confirmed by the works of Meyer [19] who illustrated, using an interactive 3D environment, the importance of users' mindset in shaping their learning experience. Additionally, the study highlighted that multiple sessions are necessary for virtual reality features to influence dependent variables related to the intention to collaborate with robots significantly. The study established enthusiasm and sense of motivation as critical in evaluating the state of mind. Hence, the research concludes that when trainees exhibit enthusiasm for learning and are appropriately motivated, their perception of human-robot teaming representation in VR improves. This finding holds both theoretical and practical implications. Theoretically, it underscores that the users' state of mind affects their perception of virtual reality features, thereby influencing their learning experience. On a practical level, it emphasizes the importance of prioritizing VR tool designs that are ergonomically compliant. Also, the correlation between VR features, usability, and learn-

ing experiences was examined. The findings indicated a substantial influence of VR features on VR usability. Usability, as defined in the study, encompasses perceived ease of use, usefulness, perceived enjoyment, and the perceived risk of physical danger when collaborating with robots in a VR setting.

The implication of the findings is that the quality of features, particularly in terms of realism, trainee control over robots, and task performance, directly influences the level of presence experienced in the virtual environment. However, intriguingly, the study disclosed no significant correlation between VR features and the learning experience. This lack of influence on learning experiences may be attributed to technical challenges encountered during the experiment, such as strap tightness, first-time virtual reality usage, motion sickness, and other factors impacting participants' experiences. Therefore, the likelihood of a positive learning experience for students may be compromised when the features lack ergonomic suitability.

The study demonstrated that when users perceive the VR environment as conducive to human-robot collaboration, it significantly influences their learning experience. Virtual usability was found to have a statistically significant impact on the VR learning experience. This implies that a more usable and user-friendly VR system positively affects the overall quality of the learning experience. While prior research suggested that human interaction with technology can shape attitudes toward digital tools, this study showed that trainees tend to have a positive disposition when they perceive the technology as easy to use, applicable, and useful. In practical terms, educators, designers, and developers can leverage these findings to design and optimize VR learning environments that prioritize usability, ultimately enhancing the quality and effectiveness of learning experiences. This, in turn, can contribute to more successful educational outcomes and improved user engagement within VR-based educational settings.

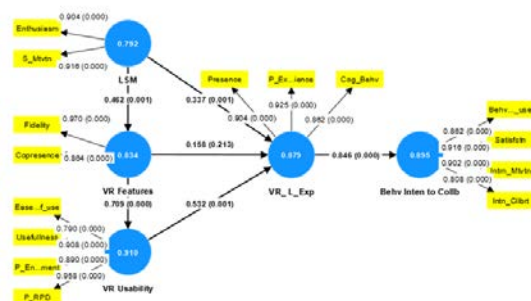


Figure 2. Virtual Reality Based Model to enhance collaboration in human-robot teaming

6 Conclusion

In conclusion, this research makes a significant contribution to the understanding of Human-Robot Teaming in the construction industry by integrating a robust theoretical framework, employing empirical methodologies, and exploring the practical implications for VR-based learning. The emphasis on usability, user mindset, and the intricate interplay between VR features and learning experiences enriches the existing knowledge in the field. The study's findings hold relevance for educators, researchers, and industry professionals seeking to harness the potential of virtual reality in improving human-robot collaboration within the construction domain. As technology continues to evolve, this research sets a foundation for further exploration and advancements in the dynamic field of Human-Robot Teaming.

7 Acknowledgements

The National Research Foundation (NRF) provided financing for this work through Grant Number 129953.

The opinions and conclusions expressed in this text belong to the authors and may not necessarily be attributed to the NRF. This work is a component of a doctoral project conducted by the Centre of Applied Research and Innovation in the Built Environment (CARINBE), University of Johannesburg.

References

- [1] Ian Ewart and Harry Johnson. Virtual reality as a tool to investigate and predict occupant behaviour in the real world: the example of wayfinding. *Journal of Information Technology in Construction*, 26:286–302, 06 2021. doi:10.36680/j.itcon.2021.016.
- [2] Alberto Jardón, Juan G. Victores, Santiago Martínez, and Carlos Balaguer. Experience acquisition simulator for operating microtunneling boring machines. *Automation in Construction*, 23:33–46, 2012. ISSN 0926-5805. doi:https://doi.org/10.1016/j.autcon.2011.12.002.
- [3] Shiyao Cai, Zhiliang Ma, Miroslaw Skibniewski, Song Bao, and Heqin Wang. Construction automation and robotics for high-rise buildings: Development priorities and key challenges. *Journal of Construction Engineering and Management*, 146: 04020096, 08 2020. doi:10.1061/(ASCE)CO.1943-7862.0001891.
- [4] Kamel S. Saidi, Thomas Bock, and Christos Georgoulas. *Robotics in Construction*. Springer International Publishing, Cham, 2016. ISBN 978-3-319-32552-1. doi:10.1007/978-3-319-32552-1_57.

- [5] Michael Decker, Martin Fischer, and Ingrid Ott. Service robotics and human labor: A first technology assessment of substitution and cooperation. *Robotics and Autonomous Systems*, 87:348–354, 2017. ISSN 0921-8890. doi:<https://doi.org/10.1016/j.robot.2016.09.017>.
- [6] Harold Soh, Yaqi Xie, Min Chen, and David Hsu. Multi-task trust transfer for human–robot interaction. *The International Journal of Robotics Research*, 39(2-3):233–249, 2020. doi:[10.1177/0278364919866905](https://doi.org/10.1177/0278364919866905).
- [7] Mustafa Demir, Nathan J. McNeese, and Nancy J. Cooke. Understanding human-robot teams in light of all-human teams: Aspects of team interaction and shared cognition. *International Journal of Human-Computer Studies*, 140:102436, 2020. ISSN 1071-5819. doi:<https://doi.org/10.1016/j.ijhcs.2020.102436>.
- [8] A. Mngadi. *The Role of Theory in Research and Practice*. GRIN Verlag, 1 edition, 2018. doi:[10.4324/9781315213033-4](https://doi.org/10.4324/9781315213033-4).
- [9] James Sorce and Raja Issa. Extended technology acceptance model (tam) for adoption of information and communications technology (ict) in the us construction industry. *Journal of Information Technology in Construction*, 26:227–248, 05 2021. doi:[10.36680/j.itcon.2021.013](https://doi.org/10.36680/j.itcon.2021.013).
- [10] Shiva Pedram, Stephen Palmisano, Richard Skarbez, Pascal Perez, and Matthew Farrelly. Investigating the process of mine rescuers’ safety training with immersive virtual reality: A structural equation modelling approach. *Computers Education*, 153:103891, 2020. ISSN 0360-1315. doi:<https://doi.org/10.1016/j.compedu.2020.103891>.
- [11] Icek Ajzen. *From Intentions to Actions: A Theory of Planned Behavior*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1985. ISBN 978-3-642-69746-3. doi:[10.1007/978-3-642-69746-3_2](https://doi.org/10.1007/978-3-642-69746-3_2).
- [12] A. Abbas. *Understanding Human Factors Issues in Immersive Virtual and Augmented Reality Applications for Construction*. Phd dissertation, The Hong Kong Polytechnic University, 2021.
- [13] Elinda Ai-Lim Lee, Kok Wai Wong, and Chun Che Fung. How does desktop virtual reality enhance learning outcomes? a structural equation modeling approach. *Computers Education*, 55(4):1424–1442, 2010. ISSN 0360-1315. doi:<https://doi.org/10.1016/j.compedu.2010.06.006>.
- [14] Adetayo Olugbenga Onososen, Innocent Musonda, and Molusiwa Ramabodu. Construction robotics and human–robot teams research methods. *Buildings*, 12(8), 2022. ISSN 2075-5309.
- [15] Chufeng Huang, Wen Zhang, and Liang Xue. Virtual reality scene modeling in the context of internet of things. *Alexandria Engineering Journal*, 61(8):5949–5958, 2022. ISSN 1110-0168. doi:<https://doi.org/10.1016/j.aej.2021.11.022>.
- [16] Sangseok You, Jeong-Hwan Kim, SangHyun Lee, Vineet Kamat, and Lionel P. Robert. Enhancing perceived safety in human–robot collaborative construction using immersive virtual environments. *Automation in Construction*, 96:161–170, 2018. ISSN 0926-5805. doi:<https://doi.org/10.1016/j.autcon.2018.09.008>.
- [17] Ned Kock. *Minimum Sample Size Estimation in PLS-SEM: An Application in Tourism and Hospitality Research*. Emerald Publishing Limited, 2018. ISBN 978-1-78756-700-9, 978-1-78756-699-6. doi:[10.1108/978-1-78756-699-620181001](https://doi.org/10.1108/978-1-78756-699-620181001).
- [18] Liyin Shen, Xiangnan Song, Ya Wu, Shiju Liao, and Xiaoling Zhang. Interpretive structural modeling based factor analysis on the implementation of emission trading system in the chinese building sector. *Journal of Cleaner Production*, 127:214–227, 2016. ISSN 0959-6526. doi:<https://doi.org/10.1016/j.jclepro.2016.03.151>.
- [19] Lars Meyer-Waarden and Julien Cloarec. “baby, you can drive my car”: Psychological antecedents that drive consumers’ adoption of ai-powered autonomous vehicles. *Technovation*, 109:102348, 2022. ISSN 0166-4972. doi:<https://doi.org/10.1016/j.technovation.2021.102348>.

Interpretation Conflict in Helmet Recognition under Adversarial Attack

He Wen, Simaan AbouRizk

Department of Civil and Environmental Engineering, University of Alberta, Canada
hwen7@ualberta.ca, abourizk@ualberta.ca

Abstract –

Humans and Artificial Intelligence (AI) may have observation and interpretation conflicts in collaborative interaction. The adversarial samples make such conflicts more likely to occur in the field of image recognition. However, few studies have been seen combining the human-AI conflict and adversarial attack. This study presents the interpretation conflict due to adversarial samples in the helmet recognition task. A simulation also has been conducted to illustrate this problem. The results show that it should be prudent for the construction industry to land AI applications due to adversarial attacks on image recognition; the adversarial samples easily trigger interpretation conflicts, for example, the logo, graffiti, sticker, and text on helmets; lean construction should be propagated for the preconditions for AI applications.

Keywords –

Human-AI conflict; Adversarial attack; Risk; Cross-entropy; Distance

1. Introduction

Artificial intelligence (AI) and machine learning have enabled numerous creative applications in construction operations [1]. Typical examples include face recognition, personnel positioning, and violation detection [2], [3], contributing to safety management and operations. More specifically, helmet detection is one of the mature experiments [4], [5], since the safety helmet is the most essential and mandated personal protective equipment (PPE), and violation is still occasional or even expected. For example, a recent survey in California shows that 36% of construction workers struggle to ensure they consistently wear PPE [6], even if they are all well-equipped at daily check-in. Similar studies also indicate that on-site supervision and enforcement are required but time/effort consuming [7], [8].

Fortunately, computer vision and image recognition with deep learning facilitate this task instead of human inspection by vision (Figure 1), integrating body detection and personnel location [9], [10], [11]. As the pioneer field of AI, image recognition of helmets has

constantly improved its accuracy in academic experiments [12], especially with the version update of the algorithm of You Only Look Once (YOLO) [13]. While the majority of research findings boast an accuracy rate exceeding 90%, the authors endeavoured to replicate these experiments utilizing algorithms outlined in published papers and publicly available construction site images, however, the accuracy was still unsatisfactory.



Figure 1. Worker location and helmet detection [10]

One significant cause of such accuracy problems is the samples in field applications often have some noise or are heavily polluted. For example, in the helmet recognition task, the logo, graffiti, sticker, and text might be considered the adversarial samples (Figure 2), or even the light and shadow may manipulate the results. This is the phenomenon of adversarial attacks [14]. The deep learning neural network misclassifies the adversarial sample by adding an imperceptible perturbation to the original image [15]. Indeed, the problem of adversarial attacks in image recognition has received long-term attention and research, and feasible countermeasures have been proposed [16], [17]. However, in the field of helmet detection, many studies do not mention this issue.



Figure 2. Adversarial samples of helmets

Once the adversarial sample misleads the AI, it occurs a typical human-AI conflict [18], both observation conflict and interpretation conflict. This can trigger false alerts to workers or false violation records. On the other hand, it may also miss the detection of the helmet. In a critical environment, such a situation might trigger prudent risks. Therefore, this study aims to present this problem, alert the practitioners about this risk, and then demonstrate it through a simulation. The novelty of this paper is:

- i. The mathematical expression of interpretation conflict in image recognition.
- ii. A measurement of the conflict based on the vector distance and cross-entropy.
- iii. The combination of helmet recognition and adversarial attack.

A reminder to the readers of this article: Section 2 describes the problem in mathematical expressions; Section 3 presents the simulation of a case; Section 4 summarizes the discussion of the simulation results and solutions to the proposed problem; Section 5 includes the remarkable conclusions, contributions, and limitations.

2. Problem statement

For computer vision, AI regards a picture as a $height \times width \times channel$ matrix, usually with a basic kernel of $3\ pixel \times 3\ pixel \times 3\ RGB$ [19], where RGB means red, green, and blue. Then the matrix is converted to a high-dimensional column vector by the three channels. On the other side, humans do not yet know how the brain works, from seeing a picture to recognizing the classification of the picture, at least not mathematically. Therefore, assume that the same is true for humans, and the human observation is also converted into a vector, then the variable of observation difference (VOD) for observation conflict can be expressed as [20]:

$$VOD = X_A - X_H \quad (1)$$

Where is X_A the AI vector and X_H is the human vector.

In addition, AI further performs deep learning with the convolutional neural network (CNN) to get the score, then applies the Softmax function to transfer the score to the classification probability \hat{y}_A . The last step is to conclude the classification result y_A through the cross-entropy function, where y_A is a one-hot vector. Naturally, when humans see an image, they estimate the probability \hat{y}_H for a limited number of classifications, and then get the result y_H , which can also be expressed as a one-hot vector. Usually, humans could recognize their classification result y_H immediately. An example of recognizing a helmet is shown in Figure 3. Thus, the variable of interpretation difference (VID) [18], which is the interpretation conflict, can be simplified as the difference between two $n \times 1$ one-hot vectors:

$$VID = Y_A - Y_H \quad (2)$$

Where n is the number of classifications. When $VID = 0_{n \times 1}$, there is no interpretation conflict; when $VID \neq 0_{n \times 1}$, there is an interpretation conflict.

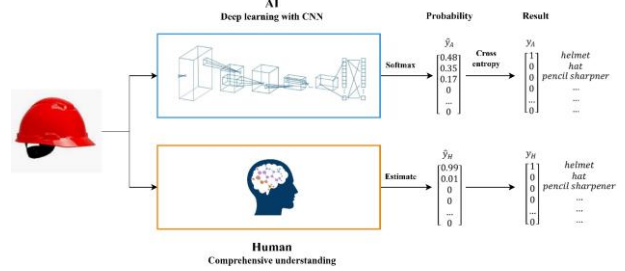


Figure 3. An example of how humans and AI recognize a picture

Then

$$VID \propto d = \text{cross entropy}(\hat{y}_A, y_H) \quad (3)$$

Where d is the distance between \hat{y}_A and y_H . This distance could be applied to measure the interpretation conflict under various noises, including adversarial samples.

Due to the improvement of AI learning ability, the accuracy rate has been increasing for the recognition task of standard samples, which is close to human cognition, reaching above 80% accuracy in 50-150ms [21], [22]. However, adversarial samples have a greater chance of interpretation conflict. As described in the Introduction, a small perturbation is added to the picture. Typically, a perturbation involves increasing or decreasing small values to/from each pixel of the image. Then humans cannot tell the difference between before and after, and get the same classification result. However, AI may give an unexpected result; for example, Figure 4 shows that AI recognizes a helmet as a pencil sharpener under adversarial attack. Here the $VID = [-1, 0, 1, 0, \dots, 0]^T$.

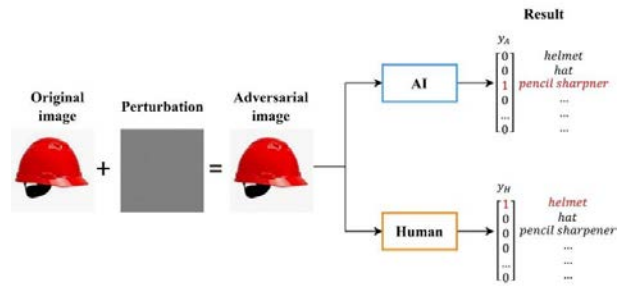


Figure 4. An example of interpretation conflict under adversarial attack

Therefore, the problems to be presented and solved in this study are: Does AI accurately detect whether workers are wearing helmets on construction sites, and do adversarial samples potentially trigger interpretation conflicts between human supervisors and AI?

3. Simulation and results

Though helmet identification is a typical application of image recognition with high accuracy, under adversarial attack, it may show unpredictable errors. Therefore, the simulation is designed as the following major steps to present this problem (Figure 5).

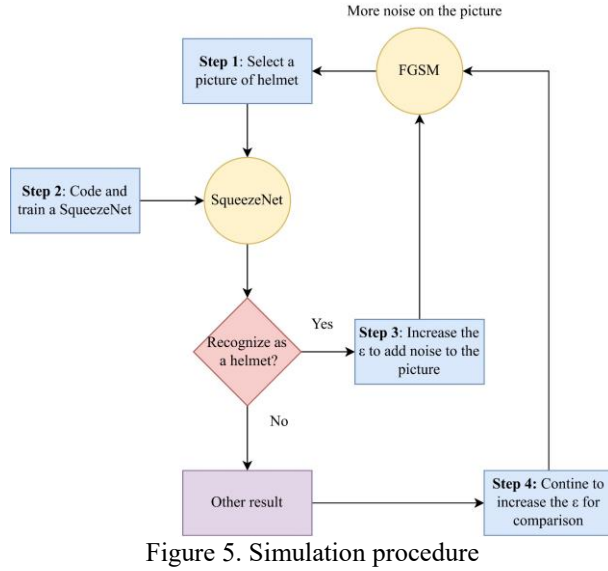


Figure 5. Simulation procedure

Step 1: Randomly search and choose a helmet picture from the public Internet as the test sample.

Step 2: Code the program in MATLAB r2022a to train and test the image with SqueezeNet and the Fast Gradient Sign Method (FGSM) [23].

SqueezeNet is a pre-trained neural network with plenty of classification labels and a relatively small computational occupation suitable for academic simulation [24]. This study added more training samples to optimize the SqueezeNet to identify the test examples.

200 training samples of safety helmets from the public Internet are added (Figure 6), and the label is marked “safety helmet” (Hereinafter referred to as “helmet”) to distinguish the “crash helmet” in the original SqueezeNet.



Figure 6. Example of training samples

The FGSM is a mature technique for generating adversarial samples [15], and it has

$$X_{adv} = X + \varepsilon * \text{sign}(\nabla_X L(X, T)) \quad (4)$$

Where X is the original image vector, X_{adv} is the adversarial image vector, $\nabla_X L(X, T)$ is the gradient of

the loss function L to the targeted label T ; ε controls the size of the push and the adversarial strength, which means that the larger the ε value, the greater the perturbation.

Step 3: Increase ε gradually until the classification result changes from “helmet” to another label. The procedure is designed to trigger an interpretation conflict.

Step 4: Continue to increase ε to generate enough conflict results for comparison and discussion. After the simulation, the results are shown in Figure 7.

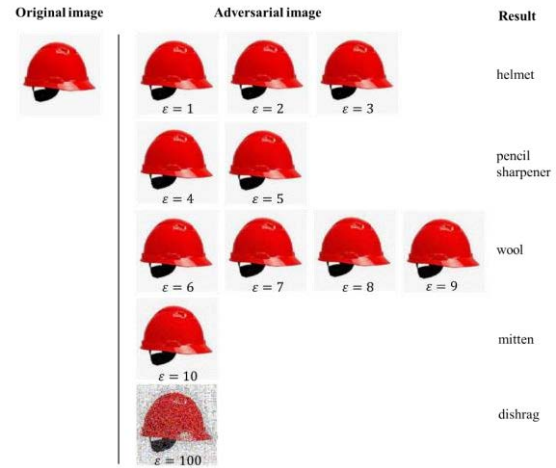


Figure 7. Original image and adversarial images

As the attack strength increases, in other words, the noise increases, it misleads the AI to recognize the helmet as “pencil sharpener”, “wool”, “mitten”, and “dishrag”. The relation between conflict measurement (distance d) and attack strength (control parameter ε) is shown in Figure 8.

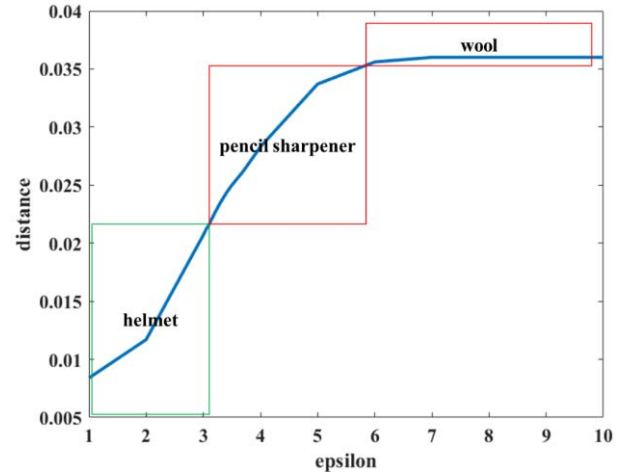


Figure 8. The relation between conflict measurement and attack strength

4. Discussion

The simulation results presented in this study shed light on the vulnerability of AI-based image recognition systems to adversarial attacks, particularly in helmet detection on construction sites. While this issue is present across all PPE detection scenarios, safety helmets are particularly vulnerable to adversarial attacks compared to safety vests and shoes. This susceptibility has been confirmed through extensive experimentation with various items, leading the authors to highlight it as a typical scenario in helmet detection.

By systematically increasing the attack strength, as controlled by the parameter ϵ in the FGSM, it is observed a notable increase in interpretation conflicts, manifested as misclassifications by the AI system. One of the key findings of this study is the progressive deterioration in the AI system's performance as the attack strength increases. Initially, the AI system accurately identifies safety helmets from the test samples, achieving high classification accuracy. However, as the adversarial perturbations intensify, the system's confidence diminishes, ultimately resulting in misclassifications. Notably, the misclassifications observed in the simulation ranged from plausible but incorrect labels such as "pencil sharpener" to more evidently erroneous labels like "wool" or "mitten." This progression highlights the escalating confusion and uncertainty introduced by adversarial attacks.

The relationship between the conflict measurement (distance d) and attack strength (ϵ) provides valuable insights into the vulnerability of the AI system. As depicted in Figure 8, there is a clear positive correlation between attack strength and conflict measurement, indicating that stronger adversarial perturbations lead to greater deviation between the AI system's classification and the ground truth. This observation underscores the sensitivity of AI systems to subtle changes in input data, which can be exploited to induce interpretation conflicts and undermine the system's reliability.

From the simulation results, the mitigation strategies can be induced. One solution is adversarial training with adversarial samples, for example, enabling recurring training based on false positive data identified from construction sites.

In addition, model robustness evaluation would encourage the model to learn robust representations that are resilient to adversarial perturbations. It tunes model sensitivity to have a higher tolerance for various types of image qualities.

