

APPLICATIONS OF COMPUTER VISION IN PRODUCTION PLANNING AND CONTROL SYSTEM FOR MODULAR AND OFFSITE CONSTRUCTION: TRENDS, CHALLENGES, AND OPPORTUNITIES

M.R.I. Chowdhury¹, M. Rahman^{1,3*}, M. Elkabalawy¹, Q.A. Wasiew¹, E. Rahman¹, S.K. Moon⁴, S.H. Han^{1,2}

¹ Dept of Building, Civil & 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, Canada

³mizanoor.rahman@mail.concordia.ca

⁴ Department of Civil Systems Engineering, College of Engineering, Ajou University, 16499 Suwon, Republic of Korea, skmoon@ajou.ac.kr

ABSTRACT: Modular and off-site construction (MOC) is gaining popularity, particularly in developed countries, due to its environmental, social, and economic advantages. MOC is recognized for improving productivity, reducing costs, and incorporating innovative technologies, making it an efficient and sustainable project delivery method. In recent years, computer vision (CV) is quickly becoming a game-changing tool in construction industry. As a result, the incorporation of CV in MOC is growing, but thorough understanding of its trends, difficulties, and potential in production planning and control system is yet to be discovered. In this respect, this paper is proposed a comprehensive framework to investigate the current developments in production planning and control system using CV applications in the MOC. The proposed framework adopted the Preferred Reporting Items for Systematic Reviews and Meta Analyses (PRISMA) method to discover, assess, and synthesize relevant literature from 2012 to 2024, providing a systematic examination of significant trends, challenges, and new prospects based on Scopus and Web of Science (WoS) database. The outcome of this study identified that the Mask RCNN is the most popular algorithms for detection, tracking and segmentation. Dataset limitations, camera variations, poor lighting, occlusions are the main challenges for CV application in MOC in production planning and control system. Future modular and off-site construction research will integrate RFID and BIM data, enhance algorithms like YOLOv7 with data augmentation and parallel processing, address site-specific challenges like occlusions and congestion, and develop intelligent systems for safety, worker tracking, and real-time monitoring of precast components and module transportation.

1. INTRODUCTION

Modular and off-site construction (MOC) is becoming a popular method in construction due to the controlled production environment. The MOC involves prefabricating building components in controlled factory settings before transporting and assembling them on-site, have gained attention due to their ability to improve quality control, customize design, waste minimization, reduce construction timelines, and minimize disruptions at the project site (Arashpour et al. 2015; Rahman and Han 2024a, 2024b; Rahman and Sobuz 2018).

Construction industry 4.0 is known as converging digital and computing techniques to redefine sustainable construction (Baduge et al. 2022). The construction industry is experiencing a major transformation, driven by the increasing adoption technologies, such as computer vision (CV), virtual reality (VR) etc. This transformation is further bolstered by the adoption of emerging technologies, with computer vision, a subset

of artificial intelligence (AI), playing a pivotal role. Computer vision harnesses digital imagery, video footage, and surveillance data to extract relevant information to facilitate decision-making as well as improved operational precision (Pham et al. 2021). Its growth stems from its ability to execute a range of visual tasks, including object detection (Kim 2020), image classification, object tracking (Wang et al. 2021), action recognition (Xiao et al. 2022), human pose estimation (Chu et al. 2020), and semantic as well as instance segmentation. These capabilities make computer vision a versatile and powerful tool for addressing complex challenges within the construction sector.

Traditionally, production planning and control in offsite construction have relied on manual inspections. However, traditional methods are often resource-intensive, time-consuming, and subject to human fault. The integration of computer vision offers a revolutionary alternative by facilitating real-time monitoring, automated activity recognition, and actionable feedback on productivity challenges. For example, surveillance camera footage analyzed using computer vision can detect material shortages, equipment downtime, and idle labor, all of which directly influence productivity. Advanced deep learning algorithms further enhance these analyses by tackling issues such as occlusion, variable lighting, and complex work environments.

This study seeks to investigate the role of computer vision in production planning and control for modular and offsite construction. Specifically, it aims to address the following research questions: (1) What are the key research trends in the use of computer vision? (2) Which computer vision algorithm(s) or method(s) are being used? (3) What challenges or limitations exist in these applications? By exploring these questions, this paper identifies emerging trends, examines the challenges associated with implementation, and highlights opportunities for future research and practical integration. Ultimately, this research provides valuable insights into how computer vision could change production planning and control, which contributes to the overall productive growth of the MOC.

