Non-Productive Time Reduction in Off-Site Construction: A Predictive Analytics Approach Using Deep Learning

Xue Chen[1], Weining Zeng[1], Runcong Liu[1], Ahmed Bouferguene[2], Mohamed Al-Hussein[1]

[1]Department of Civil and Environmental Engineering, University of Alberta, Edmonton, Alberta, Canada
[2]Campus Saint-Jean, University of Alberta, Edmonton, Alberta, Canada
xue15@ualberta.ca, weining3@ualberta.ca, runcong@ualberta.ca, ahmedb@ualberta.ca, malhussein@ualberta.ca

Abstract –

This study proposes a data-driven framework to reduce Non-Productive Time (NPT) in Off-Site Construction (OSC) by integrating predictive modeling with lean principles. Through comprehensive computer vision (CV) techniques, large-scale video data are analyzed to identify key operational states, worker positions, and material flows. Two primary classification approaches—a Random Forest (RF) model and a Deep Learning (DL) network—are employed to detect NPT episodes. Experimental results reveal that the DL model achieves substantially higher accuracy and F1 scores, demonstrating its capability to handle complex spatial–temporal interactions. Feature-weight analysis further indicates that task transitions, assembly processes, and coordinate-based congestion zones are critical drivers of NPT. Building on these insights, targeted lean interventions—such as just-in-time material delivery, proactive maintenance, and optimized workstation layouts—can streamline operations, minimize idle intervals, and enhance overall productivity. The conclusions underscore the effectiveness of merging data analytics with lean strategies in uncovering hidden inefficiencies within OSC, providing a clear direction for future research and industry adoption.

Keywords –

Non-Productive Time, Off-Site Construction, Deep Learning, Lean Principles, Predictive Analytics

1 Introduction

As the construction industry increasingly prioritizes efficiency and quality, OSC has emerged as a prominent solution due to its advantages in factory prefabrication and rapid on-site assembly. These features not only shorten construction timelines and enhance quality control but also significantly reduce reliance on on-site labor and environmental impact [1,2]. Building on this foundation, lean construction emphasizes systematic workflow management and continuous improvement to minimize waste and maximize value-adding activities, thereby improving overall productivity [3]. However, despite its advantages, practical challenges such as delays, equipment idling, and communication inefficiencies often result in NPT during manufacturing and assembly processes, which substantially undermine the potential benefits of OSC [4,5].

Recent research trends highlight the application of CV and artificial intelligence (AI) techniques in construction settings as a promising means to efficiently detect, track, and interpret the behavior of workers, machinery, and materials. These technologies enable comprehensive monitoring of OSC workflows and the identification of non-value-adding activities [6,7]. Compared to traditional manual observations or documentation, vision-based systems can process large-scale image and video data in significantly less time, thereby achieving greater precision in identifying waste [8]. For instance, some studies have employed trajectory analysis of workers and machinery to pinpoint repetitive idle zones, providing actionable insights for optimizing workflow layouts and material transport. Others have utilized convolutional neural networks (CNNs) or enhanced You Only Look Once (YOLO) models to rapidly detect prefabricated components and issue real-time alerts for anomalies or defects in production processes [9,10]. Additionally, advanced techniques such as semantic segmentation and pose estimation have facilitated the precise classification and quantification of work behaviors, offering enriched data support for productivity enhancement and safety management [11].

In the context of productivity management, automated monitoring and data-driven analytics are being increasingly integrated into OSC practices. Numerous studies have focused on using vision-based detection and tracking to monitor crane operations, worker pathways, and component statuses. These datasets are often synchronized with BIM models, digital twins, or project schedules to identify deviations from planned progress and supply chain inefficiencies [12,13].

However, much of the existing literature centers on quantifying discrepancies between actual progress and planned schedules, with limited emphasis on systematically identifying the root causes of process delays and resource inefficiencies [14]. Integrating the principles of lean construction—specifically its focus on reducing "process waste" and driving "continuous improvement" — with the automated detection and predictive analytics capabilities of CV offers a pathway to achieve significant breakthroughs in reducing NPT and enhancing construction efficiency [15].