Furthermore, implementing defense mechanisms with adversarial sample detection can help mitigate adversarial attacks, since input data pre-processing can improve the sample quality.

5. Conclusions

This study points out the problem of helmet recognition under adversarial attack in the construction industry, which is a matter of deep concern with observation and interpretation conflict. The distance between AI prediction and human cognition could measure the human-AI conflict. This reminds practitioners not to mindlessly launch new AI applications and ignore the weaknesses and defects of AI technology itself.

This study underscores the discrepancy between AI-based image recognition systems and human perception. This interpretation conflict raises important questions regarding the limitations of current AI technologies and the need for further research to bridge the gap between AI and human cognition. Moreover, the findings of this study have significant implications for the deployment of AI-based safety monitoring systems in real-world contexts. The susceptibility of these systems to adversarial attacks underscores the importance of rigorous testing and validation procedures to assess their robustness and reliability. Lastly, promoting education and awareness initiatives of adversarial attacks can increase understanding of the capabilities, limitations, and risks associated with AI technologies. This is also the main intention of this study.

Thus, this study serves as a reminder to both industry and academia to consider the diverse array of environmental disturbances present at construction sites when employing AI technology. It underscores the importance of the construction site environment, since dirtiness, dim lighting, and outdated equipment/tools can potentially create adversarial samples. As a result, it suggests the preconditions for AI application, for example, maintaining cleanliness, ensuring adequate lighting, regularly maintaining equipment/tools/PPE, and adhering to standardization protocols for safety signs. They are equally vital for enhancing the precision of image recognition. Moreover, these are also the basic requirements for lean construction management.

Despite the contributions and insights, several limitations must be acknowledged to ensure a comprehensive understanding. Firstly, the simulation environment employed in this study inherently simplifies the complexity of real-world scenarios, for example, the training and test samples are from the public Internet, not real construction sites. Moreover, the generalizability of the findings and proposed solutions may be constrained by the specific characteristics of the AI models, datasets, and application domains. Also, the efficacy of the proposed solutions may vary depending on factors such as the architecture of the AI system, the nature of the adversarial attacks, and the diversity of the input data.

Therefore, future research should aim to address these limitations and explore new approaches to enhance the

robustness and reliability of AI systems in safety-critical applications, such as helmet recognition in this study. The research and practice of AI reliability are full of challenges and encourage further exploration.

References

- [1] M. Regona, T. Yigitcanlar, B. Xia, and R. Y. M. Li, "Opportunities and adoption challenges of AI in the construction industry: a PRISMA review," *Journal of open innovation: technology, market, and complexity*, vol. 8, no. 1, p. 45, 2022.
- [2] J. Chai, H. Zeng, A. Li, and E. W. T. Ngai, "Deep learning in computer vision: A critical review of emerging techniques and application scenarios," *Machine Learning with Applications*, vol. 6, no. August, p. 100134, 2021.
- [3] W. Lin, J. Xue, C. Zhu, F. Jiang, and C. Ding, "An integrated application platform for safety management and control of smart power plants based on video image technology," in *2022 IEEE 10th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, IEEE, 2022, pp. 1436–1439.
- [4] H. Peng and Z. Zhang, "Helmet Wearing Recognition of Construction Workers Using Convolutional Neural Network," *Wirel Commun Mob Comput*, vol. 2022, 2022.
- [5] K. Li, X. Zhao, J. Bian, and M. Tan, "Automatic safety helmet wearing detection," *arXiv preprint arXiv:1802.00264*, 2018.
- [6] A. N. Gattuso, "Common Issues of Compliance with Personal Protective Equipment for Construction Workers," 2021.
- [7] T. K. M. Wong, S. S. Man, and A. H. S. Chan, "Critical factors for the use or non-use of personal protective equipment amongst construction workers," *Saf Sci*, vol. 126, p. 104663, 2020.
- [8] A. D. Rafindadi, M. Napiiah, I. Othman, H. Alarifi, U. Musa, and M. Muhammad, "Significant factors that influence the use and non-use of personal protective equipment (PPE) on construction sites—Supervisors' perspective," *Ain Shams Engineering Journal*, vol. 13, no. 3, p. 101619, 2022.
- [9] L. Wang *et al.*, "Investigation Into Recognition Algorithm of Helmet Violation Based on YOLOv5-CBAM-DCN," *IEEE Access*, vol. 10, pp. 60622–60632, 2022.
- [10] J. Shen, X. Xiong, Y. Li, W. He, P. Li, and X. Zheng, "Detecting safety helmet wearing on construction sites with bounding-box regression and deep transfer learning," *Computer-Aided Civil and Infrastructure Engineering*, vol. 36, no. 2, pp. 180–196, 2021.
- [11] L. Wei, M. Cheng, M. Feng, and Z. Lijuan, "Research on recognition of safety helmet wearing of electric power construction personnel based on artificial intelligence technology," *J Phys Conf Ser*, vol. 1684, no. 1, 2020.
- [12] N. D. Nath, A. H. Behzadan, and S. G. Paal, "Deep learning for site safety: Real-time detection of personal protective equipment," *Autom Constr*, vol. 112, p. 103085, 2020.
- [13] M. I. B. Ahmed *et al.*, "Personal protective equipment detection: A deep-learning-based sustainable approach," *Sustainability*, vol. 15, no. 18, p. 13990, 2023.
- [14] C. Szegedy *et al.*, "Intriguing properties of neural networks," *2nd International Conference on Learning Representations, ICLR 2014 - Conference Track Proceedings*, pp. 1–10, 2014.
- [15] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pp. 1–11, 2015.
- [16] S. Qiu, Q. Liu, S. Zhou, and C. Wu, "Review of artificial intelligence adversarial attack and defense technologies," *Applied Sciences*, vol. 9, no. 5, p. 909, 2019.
- [17] H. Xu *et al.*, "Adversarial attacks and defenses in images, graphs and text: A review," *International Journal of Automation and Computing*, vol. 17, pp. 151–178, 2020.
- [18] H. Wen, Md. T. Amin, F. Khan, S. Ahmed, S. Imtiaz, and E. Pistikopoulos, "Assessment of Situation Awareness Conflict Risk between Human and AI in Process System Operation," *Ind Eng Chem Res*, vol. 62, no. 9, pp. 4028–4038, Mar. 2023.
- [19] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pp. 1–14, 2015.
- [20] H. Wen, Md. T. Amin, F. Khan, S. Ahmed, S. Imtiaz, and S. Pistikopoulos, "A methodology to assess human-automated system conflict from safety perspective," *Comput Chem Eng*, vol. 165, p. 107939, Jul. 2022.
- [21] R. Wu, S. Yan, Y. Shan, Q. Dang, and G. Sun, "Deep image: Scaling up image recognition," *arXiv preprint arXiv:1501.02876*, vol. 7, no. 8, p. 4, 2015.
- [22] S. Thorpe, D. Fize, and C. Marlot, "Speed of processing in the human visual system," *Nature*, vol. 381, no. 6582, pp. 520–522, 1996.

- [23] M. H. Beale, M. T. Hagan, and H. B. Demuth, *Deep Learning Toolbox™ User's Guide*. The MathWorks, Inc., 2023.
- [24] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," *5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings*, pp. 1–13, 2017.

Enhancing Decision-Making for Human-Centered Construction Robotics: A Methodological Framework

Marc Schmailzl¹, Anne-Sophie Saffert¹, Merve Karamara¹, Thomas Linner¹, Friedrich Anton Eder¹, Simon Konrad Hoeng¹ and Mathias Obergruesser¹

¹Faculty of Civil Engineering, Ostbayerische Technische Hochschule Regensburg, Germany
marc.schmailzl@oth-regensburg.de, annesophie.saffert@oth-regensburg.de, merve.karamara@oth-regensburg.de,
thomas.linner@oth-regensburg.de, friedrich.eder@oth-regensburg.de, simon.hoeng@oth-regensburg.de,
mathias.obergruesser@oth-regensburg.de

Abstract –

While the Architecture, Engineering, and Construction (AEC) industry is increasingly aware of the rising demands for productivity and human-centered construction improvements, the holistic adoption of robotics as a fundamental strategy to address these challenges has not yet reached comprehensive fruition. This paper therefore introduces a methodological framework aiming to address the industry's pressing need for a systematic approach for assessing the feasibility of integrating robotics into human-centered construction processes. It aims to enhance decision-making regarding the degree of automation in human-centered construction processes, ranging from partial to full robotization or non-robotization. The framework is characterized by a more holistic end-to-end data-/workflow and therefore adopts a multifaceted approach, leveraging BIM-based planning methodologies and integrating new technologies [e.g., Motion Capturing (MoCap), work process simulation software incorporating Digital Human Models (DHM), self-developed conversion/interfaces software and more] that have not been widely used in the industry to date. Subsequently, the framework is evaluated in a real-life bricklaying construction process to ensure a more application-based approach. Overall, the framework advances current construction processes with a more inclusive and conscious technology infill to empower construction professionals with the workflow and corresponding tools necessary for the practical integration of robotics into human-centered construction processes.

Keywords –

Decision-Making; Framework; Workflow; AEC Industry; Robotics; Building Information Modeling (BIM); Human-Centered; Motion Capturing

1 Introduction

The AEC industry is a major contributor to the global economy facing a variety of challenges. Socially, it involves a lack of skilled labor and challenging work conditions [1]. Economically, it lags in productivity, digitization, and faces obstacles in high inflation environments [2]. Moreover, the housing demand outpaces supply, causing significant increases in rental prices [3]. Ecologically, the industry is one of the most resource-intensive industries and therefore significantly affecting climate targets [4].

Considering these challenges, while recognizing the need for transformative measures (especially regarding Industry 4.0 to 5.0/6.0 shift), the industry is advised to explore new technologies and create more human-centered construction processes to counteract aforementioned challenges [5, 6]. Therefore, robotics as one of the most disruptive technologies across a multitude of industries offers the potential to address several challenges accordingly.

However, to introduce robotics in construction and at the same time ideally foster more human-centered construction processes (incl. better, ergonomically favorable work conditions), it is crucial to address the inherent industry limiting factors (e.g., weak business/use cases, low research budgets, high-risk with partly immature technology, data inconsistency and lacking interfacing capabilities, dynamically changing on-site conditions, lack of upstream feasibility assessment etc.) that impede an appropriate transition and subsequent integration [5]. In this regard, research has already been undertaken to address the aforementioned problems. However, most research lacks in terms of formulating a systematic and holistic approach for integrating robotics into human-centered construction processes. This deficiency is compounded by inadequate technology utilization, insufficient attention to human factors, coupled with a lack of cross-domain knowledge

utilization, which further exacerbates the problems and ultimately leading to insufficient workflows.

Therefore, a critical research question emerges: How can the industry assess the feasibility and subsequently enhance decision-making regarding the integration of human-centered construction robotics?

This paper aims to give a preliminary answer to this question with a structured approach (methodological framework) characterized by a more holistic end-to-end data-/workflow. This enables enhanced decision-making, strategically incorporating technologies to assure an appropriate and more future-proof adoption as well as feasibility assessment of robotics in construction. By doing so, the industry can achieve a higher efficiency, while acting human-centered, ensuring its continued resilience in the face of evolving economic, social, and environmental landscapes.

2 Related Work

The AEC industry and affiliated research features a variety of methodologies aiming to counteract aforementioned industry challenges and associated limiting factors (see section 1). One example is the Building Information Modeling (BIM) methodology utilizing 3D-models (incl. geometric and correlating semantic data/information) to streamline multidisciplinary work processes throughout the entire building life cycle. Building up on the BIM-methodology, there are further ones based on it and aiming to enhance decision-making towards a higher adoption of robotics in construction through various improvements (e.g., enhanced data/information flow incl. added process data/information leveraging BIM-based planning methodologies etc.) [7, 8, 9, 10].

Despite those methodologies, it is recognizable that the adoption of (industrial) robotics in construction is still significantly lower compared to other industries (see Figure 1, “Others”) [11].

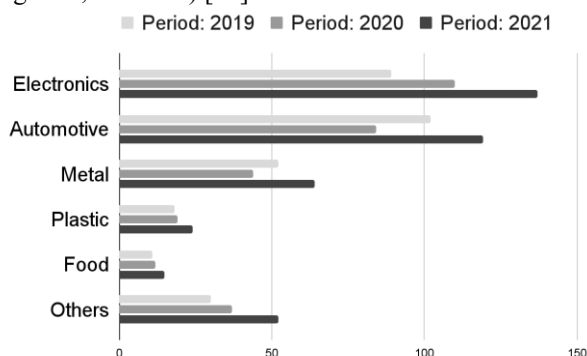


Figure 1. Annual installations of industrial robots by customer industry worldwide (1,000 units)

In this context, one derivation the related work

emphasized is the potential of adequate data preparation, accumulation, and interfacing capabilities already in early planning phases to enable enhanced decision-making and more holistic end-to-end data-/workflows.

In addition, Industry 5.0 proclamations call for a necessity towards more human-centered construction processes where ergonomic assessments come into play and are mutually respected in the digital as well as physically applied context to improve ergonomically unfavorable work conditions. This aspect is currently hardly considered in traditional construction processes (especially regarding the associated data-/workflow).

Subsequently, a structured data-driven approach (methodological framework) needs to be established, that incorporates work and machine process requirements in early planning phases, aiming to holistically address aforementioned industry challenges and associated issues (respecting Industry 5.0 proclamations) through consecutive economic and ergonomic assessments of construction processes.

3 Methodology

This paper builds upon a comprehensive research methodology (RM, see Figure 2) considering diverse

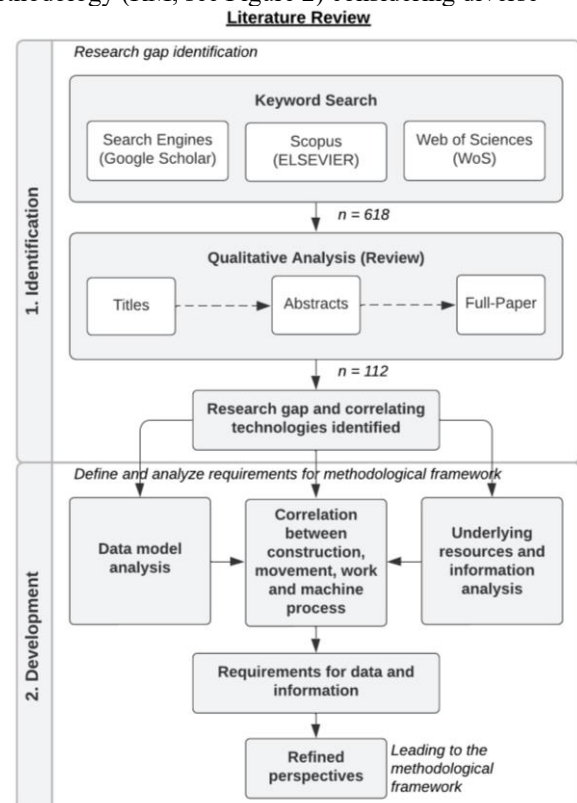


Figure 2. Research methodology (RM)

sources and analytical approaches to identify technologies which ideally could enable enhanced

decision-making for integrating robotics in human-centered construction processes.

Therefore, the RM starts with a literature review through search engines focusing on peer-reviewed high impact articles (2018 - mid 2023) ensuring a meaningful, thorough, and differentiated exploration. The selected period is chosen to encompass the latest literature considering more application-oriented and ideally already (empirically) validated literature.

The initial search used specific keywords (“advanced search”, incl. boolean operators, see Table 1) aiming to find relevant titles (after deleting duplicates = $n = 618$) before the abstract and subsequently full paper for a more in-depth selection was reviewed in an incremental approach (qualitative analysis).

Table 1. Keywords used for the keyword search (“Advanced Search”, Accessed: December 2023)

Framework AND Production AND Construction		
“framework”	“production”	“construction”
OR	OR	OR
“methodology”	“robotic”	“architecture”
OR	OR	OR
“decision”	“automation”	“building”

The identified literature derived from the qualitative analysis ($n = 112$), revealed that certain technologies (total = $n = 16$, e.g., BIM-software, MoCap, work process simulation software incorporating DHM, conversion/interfaces software, Artificial Intelligence algorithms, agnostic robotic frameworks etc.) offer the potential for enhanced decision-making regarding the integration of robotics into human-centered construction processes. However, a lot of identified technologies are currently largely unutilized in the AEC industry (e.g., MoCap).

The subsequent development of the methodological framework incorporated a final analysis of the identified technologies regarding their data models (e.g., tabular, hierarchical, entity-relationship etc.), alongside with underlying resources and information (e.g., Level Of Information Need, LOIN etc.) to consecutively weave them into a data-driven methodological framework respecting the correlations and interplay between construction, movement, work and machine processes which was particularly highlighted in literature as an essential approach. Concluding the overall literature review, the RM ended with the synthesis of necessary requirements for data and information needed to develop the methodological framework accordingly and derive associated relationships from it.

4 Methodological Framework

Through the literature review (see section 3, incl. associated qualitative analysis) and assimilation of

insights gained, a multitude of technologies were identified as pivotal contributors or enablers for enhanced decision-making for integrating robotics into human-centered construction processes. Subsequently, the overarching perspectives ($n = 5$) were derived from the analyzed technology potentials for robotic integration (weightings) and their number of mentions (see literature review and RM). Accordingly, these perspectives as well as corresponding technologies are used in the methodological framework and listed below:

- **Construction perspective:** Utilizing industry-standard BIM-based planning methodologies and associated software (e.g., Autodesk Revit, Graphisoft Archicad, BlenderBIM etc.) with correlating data models (Industry Foundation Classes, IFC, standard developed by buildingSMART for vendor-neutral data sharing) can enable the adoption of more holistic planning respecting the entire life cycle of a building and subsequently enhancing planning, coordination, and execution of robotic tasks by offering a detailed building representation (incl. geometric and semantic data/information).
- **Movement perspective:** Utilizing MoCap systems (e.g., Xsens, Vicon, OptiTrack, Theia Markerless etc.) can enable the integration of explicit construction process data (Biovision Hierarchy, BVH) derived from real-world applications and hereby enhance human-centered (especially ergonomic) planning potentially enhancing construction processes with human-robot-collaboration (HRC) based processes.
- **Work process perspective:** Utilizing work process simulation software incorporating DHM (e.g., ema Work Designer etc.) can enable the integration of economic and ergonomic assessments within a simulated work process environment (incl. robotic and HRC processes) to enhance human-centered planning ideally enhancing construction processes with economically and ergonomically assessed processes.
- **Machine perspective:** Utilizing (vendor-neutral or agnostic) robot frameworks for programming, simulating, controlling, and monitoring production processes (e.g., ROS, RoboDK, HAL Robotics Framework etc.) can provide an intuitive and user-friendly way for the subsequent utilization of robotics in an end-to-end data-/workflow.
- **Interface perspective:** Utilizing conversion or interfacing software (e.g., based on Application Programming Interface programming etc.) can enable the integration of multidisciplinary data

sets within a construction process.

Despite the added value these technologies offer, they are currently still largely applied independently from each other and not interlinked in any tangible, overarching methodology within the AEC industry. Therefore, data and correlating information is often siloed, not appropriately prepared (e.g., fragmented due to different naming conventions etc.), accumulated and interfaced regarding downstream processes. Subsequently this methodological framework aims to integrate and interface aforementioned technologies and associated perspectives with traditional construction processes incorporating adequate data preparation and accumulation possibilities through enriched data sets (semantic enrichment) adding additional contextual information.

Considering the identified technologies, derived perspectives and traditional construction processes, the methodological framework starts with defining the project requirements (see Figure 3). In the construction context, this involves conveying the client's design intent via documentation to the planner (construction perspective). This sets the stage for the subsequent CAD-/BIM-based modelling of the desired building, building component or product design. During this stage, the work and machine process perspective become crucial in establishing the requirements for incorporating related process data/information in early planning phases to mitigate potential correction loops. Here, the utilization of semantic enrichment via custom property sets (related to domain-specific entities) leveraging BIM-based planning methodologies becomes essential, offering the capability to automate the retrieval and inference of missing or needed data/information (LOIN) for downstream processes. Custom property sets are generally defined using several inputs (e.g., name, instances, entities/classes and the underlying properties which can be freely chosen and described using strings, floats, Booleans, integers etc. depending on the desired property). The custom property sets are part of the IFC data model (based on the EXPRESS data modeling language and overarching STEP data model; incl. attributes, relationships, property sets etc.) which can inherent geometric and semantic data/information describing the meaning of its instances. Therefore, the IFC data model can serve as a repository for the infused work and machine process requirements. The work and machine process requirements can involve data/information concerning building components (e.g., dimensions/scales, weights, assembly orders etc.) which are necessary to align consecutive process requirements with the intended design established in the CAD/BIM-based planning phase (construction perspective). For instance, can infused machine requirements regarding the maximum layer height in a concrete 3d-printing use-case

or a maximum gripper width (robotic end effector) in a brick-laying use-case inform the planner about the feasibility of a robotic execution and possibly enhanced economic and ergonomic construction process. In addition, these requirements are not only embedded into the IFC data model via custom property sets but also in an Information Delivery Specification (IDS) to define exchange requirements within an openBIM process enabling a consecutive IDS-rule/specification-based model checking regarding related work and machine process requirements (so-called "specifications"). The IDS allows the definition of certain specifications described by "facets" (related 6 types: entity/class, attribute, classification, property, material, and relation/part of). To define a specification or facet type (in this case property) there is a so-called description (human readable information to elaborate on the requirement), applicability (type of object the specification applies to) and requirement (which information is required) inputs needed. These specifications can refer to the work and machine process requirements subsequently be used with a model checker (e.g., established with IfcOpenShell as open-source software library for processing the IFC data model) to check if the requirements in the IFC are matching the ones defined in the IDS. This step can improve communication, avoid errors, and facilitate better collaborations throughout the entire construction process by establishing data/information consistency and retrieval when needed.

However, this model checking routine should not limit the planner in terms of their design freedom. Therefore, a data exchange to the next planning phase (work process perspective) is still possible even when the requirements are not matched due to the approach that this framework should establish an enhanced decision-making but not limit the design. Subsequently the planner could intentionally design a building which does not match with any of the work or machine process requirements but consciously deciding to continue with it due to their individual design intention.

Afterwards the IFC (incl. the enriched semantics) is imported into a work process simulation software incorporating DHM (CAPP, work process perspective) and enabling work process planning, simulation, analysis, and optimization regarding human-centered construction processes (e.g., bricklaying, concreting, painting, plastering etc.). However, due to interoperability issues regarding the IFC (enriched semantics), the file is converted through a conversion/interfacing software into a software-readable XML-format (structured data). The conversion/interfacing software makes it possible to check for specific relations and properties among the IFC entities. In addition to the advantages a CAPP software incorporating DHM can offer (e.g., enhancing task

allocations, capability of assessing economic and ergonomic aspects etc.) there is also the possibility to import and work with MoCap recordings, enabling an explicit representation of real-life construction processes to plan more application-based, human-centered, and inclusively.