2. METHODOLOGY

To accomplish the objectives of this research, a systematic review (SR) method is utilized. This approach includes defining the research framework, developing a search plan to gather relevant information, extracting essential data, conducting data analysis, and presenting the results. The study follows the guidelines outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) as described by Page et al. (2021).

The systematic review process, as depicted in Figure 1. The first step is selecting the database and relevant research articles. The study utilized the Scopus and Web of Science (WoS) databases due to their reputation for high-quality articles, particularly in the domain of modular construction (Afzal et al. 2023). Using specific search codes tailored to each database (Table 1), 98 articles related to computer vision were initially retrieved, covering the period from 2012 to 2024. To ensure relevance, articles that were irrelevant, literature reviews, published before 2012, or non-research in nature were excluded, narrowing the selection to 64 publications. Given the overlap between the two databases, duplicate entries were identified and removed, leaving a final set of 58 articles imported into bibliography software for further processing. After thoroughly reviewing the abstracts, 18 articles were ultimately selected for detailed analysis. The 12-year time frame (2012–2024) was deliberately chosen to focus on the most recent advancements and emerging trends, ensuring the study's findings reflect novel and relevant developments in the field.

Bibliometric analysis was employed to uncover key characteristics and trends in the literature by analyzing dimensions such as authorship, institutions, citations, and more. A variety of tools are available for performing bibliometric analysis, such as VOSviewer®, Gephi®, CiteSpace®, SciMAT®, Bibliometrix®, and HistCite® (Junjia et al. 2023). For this study, Bibliometrix® was chosen for bibliographic data analysis because of its extensive functionality and robust data visualization capabilities (Aria and Cuccurullo 2017). To identify current research trends, co-occurrences of keywords within primary publication sources were examined. using Bibliometrix®, an open-source software built on the R programming language. The articles included in this review adhered to specific inclusion criteria, emphasizing research focused on modular construction and the application of deep learning (DL) algorithms.

Table 1: Code used for literature search.

Database	Applied Code
Web of Science (WoS)	(TI= (off site construction OR modular construction OR panelized construction) AND TI= (Deep learning OR computer vision)
Scopus	(TITLE-ABS-KEY (“off site construction” OR “modular construction” OR “panelized construction”) AND (“Deep learning” OR “computer vision”))