In summary, this study builds on the high standardization and modularity characteristics of OSC by integrating lean construction principles with CV technologies to automatically identify and quantify non-productive time. The paper begins with a review of advancements in the application of OSC, lean construction, and vision algorithms in the construction sector. Subsequently, a DL model is developed using real-world data to identify and predict NPT in OSC processes. The findings provide empirical evidence for both academia and industry practitioners seeking to enhance productivity in OSC under the backdrop of digitization and automation.

2 Methodology

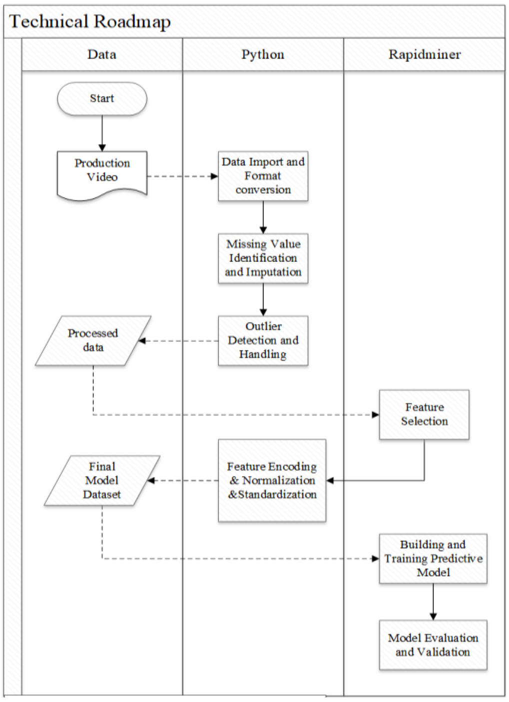

Figure 1. Technical Roadmap for NPT Prediction of CV Acquired Data Using DL model

Figure 1 shows the technical path for NPT prediction using DL. It consists of the following seven components:

1. Data Import and Format Conversion: Diverse data formats, including CSV and Excel files, are loaded and merged into consistent tables to enable unified analysis. YOLO outputs, which have undergone pre-processing, are incorporated at this stage to ensure compatibility with the analysis pipeline.
2. Missing Value Identification and Insertion: Context-driven imputation methods address missing values by filling short vacant intervals with information from adjacent frames. This strategy helps maintain data continuity while minimizing the risk of introducing artificial trends.
3. Outlier Detection and Handling: Outliers, such as sudden coordinate jumps or unrealistic worker counts, are detected and removed using custom algorithms. Thresholds for identifying anomalies are visually validated with graphical tools to ensure robustness.
4. Feature Selection: The influence of each attribute on NPT is assessed through statistical techniques such as ANOVA and Logistic Regression. Attributes with minimal contribution are excluded to optimize the predictive feature set.
5. Feature Encoding, Normalization, and Standardization: Numerical attributes are standardized using Min-Max scaling or Z-score normalization to achieve uniform magnitudes. Categorical variables, such as “work conditions,” are encoded using One-Hot encoding to eliminate potential issues with implicit ordering.
6. Build and Train Predictive Models: Processed features are used to train classifiers, including RF and DL. Iterative tuning of hyperparameters, such as tree depth and layer size, is conducted to enhance model performance and accuracy.
7. Model Evaluation and Validation: Model performance is validated through multiple holdout experiments, evaluating metrics such as Recall, F1 score, and AUC-PR to address class imbalance. Iterative improvements are made until the model meets predefined performance targets, ensuring reliability and robustness.

3 Implemented Details of case study

3.1 Data Acquisition Process and Database

3.1.1 Camera Deployment and Data Collection

The research team strategically installed cameras within an off-site construction factory that specializes in manufacturing prefabricated residential building components. The collected videos focused on a specific workstation dedicated to producing wooden floors with an example shown in Figure 2. Videos were recorded at

1080 × 1920 pixels across multiple production cycles, capturing diverse working conditions and work resources (such as workers, materials, and machines).

Figure 2. Screenshot of Video

3.1.2 Object Detection and Data Acquisition

The methods for tracking, detecting, and identifying objects align with those proposed by Chen, with extensions in coverage and quantity [16]. After video collection, the YOLO V8 algorithm was employed to detect objects such as workers, machinery, and material racks, generating bounding box coordinates for each frame. ByteTrack was then used to assign and maintain tracking IDs, ensuring object consistency across frames. Spatial density analysis facilitated the assessment of worker positions relative to workstations, enabling the identification of potential idle times.