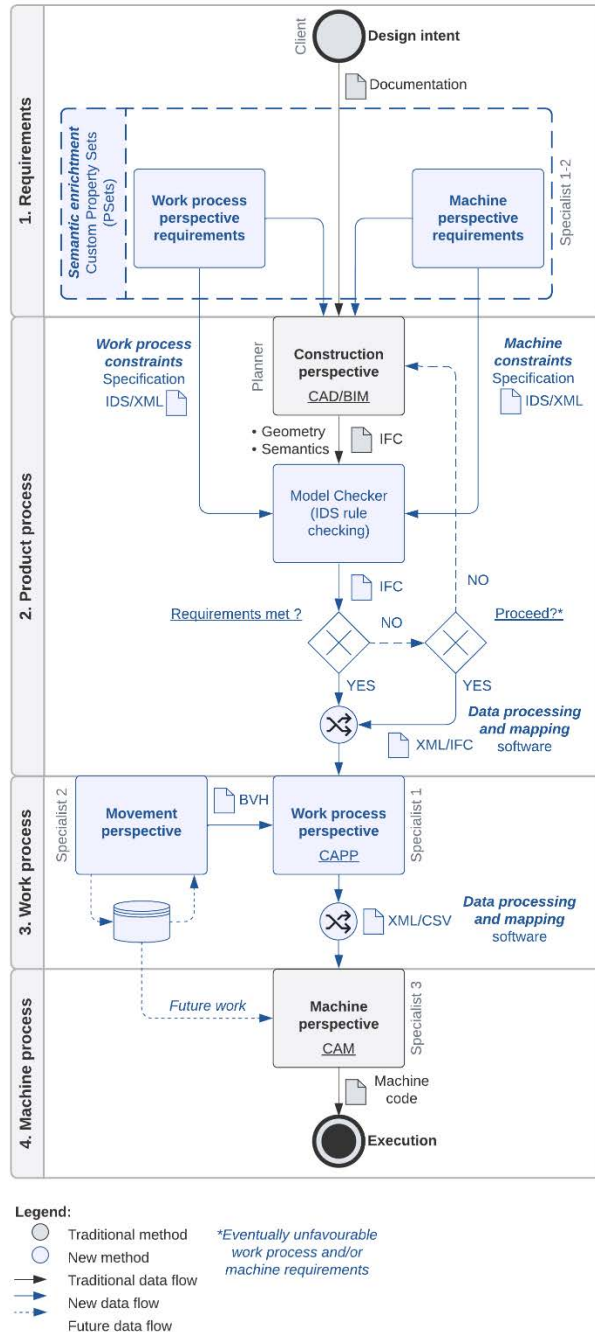


Figure 3. Methodological framework

Therefore, the proposed methodological framework emphasizes on using this capability of integrating MoCap data (BVH) strategically to mitigate diverse economic and ergonomic risks (e.g., derived from repetitive,

mundane, and physically intense tasks etc.). In this context the methodological framework implemented a database (movement perspective) utilizing Python and SQLite, as well as a necessary database management system (DBMS) to subsequently be used to store, but also push and pull data/information as needed. This database can be used as a MoCap recording repository for generic as well as replicable human-centered construction processes and therefore future application. Utilizing MoCap data/information of a human-centered construction process as well as the IFC (incl. enriched semantics), the CAPP software can make a work process assessment considering economic (e.g., timely aspects via MTM-UAS) and ergonomic (e.g., ergonomic aspects via EAWS) guidelines to support with the decision-making regarding the integration of robotics into human-centered construction processes. Furthermore, a robotic construction process can be modelled and compared in the CAPP software leading to a robot simulation which can be exported as a CSV (non-proprietary open format) and subsequently converted/interfaced with the data compartments of the robotic framework. This step includes the conversion of all prepared and accumulated data/information (coordinates/orientation frames for automated toolpath/trajectory planning; action states; task allocations; component information such as sizes, weight etc.) derived from previous processes and the robot simulation in the CAPP software to interface it accordingly for simulation (e.g., checking on singularities or orientation frames being out of reach etc.), programming, control and monitoring. Subsequently the agnostic robotic framework (CAM, machine perspective) enables the program execution and production as desired while building up on prepared, accumulated, and interfaced data/information and thereby accelerating the robot programming.

5 Evaluation and Results

The methodological framework is evaluated with a bricklaying construction process since the use-case is globally applied and holds a pivotal position within the AEC industry owing to its historical importance and extensive utilization. Furthermore, bricklaying as a construction process is one of the most repetitive as well as physically demanding tasks, therefore tends to be economically and ergonomically unfavorable which leads to a need for optimization.

In this context the evaluation of the methodological framework focusses on a load-bearing brick wall assembly (brick size: 24 x 11,5 x 11,3 cm) which is derived from the planner's building model (CAD/BIM, IFC, Autodesk Revit) and a result of the previously conveyed design intent via documentation from the client. In the next step the model (load-bearing brick wall) is

semantically enriched through custom property sets. The infused properties refer to work and machine process requirements [e.g., “assembly order” for task allocations; “brick weight” for ergonomic assessment and consideration of robotic payloads as well as moment of inertia and therefore appropriate toolpath/trajectory planning; “brick dimensions” (length x width x height) for ergonomic assessment and robotic end effector conformity checks e.g. based on the gripper width etc.] to mitigate potential correction loops and enable a consecutive feasibility assessment regarding a robotic construction process. The aforementioned properties are also implemented as part of an IDS (specification) which consecutively is used with an implemented model checker (utilizing IfcOpenShell) to check the IFC if the custom property sets are created accordingly or not.

Afterwards the planner can decide (based on the issue tracking) if he/she wants to continue with the IFC or in case the properties are not set up yet, to return to the CAD/BIM software and adjusting it accordingly.

Afterwards the implemented conversion/interfaces software enables an appropriate data processing and mapping (based on the IFC data model) towards the work process planning (CAPP).

In the next step the model (IFC) is imported into the work process simulation software which incorporates DHM (CAPP, ema Work Designer) and can be used as basis for economic and ergonomic assessments. In this use-case the simulation environment and procedural accuracy is enhanced through explicit movement data/information stored in the BVH-database serving as a human-centered construction process repository. For this purpose, the bricklaying construction process is recorded using an inertial measurement unit (IMU) based MoCap system (Xsens Awinda by Movella). The MoCap recordings are done with a construction/masonry worker (male, 186cm body height) with 13 years of experience within a laboratory setting to establish as viable results as possible.

In this context two process variants were recorded: the first one (see Figure 4 and 5) utilizes a pallet with varying brick positions (4 corners of the pallet with 8 bricks on each corner), whereas the second variant refers to various pick and place heights to draw assumptions based on different heights in relation to the construction worker's body height (see first process variant in Figure 4 and 5).

The process variants are therefore abstracted based on standardized bricklaying construction process practices (incl. a pallet as pick-up setup, plank referring to the insulation offset layer needed and maximum bricklaying height of 170cm referring to the eye-level of the construction worker). After the MoCap recordings are stored in the BVH-database and imported into the CAPP software, the evaluation of the bricklaying construction

process based on the model (IFC) can be initiated.



Figure 4. First process variant, human-based, manual (left: setup, right: MoCap recordings)

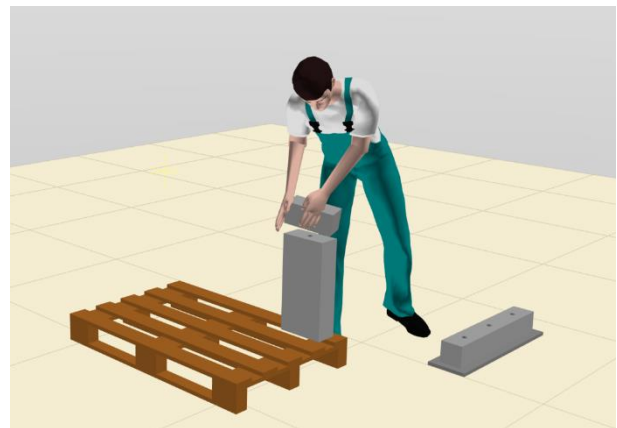


Figure 5. First process variant (CAPP software incorporating DHM, ema Work Designer)

Subsequently, the virtual ergonomics of aforementioned process variants can be evaluated by utilizing the Ergonomic Assessment Work Sheet (EAWS). As the EAWS score recommends a redesign of working processes or environments within a score of 25-50, processes with assigned scores more than 50 shall be urgently modified. Regarding the EAWS analysis of the first process variant, the following table (see table 2) summarizes the influence of the brick positions on human ergonomics as well as required process times (according to the 4 corner positions of the bricks).

Table 2. EAWS Score based on MoCap recordings from the four subprocesses in the first process variant

Brick position/ subprocess	EAWS Score for 8-hour shift
front left	128,5
front right	154
back left	123
back right	116

The comparison of the four subprocesses correlating to the four different brick positions on the pallet, results in a respective EAWS score of more than 100, which indicates an urgent need for a work process redesign.

The second process variant investigates the influence

of the bricklaying height in relation to the construction worker's body height. The worker picks 15 bricks one by one from the front left brick position of the pallet, whereby the maximum stacking height on the pallet comes to 8 bricks. The resulting EAWS score for each subprocess is presented in Table 3.

Table 3. EAWS Score – MoCap results from the subprocesses of the second process variant

Brick ID/ subprocess	EAWS Score for 8-hour shift
1	90
2	71,5
3	99
4	59
5	74
6	66,5
7	81,5
8	66,5
9	36,5
10	35
11	47
12	47
13	79,5
14	58
15	77

Table 3 reveals an EAWS score under 50 in brick 9, 10, 11 and 12, which relates to subprocesses with predominantly straight postures of the human worker without bending or working near the ground.

Furthermore, the table 4 shows the required execution time of the human-based manual process (referring to the first process variant) and the same process but executed by a collaborative robot (Yaskawa HC10, incl. a process variant with adapting the robot velocity with a maximum of 1m/s).

Table 4. Required execution times of various processes

Executor	Time [s]
Human	69
Robot	44
Robot (velocity adoption)	60

These ergonomic and economic assessments based on virtual ergonomics via EAWS score and execution times are derived from CAPP software (ema Work Designer) and subsequently enable an enhanced decision-making concerning the applicability of robotics in human-centered construction processes. In this use-case, the EAWS score indicates (see table 2 and 3) a need for a work process redesign and therefore a new task allocation which is ergonomically more favorable. Therefore, the process was opted with a second one (see Figure 6) where only the collaborative robot (Yaskawa HC10) was executing the construction process accordingly. This led to an optimization in execution time (see table 4) and no

unfavorable ergonomic processes since only the robot is executing the process without human intervention.

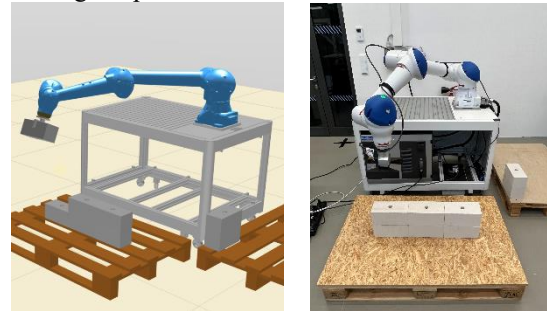


Figure 6. Second process variant, robot based (left: CAPP software incorporating DHM, ema Work Designer, right: real-world execution)

However, the EAWS score also indicates that certain subprocesses (see table 3, grey markings) are not as unfavorable as others, which means that the human could still overtake certain subprocesses in a HRC process variant. However, in the second process executed by the robot, but also in the suggested process variant (HRC), the human could be freed from ergonomically unfavorable as well as repetitive processes and overtake more value adding processes. Moreover, the human could still be part of the process for monitoring purposes making sure that the brick wall is straight (e.g., utilizing a water scale etc.). Furthermore, nowadays the human still offers the highest degree of flexibility since no reprogramming for tasks is needed and the human can react immediately to unforeseen circumstances and is therefore in many use-cases irreplaceable.

However, at the end the decision-maker must decide for a task allocation supported by the economic (execution time) and ergonomic (EAWS) assessments based and optimized through the methodological framework.

Finally, the machine process was evaluated and established using an agnostic robotic framework (HAL Robotics Framework, Grasshopper Plugin). In this step the simulated robot tasks from the previous CAPP software (ema Work Designer) were processed and mapped using the implemented conversion/interfacing software. Here, the robot simulation results including necessary orientation frames etc. could be exported and mapped to a CSV to consecutively weave them into a visual algorithm editor (Grasshopper) where the corresponding coordinates could be automatically interpolated to generate a curve used for the robot toolpath eventually resulting in a more efficient execution.

6 Discussion and Outlook

The IMU-based MoCap system of the methodological framework is commonly used for

applications with a required setup simplicity. However, this simplicity comes with a loss in accuracy. Due to external and internal influences, calculation-driven stochastic errors accumulate over time and lead to a drift and therefore inaccuracy in position determination. That is why in industrial MoCap applications, mostly camera-based (optical) systems are used due to lower error rates (incl. lower time-depending drifts). Considering the accuracy and reliability, a camera-based (optical) MoCap system should be considered in the future.

Moreover, the methodological framework currently works with an implemented IDS-rule/specification-based model checker which works for all IFC entities. Therefore, a direct checking routine regarding a certain geometric representation (e.g., IfcSweptSolid, IfcFacetedBRep etc.) would be possible (currently only checking for custom property sets). However, due to a multitude of geometric representations within an IFC, this possible improvement has to be thoroughly evaluated before implementation.

Finally, the methodological framework could be further enhanced with a software extracting specific orientation frames from the MoCap recording (incl. AI algorithms for semantic action segmentation and classification of demonstrations) which correspond to a desired position and can be utilized as toolpath/trajectory planning basis. This could serve as an alternative kinesthetic and user-friendly programming method enabling users to reprogram without specialized expertise. In this context one could benefit directly from the construction worker's experience which forms the basis for the robot program and consecutive execution.

7 Conclusion

The methodological framework offers a systematic and holistic approach to enable enhanced decision-making regarding the integration of robotics into human-centered construction processes, strategically incorporating technologies while assuring an appropriate data/information preparation, accumulation, and interfacing capabilities. By doing so, it aims to promote a higher adoption of robotics in construction. Compared to existing studies, this approach presents significant contributions by providing a more holistic end-to-end data-/workflow that not only addresses decision-making but also emphasizes the strategic incorporation of cross-domain knowledge (e.g., utilization of software mainly used in the automotive industry, ema Work Designer etc.), human factors and data management. Moving forward, further refinement and implementation of the critical aspects (see section 6) will be essential for maximizing the effectiveness and applicability of the proposed framework.

References

- [1] Tender M. Couto J. P. Gibb A. Fuller P. and Yeomans S. Emerging Technologies for Health, Safety and Well-being in Construction Industry. *Industry 4.0 for the Built Environment*, 20:369-390. Structural Integrity, Springer, Springer, Cham, 2021.
- [2] Agarwal R. Chandrasekaran S. and Sridhar M. Imagining construction's digital future. Online: <https://www.mckinsey.com/capabilities/operations/our-insights/imagining-constructions-digital-future>, Accessed: 08/12/2023.
- [3] Organization for Economic Co-Operation and Development. Analytical house prices indicators. Online: https://stats.oecd.org/Index.aspx?DataSetCode=HOUSE_PRICES, Accessed: 08/12/2023.
- [4] Benachio G. L. F. Freitas M. do C. D. and Tavares S. F. Circular economy in the construction industry: A systematic literature review. *Journal of Cleaner Production*, 260(1):121046, 2020.
- [5] Delgado J. M. D. Oyedele L. Ajayi A. Akanbi L. Akinade O. Bilal M. and Owolabi H. Robotics and automated systems in construction: Understanding industry-specific challenges for adoption. *Journal of Building Engineering*, 26:100868, 2019.
- [6] Zizic M. C. Mladineo M. Gjeldum N. and Celent L. From Industry 4.0 towards Industry 5.0: A Review and Analysis of Paradigm Shift for the People, Organization and Technology. *Energies*, 15(14):5221, 2022.
- [7] Slepicka M. Vilgertshofer S. and Borrmann A. Fabrication information modeling: interfacing building information modeling with digital fabrication. *Construction Robotics*, 6:87–99, 2022.
- [8] Schmailzl M. Spitzhirn M. Eder F. Krüll G. Linner T. Obergrießer M. Albalkhy W. and Lafhaj Z. Towards interfacing human centered design processes with the AEC industry by leveraging BIM-based planning methodologies. In *Proceedings of the 40th ISARC*, pages 325-331, Chennai, India 2023.
- [9] Spengler A. *Digitalisierung der Baustelle - Einstieg in die Robotik im Bauwesen*. DIN/Beuth, Berlin, Wien, Zürich, 2021.
- [10] Shan Ng M. Hall D. Schmailzl M. Linner and T. Bock T. Identifying enablers and relational ontology networks in design for digital fabrication. *Automation in Construction*, 144, 2022.
- [11] Bill M. Müller C. Kraus W. and Bieller S. World Robotics. IFR - International Federation of Robotics, page 12, 2022. Online: https://ifr.org/downloads/press2018/2022_WR_extended_version.pdf, Accessed: 15/12/2023.

Interpretable Machine Learning Approaches for Assessing Maximum Force in Fiber-Reinforced Composites

Soheila Kookalani¹, Erika Parn¹, Ioannis Brilakis²

¹Department of Engineering, University of Cambridge, Cambridge, UK.

²Laing O'Rourke Professor, Department of Engineering, University of Cambridge, Cambridge, UK.

sk2268@cam.ac.uk, eap47@cam.ac.uk, ib340@cam.ac.uk

Abstract

This paper investigates the accurate prediction of the maximum force in fiber-reinforced composites using the CatBoost machine learning algorithm. The study incorporates the Shapley additive explanations technique to enhance interpretability, revealing the significance of the impact of each variable on the output at both local and global scales. The research demonstrates that Shapley additive explanations provide valuable insights into the decision-making process of the machine learning model, identifying influential variables for specific instances and contributing to a comprehensive understanding of the overall model predictions. Notably, the alignment between the feature importance analyses from the machine learning model and Shapley additive explanations reinforces the significance of certain parameters in predicting maximum force as an interfacial property. The study advances the prediction of interfacial properties in fiber-reinforced composites and underscores the value of interpretable machine learning methods in offering insights into complex predictive models.

Keywords –

Machine learning; Interfacial properties; Maximum force; Regression; Fiber-reinforced composites; Interpretability methods.

1 Introduction

Fiber-reinforced composites have become integral materials in civil engineering applications, owing to their remarkable combination of stiffness, strength, and lightweight properties [1]–[3]. The mechanical performance of these composites is primarily dictated by their interfacial properties [4], [5]. The determination of interfacial properties through fiber pullout tests involves labor-intensive and time-consuming experimental and numerical methods. Hence, there is a pressing need for an accurate and efficient alternative for predicting interfacial properties, essential for the design and

customization of composite materials.

In recent years, machine learning (ML) techniques have emerged as promising substitutes for time-consuming simulation processes that offer the advantage of low computational cost and high accuracy [6]–[10]. For instance, Mangalathu and Jeon [11] employed lasso regression for beam-column joints. Yao et al. [12] demonstrated the superiority of two-class support vector regression (SVR) over one-class SVR and logistic regression in mapping landslide susceptibility. Chopra et al. [13] investigated the efficiency of ML models such as decision trees (DT), random forests (RF), and neural networks in estimating concrete compressive strength. Their findings revealed the superior efficiency of the neural network model, followed by the RF method. Additionally, Das et al. [14] introduced a data-driven physics-informed approach for concrete crack estimation, showcasing the capability to predict infrastructure service life based on real-time monitoring data.

In this study, the CatBoost algorithm is employed to predict the maximum force in fiber-reinforced composites. A grid search approach and K-fold cross-validation are employed, utilizing a dataset comprising 922 samples to identify the optimum parameters of the ML model. Understanding why an ML model produces specific estimations and identifying the features influencing those estimations is crucial. Therefore, the Shapley additive explanations (SHAP) method is applied to comprehend the behavior of the ML model. The paper is organized as follows: Section 2 introduces SHAP as an interpretable ML approach; Section 3 presents a numerical example for maximum force prediction, and finally, Section 4 offers concluding remarks and future directions.

2 Interpretable ML approach

In addressing the inherent black-box nature of ML models, particularly in the context of predicting the maximum force in fiber-reinforced composites, this study employs an interpretable ML approach. The opaqueness of such black-box models can lead to a

diminished level of trust and understanding, impeding their broader acceptance and applicability.

In this study, the SHAP approach is employed for model interpretation to mitigate this challenge. The SHAP technique draws inspiration from conditional expectation and game theory [15]. It provides a systematic framework for assessing the impact of individual input features on the predictions of the model, thereby offering a clearer understanding of the decision-making process.

The central idea behind SHAP lies in evaluating the significance of each feature by assessing the increase in estimation error after modifying the values of a given factor. A feature is deemed significant if its manipulation results in a substantial increase in prediction error; conversely, a feature is considered non-significant if its alteration has little impact on the error. This approach allows us to identify and prioritize the input features that contribute most significantly to the predictions of the model, thereby enhancing interpretability.

In essence, SHAP aids in ranking the features based on their contribution to the decision-making process, highlighting the interactions and relationships among these features. The interpretability framework is constructed on the principles of additive feature attribution, wherein the output model is represented as a linear function comprising the sum of the actual values associated with each parameter.

The interpretable framework developed through SHAP provides a transparent and comprehensible representation of the decision logic of the ML model. This approach not only enhances the trustworthiness of predictions by elucidating the role and influence of individual features but also contributes to the broader adoption of ML techniques in assessing the maximum force in fiber-reinforced composites. The resulting interpretable model serves as a valuable tool for practitioners and researchers seeking to bridge the gap between advanced ML capabilities and a clear understanding of the underlying mechanics in composite material.

3 Numerical examples

The assessment of F_{\max} in fiber-reinforced composites

can be effectively carried out using dependable quantitative methods such as the fiber pullout test and simulation. In this study, a thorough compilation of 922 fiber pullout outcomes from existing literature forms a comprehensive dataset [16]–[27].

The input data comprises 11 distinct features, providing details on fiber characteristics, sample preparation environment, and testing conditions. The resulting output, F_{\max} , represents an interfacial property. Table 1 outlines the specified ranges for these parameters.

Table 1. Statistical attributes of dataset.

Attribute	Unit	Minimum	Maximum
Type of fiber	-	1	10
Fiber diameter	μm	5	300
Embedded length	μm	24.5	2018
Young's modulus of fiber	GPa	3.39	294
Poisson's ratio of fiber	-	0.17	0.37
Type of matrix	-	1	10
Young's modulus of matrix	GPa	1.2	3.96
Poisson's ratio of matrix	-	0.31	0.37
Loading rate	m/s	0.0017	6
Prepare temperature	$^{\circ}\text{C}$	20	370
Test temperature	$^{\circ}\text{C}$	-196	120
F_{\max}	N	0.005	1.902

Figure 1 illustrates the correlation matrix of the input parameters, where each correlation coefficient signifies the degree of interaction between two parameters. In this paper, any correlation exceeding 0.70 is considered a significant dependency. The matrix reveals a significant association, with a correlation of 0.95, between embedded length and fiber diameter. Additionally, there is a correlation coefficient of 0.79 between embedded length and fiber type. Moreover, the correlation coefficient between the type of fiber and fiber diameter stands at 0.73. Notably, there are no evident correlations observed for the remaining parameters.