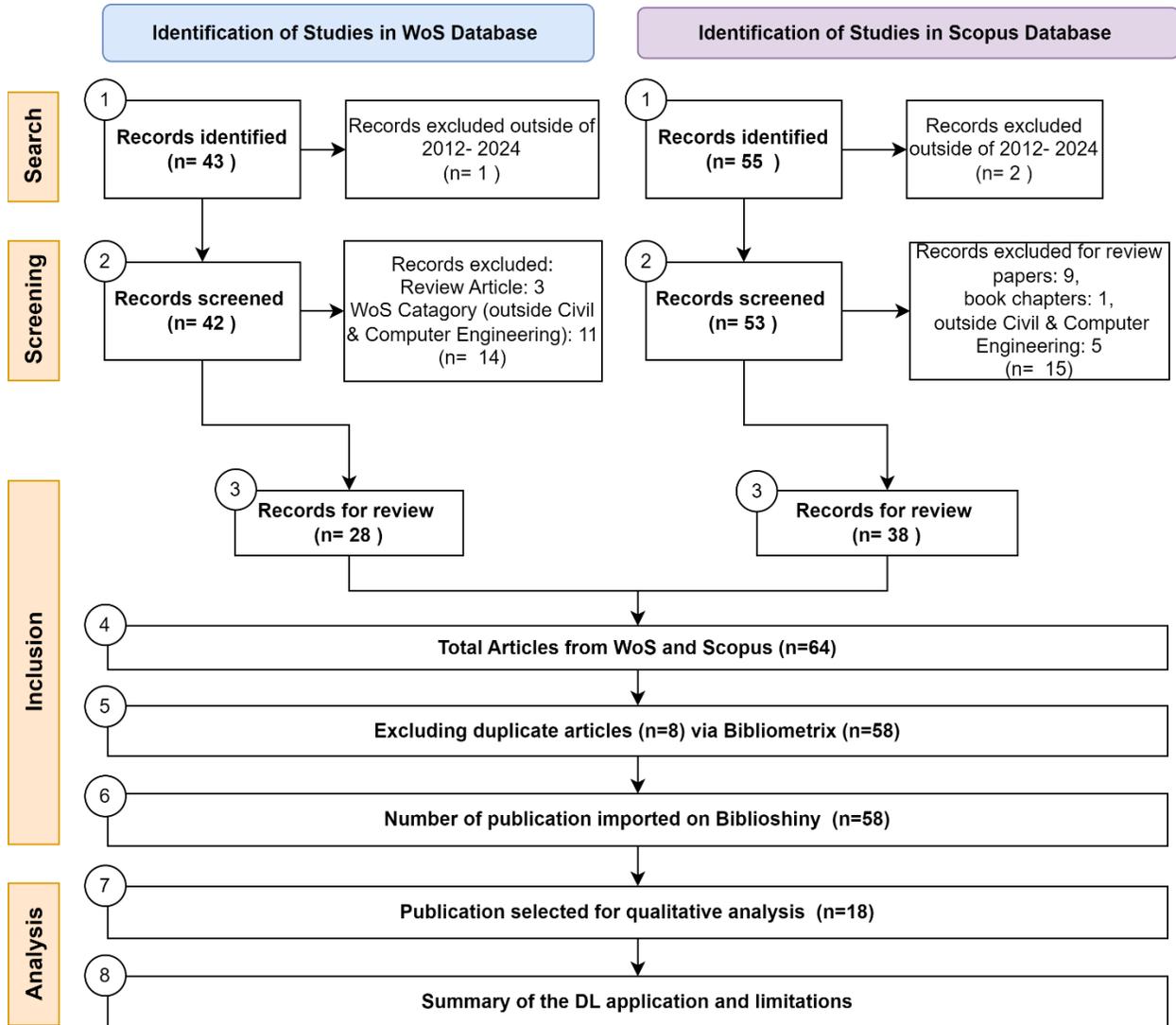


Figure 1: Research Framework

3. RESULTS

3.1 Top Keywords Cluster

The finalized 18 articles are analyzed in Biblioshiny to elevate the research hotspot. The analysis reveals four key clusters within the construction industry network. Metrics such as Betweenness, Closeness, and PageRank highlight the significance and centrality of each node, helping to identify their roles within their respective clusters. Cluster 1 focuses on the integration of computer vision technologies like object detection, neural networks, and deep learning. The construction industry node holds a pivotal position with a Betweenness of 27.837, reflecting its role as a key connector in the network. Similarly, off-site construction has a Betweenness score of 20.294, underscoring its importance in modern modular construction approaches. Technologies such as neural networks (3.859) and deep learning (2.183) further highlight the growing adoption of AI-based methods in improving construction efficiency. Nodes like construction activities (0.181) and network cnn (0.219), while relevant, exhibit lower scores, indicating their limited connectivity within this cluster. Similarly, cluster 2 emphasizes production processes and the enhancement of safety and quality in construction. Cluster 3 centers on modular construction, highlighting its pivotal role in modern construction methods. Cluster 4 explores computational tools, including computer vision, and their application in project management and progress monitoring. Table 2 categorizes nodes into four distinct clusters, each representing a specific focus area in the construction and technology network.

Table 2: Keyword clustering results.

Node	Cluster	Betweenness	Closeness	PageRank
Construction industry	1	27.837	0.019	0.062
Object detection	1	0.923	0.016	0.023
Off-site construction	1	20.294	0.017	0.047
Neural network	1	3.859	0.016	0.038
Deep learning	1	2.183	0.016	0.034
Construction activities	1	0.181	0.015	0.016
Network cnn	1	0.219	0.015	0.021
Production line	2	13.272	0.017	0.039
Production process	2	0	0.012	0.011
Safety quality	2	19.319	0.018	0.045
Vision-based methods	2	4.712	0.017	0.037
Construction methods	2	2.808	0.017	0.033
Integrated construction	2	2.929	0.017	0.036
Modular integrated	2	2.929	0.017	0.036
Surveillance videos	2	8.406	0.017	0.034
Modular construction	3	330.656	0.026	0.145
Quality control	3	0.615	0.015	0.017
Assembly process	3	0.503	0.015	0.016
Construction activity	3	0	0.014	0.007
Detection algorithms	3	0	0.014	0.007
Automatically generate	3	0	0.014	0.008
Building information	3	0	0.014	0.01
Computer vision	4	1.317	0.016	0.025
Construction projects	4	39.451	0.019	0.063
Progress monitoring	4	0	0.016	0.023
Proposed framework	4	0	0.015	0.014
Mask r-cnn	4	33.767	0.016	0.032
Reinforcement learning	4	0	0.012	0.011

operations, Mask R-CNN in conjunction with DeepSORT for tracking precast wall progress, and the Frame Inspection Algorithm (FIA) for monitoring light-gauge steel frame pre-manufacturing.