With the vision-based analysis, non-productive time, including waiting time for machine, waiting time for worker, waiting time for material, empty time of a workstation, and delayed movement of completed products, are automatically identified from off-site construction processes. For example, waiting time for machine, materials and workers are defined as the time difference between the actual acquisition time of the required resource (observed from the object detection results) and the expected timepoint to obtain this resource (predicted with the activity recognition results since it's usually the endpoint of the preceding activity). The specific non-productive time identification algorithms are developed in the author's previous work [16]. Using the outputs from the computer vision models, the following sections work on the data-driven non-productive time analysis and prediction with predictive analytics approaches.

3.1.3 Data Structure

The video dataset comprised 19,828 frames, later expanded into 50,450 records. The final dataset includes 11 columns, encompassing numerical, categorical, and Boolean data types. Table 1 provides an overview:

Table 1. Overview of Dataset

Column Name	Data Type	Completeness
Floor ID_x	Integer	No missing values
Video Frame ID	Integer	No missing values
Original Worker Count	Integer	No missing values
Real Number of Workers	Integer	No missing values
Work Status	Categorical	No missing values
Current Time	Floating-point	No missing values
Non-productive Time	Boolean	No missing values
Floor ID_y	Integer	50446 of 50450
X_final	Floating-point	50446 of 50450
Y_final	Floating-point	50446 of 50450
Class	Floating-point	50446 of 50450

3.2 NPT Identification and Screening

3.2.1 Initial Rules and Manual Annotation

A subset of data samples was initially reviewed and manually labeled to identify typical instances of "idle workers" or "unattended workstations." This step involved defining key thresholds for "prolonged inactivity" or "unmanned stations," along with minimum durations and corresponding workstations.

3.2.2 Rule Automation and Boundary Validation

The rules derived from manual annotation were then automated and applied to the entire dataset. For edge cases, such as brief task switches or discontinuous video frames, manual re-validation was performed to ensure the accuracy of NPT labels. Periods with active workers or operational equipment were marked as productive, while others were classified as NPT and stored as Boolean variables in the database.

3.3 Data Cleaning

3.3.1 Handling Missing Values

Coordinate and work status gaps: A context-aware imputation method was applied. For example, if a worker temporarily moved out of the camera's field of view, their coordinates were estimated using adjacent frames. Similarly, during task transitions where labels were briefly unidentifiable, transitional states were inferred based on operational sequences.

This approach minimized distortions from blind interpolation and reduced misjudgments of spatial-temporal relationships [17].

3.3.2 Removing Outliers

Threshold detection was applied to worker counts and coordinates. Records were flagged and removed if coordinate shifts exceeded a predefined max value and did not align with workstation entry zones.

For static objects, such as machinery or material racks, abrupt positional changes were cross verified against predefined spatial boundaries and cleaned accordingly. This process adhered to the physical layout of the factory and the logical continuity of OSC activities, effectively eliminating noise without compromising valid samples.

3.4 Feature Selection and Processing

3.4.1 Logistic Regression and ANOVA Integration

Logistic regression was used to analyze the relationship between features and the target variable, NPT, while ANOVA assessed whether numerical attributes exhibited significant differences between productive and non-productive periods ($p<0.05$). The results highlighted five significant attributes- “number of workers,” “work status,” “X_final,” “Y_final,” and “class”-while “current time” showed weak correlation and was excluded. Table 2 shows the results of the individual correlation analyses (W stands for work conditions in attribute.). Table 3 shows the results of the correlation analysis for the ANOVA matrix model.

Table 2. Correlation analyzing results for logistic regression model

Attribute	Coefficient	Std. Coefficient	Weight
W: placing joist	-0.56	-0.56	-0.56
W: connecting joist	0.35	0.35	0.35
W: placing panel	0	0	0
W: connecting joist and panel	0	0	0
W: lifting the completed panel	-0.057	-0.057	-0.057
W: cleaning the workstation	-0.316	-0.316	-0.316
W: preparing for the new work	2.153	2.153	2.153
W: idle	0	0	0
Number of workers	-0.337	-0.303	-0.303
Current time	0	0.306	0.306
x_final	0	0	0
y_final	0.001	0.041	0.041
Class	0.168	0.164	0.164
Intercept	-3.452	-1.942	-

Table 3. Correlation analyzing results for ANOVA matrix model

ANOVA Attribute	Group non-productive time
x_final	0
y_final	0.004
class	0

3.4.2 Normalization and Encoding

Number of workers: Min–Max normalization was applied to facilitate DL processing.