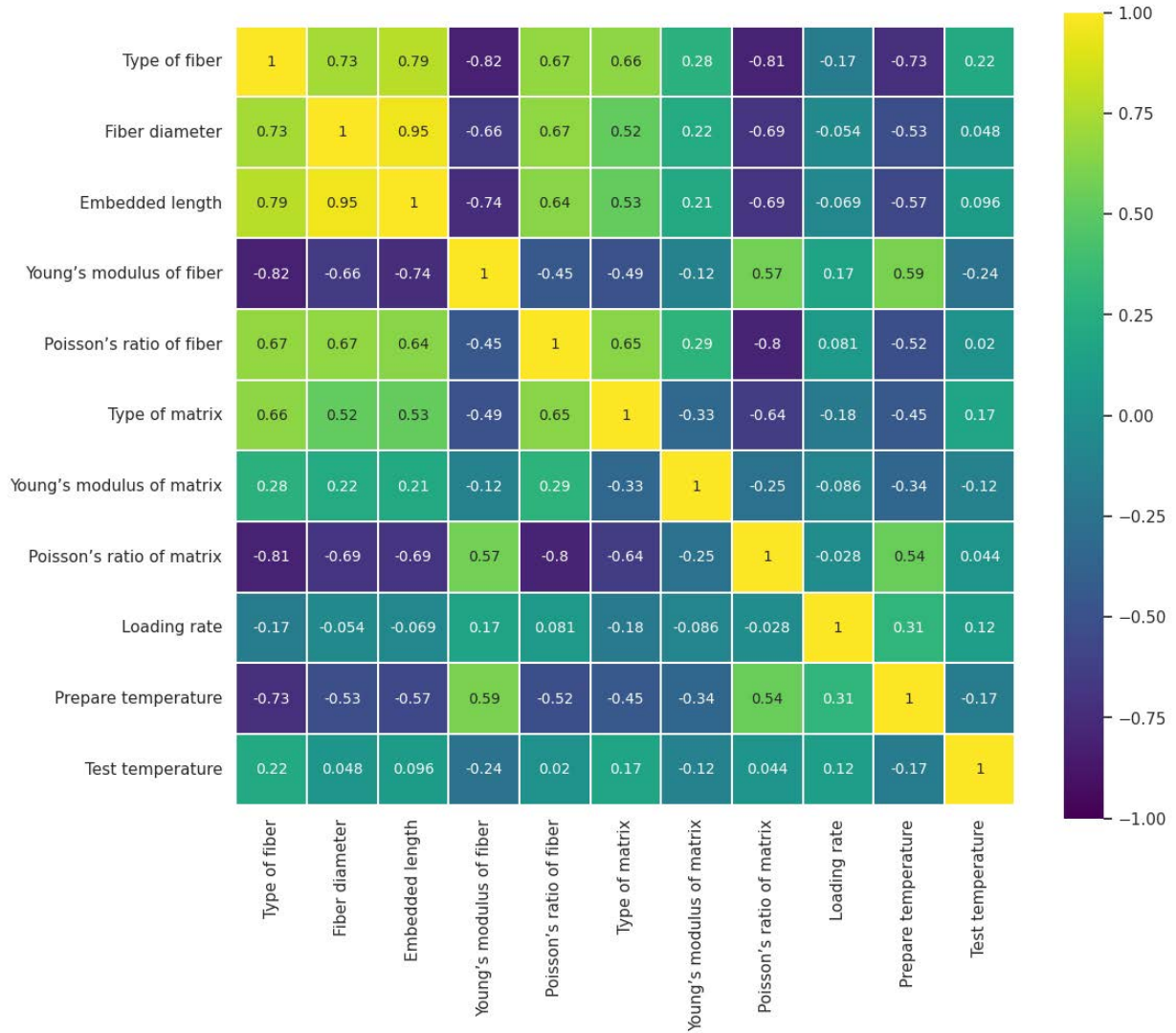


Figure 1. Correlation matrix for input variables

3.1 Regression model

In this paper, a comparative study has been conducted to determine the best regression model. The investigation encompasses several regression algorithms, namely decision tree (DT), AdaBoost, and CatBoost. Through rigorous analysis and comparison, the aim is to identify the most suitable regression model for the given dataset.

The process of fine-tuning the hyperparameters for ML models involves utilizing a grid search approach combined with 10-fold cross-validation to prevent overfitting. Consequently, the hyperparameter values that yield the best performance are chosen for the ML approaches. The optimal hyperparameter values for the regression algorithms are presented in Table 2. The average R^2 and RMSE values resulting from 10-fold cross-validation are presented in Table 3.

Table 2. Optimal hyper parameters.

Model	Optimal configuration
DT	Max_depth=6, min_samples_leaf=1, min_samples_split=2, random_state=3
AdaBoost	Learning_rate=1, n_estimators=50, random_state=0
CatBoost	Depth=8, iterations=1000, learning_rate=0.01

Table 3. Performance of regression models.

Model	Average R^2	Average RMSE
DT	0.988	0.028
AdaBoost	0.982	0.039
CatBoost	0.997	0.018

As a result of the comparative study, the CatBoost model exhibits the highest accuracy. Figure 2 displays the regression plot of the CatBoost model, confirming its high accuracy. The x-axis illustrates the actual data, whereas the y-axis denotes the predicted data. The plot illustrates the correlation between actual and predicted values. A notable observation is that a significant portion of the data points closely align with the regression line, indicating a strong predictive accuracy. In the subsequent subsection, the CatBoost model is further explored as an interpretable method.

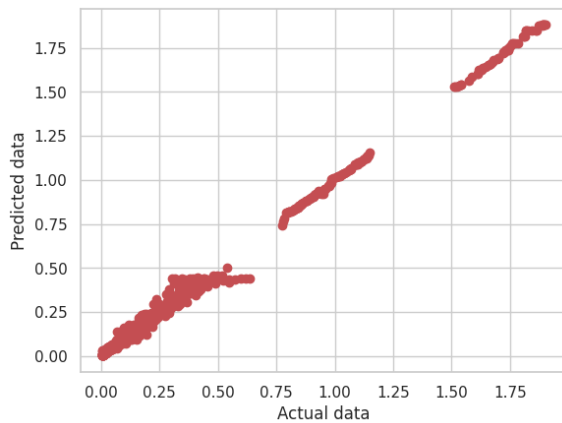


Figure 2. Regression plot

3.2 Interpretable method

Illustrated in Figure 3 is the significance of parameters in shaping the CatBoost model, taking into account their contributions to each tree. The most influential feature is the fiber diameter, followed by the embedded length. Conversely, the loading rate holds the least importance, with the Poisson's ratio of the matrix following as the next less critical variable. Notably, discerning whether an input variable exerts positive or negative effects on the relative importance plots is unfeasible.

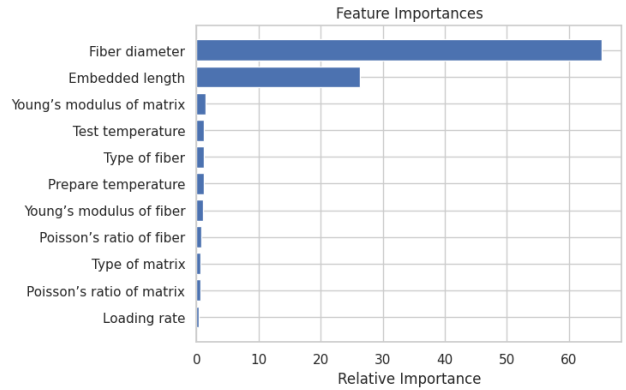


Figure 3. CatBoost importance factor

Figure 4a illustrates the SHAP summary plot, where individual points correspond to Shapely values for the parameters. Each row in the plot contains an equal number of samples. The Shapely values and input variables are represented on the x-axis and y-axis, respectively. The variables are arranged in descending order of importance, with the most crucial variable positioned at the top. Samples with the same SHAP value for a factor are dispersed along the horizontal axes. High variable values are depicted in red, while low values are in blue. The red color signifies the range of values that elevate the SHAP value and, consequently, the associated estimation.

Observations indicate that an increase in the Young's modulus of the fiber, preparation temperature, and Poisson's ratio of the matrix results in a reduction of the SHAP value, leading to a decrease in F_{\max} . Conversely, an increase in embedded length, fiber diameter, and Poisson's ratio of the fiber contributes to an elevation in the F_{\max} value. Figure 4a maintains an equal distribution of points in each row.

In Figure 4b, the global significance factor is presented as the mean of the absolute SHAP value per factor. According to SHAP analysis, the input parameters of embedded length and fiber diameter are identified as the most crucial, aligning with findings from the CatBoost significance variable.

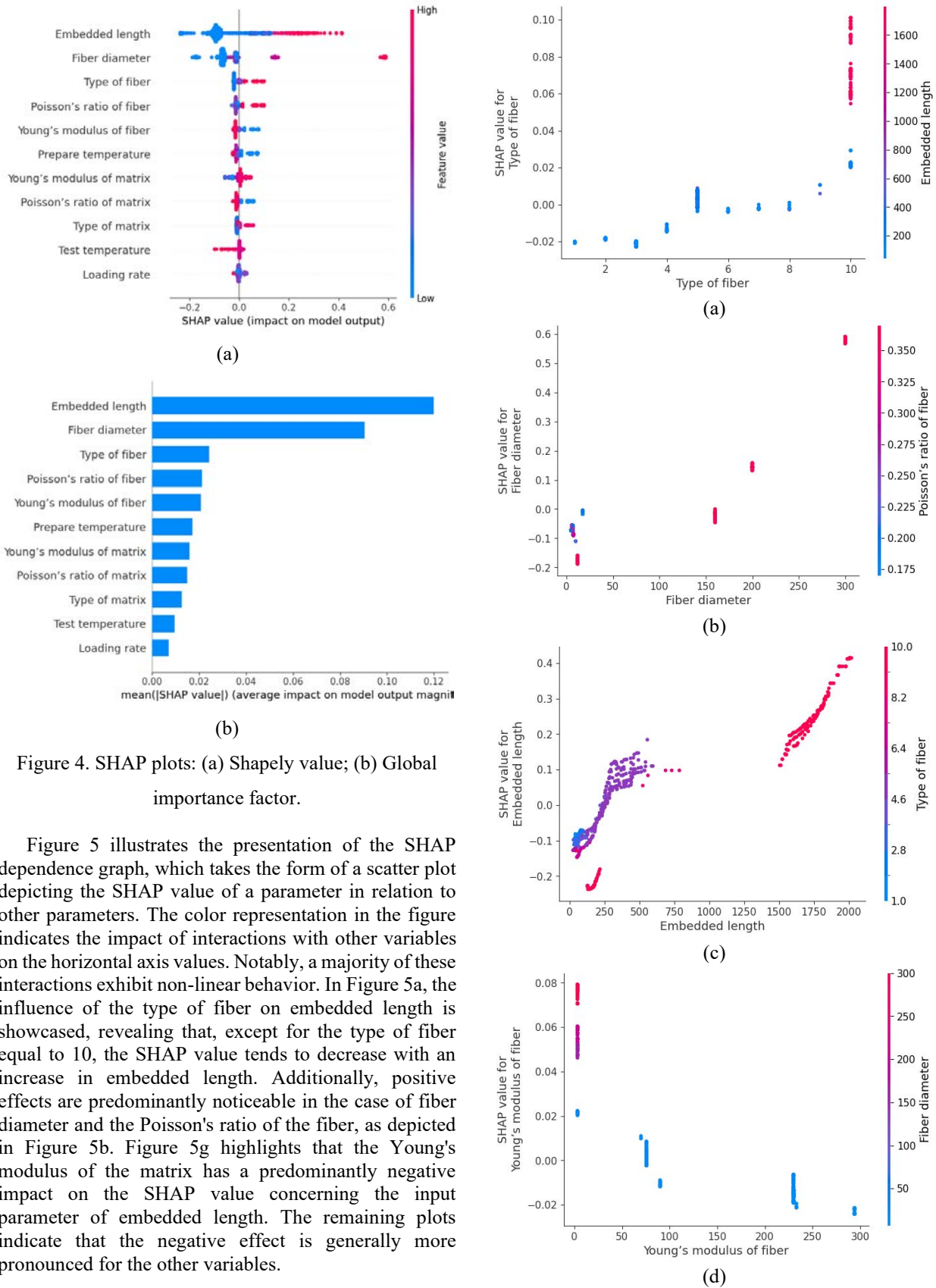


Figure 4. SHAP plots: (a) Shapely value; (b) Global importance factor.

Figure 5 illustrates the presentation of the SHAP dependence graph, which takes the form of a scatter plot depicting the SHAP value of a parameter in relation to other parameters. The color representation in the figure indicates the impact of interactions with other variables on the horizontal axis values. Notably, a majority of these interactions exhibit non-linear behavior. In Figure 5a, the influence of the type of fiber on embedded length is showcased, revealing that, except for the type of fiber equal to 10, the SHAP value tends to decrease with an increase in embedded length. Additionally, positive effects are predominantly noticeable in the case of fiber diameter and the Poisson's ratio of the fiber, as depicted in Figure 5b. Figure 5g highlights that the Young's modulus of the matrix has a predominantly negative impact on the SHAP value concerning the input parameter of embedded length. The remaining plots indicate that the negative effect is generally more pronounced for the other variables.

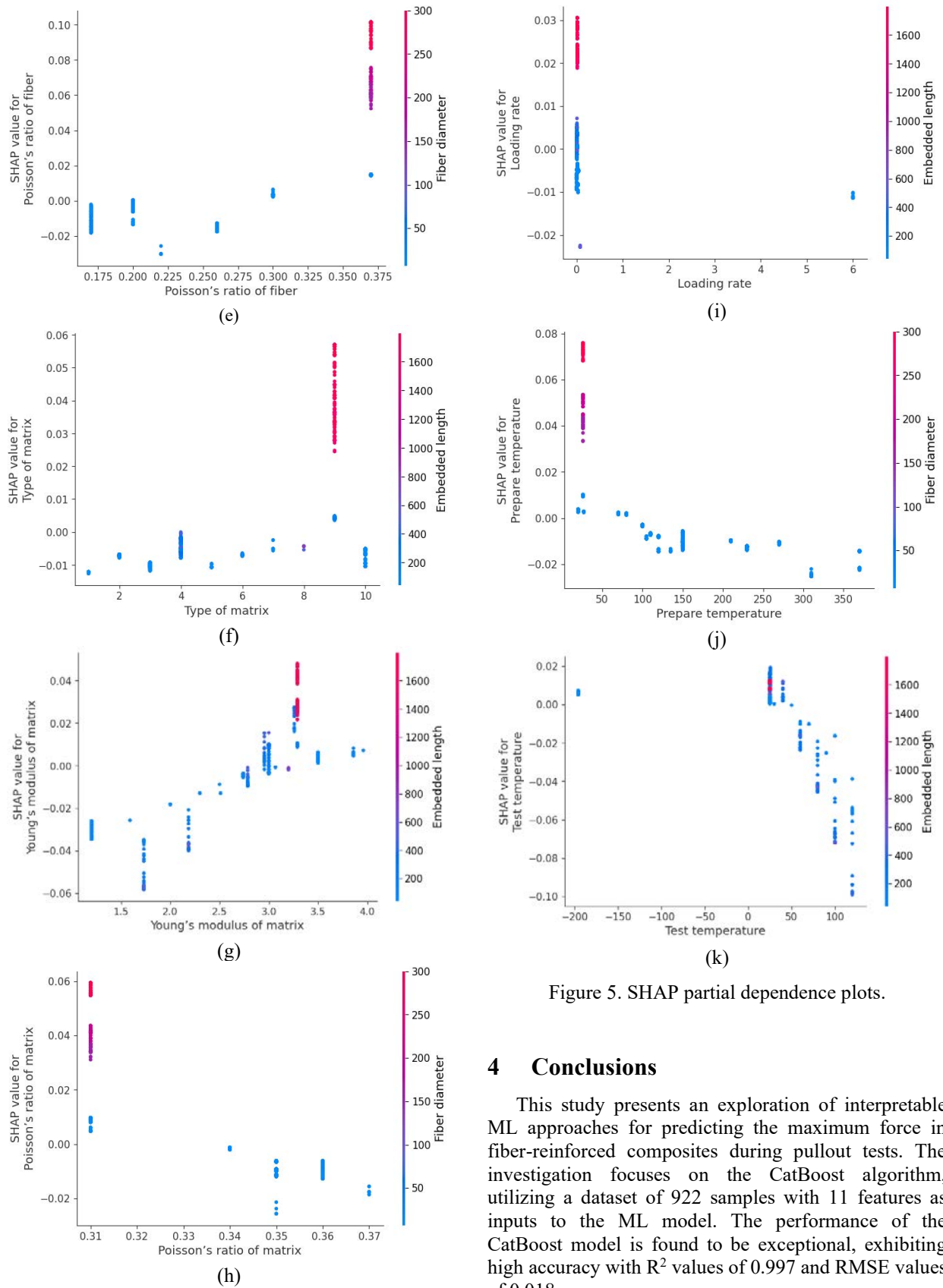


Figure 5. SHAP partial dependence plots.

4 Conclusions

This study presents an exploration of interpretable ML approaches for predicting the maximum force in fiber-reinforced composites during pullout tests. The investigation focuses on the CatBoost algorithm, utilizing a dataset of 922 samples with 11 features as inputs to the ML model. The performance of the CatBoost model is found to be exceptional, exhibiting high accuracy with R^2 values of 0.997 and RMSE values of 0.018.

The CatBoost algorithm is coupled with the SHAP interpretable approach to enhance the transparency and interpretability of the ML model. This step is crucial for elucidating the predictions made by the ML model, providing insights into the relative importance of input features. SHAP, chosen for its comprehensive explanatory capabilities, proves to be a valuable tool in understanding the trends and contributions of individual parameters. Remarkably, the SHAP values highlight embedded length and fiber diameter as the most influential variables, positively impacting the estimation of the maximum force.

Furthermore, the congruence between the features identified as significant by both the CatBoost method and SHAP underscores the reliability and consistency of the proposed approach. This alignment strengthens confidence in the interpretability of the ML model and the robustness of the identified influential parameters.

While this study focuses on fiber-reinforced composites, the methodology developed herein holds promise for broader applications in structural performance estimation across diverse systems. The successful implementation of the proposed approach creates opportunities for future research, which will extend the methodology to predict other interfacial properties of fiber-reinforced composites. This expansion not only contributes to the versatility of the developed methodology but also underscores its potential impact on advancing the understanding and prediction of structural behavior in various engineering applications.

Acknowledgment

The financial support of EPSRC via grant number EP/W018705/1 is gratefully acknowledged.

References

- [1] K. M. Liew, Z. X. Lei, and L. W. Zhang, "Mechanical analysis of functionally graded carbon nanotube reinforced composites: A review," *Compos Struct*, vol. 120, pp. 90–97, Feb. 2015, doi: 10.1016/J.COMPSTRUCT.2014.09.041.
- [2] K. M. Liew, Z. Z. Pan, and L. W. Zhang, "An overview of layerwise theories for composite laminates and structures: Development, numerical implementation and application," *Compos Struct*, vol. 216, pp. 240–259, May 2019, doi: 10.1016/J.COMPSTRUCT.2019.02.074.
- [3] K. M. Liew, Z. Pan, and L.-W. Zhang, "The recent progress of functionally graded CNT reinforced composites and structures," *Sci. China-Phys. Mech. Astron*, vol. 63, no. 3, p. 234601, 2020, doi: 10.1007/s11433-019-1457-2.
- [4] E. B. Callaway, P. G. Christodoulou, and F. W. Zok, "Deformation, rupture and sliding of fiber coatings in ceramic composites," *J Mech Phys Solids*, vol. 132, p. 103673, Nov. 2019, doi: 10.1016/J.JMPS.2019.07.016.
- [5] W. H. Liu, L. W. Zhang, and K. M. Liew, "Modeling of crack bridging and failure in heterogeneous composite materials: A damage-plastic multiphase model," *J Mech Phys Solids*, vol. 143, p. 104072, Oct. 2020, doi: 10.1016/J.JMPS.2020.104072.
- [6] S. Kookalani and B. Cheng, "Structural Analysis of GFRP Elastic Gridshell Structures by Particle Swarm Optimization and Least Square Support Vector Machine Algorithms," *Journal of Civil Engineering and Materials Application*, vol. 5, no. 3, pp. 139–150, Sep. 2021, doi: 10.22034/JCEMA.2021.304981.1064.
- [7] S. Kookalani and B. Cheng, "Structural Analysis of GFRP Elastic Gridshell Structures by Particle Swarm Optimization and Least Square Support Vector Machine Algorithms," *Journal of Civil Engineering and Materials Application* is published by Pendar Pub, vol. 5, no. 3, pp. 139–150, 2021, doi: 10.22034/jcema.2021.304981.1064.
- [8] S. Kookalani, B. Cheng, and S. Xiang, "Shape optimization of GFRP elastic gridshells by the weighted Lagrange ϵ -twin support vector machine and multi-objective particle swarm optimization algorithm considering structural weight," *Structures*, vol. 33, pp. 2066–2084, 2021, doi: 10.1016/j.istruc.2021.05.077.
- [9] S. Kookalani, S. Nyunn, and S. Xiang, "Form-finding of lifting self-forming GFRP elastic gridshells based on machine learning interpretability methods," *STRUCTURAL ENGINEERING AND MECHANICS*, vol. 84, no. 5, pp. 605–618, 2022.
- [10] S. Kookalani, B. Cheng, and J. L. C. Torres, "Structural performance assessment of GFRP elastic gridshells by machine learning interpretability methods," *Frontiers of Structural and Civil Engineering*, vol. 16, no. 10, pp. 1249–1266, 2022, doi: 10.1007/s11709-022-0858-5.
- [11] S. Mangalathu and J. S. Jeon, "Classification of failure mode and prediction of shear strength for reinforced concrete beam-column joints using machine learning techniques," *Eng Struct*, vol. 160, pp. 85–94, 2018, doi: 10.1016/j.engstruct.2018.01.008.
- [12] X. Yao, L. G. Tham, and F. C. Dai, "Landslide susceptibility mapping based on Support Vector

- Machine: A case study on natural slopes of Hong Kong, China,” *Geomorphology*, vol. 101, no. 4, pp. 572–582, 2008, doi: 10.1016/j.geomorph.2008.02.011.
- [13] P. Chopra, R. K. Sharma, M. Kumar, and T. Chopra, “Comparison of Machine Learning Techniques for the Prediction of Compressive Strength of Concrete,” *Advances in Civil Engineering*, 2018, doi: 10.1155/2018/5481705.
- [14] S. Das, S. Dutta, C. Putcha, S. Majumdar, and D. Adak, “A Data-Driven Physics-Informed Method for Prognosis of Infrastructure Systems: Theory and Application to Crack Prediction,” *ASCE ASME J Risk Uncertain Eng Syst A Civ Eng*, vol. 6, no. 2, p. 04020013, 2020, doi: 10.1061/ajrua6.0001053.
- [15] S. M. Lundberg and S. I. Lee, “A unified approach to interpreting model predictions,” *Adv Neural Inf Process Syst*, vol. 30, 2017.
- [16] L. Wenbo *et al.*, “Interfacial shear strength in carbon fiber-reinforced poly(phthalazinone ether ketone) composites,” *Polym Compos*, vol. 34, no. 11, pp. 1921–1926, Nov. 2013, doi: 10.1002/PC.22599.
- [17] C. Zhi, H. Long, and M. Miao, “Microbond testing and finite element simulation of fibre-microballoon-epoxy ternary composites,” *Polym Test*, vol. 65, pp. 450–458, Feb. 2018, doi: 10.1016/J.POLYMERTESTING.2017.12.029.
- [18] B. Liu, Z. Liu, X. Wang, G. Zhang, S. Long, and J. Yang, “Interfacial shear strength of carbon fiber reinforced polyphenylene sulfide measured by the microbond test,” *Polym Test*, vol. 32, no. 4, pp. 724–730, Jun. 2013, doi: 10.1016/J.POLYMERTESTING.2013.03.020.
- [19] Q. Li, G. Nian, W. Tao, and S. Qu, “Size effect on microbond testing interfacial shear strength of fiber-reinforced composites,” *Journal of Applied Mechanics, Transactions ASME*, vol. 86, no. 7, Jul. 2019, doi: 10.1115/1.4043354/726015.
- [20] M. Sato, J. Koyanagi, X. Lu, Y. Kubota, Y. Ishida, and T. E. Tay, “Temperature dependence of interfacial strength of carbon-fiber-reinforced temperature-resistant polymer composites,” *Compos Struct*, vol. 202, pp. 283–289, Oct. 2018, doi: 10.1016/J.COMPSTRUCT.2018.01.079.
- [21] Q. Li, G. Nian, W. Tao, and S. Qu, “Temperature-Dependent Interfacial Debonding and Frictional Behavior of Fiber-Reinforced Polymer Composites,” *Journal of Applied Mechanics, Transactions ASME*, vol. 86, no. 9, Sep. 2019, doi: 10.1115/1.4044017/955633.
- [22] M. Yan *et al.*, “Simulation and measurement of cryogenic-interfacial-properties of T700/modified epoxy for composite cryotanks,” *Mater Des*, vol. 182, p. 108050, Nov. 2019, doi: 10.1016/J.MATDES.2019.108050.
- [23] R. Dsouza *et al.*, “3D interfacial debonding during microbond testing: Advantages of local strain recording,” *Compos Sci Technol*, vol. 195, p. 108163, Jul. 2020, doi: 10.1016/J.COMPSCITECH.2020.108163.
- [24] M. Nishikawa, T. Okabe, K. Hemmi, and N. Takeda, “Micromechanical modeling of the microbond test to quantify the interfacial properties of fiber-reinforced composites,” *Int J Solids Struct*, vol. 45, no. 14–15, pp. 4098–4113, Jul. 2008, doi: 10.1016/J.IJSOLSTR.2008.02.021.
- [25] N. S. Choi and J. E. Park, “Fiber/matrix interfacial shear strength measured by a quasi-disk microbond specimen,” *Compos Sci Technol*, vol. 69, no. 10, pp. 1615–1622, Aug. 2009, doi: 10.1016/J.COMPSCITECH.2009.03.012.
- [26] H. Wang, X. Zhang, Y. Duan, and L. Meng, “Experimental and Numerical Study of the Interfacial Shear Strength in Carbon Fiber/Epoxy Resin Composite under Thermal Loads,” *Int J Polym Sci*, vol. 2018, 2018, doi: 10.1155/2018/3206817.
- [27] X. Wang *et al.*, “Effects of thermal residual stress on interfacial properties of polyphenylene sulphide/carbon fibre (PPS/CF) composite by microbond test,” *J Mater Sci*, vol. 51, no. 1, pp. 334–343, Jan. 2016, doi: 10.1007/S10853-015-9251-2/TABLES/2.