Table 3: Algorithm used in OSC Monitoring and Productivity Analysis.

Authors	Goal of the study	Used Algorithms
Chen et al. (2024)	Real time process monitoring	YOLO v8
Baek et al. (2024)	Automated productivity monitoring	YOLO v4
Alsakka et al. (2023)	Automated collection of productivity data	YOLO v4
Akbar et al. (2023)	Concrete crack segmentation and quantification	YOLO v7
Panahi, Louis, Podder, Pless, et al. (2023)	Monitoring the assembly process	Mask R-CNN with Resnet-50 backbone
(Panahi, Louis, Podder, Swanson, et al. 2023)	Identifying the bottleneck station on the production line	Mask R-CNN with Resnet-50 backbone
Panahi et al. (2024)	Monitoring production line assembly progress of modular units	Mask R-CNN with Resnet-50 backbone
Xiao et al. (2022)	Tracking workers in off-site construction using instance segmentation	Mask R-CNN with segmentation
Panahi, Louis, Podder, et al. (2022)	Monitoring the construction processes in OSC factories.	OpenCV
Xiong et al. (2022)	Activity recognition using sound via deep learning	CRNN
Liu et al. (2021)	Evaluating the applicability and performance of object detection algorithms	Faster RCNN and SSD
Park et al. (2021)	Generating installation datasets in virtual environments and evaluating the performance of AI models	MV-CNN
Panahi, Louis, Aziere, et al. (2022)	Recognizing the worker activities	CNN Resnet 50
Rashid and Louis (2021)	Measurement of active and idle time at various workstations	RNN (LSTM)
Martinez et al. (2021)	Automatically detect and track the progress of construction operations	Faster R-CNN
Zheng et al. (2020)	Virtual prototyping and transfer-learning techniques for detection	Mask R-CNN
Wang et al. (2021)	Automatic Progress Monitoring of Precast Walls	Mask R-CNN and DeepSORT
Martinez et al. (2019)	Automatic supervision of the light-gauge steel frame pre-manufacturing stage.	Frame Inspection Algorithm (FIA)

3.4 Challenges and Limitations of Computer Vision Application

While applying the CV in OSC, researchers have highlighted challenges in construction monitoring and productivity analysis. For example, Chen et al. (2024) identified dataset diversity limitations, while Baek et al. (2024) noted issues with training dataset size and camera variation. Additionally, occlusions and subprocess detection difficulties were reported by Panahi, Louis, Podder, Pless, et al. (2023) and Alsakka et al. (2023). Akbar et al. (2023) encountered constraints in camera-object distance, and Xiao et al. (2022) faced difficulties in handling mixed sound environments. Overfitting due to static backgrounds Panahi, Louis, Podder, et al. (2022) and reliance on virtual datasets (Park et al. 2021) added to the challenges. Martinez et al. (2021) observed detection issues in wall panel designs and simplifying assumptions affecting task tracking. Moreover, Wang et al. (2021) faced obstacles with poor lighting and camera shaking, highlighting the need for further refinements in construction automation systems.