Coordinates (X_final, Y_final): Z-score standardization was used to mitigate scale differences and outlier impacts.

Work status and class: One-Hot encoding converted these attributes into binary indicators, avoiding the introduction of pseudo-order relationships.

The final dataset comprised 17 processed features, enabling robust predictive modeling.

3.5 Predictive Model Development

3.5.1 Data Partitioning and Model Selection

The structured dataset, created through data cleaning, feature engineering, and encoding, was divided into training and validation subsets to ensure unbiased evaluation. RF and DL were selected to address both linear and non-linear relationships with the target variable, NPT.

RF: An ensemble of decision trees trained on bootstrapped samples and random feature subsets, RF enhances generalization and mitigates overfitting. It also provides feature importance insights for NPT prediction.

DL leverage flexible architectures with hidden layers and non-linear activation functions, such as ReLU, to capture complex feature interactions. Regularization techniques, such as dropout and limited epochs, prevent overfitting and ensure computational efficiency and repeatability.

Random forest: Key hyperparameters such as the number and depth of trees were optimized.

DL: The focus was on tuning hidden layers and learning rates. If performance metrics (e.g., Recall, F1, AUC-PR) were unsatisfactory, feature adjustments or model structure refinements were iteratively performed.

3.5.2 Iterative and Validation

Hyperparameters were iteratively tuned to optimize both models:

- RF: Parameters like the number of trees and maximum depth were adjusted.
- DL: Network depth, neuron count, and learning rate were fine-tuned, with adaptive learning rates ensuring stable convergence.

Validation used a multiple holdout strategy, averaging performance metrics (recall, F1 score, and AUC-PR) across several data splits. This robust evaluation minimized bias and ensured generalizability to unseen data.

4 Results and discussions

4.1 Overall Performance and Confusion Matrices

This study employs both RF and DL models to classify NPT. Table 4Table 5Table 6 sequentially present the confusion matrices of the two models and their key evaluation metrics.

Table 4. Deep Learning Model Confusion Matrix

	True FALSE	True True	Class precision
Pred FALSE	9119	2154	80.89%
Pred Ture	4106	11081	72.96%
Class recall	68.95%	83.72%	

Table 5. Random Forest model confusion matrix

	True FALSE	True True	Class precision
Pred FALSE	5424	2033	72.74%
Pred Ture	7806	11197	58.92%
Class recall	41.00%	84.63%	

Table 6. Comparison of key indicators of different models

	Deep Learning	Random Forest
Accuracy	76.34% ± 0.16%	62.82% ± 0.16%
Classification Error	23.66% ± 0.16%	37.18% ± 0.16%
F1	77.97% ± 0.13%	69.48% ± 0.13%
Sensitivity (Ture)	83.72% ± 0.21%	84.63% ± 0.16%

With the construction data collected from off-site construction factories, the performances of both DL model and RF model as well as their underlying structural reasons are validated and compared in actual construction cases.

1. Accuracy: The DL model achieves an accuracy of 76.34%, which is over 13% higher than Random Forest's 62.82%. This indicates that the DL model is more precise in distinguishing between productive and non-productive states under complex working conditions. The relatively lower accuracy of RF can be attributed to its limitations in handling complex decision boundaries, particularly when distinguishing short pauses from genuine NPT, where RF tends to oversimplify feature interactions.
2. Classification Error: The DL model's classification error is 23.66%, significantly lower than RF's 37.18%, demonstrating that DL is more balanced in classifying negative (productive) and positive (NPT) classes. RF's high error rate is mainly due to its low Specificity (only 41%), implying that many productive cases (FALSE) are misclassified as NPT (TURE), potentially causing unnecessary disruptions in practical scenarios.
3. F_measure: F_measure combines Precision and Recall, serving as a crucial indicator of model performance in real-world settings. The DL model's F_measure is 77.97%, surpassing RF's 69.48%, reflecting that DL not only accurately identifies NPT but also significantly reduces false alarms. RF's lower F_measure indicates a higher misclassification rate in productive states, undermining its practical utility.
4. Recall: For the positive class (NPT), both RF and DL exhibit relatively high recall rates, 84.63% and 83.72% respectively. While RF's Sensitivity is slightly higher, its overall effectiveness remains limited due to low Specificity.