An OpenBIM-based Framework for Lifecycle Management of Steel Reinforcement

Mingkai Li¹, Yuhan Liu², Boyu Wang² and Jack C. P. Cheng¹

¹Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong SAR

²China State Construction Science and Technology Limited, Hong Kong SAR

mlicj@connect.ust.hk (Mingkai Li), ceicheng@ust.hk (Jack C. P. Cheng)

Abstract –

Reinforced concrete (RC) is commonly used in modern building construction. Steel reinforcement (or rebar) design and fabrication are important tasks for the construction of RC buildings. However, these processes are currently conducted semi-automatic, requiring large human effort and processing time. Besides, there is a lack of efficient approach for progress monitoring of onsite rebar installation, and the deviation between actual rebar layout and as-design model can not be accurately analyzed and recorded. Therefore, this study aims to develop an openBIM-based framework for rebar lifecycle management, concerning the design, manufacturing, assembly, renewal and demolition. Firstly, an automatic clash-free rebar design optimization approach is proposed integrating optimization algorithms and AI techniques. Secondly, A open digital workflow based on information delivery manual (IDM), model view definition (MVD), and Industry Foundation Classes (IFC), and the integration with machine fabrication code is developed to support rebar prefabrication. Thirdly, a BIM-based progress monitoring approach is proposed integrating laser scanning, while the rebar BIM model is updated with as-built data. Moreover, the applications of as-built rebar BIM model in maintenance, renewal and demolition are discussed.

Keywords –

Steel reinforcement design; Design optimization; Building information Modelling (BIM); Progress monitoring; Prefabrication; Graph Neural Network

1 Introduction

Reinforced concrete (RC) stands as a prevalent construction material in modern building construction. The design and fabrication of steel reinforcement (commonly known as rebar) are important and necessary tasks for the construction of RC buildings, significantly

affecting the construction time, material consumption and waste generation. Current rebar design process relies heavily on manual participation, which is time-consuming and error-prone. Structural engineers usually conduct rebar design using structural analysis results, and the optimality of rebar design is largely influenced by the experience and expertise of engineers. Besides, the existing rebar design approach operates at the individual member level, often overlooking potential clashes and buildability challenges arising from the interaction between RC members. The identification and resolution of rebar clashes typically rely on manual efforts at the construction site, resulting in time-intensive processes. Sometimes, achieving a clash-free layout is challenging within the given rebar design. Moreover, the manual conversion of rebar detailing into machine codes for automated rebar prefabrication necessitates additional time and labor. As the industry promotes prefabrication and low-carbon construction, there is a pressing need to devise a novel approach for automated rebar design optimization and prefabrication.

There have been some related research efforts in the past attempting to solve some aspects of the above problems. Mangal et al. [1] proposed an automatic rebar design optimization approach for RC frames using two-stage genetic algorithm and BIM. Eleftheriadis et al. [2] developed an automated rebar specification method for flat slabs using BIM and finite element analysis (FEA). Some customized constructability considerations are integrated into the method to generate realistic rebar layout. Genetic algorithms were also incorporated into other studies for rebar design optimization of RC components [3-4]. Vaez and Qomi [5] presented a continuous optimization approach for the rebar design of RC shear walls. However, all the above methods are time-demanding and cannot accumulate the design experience from previous design cases like a human engineer does, hence they are generally computationally inefficient. Some studies tried to identify and resolve the rebar clashes. Mangal et al. [1] defined an indicator called clash number to count the number of clashes in

beam-column joint and leveraged metaheuristic algorithm (MA) to reduce this indicator and thereby resolve clashes. However, this method assumes that beam and column rebar cannot be directly adjacent to each other, and produces a loose arrangement, which may not be applicable to those joints with large amount of steel. Liu et al. [6] proposed to formulate rebar clash avoidance as a path planning problem. However, the generated path may not satisfy the rebar shape code, undermining its practicality. Considering the above shortcomings, it is necessary to develop an automatic clash-free rebar design optimization approach.

To realize automatic rebar prefabrication, a rebar BIM model is the first step, with which the prefab-related information in it can be further extracted and transformed to fabrication machine codes for rebar fabrication. Following this, other materials like rebar detailed drawing can also be automatically generated with customized plugins to facilitate the assembly. However, currently in industrial practice, there is no a universal program/tool to automatically import rebar design results into commonly used BIM software like Revit or Tekla. Worse still, though the rebar design and prefabrication involve a lot of software and customized tools, there is a lack of mature software ecosystem or pipeline to ensure smooth and accurate information passing among them.

As for the progress monitoring of rebar installation, Turkan classified rebars as secondary construction objects and conducted preliminary experiments on automated rebar recognition and tracking for RC

columns using laser scanning [7]. However, the approach has not been validated in other types of components and could not satisfy the practical use. Moreover, research on the applications of as-built rebar BIM model is still lacking.

The objective of this paper is to develop an openBIM-based framework for rebar lifecycle management. An automatic clash-free rebar design optimization approach is proposed integrating MAs and graph neural networks (GNN). Besides, an open digital workflow based on information delivery manual (IDM), model view definition (MVD), and Industry Foundation Classes (IFC), and the integration with BundesVereinigung der Baustoffe (BVBS) code is developed to support rebar prefabrication. In addition, a rebar installation progress monitoring method is developed using 3D laser scanning. Last but not least, the applications of the as-built rebar BIM model in building maintenance, renewal and demolition are discussed.

2 Methodology

This paper proposes a BIM-based framework to support the automation of clash-free rebar design optimization and prefabrication, as shown in Figure 1, which consists of 4 main modules, namely (1) automatic clash-free rebar design optimization, (2) BIM-based rebar prefabrication automation, (3) BIM-based progress monitoring, and (4) BIM-supported building renewal and demolition.

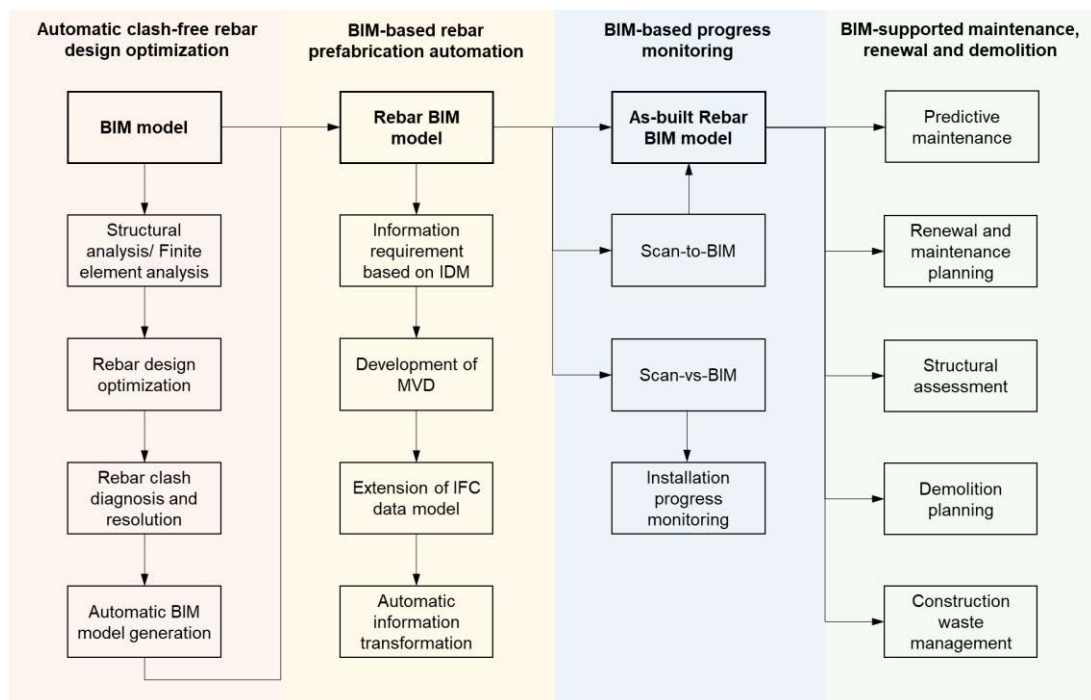


Figure 1. Proposed openBIM-based framework

2.1 Automatic Clash-free Rebar Design Optimization

To realize rebar design optimization, the first step is the establish the optimal formulation of rebar design, based on which optimization algorithms can then be leveraged to achieve the optimal design for a given design case. We also propose to adopt GNN to accelerate the design process and realize rebar clash avoidance.

2.1.1 Optimal Design Formulation

The optimal formulation for rebar design specifically includes the identification of design variables, searching space and design constraints, and the establishment of objective function.

Design variables: The general process to identify the design variables for one type of RC components include 3 steps: (1) collecting the typical rebar layouts, (2) dividing the rebar layouts into rebar groups, and (3) assigning design variables to each rebar group to describe the layout. For elongated RC components, the division of rebar groups are usually more intuitive. For example, the typical longitudinal rebar layout for a two-span continuous beam is shown in Figure 2(a), and 5 rebar groups can be identified. Each number in the figure refers to a rebar group. For each rebar group, since the number of rebar layers and number of rebars are variable, it is inconvenient to determine the number of design variables if each rebar is treated as a design variable. Therefore, we simplify the problem by assigning only 1 variable called rebar combination index to each rebar group, which refers to a detailed rebar layout in our rebar combination dataset. On the other hand, strip method is usually used for the rebar design of flat RC components like slab and walls, which separates components into several strips (as shown in Figure 2(b) with different colors), and assigns rebar layout to each strip.

Searching space: To ensure the optimized rebar design can be found on the market and comply with the industry practice, we establish a dataset containing all practical combinations of rebars with market-supplied diameters. This dataset is set as the searching space for optimization, and can be customized according to the need and preference of users. For example, if the user prefers using only one kind of rebar diameter in one rebar layer, those rebar layouts considering two or more kinds of diameters can be removed to generate a new dataset.

Design constraints: First of all, the steel area of selected rebar layout in each section should not be less than the required steel area. Besides, the rebar layout should follow requirements specified in building codes. Other code-stipulated requirements are also needed to be incorporated into the formulation, such as lapping and anchorage length, which are not all listed here.

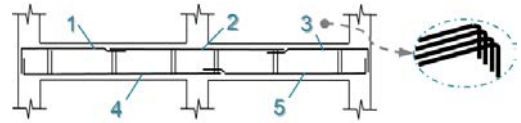
Objective function: Most existing optimization

methods take material consumption or total weight as the objective. However, as the construction industry faces labor shortages and increasing labor costs, buildability has become a increasingly important metric. We incorporate the evaluation of the buildability of rebar design into optimization, which is quantified by calculating the time required to install a specific rebar design. Therefore, the multi-objective optimization is formulated as follows.

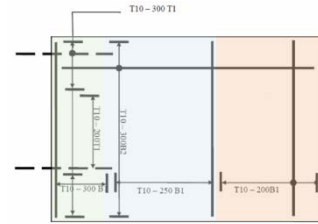
$$C = w \times M_s \times c_s + (1 - w) \times T_i / \tau \times c_l \quad (1)$$

$$T_i = T_p + T_t = N_{bar}/p_p + N_{tie}/p_t \quad (2)$$

where w is the weighting factor to take account of two objectives – material consumption and buildability, M_s refers to the mass of steel rebar (in ton), c_s refers to the unit price for steel (in \$/ton), T_i refers to the time required for installation, τ is a constant 480 min/man-day for unit conversion considering 8-hour work for one day, c_l refers to the labor cost per man-day (in \$/man-day). In Equation (2), T_p and T_t refer to the time (in min) of rebar placing and rebar tying, N_{bar} refers to the number of rebars, N_{tie} represents the number of ties, which can be easily counted for the given longitudinal and transverse rebar layouts, p_p and p_t denote the production rates for rebar placing and rebar tying, respectively [8].



(a) Typical longitudinal rebar layout for continuous beams



(b) Typical rebar layout for RC slabs based on strip method

Figure 2. Typical rebar layouts for common RC components

2.1.2 GNN-based Design Proposal

GNNs are a class of machine learning models designed to analyze and make predictions on graph-structured data, such as social networks, molecular structures, or citation networks [9]. GNNs operate by learning representations of nodes and edges in a graph, capturing relational information and structural dependencies within the data. These models have

demonstrated significant success in tasks like node classification, link prediction, and graph classification, making them a powerful tool for analyzing complex interconnected data sets. In this regard, GNN can be leveraged to improve the computational efficiency of rebar design, which involves a lot of rebar groups interacting with each other.

The workflow of GNN-based design proposal and MA-based design optimization are presented in Figure 3. To unlock the potential of GNN, the first step is to represent the rebar design problem in the format of graph. For elongated RC components like beams and columns, each rebar group in Section 2.1.1 can be treated as a node in the graph representation, while the edge can be established according to the interrelationship between rebar groups, as shown in Figure 4(a). Each node has attributes containing geometrical, positional and design information for the corresponding rebar group, while each edge indicates that the connected rebar groups will affect each other in term of rebar number. GNN can then aggregate information on the graph and provide predictions for each node, which will be the initial design for each rebar group. For flat RC components like slabs and walls, the graph representation can be established based on the finite element mesh, as shown in Figure 4(b). Each node refers to a vertex of the finite element, containing three attributes: x-coordinate, y-coordinate and required steel area. The edges are established using full connection among the vertexes of each finite element, showing that they together determine the distribution of required steel area in the element. GNN is then leveraged to conduct graph-level prediction, providing the positions of strip segmentations and supporting rebar design in each strip.

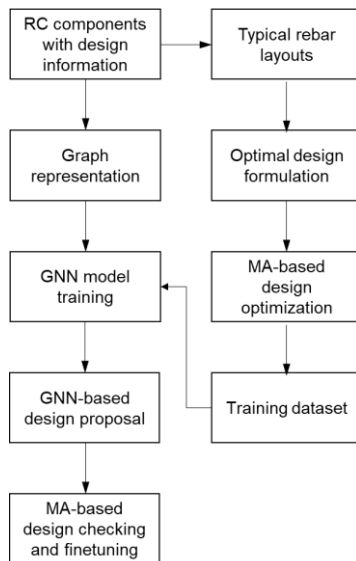
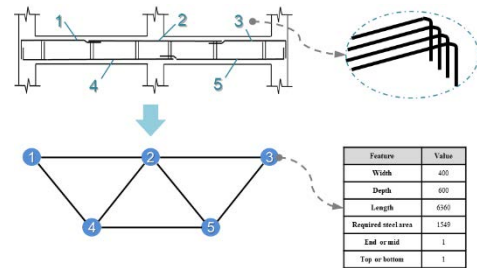
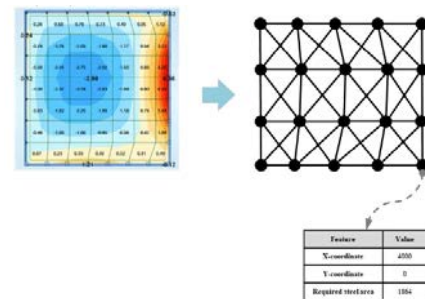


Figure 3. GNN-based design proposal

Several MAs including genetic algorithm (GA), particle swarm optimization (PSO), artificial bee colony (ABC) and ant colony optimization (ACO) are adopted to generate optimal design for thousand of collected design cases based on the optimal formulation in Section 2.1.1. The best design among the optimized results from different MAs are selected as the optimal solution to the design case. In this way, these design cases with optimal designs are used to train GNN. Since GNN aims to uncover patterns and relationships within data, it may not guarantee the generated design is optimal. Therefore, MA is used to further optimize the initial design from GNN and conduct code-compliance checking. This hybrid mechanism not only makes use of the advantage of GNN to accumulate design experience to accelerate the design process, but also exerts the searching ability of MA to ensure the optimality of final design.



(a) Elongated components (RC beam)



(b) Flat components (RC slabs)

Figure 4. Graph representations for elongated and flat RC components

2.1.3 GNN-based Clash Avoidance

The framework for clash avoidance is presented in Figure 5. Rebar clashes are classified as solvable or unsolvable clashes, depending on whether the rebar clash can be resolved without changing the rebar design. Unsolvable clashes can be found in areas with large amount of rebars, where there are few spaces for repositioning or reshaping rebar to avoid clash. When an unsolvable clash is identified, rebar design adjustment is necessary, and the revised rebar layout will go through the classification again until a clash-free layout or a

solvable clash is achieved. On the other hand, a solvable clash can be resolved by only repositioning or/and reshaping some rebars.

To identify and classify rebar clash, we developed the vector representation of rebar clash and formulate the clash avoidance as an optimization problem [8]. Graph representation was further proposed to enable the adoption of GNN to realize efficient clash diagnosis [10]. Thousands of rebar clash cases were collected and presented using the vector representation, and MAs were leveraged to determine the class of each clash case. These clash cases with labels are then used to train the GNN for clash diagnosis.

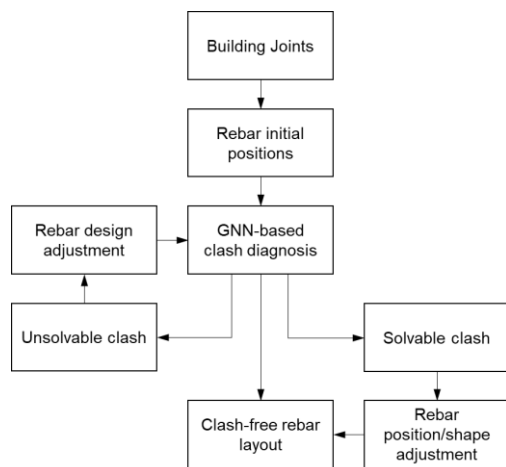


Figure 5. GNN-based clash avoidance framework

2.2 Automatic Rebar Prefabrication

As the construction industry embraces digitalization and industrialization, the significance of BIM and IFC is increasing as collaborative approaches to support data interoperability and off-site prefabrication. BVBS is a type of machine code for rebar fabrication, containing rebar geometrical information for fabrication machine to conduct automatic rebar processing [11]. To realize automatic rebar prefabrication, an openBIM-based framework for BIM-BVBS integration is proposed with extended IFC data model, as shown in Figure 6.

2.2.1 Information Requirements based on IDM

IDM provides detailed guidance on how information should be exchanged and delivered throughout the lifecycle of a construction project, particularly in the context of BIM. The IDM process includes identifying the information requirements for prefabrication of rebars, and creating the process map to display the information flow and data exchanged among engaging stakeholders. The former is realized following BVBS specification and rebar shape code BS8666:2005, while the latter is developed by linking the product (prefabricated steel reinforcement), resources (material and labor) and

control (schedule) in different phases of building lifecycle. The detailed information flow in the off-site manufacturing and prefabrication of rebars should be emphasized and integrated in the process map to support rebar prefabrication.

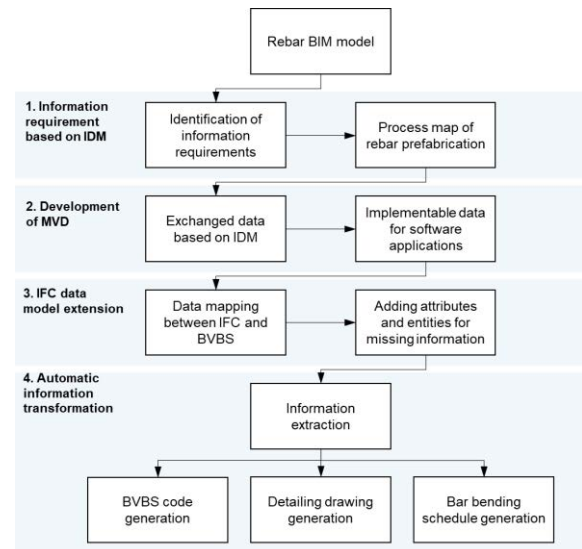


Figure 6. Proposed methodology for BIM-BVBS integration and BIM-based applications

2.2.2 Development of MVD

MVD refers to predefined subsets of information and views within a BIM model, which helps standardize and streamline information exchange in the BIM process, ensuring that stakeholders receive the relevant data they need for their specific roles or tasks. These definitions specify how information should be structured, exchanged, and displayed for a particular purpose or stage of the building lifecycle. Following the IDM process, MVD is leveraged to display the necessary information for automatic prefabrication, which also indicates the needful extension of existing IFC data model. New attributes related to rebar prefabrication are added to different aspects of existing MVD structure:

- “Reinforcing bar attributes”: mark number, rebar length, rebar quantity, rebar diameter;
- “Owner and status information”: project number, drawing number, revision index;
- “Steel material association”: steel grade;
- “Rebar bending attributes”: segment angle.