3.5 Proposed Future Research Directions

Future research trends in modular and offsite construction are centered on advancing computer vision systems through integration with multi-modal data sources and emerging technologies. Figure 5 illustrates a word map of future research trend. Currently, a key focus is enhancing data collection using multiple high-resolution cameras and CCTV systems, while incorporating complementary data such as RFID and Building Information Modeling (BIM) to provide more comprehensive insights. Additionally, efforts are also directed at improving algorithmic performance through advanced models like YOLOv7, data augmentation for segmentation, and parallel video processing to enable real-time applications. Apart from that, adapting systems to site-specific challenges, such as congestion, occlusions, and variable distances, remains critical, alongside optimizing camera placement and addressing occlusion issues are also challenging. A comprehensive framework (Figure 4) is proposed to advance the development and application of DL techniques in productivity management within the MOC industry, with a focus on achieving more accurate object detection, efficient tracking, and precise segmentation. Furthermore, research is expanding to integrate machine hearing techniques for mixed sound environments, develop communication strategies among camera networks, and include blurry samples for better dataset robustness. Applications are also becoming more specific, focusing on monitoring precast components, optimizing module transportation, and addressing unconventional geometries. Safety and ergonomics monitoring using CCTV footage, scalability through worker tracking, and validation via long-duration videos highlight the growing emphasis on creating adaptable, intelligent, and sustainable solutions for production planning and control in modular construction.

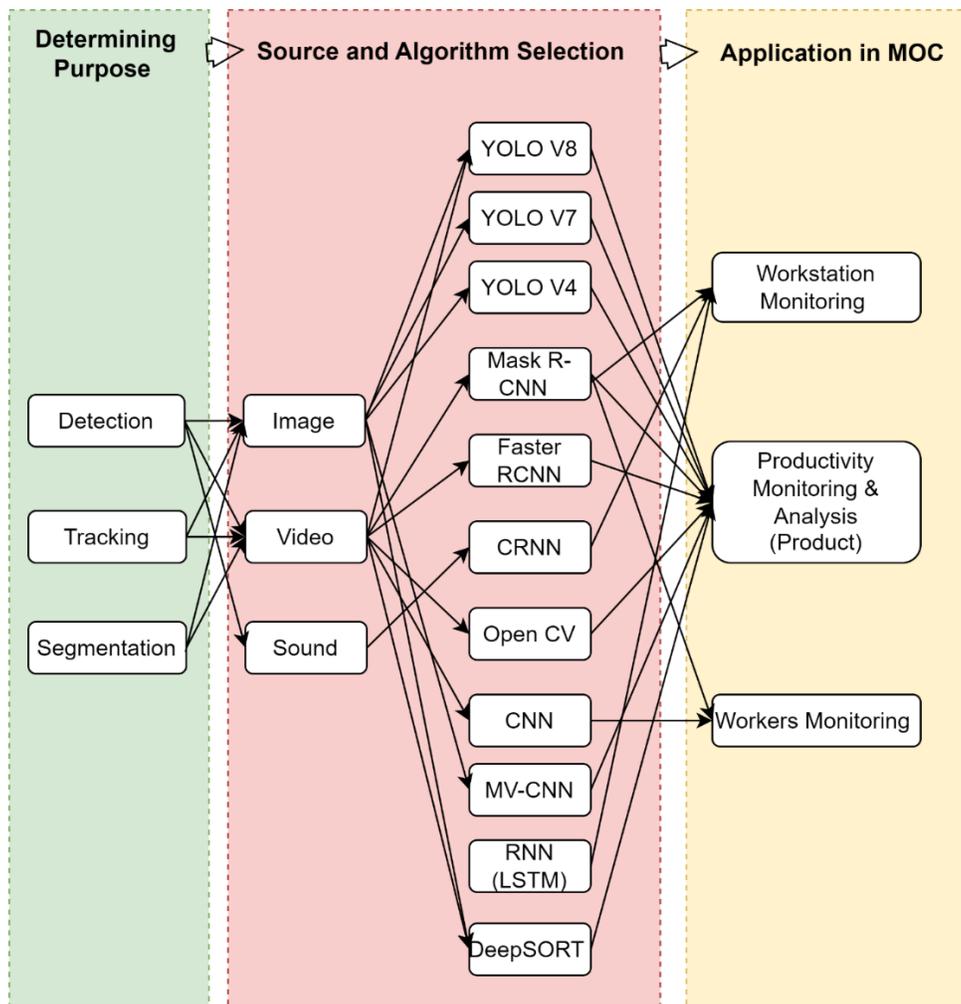


Figure 4: Synthesized framework for a DL application paradigm for MOC



Figure 5: Word map of future research trend using Biblioshiny.