Compared to the RF model, the good performance of the DL model with an accuracy of 76.34% and a recall of 83.72% indicates its capacity in predicting non-productive time for off-site construction processes based on the given information such as current work activities, worker locations and current time. Using the prediction results, adaptive measures in construction management plans can be developed in advance, thereby mitigating construction delays, and ensuring on-time project delivery.

4.2 Analysis of Model Characteristics

RF splits features using multiple decision trees and performs well for single-dimension or discrete variables. However, when facing high-dimensional, nonlinear, or spatiotemporal interactions (e.g., worker coordinates over time, complex assembly processes), insufficient tree depth or limited feature engineering can hamper its ability to capture subtle, dynamic patterns.

DL leverages multi-layer network architectures to learn higher-order nonlinear relationships, which proves advantageous for modeling worker trajectories, equipment interactions, and process coordination. In scenarios involving workflow transitions or multi-person collaboration, the representational power of DL leads to more balanced, stable, and accurate classification in OSC contexts characterized by multi-station parallel production and high spatiotemporal complexity.

4.3 Feature Weight Discussion

4.3.1 Significance of High-Weight Features

'Preparing for the new work' (highest weight, 0.263): When the DL model detects a transition or preparation phase, it becomes more inclined to classify the process as "non-productive or inefficient." This is because waiting and non-value-added activities (e.g., unprepared materials or unfamiliar workflow sequences) can easily occur [1]. In this study, delayed start times or repeated waiting within such periods often lead to NPT classification.

'Connecting joist and panel' & 'Connecting joist' (weights - 0.179, 0.156): These tasks involve more intricate or tedious assembly processes requiring multi-person coordination, alignment, cutting equipment, and fasteners. Any delay or error in one step can halt the entire workstation. The deep network captures the high complexity of these phases, thus quickly labeling them as NPT upon detecting anomalies (e.g., lack of materials, worker misalignment). •

'Coordinate Position' (x_final_normalized, 0.131): Conventional logistic regression or ANOVA analyses may not consistently highlight the significance of spatial coordinates. Nonetheless, in multi-layer network learning, coordinate data can indicate congestion or interference in specific areas—e.g., storage zones, material stacking areas, or transfer pathways. The model's hidden layers may capture an implicit pattern that "frequent stays in certain locations correlate with higher NPT probability."

'Class_0.0' (One-hot encoding, 0.107): Different object types (e.g., worker, machine, material rack) can influence the likelihood of pauses in a working environment:

4.3.2 Possible Reasons for Medium-Low Weight Features

'Number_of_workers_normalized' (0.097): Although the match between labor availability and task demands is often considered crucial, it did not rank among the highest-weight features in this study. Potential reasons include:

In scenarios where complex procedures (such as connection or prep work) become bottlenecks, the effect of workforce size diminishes.

In OSC production settings, workforce numbers typically remain fairly constant, with significant variations appearing only under extreme conditions.

'Idle, cleaning the workstation, etc.' (low-weight features): Such auxiliary or short-duration tasks comprise a minor share of total production time. They are usually simpler in nature and do not necessarily trigger major waiting periods. The DL model is more attentive to contexts such as "workflow transitions" or "insufficient material/equipment readiness," which have a greater impact on overall efficiency. Consequently, these auxiliary states are assigned lower weights.

4.3.3 Adaptability for effective deployment of model

In addition to the application scenarios in this study, the framework in this study also has the potential to help capture and predict waiting and idle time in the field of steel bar processing, where multi-camera monitoring combined with deep learning recognition can help locate snags in time, such as during steel bar cutting, bending, welding or bundling. At the same time, it also has excellent potential for producing precast concrete components, especially the fabrication and assembly of steel cages. It can collect real-time interaction data on personnel, materials and equipment at workstations and identify and alarm for lengthy waiting or cumbersome operations during bundling, positioning and welding.