2.2.3 Extension of IFC Data Model

The conversion of IFC data to BVBS needs to be precise enough to guarantee that all attributes needed for cutting and bending are covered, which highlights the necessity to extend the existing IFC schema.

The first step is the decomposition and relations of *IfcProcess*. Based on the proposed process map and

MVD, the process of rebar prefabrication can be classified using *IfcTask*, which is the sub-entity of *IfcProcess*. Four main tasks are included, including planning, manufacturing, delivery and assembly. Next, the process information model of steel reinforcement fabrication is established by connecting the IFC entities with relationships among the specific tasks. The entities are located through the entity inheritance of those abstract entities according to the nature of the product, resource or control. Taking the manufacturing process for illustration, it operates on the material resource “metal”, which can be found in the predefined type of the *IfcConstructionMaterialResource*, and the product related to this process is indicated as *IfcReinforcingBar* connected by *IfcRelAssignsToProcess*. Linked to the sub-entities of *IfcControl*, all the related products, resources, and the manufacturing process are under its supervision. The product and the process are included in the work plan on the manufacturing arrangement, while the related progress of the resource “metal” is recorded in the work calendar through *IfcRelAssignsToControl*.

In order to support seamless data conversion, the IFC data model needs to be further extended. IFC data model extension includes the data mapping between IFC and BVBS, and the addition of new IFC entities with reference to information requirements in IDM. Figure 7 shows the framework for data mapping between IFC and BVBS, in which the information requirements for generating the BVBS machine codes is clarified. At the data collection level, the list of attributes required by BVBS is extracted. IDM can be utilized as well to provide reference for the information requirements needed for the development of the rebar prefabrication and offers a universal and standardized method for information exchange.

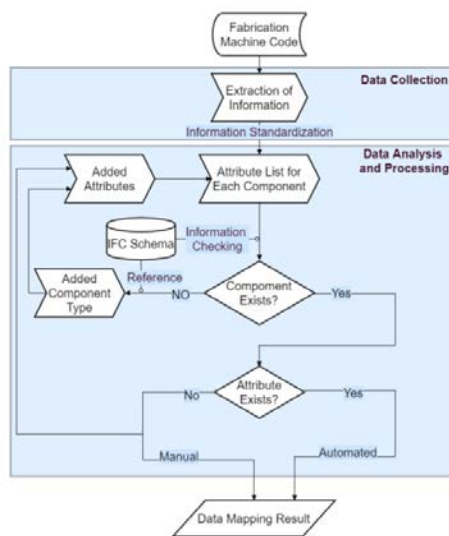


Figure 7. Framework for data mapping between IFC and BVBS

2.2.4 Automatic Information Transformation

Based on the extended IFC schema, automatic algorithms and customized programs can further be developed to extract necessary information including identity, geometrical, material and semantic information, to support the automatic generation of BVBS code, detailing drawing and bending schedule to enable automatic rebar prefabrication.

2.3 Rebar Installation Progress Monitoring

The onsite installation of rebars is a time-consuming and error-prone step of the RC structure construction, which includes rebar placing and tying. Therefore, it is necessary to develop a monitoring approach for project manager to grasp the installation progress of rebars timely. To be specific, the progress monitoring for rebar installation should provide the project manager with detailed information for the following aspects.

1. What is the progress of rebar installation for each RC component?
2. Are the installed rebars in the correct position?
3. Are the diameters of rebars installed correctly?

Since the rebars will be embedded in RC structures, the acquisition of rebar installation progress should be minutely integrated with the construction schedule. For example, the data of rebars of horizontal RC components (like slabs and beams) should be obtained before concrete pouring, while those of vertical components (like columns and walls) should be obtained before the installation of their formworks. This may involve the adjustment and integration of existing construction procedures, which means realizing progress monitoring of rebar installation not only relies on the development of technical solution, but also emphasize the improvement of project management level.

As the construction industry promotes prefabrication of RC components, the onsite rebar installation work will be reduced and gradually moved to factories. Progress monitoring and indoor prefabrication with computer numerical control (CNC) machines are a natural fit, since the indoor environment provides a stable environment for the layout and parameter setting of reality capturing devices. As the use of precast components has not been fully popularized yet, and partial cast-in-situ is necessary to improve the integrity of structures, the progress monitoring approach proposed in the study is mainly aimed at the onsite construction, and can be adapted to indoor scenario with necessary technical adjustments.

2.3.1 Onsite Reality Capturing

3D sensing is a type of techniques that has been widely used to capture and represent 3D information about the objects or environments. These techniques

include ultrasonic tomography (UT), photogrammetry, laser scanning, etc. Ultrasonic tomography is a non-destructive testing method, which allows for the inspection of materials and structures without causing any damage. However, when moving to onsite rebar installation, UT can be leveraged to inspect the rebars embedded in solidified concrete, but may not directly adopted to track the installation progress of rebars that has not been covered by concrete. Photogrammetry is a technique used to extract geometric information from photographs. It involves the process of capturing multiple images of an object or a scene from different angles and then using computational methods to analyse and reconstruct the 3D structure of the photographed subject. However, the accuracy of photogrammetry is limited and may not satisfy Aspect 2 and 3. Laser scanning turns out to be a better solution because of its high precision and versatility for various applications, especially with the rapid development of deep learning techniques based on point cloud [12]. Laser scanning techniques can be classified into two categories according to the movability of the scanning platform: terrestrial laser scanning (TLS) and mobile laser scanning (MLS). TLS involves stationary laser scanners that are placed on tripods or fixed mounts, and has limited coverage, which means for complex environments like construction sites, a large number of scanning stations are required to cover all the concerned areas. In addition, the setting up of scanners on construction site when installing rebars could be challenging. On the other hand, MLS relying on mobile platforms such as drones and cars, can be an alternative solution to rapidly capture onsite information. Current RC buildings are built storey by storey, which means that there are generally no buildings above when installing rebars, making the use of drones for laser scanning promising. For low-profile RC buildings, the delta style steel truss or gantry may also be used to carry the laser scanner. In recent years, handheld laser scanning (HLS) appears to be a more portable and flexible approach for rebar installation sites.

2.3.2 Vision-assisted BIM Reconstruction using 3D Point Cloud

To improve the efficiency and accuracy of BIM reconstruction based on point cloud, and make full use of the rich information provided by images, a vision-assisted BIM reconstruction method is proposed [13]. (1) RGB images of onsite rebar installation scenarios can be collected to train a deep learning model for semantic segmentation. Visual foundation models (VFM) can be leveraged to accelerate the development process [14]. (2) The segmented images and corresponding depth images can be used to generate a semantic-rich 3D map. To track rebar installation, the pixels referring to rebar should be identified. (3) The semantic-rich 3D map is then used to

guide the 3D point cloud segmentation. Specially, extracting the rebars from the 3D point cloud obtained by laser scanner. (4) Central lines of rebars are then extracted from the point cloud for BIM reconstruction and progress monitoring, while diameter or rebars are estimated and matched with market-supplied diameters. (5) Customized program based on Dynamo can be developed to establish rebars in Revit BIM model.

2.3.3 Progress Determination

The central lines and diameters of rebars in the as-design BIM model are extracted and compare to those extracted in Section 2.3.2 to track the progress of installation. The central lines and diameters from as-design and as-built model can be matched and verified with a given tolerance. Each rebar that has correct diameter and is located at the correct position will be counted as successful installation. Equation (3) shows the formulate the determine the rebar installation progress.

$$P = N_{ab}/N_{ad} \quad (3)$$

where P is the installation progress, N_{ab} refers to the number of rebars that have been successfully installed according to the as-built model, while N_{ad} refers to the number of rebars in the as-design model.

2.4 BIM-supported Maintenance, Renewal and Demolition

A BIM model with rebar information could unlock a range of applications throughout lifecycle of buildings, including maintenance, renewal and demolition.

Predictive maintenance: With the rebar information, when damage on waterproof layer or concrete cover layer is identified, rebar corrosion simulation can be conducted to analyze possible disasters, and help schedule maintenance activities proactively.

Optimized renewal and maintenance planning: Building maintenance and renewal, which can include modifications to the building's structure, such as adding new components, reinforcing existing components, or altering the layout, sometimes require excavating RC components. Knowing the specific positional and geometrical information of each rebar in the building allows architects and engineers to visualize the existing rebar layout and plan maintenances or renovations more efficiently, while minimizing the clashes with rebars and avoiding unnecessary rebar cutting.

Reliable structural assessment: A BIM with rebars could provide detailed information about the structural components of the building. This information is crucial for conducting structural assessments before renewal or demolition. Engineers can analyze the integrity of the existing structure and determine whether it can be safely renovated or if demolition is necessary.

Efficient demolition planning: The rebar BIM model

can provide insights into the structure's composition, allowing demolition teams to strategize the removal of materials in a safe and efficient manner. For building demolition with explosive, the BIM with rebars could provide accurate information for demolition simulation to optimize the layout of explosive.

Accurate construction waste management: A rebar BIM can not only assist in evaluating the construction waste and environmental impact of building renewal or demolition projects, but also help identify the quantities and positions of salvageable materials. The steel is usually the most valuable material for recycling, and a rebar BIM model could definitely support the estimation of steel bars that can be recycled. This can reduce waste and contribute to sustainable construction practices.

3 Conclusion

This study proposes an openBIM-based framework for lifecycle management of rebars. An automatic clash-free rebar design optimization approach is proposed integrating MAs and GNNs, with detailed optimal formulation. Besides, an integrated digital workflow is invented based on IDM, MVD, and IFC to support BIM-BVBS integration to support automatic rebar prefabrication. In addition, we propose to adopt laser scanning and vision-assisted BIM reconstruction for onsite rebar installation progress monitoring. The applications of the as-built rebar BIM model in building maintenance, renewal and demolition are discussed. Future directions of this work include: (1) validation of the proposed method on more real-life projects, and (2) software development for rebar lifecycle management.

References

- [1] Mangal M., Li M., Gan V., Cheng J. Automated clash-free optimization of steel reinforcement in RC frame structures using building information modeling and two-stage genetic algorithm. *Automation in Construction*, 126 (2021) 103676.
- [2] Eleftheriadis S., Duffour P., Stephenson B., Mumovic D., Automated specification of steel reinforcement to support the optimisation of RC floors. *Automation in Construction*, 96 (2018) 366-377.
- [3] Vaez S. and Qomi H. Bar layout and weight optimization of special RC shear wall. *Structures*, Vol. 14, Elsevier, 2018, pp. 153-163.
- [4] Lee C. and Ahn J. Flexural design of reinforced concrete frames by genetic algorithm. *Journal of structural engineering*, 129 (6) (2003) 762-774.
- [5] Lepš M. and Šejnoha M. New approach to optimization of reinforced concrete beams, *Computers Structures*, 81 (18-19) (2003) 1957-1966.
- [6] Liu, J., Liu, P., Feng, L., Wu, W., Li, D., Chen, Y. F. Automated clash resolution for reinforcement steel design in concrete frames via Q-learning and building information modeling. *Automation in Construction*, 112 (2020), 103062.
- [7] Turkan, Y., Bosche, F., Haas, C. T., Haas, R. Automated progress tracking using 4D schedule and 3D sensing technologies. *Automation in construction*, 22 (2012), 414-421.
- [8] Li, M., Wong, B., Liu, Y., Chan, C., Gan, V., Cheng, J. DfMA-oriented design optimization for steel reinforcement using BIM and hybrid metaheuristic algorithms. *Journal of Building Engineering*, 44 (2021), 103310.
- [9] Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., Philip, S. A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems*, 32(1) (2020), 4-24.
- [10] Li, M., Liu, Y., Wong, B., Gan, V., Cheng, J. Automated structural design optimization of steel reinforcement using graph neural network and exploratory genetic algorithms. *Automation in Construction*, 146 (2023), 104677.
- [11] Liu, Y., Li, M., Wong, B., Chan, C., Cheng, J., Gan, V. BIM-BVBS integration with openBIM standards for automatic prefabrication of steel reinforcement. *Automation in Construction*, 125 (2021), 103654.
- [12] Wang, B., Yin, C., Luo, H., Cheng, J., Wang, Q. Fully automated generation of parametric BIM for MEP scenes based on terrestrial laser scanning data. *Automation in Construction*, 125 (2021), 103615.
- [13] Wang, B., Wang, Q., Cheng, J. C., Song, C., Yin, C. Vision-assisted BIM reconstruction from 3D LiDAR point clouds for MEP scenes. *Automation in Construction*, 133 (2022), 103997.
- [14] Wu, C., Yin, S., Qi, W., Wang, X., Tang, Z., Duan, N. Visual chatgpt: Talking, drawing and editing with visual foundation models. *arXiv preprint arXiv:2303.04671* (2023).

A Digital Twin Based Approach to Control Overgrowth of Roadside Vegetation

Varun Kumar Reja¹, Diana Davletshina¹, Mengtian Yin¹, Ran Wei¹, Quentin Felix Adam¹, Ioannis Brilakis¹ and Federico Perrotta²

¹ Department of Engineering, University of Cambridge, UK

² AECOM, UK

varunreja7@gmail.com

Abstract -

Roadside vegetation poses a significant risk to road safety, often contributing to traffic accidents by obstructing drivers' views and impeding visibility. This study investigates existing control methods, assesses their limitations, defines technology-agnostic information requirements for vegetation control, and proposes a Digital Twin-based solution. The methodology involves expert interviews, a literature review, and a real-world case study, demonstrating the solution's applicability. The study contributes to the field by offering insights into current practices, defining information needs, and presenting a novel approach for more effective roadside vegetation management. This research contributes to advancing road safety practices by harnessing the capabilities of digital twin technology to proactively manage and mitigate the risks associated with overgrown roadside vegetation.

Keywords -

Roadside Vegetation; Road Safety; Digital Twin; Asset Management; Traffic accidents; Visibility obstruction; Non-Pavement Assets; Data-driven methodology; Vegetation Removal; Transportation Infrastructure

1 Introduction

This paper is about a digital twin-based approach to control the overgrowth of roadside vegetation. Digital twins are "an up-to-date digital representation of a system's physical and functional properties" [1]. Roadside vegetation overgrowth refers to the excessive and uncontrolled growth of plants, shrubs, or trees along roadsides. More than 400 casualties, including three deaths, have been reported in national figures provided by the Department for Transport, UK, where overgrown vegetation was given as a contributory factor [2]. It occurs when vegetation surpasses desirable or safe levels, impeding driver's visibility, encroaching onto road surfaces, obstructing traffic signs, and posing safety hazards [3]. Consequently, there is a higher risk of wrong turns or missed exits, increasing the potential for accidents.

The existing approach to managing vegetation overgrowth relies heavily on manual methods, where visual

inspections are the primary means of identifying and addressing overgrowth along roads [4]. This manual approach extends to the labour-intensive trimming process, and the information needed for effective vegetation control remains undefined. The reliance on manual methods for managing vegetation overgrowth along roads, as highlighted by visual inspections, presents challenges in efficiency and effectiveness. Firstly, the time required for visual inspections can lead to delays in identifying overgrowth issues promptly. This delay can result in increased vegetation encroachment, potentially escalating safety risks for road users. Secondly, the labour-intensive nature of the trimming process, stemming from the manual identification of overgrowth, can contribute to higher operational costs. KPIs related to cost-effectiveness and resource optimization are adversely affected as a consequence.

Additionally, the lack of well-defined information needs further hinders the development and implementation of targeted strategies for vegetation control, making it challenging to measure and improve KPIs related to the precision and efficacy of control measures. This situation underscores a critical demand for a paradigm shift in vegetation control methodologies. There is an urgent requirement to transition from predominantly manual methods to an automated and efficient system that addresses specific information needs.

The research design followed a structured methodology, first analysing current roadside vegetation control processes through expert interviews and a literature review. Identified drawbacks were scrutinized, and information requirements were established by analysing control parameters. Then, a Digital Twin-based solution was proposed to address these requirements and overcome identified drawbacks. A real-world case study demonstrated the solution's applicability, validating its effectiveness in roadside vegetation control. This approach ensures a thorough investigation integrating expert insights, critical analysis, and practical validation.

The expected contributions include an investigation into current control methods and their limitations, the defini-

tion of information requirements for vegetation control, and the illustration of how these requirements can be met through a Digital Twin technology-based solution. The study aims to provide valuable insights and a practical solution for more effective roadside vegetation management.

The paper is organized as follows: It commences with an introduction in Section 1, followed by background information in Section 2. The proposed concept and solution are presented in Section 3, and the research methodology is outlined in Section 4. Subsequently, Section 5 delves into the findings, emphasizes limitations, and proposes directions for future research. Ultimately, conclusions are drawn in Section 6.

2 Background

There are several advantages of roadside vegetation, which multiple studies have highlighted [5]. Key to these are improvements in air quality, temperature regulation, and noise reduction [6]. However, roadside vegetation along road networks can bring some drawbacks, including:

- **Reduced Visibility:** Overgrown vegetation can limit the visibility of road signs, signals, and other essential traffic control elements, increasing the risk of accidents to users and wildlife [7].
- **Infrastructure Damage:** Tree branches can damage road infrastructure, such as speed camera systems and road lights, increasing maintenance costs.
- **Obstructed Sightlines:** Tall or dense vegetation can obstruct sightlines at intersections, driveways, and pedestrian crossings, making it difficult for drivers to anticipate and react to potential hazards [8].
- **Encroachment on Road Space:** Uncontrolled growth of vegetation may encroach upon the road space, narrowing lanes and reducing the overall capacity of the road network.

Most of these pose risks to public safety and cannot be ignored; therefore, roadside vegetation maintenance must happen continuously to avoid such incidents.

The current body of literature concerning the automated detection and removal of overgrown vegetation is notably limited. Hyypä et al.'s study [9] showcases the potential of utilizing autonomous vehicle-based perception data to transform map updating processes, particularly in the context of city tree registers. The research findings highlight that, through appropriate post-processing, autonomous perception data can accurately estimate key parameters, such as the diameter at the breast height of city trees, showcasing lower errors compared to specially planned mobile mapping laser scanning surveys. This emphasizes the critical role of time-based filtering for precise point cloud analysis in autonomous perception data. Another relevant study by Harbas et al. [10] introduces a method for detecting roadside vegetation in traffic infrastructure maintenance

using visible spectrum images captured by standard colour cameras on vehicles. The study underscores the effectiveness of Fully Convolutional Neural Networks in real-world applications. However, it is noteworthy that existing studies have primarily concentrated on detecting the presence of roadside vegetation cover rather than specifically detecting their overgrowth, which poses a risk of obstruction.

The integration of digital twin and computer vision technologies has the potential to revolutionize conventional roadside vegetation control methods, as demonstrated in other road management applications [11, 12, 13]. By combining these technologies, a virtual counterpart capable of capturing real-time vegetation data and its impact on visibility is created, offering a more accurate representation of the roadside landscape. Computer vision algorithms integrated within the digital twin framework enable autonomous identification and precise assessment of overgrown vegetation, leveraging visual data from sources like surveillance cameras or drones [14, 15]. This integration represents a paradigm shift in roadside vegetation control, introducing a dynamic, data-driven, and precise approach compared to traditional methods.

Therefore, three primary gaps have been recognised, reflecting the analysis of the state of practice and research presented. Firstly, the existing documentation of practices for controlling overgrowth lacks specific and detailed information. Secondly, the technology-agnostic information requirements for vegetation control processes are not defined. Lastly, research on applying innovative technologies in this domain is scarce.

Recognizing this potential, the paper proposes a digital twin-centred strategy, aided by computer vision technique, to manage overgrown roadside vegetation and meet the identified information needs. The research aims to: (a) investigate existing methods for controlling roadside vegetation, identifying their limitations, and (b) assess the information requirements for vegetation control, illustrating their fulfilment through a solution based on Digital Twin technology.

To achieve the above objectives, the following specific research questions should be answered:

- RQ1: What is the current method (in practice) of controlling roadside overgrown vegetation?
- RQ2: How efficient and effective is this method in the current era of technology?
- RQ3: What are the technology-agnostic information requirements for vegetation control?
- RQ4: How can Digital Twin technology help effectively control vegetation overgrowth?

The subsequent section delves into the proposed concept and solution.

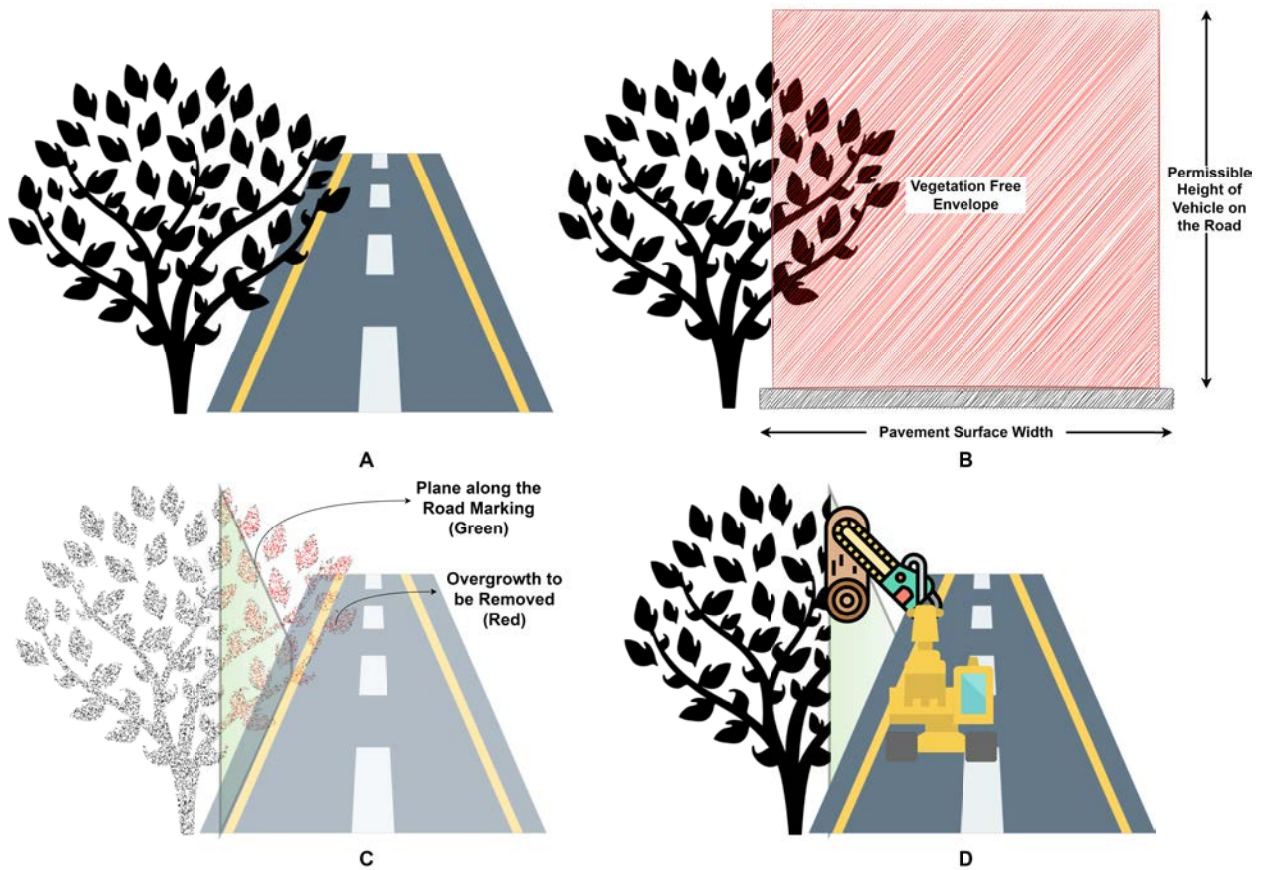


Figure 1. (A) Roadside overgrowth from driver's perspective, (B) vegetation-free envelope, (C) The segmented overgrowth in a digital twin environment as cut by the plane along road marking, (D) the vegetation clearing autonomous vehicle accomplishing the task

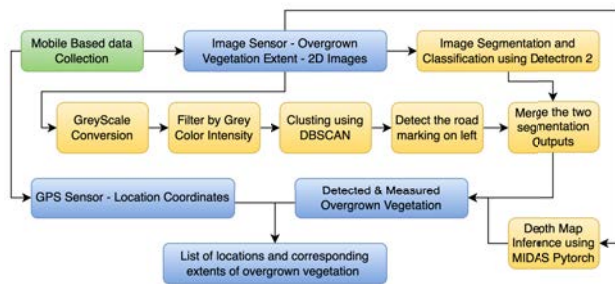


Figure 2. Proposed Solution for overgrowth detection

3 Proposed Concept and Solution

This work is confined to presenting and validating a proof of concept that addresses the identified information requirements for detecting vegetation overgrowth using digital twin technology. It is important to note that the actual removal of vegetation falls outside the scope of this

study. The technology-agnostic information requirements for vegetation control are identified in Table 1. The following concept is developed to fulfil these information requirements.