4. CONCLUSIONS

In conclusion, this research highlights the transformative impact of computer vision on production planning and control in modular and offsite construction (OSC). The findings of this review indicate that while computer vision technologies have demonstrated significant promise in addressing key inefficiencies in modular and offsite construction, their practical implementation is still in its early stages. Although the number of included studies is relatively limited, this reflects the application of stringent inclusion criteria designed to ensure the relevance, quality, and methodological rigor of the selected literature. By reviewing literature from 2012 to 2024, the study reveals an increasing use of deep learning algorithms like YOLO models (v4, v7, v8), Mask R-CNN, Faster R-CNN. These algorithms are particularly valued for their ability to perform tasks such as object detection, segmentation, and tracking in real-time, making them crucial for monitoring construction progress, worker activity, assembly line tracking, and quality control. Among them, YOLO and Mask R-CNN are the most frequently applied due to their efficiency and accuracy in addressing a broad spectrum of challenges in construction monitoring.

However, several challenges remain in the application of these technologies. Issues such as limited and non-diverse datasets, occlusions, camera variability, poor lighting, and subprocess detection complicate the real-world deployment of these algorithms. Additional obstacles include static backgrounds, reliance on virtual data, and handling mixed sound environments, all of which hinder the accuracy of activity recognition and progress monitoring. To address these challenges, future research should focus on integrating multi-modal data, such as RFID and BIM, and improving algorithms with techniques like data augmentation and parallel video processing. Optimizing camera placement, reducing occlusions, and adapting systems to site-specific conditions will be essential in enhancing the practical application of computer vision, ultimately driving greater efficiency and sustainability in the MOC industry. The contribution of this study lies in its systematic synthesis of computer vision applications specific to production planning and control in MOC— an area that, to date, has not been comprehensively reviewed.

ACKNOWLEDGMENTS

The authors acknowledge the support of Brain Pool program funded by the Ministry of Science and ICT through the National Research Foundation of Korea (RS-2023-00257444).

REFERENCES

- Afzal, M., Li, R. Y. M., Shoaib, M., Ayyub, M. F., Tagliabue, L. C., Bilal, M., Ghafoor, H. and Manta, O. 2023. Delving into the Digital Twin Developments and Applications in the Construction Industry: A PRISMA Approach. *Sustainability*, 15: 16436.
- Akbar, A., Njoroge, J. M., Lee, S., Chang, Y. and Kwon, S. 2023. CNN-Based Automatic Mobile Reporting System and Quantification for the Concrete Crack Size of the Precast Members of OSC Construction. *Engineering Proceedings*, 36: 15.
- Alsakka, F., El-Chami, I., Yu, H. and Al-Hussein, M. 2023. Computer vision-based process time data acquisition for offsite construction. *Automation in Construction*, 149: 104803.
- Arashpour, M., Wakefield, R., Blismas, N. and Maqsood, T. 2015. Autonomous production tracking for augmenting output in off-site construction. *Automation in Construction*, 53: 13-21.
- Aria, M. and Cuccurullo, C. 2017. bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11: 959-975.
- Baduge, S. K., Thilakarathna, S., Perera, J. S., Arashpour, M., Sharafi, P., Teodosio, B., Shringi, A. and Mendis, P. 2022. Artificial intelligence and smart vision for building and construction 4.0: Machine and deep learning methods and applications. *Automation in Construction*, 141: 104440.
- Baek, J., Kim, D. and Choi, B. 2024. Deep learning-based automated productivity monitoring for on-site module installation in off-site construction. *Developments in the Built Environment*, 18: 100382.
- Chen, X., Wang, Y., Wang, J., Bouferguene, A. and Al-Hussein, M. 2024. Vision-based real-time process monitoring and problem feedback for productivity-oriented analysis in off-site construction. *Automation in Construction*, 162: 105389.
- Chu, W., Han, S., Luo, X. and Zhu, Z. 2020. Monocular Vision-Based Framework for Biomechanical Analysis or Ergonomic Posture Assessment in Modular Construction. *Journal of Computing in Civil Engineering*, 34: 04020018.
- Junjia, Y., Alias, A. H., Haron, N. A. and Abu Bakar, N. 2023. A Bibliometric Review on Safety Risk Assessment of Construction Based on CiteSpace Software and WoS Database. *Sustainability*, 15: 11803.
- Kim, A. 2020, 2020/10/14. Automated On-Site Quality Inspection and Reporting Technology for Off-Site Construction (OSC)-based Precast Concrete Members *Proceedings of the 37th International Symposium on Automation and Robotics in Construction (ISARC)*, 1152-1159.
- Liu, C., M.E. Sepasgozar, S., Shirowzhan, S. and Mohammadi, G. 2021. Applications of object detection in modular construction based on a comparative evaluation of deep learning algorithms. *Construction Innovation*, 22: 141-159.
- Martinez, P., Ahmad, R. and Al-Hussein, M. 2019. A vision-based system for pre-inspection of steel frame manufacturing. *Automation in Construction*, 97: 151-163.
- Martinez, P., Barkokebas, B., Hamzeh, F., Al-Hussein, M. and Ahmad, R. 2021. A vision-based approach for automatic progress tracking of floor paneling in offsite construction facilities. *Automation in Construction*, 125: 103620.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P. and Moher, D. 2021. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery*, 88: 105906.
- Panahi, R., Louis, J., Aziere, N., Podder, A. and Swanson, C. 2022. Identifying Modular Construction Worker Tasks Using Computer Vision. *Proceedings Computing in Civil Engineering 2021*, 959-966.
- Panahi, R., Louis, J., Podder, A., Pless, S., Swanson, C. and Jafari, M. 2023, 2024-01-10. Automated Progress Monitoring in Modular Construction Factories Using Computer Vision and Building Information Modeling. *40th International Symposium on Automation and Robotics in Construction (ISARC)*, Chennai, India,
- Panahi, R., Louis, J., Podder, A. and Swanson, C. 2022. Tracking Volumetric Units in Modular Factories for Automated Progress Monitoring Using Computer Vision. *Proceedings Construction Research Congress 2022*, 822-829.