Extending the framework to a new environment faces some predictable technical and data challenges. First, to achieve complete monitoring in a larger or more irregular space, the number of cameras needs to be increased, and a multi-angle layout should be adopted to reduce visual blind spots and significantly overlapping views that help solve occlusion problems. If applied to outdoor construction environments, additional factors can be introduced to complement visual data so workers and equipment can be accurately tracked in light and weather conditions. Second, to maintain the prediction model's accuracy in various scenarios, it is necessary to train it using a richer dataset that covers more workflows, combined with algorithm enhancements such as deeper network structures or attention mechanisms to adapt to complex temporal and spatial changes.

5 Conclusion and limitations

5.1 Conclusion

This study combines predictive modeling with lean principles to effectively reduce NPT in OSC. Leveraging multi-dimensional data, including task status, object location, and production conditions, the developed model—especially the DL—demonstrates high precision and accuracy in detecting and quantifying periods of non-productive activities. The findings not only illustrate the impact of various task stages (e.g., "preparing for new work" or "connecting joists and panels") on overall productivity but also provide a basis for implementing targeted process improvements.

Compared to traditional manual observations or single-metric evaluations, the data-driven approach presented here more accurately captures potential efficiency gaps and helps decision-makers focus on high-impact scenarios. For instance, the model's ability to identify such issues as insufficient material preparation, improper assembly coordination, and delayed maintenance directly supports the adoption of JIT material delivery and preventive maintenance. By intervening at key points such as task transition, possible measures (such as dynamic job scheduling and resource allocation) can be explored in future work to effectively reduce idle time and resource waste while enhancing worker satisfaction and safety. Furthermore, integrating these predictive insights with lean principles enables continuous evaluation and optimization of production workflows. As data accumulates and feature engineering becomes more refined, the DL recall is expected to further improve, capturing a broader scope of inefficiencies and forming a "data–improvement–feedback" loop that drives production systems toward greater agility, efficiency, and human-centric operation.

5.2 Limitations

While the study primarily focuses on key variables such as work status, coordinate positions, and basic staffing information, the current set of features does not fully capture the intricate spatial–task interactions inherent in OSC. For example, more detailed demarcations of "material storage" versus "assembly" zones or richer descriptors of specialized tasks (e.g., more granular role assignments or real-time equipment load data) could offer deeper insights into production bottlenecks.

Although the DL outperforms traditional methods in detecting NPT, it has not yet attained an ideal recall level. Certain rare or brief forms of inefficiency may remain undetected. Future research should explore more sophisticated architectures (e.g., temporal modeling or attention mechanisms) and a tighter integration of spatiotemporal features to improve the model's ability to capture elusive inefficiency events.

The framework's validation primarily focuses on standardized prefabrication workflows, particularly the panelized wood floor production process, which includes material transport, joist and panel layout, mechanical fastening, and hoisting operations. Workers and machines responsible for placing floor joist and panel, nailing and cutting, and lifting the completed building components are monitored with surveillance camera. This off-site construction scenario can extend to other construction processes such as steel and concrete prefabrication or curtain wall fabrication, as well as repetitive on-site tasks including mechanical, electrical, and plumbing assembly or rebar cage tying. Further implementation may require adaptations for each specific task since every new application scenario demands customized camera placements suited to the spatial workflows at hand, as well as model retraining with domain-specific data.

References

[1] M. R. Hosseini, I. Martek, E. K. Zavadskas, A. A. Aibinu, M. Arashpour, and N. Chileshe, "Critical evaluation of off-site construction research: A Scientometric analysis," *Autom. Constr.*, vol. 87, pp. 235–247, Mar. 2018, doi: 10.1016/j.autcon.2017.12.002.

[2] Z. Liu, B.-G. Hwang, and J. M. Lim, "Prefabricated and Prefinished Volumetric Construction: Assessing Implementation Status, Perceived Benefits, and Critical Risk Factors in the Singapore Built Environment Sector," *J. Manag. Eng.*, vol. 39, no. 6, p. 04023045, Nov. 2023, doi: 10.1061/JMENEA.MEENG-5455.

[3] A. Hadi, F. Cheung, S. Adjei, and A. Dulaimi, "Evaluation of Lean Off-Site Construction Literature through the Lens of Industry 4.0 and 5.0," *J. Constr. Eng. Manag.*, vol. 149, no. 12, p. 03123007, Dec. 2023, doi: 10.1061/JCEMD4.COENG-13622.