Figure 1 shows the proposed concept for controlling roadside vegetation. Image (A) shows the view from the driver's perception; the tree is overgrown and growing over the pavement. The image (B) presents a sectional view of the pavement. Ideally, there should be a vegetation-free envelope to allow for safe passage of traffic, as highlighted in red. The size of this envelope is determined by the permissible height of the vehicle on the road as well as the road width calculated from the exterior ends of road markings.

The first task is determining the instances along the road network where the vegetation overgrowth breaches this envelope. To obtain this, the idea is to detect roadside vegetation using a segmentation algorithm and then segment it using a plane along the roadside marking. Therefore, detecting roadside vegetation and roadside markings

Table 1. Technology agnostic information requirements for controlling roadside vegetation overgrowth

Sl. No.	Information Requirement	Description	Example
1	Location of Vegetation Overgrowth	Identification of overgrown areas and their specific locations.	GPS coordinates
2	Size/Dimensions of Overgrowth	Measurement of the size and dimensions of overgrown vegetation.	Square meters, Length x Width
3	Appearance of Overgrowth / Type of Vegetation	Description of the overgrown vegetation's appearance and type.	Tall grass, Shrubs, Trees

is important. Figure 1 (C) shows utilising a plane along the road markings in a digital twin visualisation to generate a vegetation-free envelope. The vegetation on the plane's (green-coloured) right is segmented as excess vegetation and measured in the next step.

Once measured, the information of this excess overgrowth, the image's GPS location, and the original image are stored on the cloud as a job order for the robots or human-operated machine. Machines will utilise this information to locate and cut the overgrown trees, as shown in Figure 1 (D).

Figure 2 presents the pipeline to demonstrate this concept. The image data is obtained and then classified into pavement and non-pavement assets. This is done by applying a computer vision-based algorithm which classifies the various assets. In this study, we have applied panoptic segmentation inference with Detectron 2 [16] and their FPN-101 (Panoptic Feature Pyramid Network) model [17], which was pre-trained on the COCO dataset [18]. Integrated seamlessly into our system architecture, the panoptic segmentation pipeline processes real-time or periodic image data captured by roadside cameras or sensors. Through this integration, the Detectron 2 model generates semantic segmentation masks and instance-level detections, facilitating a thorough analysis of roadside vegetation and other objects, thereby enhancing road safety assessments.

As this neural network segments the whole road, including the roadside markings/sidewalks, as one entity, we need to detect the road edge using the line markings. For that, we use a separate road-marking detection algorithm. First, we convert the original image frame to a greyscale image, partition the image around the detected road and filter this by grey colour intensity since road markings feature high brightness. Once that is done, we cluster them

using the DBSCAN algorithm [19] as it does not require specifying the number of clusters in advance and can deal with the noisy input. Then, we detect the largest left marking, corresponding to the left road edge, and merge this marking with the segmented image obtained from Detectron 2 in the previous step.

We then apply depth map inference on the cut input frame with a pre-trained model from MIDAS Pytorch [20]. This enables us to have the scene in 3D. Finally, with the road marking edge and vegetation depth detected, we obtain overgrown vegetation points above the road. As a result, the relative dimensions are measured to estimate the overgrown size.

Our hypothesis is that fulfilling this digital twin-based method's information requirements is much more efficient than the current method of practice regarding timeliness, safety and quality.

4 Research Methodology

Figure 3 presents the research methodology. First, the existing process of controlling roadside vegetation was outlined by interviewing five domain experts (each with at least ten years of on-field experience) from different highway agencies and reviewing the existing literature on vegetation control. Next, this process was analysed, and the drawbacks were outlined. The information requirements for vegetation control were also identified by critically identifying and studying the control parameters required for the survey and removal tasks. A DT-based solution presented in the previous section is proposed to target these information requirements and address the drawbacks. The applicability of this solution was demonstrated using a real-world case.

We collected 200m road length data to demonstrate the



Figure 3. Research Methodology

proposed approach using an iPhone 12 Pro Max mobile camera sensor as a .MOV video file. We later processed it by extracting image frames from the video. Our method assumes that the road marking is alongside the road to construct the vegetation-free envelope. Next, all the steps presented in Figure 2 were applied to this data. The final relative encroachment width for overgrown vegetation was estimated in percentage to the road width.

5 Results and Discussion

5.1 Current Process and Drawbacks

The current process of controlling roadside vegetation is shown in Figure 4. The process starts with collecting data about the roadside vegetation status, i.e., locations at which the vegetation has overgrown along the road, the size of the overgrowth, and the type of vegetation. This is frequently done by routine surveys by pavement inspectors alongside a particular stretch of road through the visual inspection. Sometimes, the information is also conveyed by the public (road users) who might encounter such instances. The access growth is also realised by analysing recent accidents or near-miss reports.

This information is passed to the road asset owners with an approximate location (generally referred to as near a landmark) of the detected overgrown and sometimes supported with mobile camera images. Next, the road management authority takes the necessary action by sending workers to the locations to remove the overgrown vegetation and clear the path to traffic. Finally, the confirmation of the task completion is provided by the pavement inspector, who goes to the location and verifies.

There are various drawbacks to the current roadside vegetation control process. As observed, these are highlighted below:

- **Subjectivity and Reliance on visual inspection:** The current process heavily depends on routine surveys conducted by pavement inspectors using visual inspections, introducing a subjective element that may lead to inconsistencies and inaccuracies in vegetation status assessment.

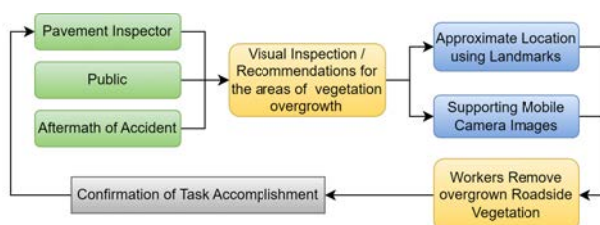


Figure 4. Current Process of Controlling Roadside Vegetation

- **Delayed Detection:** The current process may lead to delayed detection of overgrown vegetation, as it relies on periodic inspections or incident reports. This delay can contribute to heightened safety risks and a longer response time to mitigate potential hazards.
- **Ambiguous Location Information:** Providing approximate location details, often described as near a landmark, may introduce ambiguity and inefficiencies in pinpointing the exact locations of overgrown vegetation. This ambiguity can hinder the timely and precise execution of removal tasks. Also, using robots/automated machinery in the removal phase requires exact coordinates of the overgrowth, which is impossible with the current method.
- **Resource-Intensive Verification Process:** Confirming task completion involves a pavement inspector physically visiting the location, representing a resource-intensive verification step. This manual confirmation may lead to delays and increased operational costs.

5.2 Results using the Proposed Process in Real-world

Figure 5 (A) shows the original image, and (B) shows the segmented and classified image with Detectron 2. It can be seen that the algorithm detects roads, trees, grass and sky in the image frame. However, we can see that the road markings are explicitly not detected. For the road marking detection algorithm, the stepwise results are shown in Figure 5, where (C) is a greyscale image, (D) shows a cut-out for the road portion, (E) shows results of filtering grey intensity, (F) shows clusters formed by DBSCAN and (G) shows detection of the largest left marking. The results after merging the segmentation obtained by detection 2 in Figure 5 (B) and road marking obtained in Figure 5 (G) are shown in Figure 5 (H). Next, Figure 5 (I) shows the depth map obtained from the MIDAS Pytorch algorithm. The detection of overgrown vegetation inside the vegetation-free envelope is shown in (J), highlighted with navy blue. Finally, we provide the approximate distance the tree takes over the ground in pixels and the percentage of the road width at approximately the same depth as shown in Figure 6b. It can be seen that the extent of the detected overgrowth is about 44% of the road width. With the location obtained from the mobile GPS sensor, the extent of this overgrowth and the original frame, this information can be shared with any automated or human-operated system to remove the overgrowth.

The concept of a vegetation-free envelope emerges as a promising strategy for ensuring the safe passage of traffic. By defining this envelope based on the permissible height of vehicles and road width, the proposed solution aims to mitigate potential hazards caused by overgrown roadside vegetation. However, further analysis and empirical studies are required to assess this approach's real-world

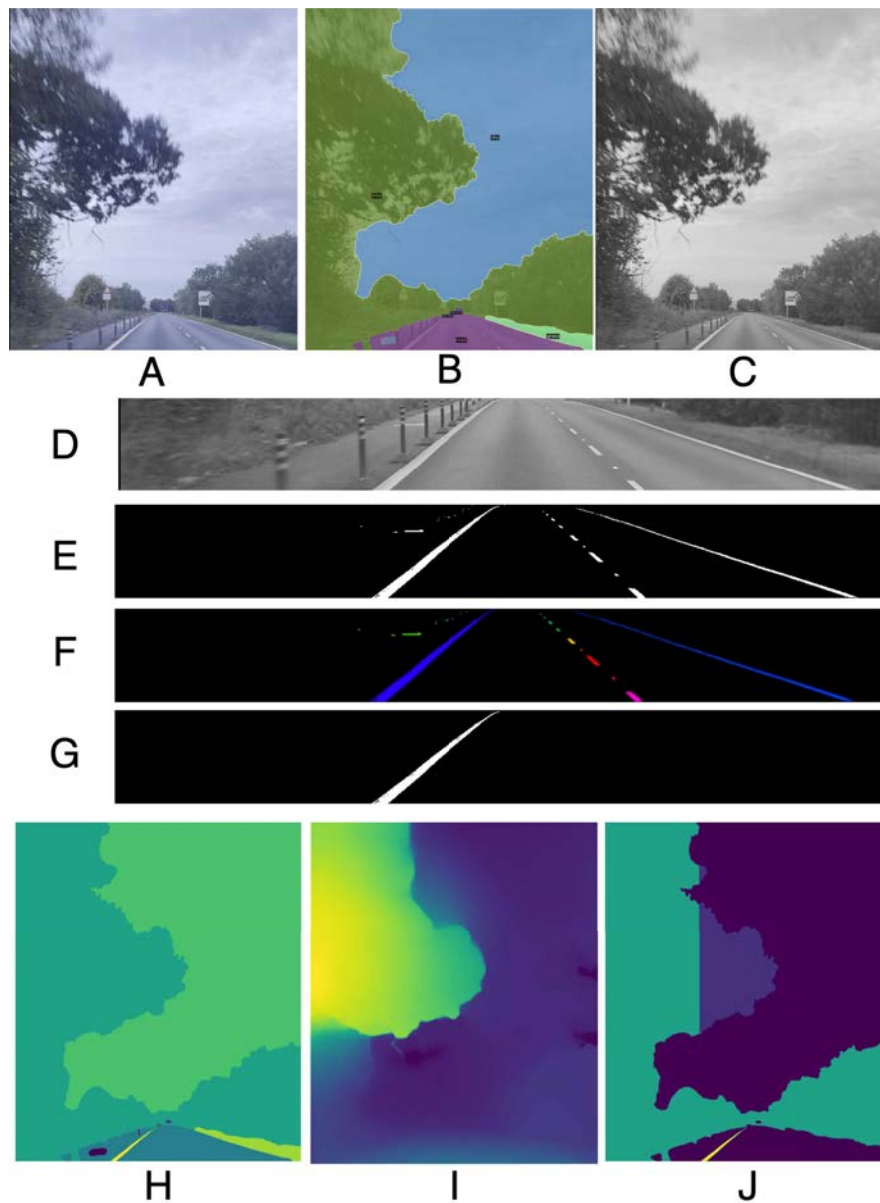


Figure 5. Stepwise Results for the Vegetation Overgrowth Detection

effectiveness and identify any possible limitations or scenarios where it may fall short.

The choice between data from public vehicle sensors and specialised survey vehicles for updating the road digital twin raises accuracy, frequency, and cost considerations. Utilising data from public vehicles offers a large-scale, crowd-sourced approach, but questions about data reliability and coverage arise. On the other hand, specialised survey vehicles provide targeted and potentially more accurate data but at an increased cost. Striking the right balance between these approaches is crucial for maintaining an up-to-date and reliable road digital twin.

The study made a dual contribution. Firstly, it delved

into the investigation and documentation of the existing control methods for roadside vegetation overgrowth (RQ1) while also analysing their performance (RQ2). Secondly, it defined the information requirements for vegetation control (RQ3) and showcased how these requirements can be met through the proposal of a Digital Twin-based solution (RQ4).

It is essential to acknowledge the limitations of the current study. Relying on road markings for envelope definition may pose challenges on roads without clear markings. In this study, we assume that the road markings are present to construct the limits of a vegetation-free envelope, which may not be the case for all types of roads. In

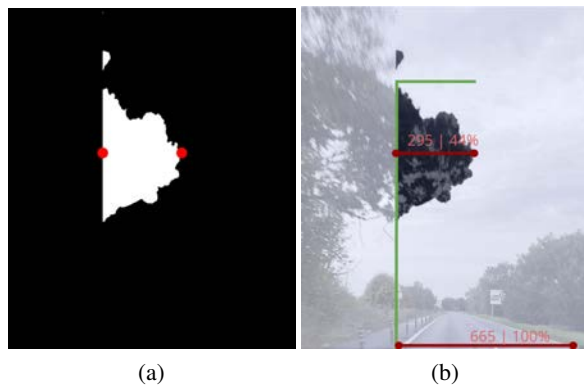


Figure 6. Detected overgrown vegetation, vegetation edge points in red and the vegetation-free envelope borders in green.

some cases, the road markings may be damaged or eroded; in that case, we will detect the pavement edge and use it to construct the vegetation-free envelope. In future, we also aim to monitor the condition of the other road assets like lamps, road signs, road markings, etc and integrate the same method to develop a combined comprehensive digital twin framework for road maintenance. In regions with different climatic conditions or vegetation profiles, adjustments to the detection algorithms and parameter tuning may be necessary to achieve optimal performance.

Compared to previously available literature [10, 9], our study specifically proposes a solution for controlling overgrown roadside vegetation rather than vegetation in general. We can also combine [10] generic approach with our method to further improve the performance of our method. Additionally, our method helps evolve conventional documentation and practices by comparing the effectiveness of these traditional methods with ours.

6 Conclusions

The practical implications of this study are substantial, as the examination and analysis of current roadside vegetation control approaches and identifying their limitations provide valuable insights for practitioners and policymakers. Understanding the inherent information requirements in vegetation control enhances the decision-making process for those involved in deciding the technology to be deployed for survey and removal. Introducing a novel Digital Twin technology-based solution addresses the limitations of traditional methods and signifies a practical advancement. Incorporating our approach into current road management practices entails adopting digital twin technology to address maintenance requirements, facilitating real-time monitoring of roadside vegetation, and facilitating data-driven decision-making. Over the long term, this integration will enhance infrastructure resilience, improve user responsiveness, reduce accidents, lower maintenance

expenses, and enhance overall road safety.

The societal implications of this study are twofold, primarily centring on the direct enhancement of road user safety. Implementing the study's findings can optimize vegetation control processes, improve visibility on roads, and diminish the likelihood of accidents arising from overgrown vegetation. Additionally, the study underscores the importance of improved infrastructure management. This contributes to transportation infrastructure's overall efficiency and resilience, fostering smoother mobility and minimizing societal disruptions. The advancement in infrastructure management elevates safety standards and fosters the development of a more sustainable and dependable transportation network.

In the future, the data collection can be fully automatic by collaborating with connected autonomous vehicle providers and will require no additional effort for DT updation. Also, a potential future research direction is predicting the overgrowth based on the growth rate of the type of vegetation detected. This can help recognise the frequently vulnerable areas to take proactive action. Also, in some cases where autonomous vegetation-trimming robots like [21] can be deployed, it can make the process completely automated.

In summary, the presented digital twin-enabled methodology offers a promising avenue for transportation authorities and infrastructure managers to address the challenges posed by overgrown roadside vegetation. By leveraging cutting-edge technology, this approach enhances safety and contributes to the evolution of smart and resilient transportation systems. As technology advances, the integration of digital twins in road infrastructure management is poised to play a pivotal role in shaping the future of transportation safety and efficiency.

7 Acknowledgement



Funded by
the European Union

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101034337. The authors would like to acknowledge the support of the academic and industry partners: Kristen MacAskill (University of Cambridge), Mike Hunter (Georgia Tech), Jeremy Morley (Ordinance Survey), Stefano Cavarazzi (Ordinance Survey), Yogesh Patel (Ringway), Matt Peck (Atkins), Nick Wang (SAP), Arnulf Hagen (SAP), Simon Hayton (Keltbary), Michael Pelken (Keltbary), Tom Tideswell (Kier), Jordan Flint (Kier), Jonathon Simons (AECOM), and James Locke (Arcadis). Their expertise and guidance have significantly enriched the content of this paper.

References

- [1] Rafael Sacks, Ioannis Brilakis, Ergo Pikas, Haiyan Sally Xie, and Mark Girolami. Construction with digital twin information systems. *Data-Centric Engineering*, 1:e14, 2020.
- [2] Amy Downward. Overgrown hedges ‘a danger to drivers’. On-line <https://www.shropshirestar.com/news>, Accessed: 20/12/2023.
- [3] Overgrown trees, branches and plants are a hazard for drivers. On-line: <https://www.edriving.com/three60/>, Accessed: 20/11/2023.
- [4] Ann M Johnson et al. Best practices handbook for roadside vegetation management. Technical report, Minnesota. Dept. of Transportation. Office of Research Services, 2008.
- [5] Richard TT Forman and Robert I McDonald. A massive increase in roadside woody vegetation: goals, pros, and cons. 2007.
- [6] Kristina Hill, Richard Ray Horner, Ray Willard, and Washington Olympia. Assessment of alternatives in roadside vegetation management. 2005.
- [7] Jeffery W Van Treese II, Andrew K Koeser, George E Fitzpatrick, Michael T Olexa, and Ethan J Allen. A review of the impact of roadway vegetation on drivers’ health and well-being and the risks associated with single-vehicle crashes. *Arboricultural Journal*, 39(3):179–193, 2017.
- [8] M Bassani, L Catani, A Salussolia, and CYD Yang. A driving simulation study to examine the impact of available sight distance on driver behavior along rural highways. *Accident Analysis & Prevention*, 131:200–212, 2019.
- [9] Eric Hyypä, Petri Manninen, Jyri Maanpää, Josef Taher, Paula Litkey, Heikki Hyyti, Antero Kukko, Harri Kaartinen, Eero Ahokas, Xiaowei Yu, Jesse Muhojoki, Matti Lehtomäki, Juho-Pekka Virtanen, and Juha Hyypä. Can the perception data of autonomous vehicles be used to replace mobile mapping surveys?-a case study surveying roadside city trees. *Remote Sensing*, 15(7), 2023. ISSN 2072-4292. doi:10.3390/rs15071790. URL <https://www.mdpi.com/2072-4292/15/7/1790>.
- [10] Iva Harbaš, Pavle Prentašić, and Marko Subašić. Detection of roadside vegetation using fully convolutional networks. *Image and Vision Computing*, 74: 1–9, 2018.
- [11] Ou Zheng. Development, validation, and integration of ai-driven computer vision and digital-twin systems for traffic safety diagnostics.
- [12] Yongkang Liu, Ziran Wang, Kyungtae Han, Zhenyu Shou, Prashant Tiwari, and John HL Hansen. Sensor fusion of camera and cloud digital twin information for intelligent vehicles. In *2020 IEEE Intelligent Vehicles Symposium (IV)*, pages 182–187. IEEE, 2020.
- [13] Diana Davletshina, Varun Kumar Reja, and Ioannis Brilakis. Automating construction of road geometric digital twins using context and location aware segmentation. 2024. URL <https://ssrn.com/abstract=4767693>.
- [14] Varun Kumar Reja, Koshy Varghese, and Quang Phuc Ha. Computer vision-based construction progress monitoring. *Automation in Construction*, 138:104245, 2022. doi:10.1016/j.autcon.2022.104245. URL <https://doi.org/10.1016/j.autcon.2022.104245>.
- [15] Varun Kumar Reja, Shreya Goyal, Koshy Varghese, Balaraman Ravindran, and Quang Phuc Ha. Hybrid self-supervised learning-based architecture for construction progress monitoring. *Automation in Construction*, 158:105225, 2024. doi:10.1016/j.autcon.2023.105225. URL <https://www.sciencedirect.com/science/article/pii/S0926580523004855?via%3Dihub>.
- [16] Yuxin Wu, Alexander Kirillov, Francisco Massa, Wan-Yen Lo, and Ross Girshick. Detectron2. <https://github.com/facebookresearch/detectron2>, 2019.
- [17] Detectron2 model zoo and baselines. On-line: https://github.com/facebookresearch/detectron2/blob/main/MODEL_ZOO.md, Accessed: 20/11/2023.
- [18] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part V 13*, pages 740–755. Springer, 2014.
- [19] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, volume 96, pages 226–231, 1996.
- [20] René Ranftl, Katrin Lasinger, David Hafner, Konrad Schindler, and Vladlen Koltun. Towards robust monocular depth estimation: Mixing datasets for zero-shot cross-dataset transfer. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2020.
- [21] A new path to safely managing vegetation. On-line: <https://www.sarcos.com/robotics-applications-and-use-cases/vegetation-management/>, Accessed: 20/11/2023.