- Panahi, R., Louis, J., Podder, A., Swanson, C. and Pless, S. 2023. Bottleneck Detection in Modular Construction Factories Using Computer Vision. *Sensors*, 23:
- Panahi, R., Louis, J., Podder, A., Swanson, C. and Pless, S. 2024. Automated Assembly Progress Monitoring in Modular Construction Factories Using Computer Vision-Based Instance Segmentation. *Proceedings Computing in Civil Engineering 2023*, 290-297.
- Park, K., Ergan, S. and Feng, C. 2021. Towards Intelligent Agents to Assist in Modular Construction: Evaluation of Datasets Generated in Virtual Environments for AI training. *ISARC*, Dubai, UAE, 327-333.
- Pham, H. T. T. L., Rafieizonooz, M., Han, S. and Lee, D.-E. (2021). Current Status and Future Directions of Deep Learning Applications for Safety Management in Construction. *Sustainability*, 13(24).
- Rahman, M. and Han, S. H. 2024a. Linear Scheduling Method-Based Multi-Objective Optimization for Off-Site Construction Manufacturing. *Automation in Construction*, 167: 105712.
- Rahman, M. and Han, S. H. 2024b. Performance Evaluation of Genetic Algorithm and Particle Swarm Optimization in Off-Site Construction Scheduling. *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*, 41: 275-282.
- Rahman, M. and Sobuz, H. R. 2018. Comparative study of IPS & PPVC precast system—A case study of public housing buildings project in Singapore. *Proceedings of the 4th International Conference on Civil Engineering for Sustainable Development (ICCESD 2018)*, KUET, Khulna, Bangladesh, 9-11.
- Rashid, K. M. and Louis, J. 2021, 12-15 Dec. 2021. Automated Active and Idle Time Measurement in Modular Construction Factory Using Inertial Measurement Unit and Deep Learning for Dynamic Simulation Input. *2021 Winter Simulation Conference (WSC)*, 1-8.
- Wang, Z., Zhang, Q., Yang, B., Wu, T., Lei, K., Zhang, B. and Fang, T. 2021. Vision-Based Framework for Automatic Progress Monitoring of Precast Walls by Using Surveillance Videos during the Construction Phase. *Journal of Computing in Civil Engineering*, 35: 04020056.
- Xiao, B., Xiao, H., Wang, J. and Chen, Y. 2022. Vision-based method for tracking workers by integrating deep learning instance segmentation in off-site construction. *Automation in Construction*, 136: 104148.
- Xiong, W., Xu, X., Chen, L. and Yang, J. 2022. Sound-Based Construction Activity Monitoring with Deep Learning. *Buildings*, 12:
- Zheng, Z., Zhang, Z. and Pan, W. 2020. Virtual prototyping- and transfer learning-enabled module detection for modular integrated construction. *Automation in Construction*, 120: 103387.