[4] I. Taifa and T. Vhora, "Cycle time reduction for productivity improvement in the manufacturing industry," *J. Ind. Eng. Manag. Stud.*, vol. 6, no. 2, pp. 147–164, Dec. 2019, doi: 10.22116/jiems.2019.93495.

[5] B. Barkokebas, P. Martinez, A. Bouferguene, F. Hamzeh, and M. Al-Hussein, "Digitalization-based process improvement and decision-making in offsite construction," *Autom. Constr.*, vol. 155, p. 105052, Nov. 2023, doi: 10.1016/j.autcon.2023.105052.

[6] A. Lekan, A. Clinton, E. Stella, E. Moses, and O. Biodun, "Construction 4.0 Application: Industry 4.0, Internet of Things and Lean Construction Tools' Application in Quality Management System of Residential Building Projects," *Buildings*, vol. 12, no. 10, Art. no. 10, Oct. 2022, doi: 10.3390/buildings12101557.

[7] H. M. Dara, A. Raut, M. Adamu, Y. E. Ibrahim, and P. V. Ingle, "Reducing non-value added (NVA) activities through lean tools for the precast industry," *Heliyon*, vol. 10, no. 7, p. e29148, Apr. 2024, doi: 10.1016/j.heliyon.2024.e29148.

[8] M. S. Bajjou and A. Chafi, "Lean construction and simulation for performance improvement: a case study of reinforcement process," *Int. J. Product. Perform. Manag.*, vol. 70, no. 2, pp. 459–487, Jul. 2020, doi: 10.1108/IJPPM-06-2019-0309.

[9] D. Roberts and M. Golparvar-Fard, "End-to-end vision-based detection, tracking and activity analysis of earthmoving equipment filmed at ground level," *Autom. Constr.*, vol. 105, p. 102811, Sep. 2019, doi: 10.1016/j.autcon.2019.04.006.

[10] H. Luo, C. Xiong, W. Fang, P. E. D. Love, B. Zhang, and X. Ouyang, "Convolutional neural networks: Computer vision-based workforce activity assessment in construction," *Autom. Constr.*, vol. 94, pp. 282–289, Oct. 2018, doi: 10.1016/j.autcon.2018.06.007.

[11] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2014, pp. 580–587. doi: 10.1109/CVPR.2014.81.

[12] S. Wang, L. Deng, J. Guo, M. Liu, and R. Cao, "Automatic Quality Inspection of Rebar Spacing Using Vision-Based Deep Learning with RGBD Camera," presented at the 41st International Symposium on Automation and Robotics in Construction, Lille, France, Jun. 2024. doi: 10.22260/ISARC2024/0009.

[13] A. Braun, S. Tuttas, A. Borrmann, and U. Stilla, "Improving progress monitoring by fusing point clouds, semantic data and computer vision," *Autom. Constr.*, vol. 116, p. 103210, Aug. 2020, doi: 10.1016/j.autcon.2020.103210.

[14] M. Arashpour, R. Wakefield, N. Blismas, and E. W. M. Lee, "Analysis of Disruptions Caused by Construction Field Rework on Productivity in Residential Projects," *J. Constr. Eng. Manag.*, vol. 140, no. 2, p. 04013053, Feb. 2014, doi: 10.1061/(ASCE)CO.1943-7862.0000804.

[15] R. Pozzi, V. G. Cannas, and T. Rossi, "Data Science Supporting Lean Production: Evidence from Manufacturing Companies," *Systems*, vol. 12, no. 3, Art. no. 3, Mar. 2024, doi: 10.3390/systems12030100.

[16] X. Chen, Y. Wang, J. Wang, A. Bouferguene, and M. Al-Hussein, "Vision-based real-time process monitoring and problem feedback for productivity-oriented analysis in off-site construction," *Autom. Constr.*, vol. 162, p. 105389, Jun. 2024, doi: 10.1016/j.autcon.2024.105389.

[17] H. Lee and S. Yun, "Strategies for Imputing Missing Values and Removing Outliers in the Dataset for Machine Learning-Based Construction Cost Prediction," *Buildings*, vol. 14, no. 4, p. 933, Mar. 2024, doi: 10.3390/buildings14040933.