Bridging the Annotation Gap: Innovating Sewer Defects Detection with Weakly Supervised Object Localization

Jianyu Yin¹, Xianfei Yin², Yifeng Sun³, and Mi Pan⁴

^{1,3,4}Department of Civil and Environmental Engineering, University of Macau, Macau, China

²Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong, China
dc12714@um.edu.mo, xianfei@ualberta.ca, mc35239@um.edu.mo, mipan@um.edu.mo

Abstract

Urban sewer systems are vital yet often neglected components of modern infrastructure system. Inspecting these systems is expensive due to labour costs and the need for manual examination by professionals. In addition to the challenges posed by traditional methods, developing deep-learning-based automatic defect detection models requires a vast number of bounding box labels, which are challenging to acquire. To address these gaps, our study introduced the application of Weakly Supervised Object Localization (WSOL) for automated defect localization in sewer pipes. WSOL is a technique that allows for the localization of objects within images using only image-level labels, without the need for precise bounding box annotations. We adopted a state-of-the-art WSOL method that mitigates feature directions with class-specific weights misalignment, enabling more accurate and complete localization of defects. By generating heatmaps from Sewer-ML's image-level annotations, bounding box labels are eliminated, rendering our approach scalable and cost-effective. The proposed WSOL-based approach was validated through five distinct classes of defects and one construction feature, demonstrating the promising localization performance. Our method achieved mean MaxBoxAccV2 scores of 64.33% and 56.89% when using ResNet-50 and VGG-16 backbones, respectively, while also attained classification accuracies of 87.00% for ResNet50 backbone and 83.00% for VGG16 backbone. As a pioneering contribution, our work established a new standard for automated sewer system maintenance, offered a benchmark for the application of WSOL methods using solely image-level annotations in defect localization for urban sewer systems, and further expanded the frontier of weakly supervised learning in critical infrastructural applications.

Keywords –

Weakly Supervised Object Localization (WSOL); Sewer Pipelines; Defect Detection; Image-Level Annotation

1 Introduction

Urban sewer system are integral to modern civilization across various urban settings, playing a pivotal role in health, sanitation, and overall well-being. Yet, the inspection and upkeep of these intricate underground networks often go overlooked. Traditional inspection methods are labour-intensive, costly, and hinge on manual evaluations by specialists. While these methods are effective, they aren't scalable for expansive urban sewers. The push towards deep-learning for defect detection introduces another challenge: the need for numerous bounding box labels. Acquiring such labels is not only tedious but also consumes significant resources. Therefore, an effective automatic localization system is urgently needed to overcome these shortcomings.

The core research question this study seeks to address is: "How can feature-level defect detection be realized with image-level annotations in the context of urban sewer system inspection?" The advent of Weakly Supervised Object Localization (WSOL) offers a promising alternative, paving the way for more automated, efficient, and cost-effective solutions as it relies solely on image-level annotations. Yet, while WSOL has seen success in other domains [1-5], its application in the realm of sewer system inspection is still quite new. To the best of our knowledge, this is the first work to apply weakly supervised learning for automatic defect detection in the sewer pipelines with evaluation of localization precision. It's essential to note that existing WSOL methods, which typically utilize Class Activation Maps (CAMs), have their limitations in accurately localizing defects, as they tend to focus on only the most prominent features [1]. This poses a concern in the context of sewer systems where the full shape and scope of defects are crucial for remedial actions. In this study, we adopted a state-of-the-art WSOL method proposed by Kim et al. [2] that mitigates feature directions with class-specific weights misalignment. Our hypothesis is that the proposed WSOL method can achieve reasonable localization accuracy for sewer defects using only image-level labels, ensuring a more holistic and accurate localization of sewer defects.

The main contributions of this study are listed below:

- 1) Applied weakly supervised learning for automated defect localization in urban sewer systems with evaluation of localization precision;
- 2) Provided a scalable solution that obviates the need for tedious and resource-intensive bounding box labels with reasonable localization accuracy;
- 3) Developed A benchmark and methodological framework for future research into weakly supervised defect localization in sewer pipelines.

The rest of the paper is structured as follows: Section 2 delves into the recent advances and the unique challenges of both sewer inspection techniques and defect detection. In Section 3, we detail our methodology, including the dataset used and our experiment set up. Section 4 offers a comprehensive evaluation, presenting our experimental results. Finally, Section 5 wraps up with the study's limitations, prospects for future research, and a concluding remark.

2 Related Work

In recent studies on sewer inspections, deep learning methods have been increasingly utilized to improve performance of defect detection [6]. For example, Cheng and Wang [7] contributed an automated defect detection approach using faster R-CNN, demonstrating how deep CNNs can identify and locate common defects in CCTV images. Li et al. [8] introduced innovations like strengthened region proposals and global context fusion to enable fine-grained defect severity grading, surpassing other methods in this capability. Yin et al. [9] applied state-of-the-art YOLOv3 for real-time automated detection in videos, showing advantages in processing speed and accuracy over manual review. And Yin et al. [10] further developed a CCTV video interpretation algorithm and sewer pipe video assessment (SPVA) system based on their previous developed deep learning-based framework.

However, the major defect inspection models rely on effective supervised learning methods that cost much time in the manual annotation process for training [11]. Therefore, researchers have begun to explore alternative methods to reduce the need for manual annotation in defect inspection models. For example, Zhang et al. [12] used a GAN framework to leverage both labelled and unlabelled images, achieving 79-81% mean IOU on a steel defect dataset with only 1/8 to 1/4 full supervision. Zhang et al. [4] used weak image-level tags rather than detailed pixel-level annotations, extracting spatial information from tags through category-aware convolutions and pooling. In another study, Wu et al. [5] relied solely on image-level labels for training, improving CAM techniques to achieve performance

surpassing some fully supervised methods. Manual labelling of feature-level data is both time-consuming and prone to human error. Thereby, our study pioneered the application of WSOL for accurately localizing sewer defects with a methodological framework used for fair comparative evaluation. Sewer pipelines inspection presents unique challenges such as varying lighting conditions, occlusions, and highly irregular object shapes. The adopted method [2] seeks to overcome the limitations of current WSOL methods by mitigating feature directions with class-specific weights misalignment, thereby ensuring more accurate defect detection in sewer pipelines. Furthermore, the completely automated systems that include automatic labelling tools need to be developed for more efficient sewer inspections [6]. The automatic generated bounding boxes from heatmaps can be served as inputs for training a supervised detection model to reduce the redundant annotation process.

3 Methodology

3.1 Dataset

The dataset utilized in this research is Sewer-ML [13], which is the first publicly available dataset dedicated to sewer defect classification. Comprising 1.3 million images sourced from 75,618 videos, the data was aggregated from three Danish water utility companies over a nine-year period. The image-level annotations were performed by licensed sewer inspectors in compliance with the Danish sewer inspection standard, Fotomanualen [14]. This contributes to the high reliability and consistency of the annotations.

For the task of WSOL, it is essential to have images with only one class of object for generating accurate heatmaps. As such, we curated a subset of Sewer-ML specifically designed for WSOL tasks. The customized dataset information is provided in table 1. The images number are the same per class for training, validation and test sets. Our training set consists of 1,000 images per category, focusing on critical defect classes such as settled deposits (AF), roots (RO), attached deposits (BE), displaced joints (FS), and notably, cracks, breaks, and collapses (RB). Among these, RB stands out as the most paramount defect due to its severe implications [15]. These defects were chosen based on both their significant impact on infrastructure integrity and their frequent occurrence. Additionally, while drilled connections (PB) is not a defect but a construction feature, we have incorporated it into our dataset to rigorously test the proposed model's discernment capabilities, given PB's unique characteristics.

For validation, we adopted a set of 10 images per defect class, each with bounding box annotations to assist

in hyperparameter tuning and model selection, but not used for training. This choice of a fixed-size validation set is influenced by Choe et al. [16], advocating that small fixed number of fully annotated images offers a robust yet adaptable baseline for comparative evaluation, thereby ensuring methodological consistency for future research, given that certain level of localization labels are inevitable for WSOL.

Our test set encompassed 50 images per selected class, designed to rigorously evaluate the model's generalization capabilities. By customizing the Sewer-ML dataset in this fashion, we provided a methodological framework that not only catered to the specialized requirements of WSOL but also set a precedent for future research in sewer defect localization.

Table 1. Dataset Information

Dataset	Sewer-ML (Haurum and Moeslund, 2021)
Classes	6
Train	1000
Validation	10
Test	50

*Image numbers are shown per class

3.2 Method

The WSOL method proposed by Kim et al. [2] bridges the gap between image classification and object localization by aligning feature directions to class-specific weights.

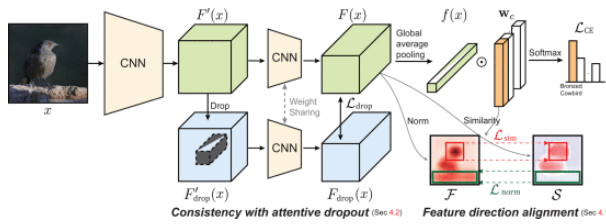


Figure 1. WSOL structure developed by (Kim et al. 2022)

Figure 1 shows the overall of their proposed method. Traditional WSOL methods using CAMs only highlight the most discriminative parts of an object [1]. In contrast, the approach by Kim et al. [2] enhances the alignment of all feature directions to the target class weights through two main strategies: 1) feature direction alignment loss and 2) consistency regularization with attentive dropout. This enables the activation of less discriminative object regions in the CAM, allowing for more accurate and complete localization. The alignment and consistency losses are incorporated into the training of a CNN-based classifier (ResNet-50 or VGG-16) on the Sewer-ML

images. This guides the model to align all feature directions to the weights of each defect class.

At inference time, CAMs are generated for each input image to produce a heatmap highlighting potential defect regions. The continuous CAM is then thresholded to obtain a bounding box localization for each predicted defect class. This allows for weakly supervised defect localization using only image-level labels, eliminating the need for manually annotated bounding boxes.

3.3 Experiment Set Up and Training Strategy

We conducted experiments using NVIDIA GeForce RTX 4080 GPU. The WSOL model was implemented in PyTorch, initialized with weights pretrained on CUB-200-2011 [17] for ResNet50 and on ImageNet-1K [18] for VGG16. We evaluated the approach using both VGG-16 and ResNet-50 backbones.

During training, we used a warm-up stage where only the classification and consistency losses are active for the first few epochs. After warm-up, the full training loss was used including the feature direction alignment losses.

We selected the best checkpoint based on localization accuracy on the validation set. The model was then evaluated on the test set to report performance for the six classes.

4 Results and Discussion

The performance of two deep learning models, ResNet50 and VGG16, was evaluated using two primary metrics: Accuracy and Ground Truth Location (GT-LOC). Accuracy measures the percentage of images for which the predicted label matches the true label. GT-LOC, assesses the accuracy of predicting the object's exact location within an image, reflecting the model's ability in object localization.

Table 2. Accuracy and GT-LOC on Validation Set

Backbone	Accuracy	GT-LOC
ResNet50	83.33%	65.00%
VGG16	81.67%	70.00%

Table 3. Accuracy and GT-LOC on Test Set

Backbone	Accuracy	GT-LOC
ResNet50	87.00%	68.33%
VGG16	83.00%	59.67%

Table 2 shows the accuracy and localization performance for both backbones on validation set.

- ResNet50: Exhibits a strong classification accuracy of 83.33%. For object localization, as indicated by the GT-LOC, it achieves 65.00%.
- VGG16: It demonstrates a solid accuracy of 81.67%.

The GT-LOC score of 70.00% shows its capacity in precisely locating objects within images.

Table 3 shows the accuracy and localization performance for both backbones on test set.

- ResNet50: The model has an accuracy of 87.00%, highlighting its consistent performance. The GT-LOC score of 68.33% further emphasizes its reliable object localization capability.
- VGG16: Achieving an accuracy of 83.00%, VGG16 maintains its robustness in object recognition. However, there's a notable drop in GT-LOC to 59.67%, suggesting some variability in its localization performance on the test set.

Additionally, we compared our MaxBoxAccV2 scores [16]. MaxBoxAccV2 measures the models' localization accuracy with ground-truth class based on multiple IOU thresholds. Table 4 and Table 5 indicate that both ResNet50 and VGG16 backbones showed strong results at the loc IOU 30 and 50 thresholds, with scores ranging from 59.67% to 93.00%. Specifically, both ResNet50 and VGG16 have good performances at the loc IOU 30 threshold, achieving as high as 93.00%. However, as the rigor increased to the loc IOU 70 threshold, both models faced challenges, with accuracies dropping to between 18.00% and 31.67%, highlighting the need for potential refinements at higher precision levels. The average localization accuracy across all thresholds placed VGG16 slightly ahead of ResNet50 on the validation set, with scores of 63.89% and 61.67% respectively, however, ResNet50 regained advantages on the test set with a score of 64.33% compared to 56.89% for VGG16. This indicates that while both models are proficient at less stringent thresholds, ResNet50 exhibits a more consistent performance across the board. Our data analysis confirmed that the proposed WSOL method achieved significant localization accuracy thereby validating our hypothesis and addressing the core research question.

Table 4. MaxBoxAccV2 Scores for ResNet50 and VGG16 on Validation Set

Backbone	ResNet50	VGG16
Loc_IOU_30	90.00%	90.00%
Loc_IOU_50	65.00%	70.00%
Loc_IOU_70	30.00%	31.67%
Mean	61.67%	63.89%

Table 5. MaxBoxAccV2 Scores for ResNet50 and VGG16 on Test Set

Backbone	ResNet50	VGG16
Loc_IOU_30	93.00%	93.00%
Loc_IOU_50	68.33%	59.67%
Loc_IOU_70	31.67%	18.00%

Mean 64.33% 56.89%

A confusion matrix visually represents the performance of a classification model by contrasting actual versus predicted classifications. The diagonal elements represent the correct classifications, while off-diagonal elements indicate misclassifications. A well-performing model would have higher numbers on the diagonal and lower numbers off-diagonal.

Figure 2 and Figure 3 show the confusion matrices for both ResNet50 and VGG16 backbones on the test set: ResNet50 showed strong diagonal values for categories BE, PB, and RO, while VGG16 exhibited good results for classes AF, PB and RO. For instance, PB consistently received a high true positive score of 50 for both backbones, indicating that both models have a firm grasp on identifying this class. Similarly, AF, BE, and RO mostly contained high values on the diagonal and minimal off-diagonal interference. However, some classes show room for improvement. For instance, FS and RB had a few off-diagonal values, indicating some misclassifications. RB, in particular, had some misclassifications where it was mistaken for other classes. Figure 4 provides some CAM examples for both backbones for each class.

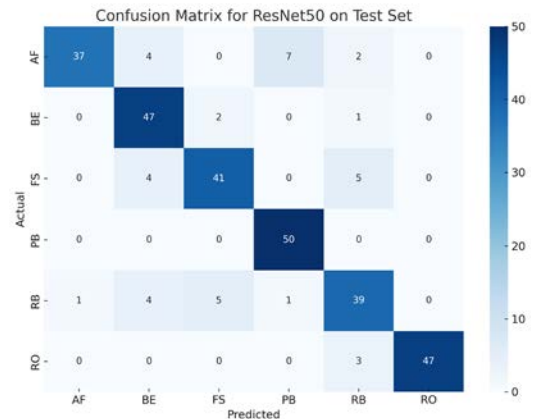


Figure 2. Confusion Matrix for ResNet50 on Test Set

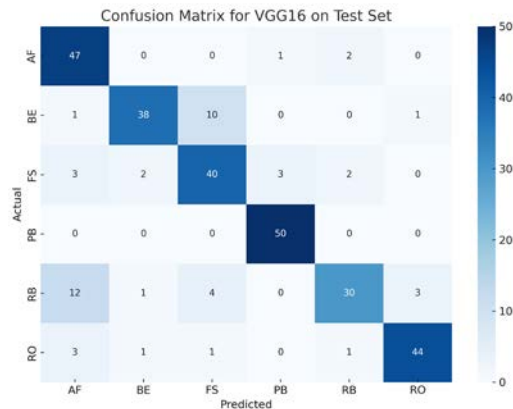


Figure 3. Confusion Matrix for VGG16 on Test Set

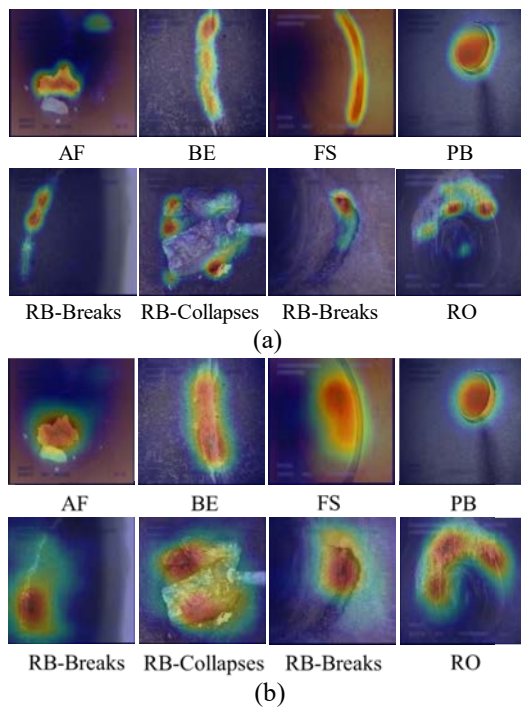


Figure 4. CAM examples for the two backbones for each class: (a) CAM examples for ResNet50; (b) CAM examples for VGG16.

Different backbones are good at localizing different classes. In our study, we found that ResNet50 is relatively better at capturing classes are smaller in scale, like cracks, breaks, and displaced joint, while VGG16 can give a more holistic localization to roots and collapses. Both backbones are good at localizing attached deposits and drilled connection feature, while both fail for settled deposits. A model based on a more recent backbone or even the combination of backbones for different classes should be developed in the future.

Our findings indicated that the VGG16 model tends to produce larger heatmaps for defect localization compared to the ResNet50 model. This difference in heatmap size can have implications on asset management decisions. Larger heatmaps may overestimate the scope of defects, resulting in unnecessary repairs or replacements. Meanwhile, smaller heatmaps could underestimate damage, failing to fully resolve infrastructure issues.

Visualizing CAM to see how and why the method is localizing the defects can be valuable, especially in critical infrastructure inspection like sewer systems where human experts need to trust and understand the model's decision-making.

5 Conclusion and Limitation

This study demonstrated the promising potential of WSOL techniques for automated defect detection in sewer pipelines. The analysis of our results revealed clear trends and relationships that affirm the effectiveness of the proposed WSOL method for sewer pipeline defect detection using image-level annotations. By generating localization heatmaps directly from image-level labels, our approach eliminated the need for tedious bounding box annotations. Across five common sewer defect classes and one construction feature, both ResNet50 and VGG16 achieved reasonable localization performance without any bounding boxes for training.

Specifically, our models attained average MaxBoxAccV2 scores of 64.33% (ResNet50) and 56.89% (VGG16) on the test set. The ResNet50 backbone demonstrated slightly more consistent localization across different IOU thresholds. Additionally, both models showed proficiency in classifying most defect types, with test accuracies reaching as high as 87.00% for ResNet50 backbone and 83.00% for VGG16 backbone.

While these initial results are encouraging, there remain limitations to be addressed. First, the localization accuracy is still far from perfect, with ample room for improvement. Second, certain defect classes like settled deposits proved more challenging for the models to precisely localize. Third, visual analysis revealed that the two backbones had relative strengths and weaknesses in localizing different defect types. An ensemble or multi-backbone approach may help mitigate these class-specific shortcomings. Additionally, transitioning to Weakly Supervised Object Detection (WSOD) approaches would be a valuable future work. WSOD removes the limitation of one object per image class, enabling multi-object defect detection learning directly from image labels. This could better handle real-world sewer images with multiple defects present. Adopting a WSOD approach would require more computational resources but may further advance automated sewer

analysis. Another important future work is to compare our WSOL model against traditional fully supervised models in terms of processing time and accuracy, to understand the trade-offs between annotation efficiency and localization precision in sewer pipeline defect detection. This comparison will help in identifying the most effective approach for automated infrastructure assessment.

References

- [1] Zhou B., Khosla A., Lapedriza A., Oliva A. and Torralba A. Learning deep features for discriminative localization. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2921–2929, Las Vegas, America, 2016.
- [2] Kim E., Kim S., Lee J., Kim H. and Yoon S. Bridging the Gap between Classification and Localization for Weakly Supervised Object Localization. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 14258–14267, New Orleans, Louisiana, America, 2022.
- [3] Wang H., Li Y., Dang L.M., Lee S. and Moon, H. Pixel-level tunnel crack segmentation using a weakly supervised annotation approach. *Computers in Industry*, 133, 103545, 2021.
- [4] Zhang J., Su H., Zou W., Gong X., Zhang Z. and Shen F. CADN: A weakly supervised learning-based category-aware object detection network for surface defect detection. *Pattern Recognition*, 109, 107571, 2021.
- [5] Wu X., Wang T., Li Y., Li P. and Liu Y. A CAM-Based Weakly Supervised Method for Surface Defect Inspection. *IEEE Transactions on Instrumentation and Measurement*, 71, 1–11, 2022.
- [6] Li Y., Wang H., Dang L.M., Song H.-K. and Moon H. Vision-Based Defect Inspection and Condition Assessment for Sewer Pipes: A Comprehensive Survey. *Sensors*, 22, 2722, 2022.
- [7] Cheng J.C.P. and Wang M. Automated detection of sewer pipe defects in closed-circuit television images using deep learning techniques. *Automation in Construction*, 95, 155–171, 2018.
- [8] Li D., Xie Q., Yu Z., Wu Q., Zhou J. and Wang J. Sewer pipe defect detection via deep learning with local and global feature fusion. *Automation in Construction*, 129, 103823, 2021.
- [9] Yin X., Chen Y., Bouferguene A., Zaman H., Al-Hussein M. and Kurach L. A deep learning-based framework for an automated defect detection system for sewer pipes. *Automation in Construction*, 109, 102967, 2020.
- [10] Yin X., Ma T., Bouferguene A. and Al-Hussein M. Automation for sewer pipe assessment: CCTV video interpretation algorithm and sewer pipe video assessment (SPVA) system development. *Automation in construction*, 125, 103622, 2021.
- [11] Czimmermann T., Ciuti G., Milazzo M., Chiurazzi M., Roccella S., Oddo C. M. and Dario, P. Visual-Based Defect Detection and Classification Approaches for Industrial Applications—A SURVEY. *Sensors*, 20(5), 1459, 2020.
- [12] Zhang G., Pan Y. and Zhang, L. Semi-supervised learning with GAN for automatic defect detection from images. *Automation in Construction*, 128, 103764, 2021.
- [13] Haurum, J. B. and Moeslund, T. B. Sewer-ML: A Multi-Label Sewer Defect Classification Dataset and Benchmark. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 13455–13467, 2021.
- [14] Dansk Vand og Spildevandsforening (DANVA). Fotoman ualen: TV-inspektion af afløbsledninger. Dansk Vand og Spildevandsforening (DANVA), 6 edition, 2010.
- [15] Dansk Vand og Spildevandsforening (DANVA). Fotoman ualen: Beregning af Fysisk Indeks ved TV-inspektion. Dansk Vand og Spildevandsforening (DANVA), 1 edition, 2005.
- [16] Choe J., Oh S.J., Lee S., Chun S., Akata Z. and Shim H. Evaluating Weakly Supervised Object Localization Methods Right. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3133–3140. 2020.
- [17] Welinder P., Branson S., Mita T., Wah C., Schroff F., Belongie S. and Perona P. Caltech UCSD birds 200. Technical Report CNS-TR-2010-001, California Institute of Technology, 2010.
- [18] Russakovsky O., Deng J., Su H., Krause J., Satheesh S., Ma S., Huang Z., Karpathy A., Khosla A., Bernstein M., Berg A.C. and Li F. Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115(3):211–252, 2015.