ISSN 2616-5767 Vol.8, Issue 3: 13-22, DOI: 10.25236/AJETS.2025.080303

Dynamic inventory management of prefabricated components using computer vision

Hong Zhang[1], Shengwei Liu[1,*]

[1]Institute of Intelligent Construction and Engineering Management, College of Civil Engineering and Architecture, Zhejiang University, Hangzhou, China
**Corresponding author: swliu@zju.edu.cn*

Abstract: *In prefabricated construction projects, the supply of prefabricated components is closely related to the project construction progress. Shortage of inventory will delay construction progress, and excessive inventory on the construction site will increase stacking costs. Therefore, on-site inventory management of prefabricated components is crucial for prefabricated construction. This paper employs computer vision to precisely identify the types and quantities of prefabricated components on construction sites. Furthermore, inventory management theory is combined to establish a dynamic inventory management optimization model for prefabricated components. Ultimately, the particle swarm optimization is utilized to determine and obtain the optimal inventory parameters. A dynamic inventory management model for prefabricated components based on real-time monitoring is proposed, and demonstrated and validated through case study. The results can facilitate the progress of prefabricated construction projects as planned, expanding the integration and application prospects of computer vision and inventory management theory in the field of construction management.*

Keywords: *Prefabricated building; Prefabricated component; Inventory management; Computer vision; Particle swarm optimization*

1. Introduction

In the context of the digital and intelligent transformation of the global construction industry, computer vision technology based on deep learning has been widely used in the field of construction engineering, focusing on the identification, monitoring and evaluation of objects, equipment and personnel on the construction site through cameras. Relevant studies include identification of workers' construction posture and behavior [1,2], estimation of earthwork productivity [3], measurement of material quantity [4], restoration and continuation of project schedule [5-7], collision detection and defect detection [8-11].

In terms of inventory management of prefabricated components, ensuring sufficient inventory of various prefabricated materials on site is the key factor to promote the steady progress of the construction schedule. Unreasonable stacking of prefabricated components may delay the lifting and installation of prefabricated components, thus prolonging the construction period of prefabricated construction projects and reducing the assembly productivity [12]. The acquisition of material quantities on traditional construction sites mainly relies on the inspection and recording of on-site construction workers and safety officers, which requires them to timely count and inspect various prefabricated components, and judge the actual lifting and use of prefabricated components. This statistical method is no longer able to respond promptly to the interference encountered during the actual construction of prefabricated components, and the transportation and on-site adjustment of prefabricated components are slow. Therefore, intelligent information monitoring methods are needed to meet the demands of dynamic inventory management of materials in prefabricated construction sites.

To address the research gap, this study uses Deep Learning Based Computer Vision technology to identify the inventory quantity of two kinds of prefabricated components, prefabricated slabs and prefabricated stairs, on the construction site in real time. Combined with the original construction schedule, the transportation and inventory quantity of prefabricated components are adjusted in real time to strengthen the supply chain coordination ability and ensure the construction progress and cost-effectiveness of prefabricated building projects. Utilizing computer vision object detection and recognition technology, prefabricated components within the material yards of prefabricated construction sites are identified and counted, enabling real-time statistics on the number of prefabricated components

ISSN 2616-5767 Vol.8, Issue 3: 13-22, DOI: 10.25236/AJETS.2025.080303

on site. Special attention is given to abnormal fluctuations in inventory, providing a foundation for dynamic adjustments to construction schedules and procurement and transportation plans. By real-time analysis of building material inventory data, a building material inventory quantity information sharing system is constructed to improve the coordination system of material production, transportation, and construction. Based on this, a collaborative plan for material procurement, transportation scheduling, and on-site stacking and installation of prefabricated building projects is constructed. This plan needs to coordinate the interests and demands of multiple parties such as suppliers, carriers, and construction parties, and meet the cooperation requirements of each stage of project construction.

2. Related Work

2.1. Computer vision and object detection in construction

With the iterative breakthroughs of deep learning technology, object detection systems have formed two major technical schools based on the differences in network training paradigms: two-stage and one-stage. This study mainly introduces single-stage object detection technology.

The single-stage detection paradigm adopts an end-to-end detection framework to achieve efficiency breakthroughs. YOLOv1 [13] proposed by Redmon et al. in 2016 synchronously completes feature extraction, boundary regression, and category determination through a unified detection network, achieving an increase in detection speed compared to Faster R-CNN. The subsequent iteration versions YOLOv2 [14]/v3 [15] continuously enhance detection accuracy through technological innovations such as anchor optimization strategy and multi-scale prediction mechanism. The YOLOv4 proposed by Bochkovskiy et al. in 2020 [16] constructs a CSP-Darknet53 backbone network to enhance feature representation capabilities, using PANet structure instead of traditional FPN to achieve multi-level feature fusion, significantly improving small object detection performance. It should be noted that the SSD algorithm proposed by Liu et al. [17] creatively combines the efficiency of single-stage detection with the advantages of multi-scale feature extraction. Its improved R-SSD [18] further enhances detection robustness by expanding the feature pyramid dimension and strengthening cross layer feature association. Subsequently, algorithms based on SSD improvements continued to emerge, using DensNet's DSOD algorithm (Deeply Supervised Object Detectors, the first detection algorithm that does not use image classification pre training models for object detection training initialization) [19], as well as FSSD (Feature Fusion Single Shot Multi box Detector, FSSD, an improved feature fusion algorithm with lightweight features) combined with FPN algorithm [20].

With the rapid development of computer hardware and software, the efficiency of machine learning has been greatly improved, gradually attracting the attention of practitioners and researchers in the construction industry. As an important branch of machine learning, CNN has also been used to solve problems related to construction management, such as monitoring construction activities of construction workers [21], evaluating worker movements [22], monitoring construction work postures [23-24], identifying workers who are not wearing safety helmets [25], conducting structural inspections of buildings [26-28], detecting quality problems in pipeline structures [29], detecting quality problems in road infrastructure [30-31], and inspecting heavy construction machinery [32-34]. Liu et al. [35] proposed an improved YOLOv5 object detection algorithm using an improved convolutional block attention mechanism module to calculate the number of dense steel bars. They also developed a relative resolution object scale measurement method to measure the scale of objects in images of different resolutions; Yan et al. [36] used multi-target tracking technology to monitor the entry of material transport trucks and make judgments on possible delays in the arrival of building materials to ensure construction progress. It can be seen that computer vision and object detection models can effectively solve construction related problems.

According to this article, target detection and classification of different building materials are required, and the quantity of each component needs to be calculated. Therefore, this article proposes an improved model for discriminating the quantity of building materials based on the YOLOv8 object detection algorithm.

2.2. Inventory management of prefabricated components

In terms of inventory management of building materials, ensuring sufficient inventory quantities of various materials on site is a key factor in steadily advancing the construction schedule. Empirical research has shown that improper layout of building material storage space will significantly increase the

ISSN 2616-5767 Vol.8, Issue 3: 13-22, DOI: 10.25236/AJETS.2025.080303

complexity of lifting operations and the risk of construction process obstruction, and the resulting delay effect will directly weaken the industrial production efficiency of prefabricated construction [37]. In dynamic construction scenarios, the Alanjari research team [38] developed a yard layout optimization strategy based on schedule constrained material flow analysis, which maximizes on-site logistics timeliness by balancing material entry timing and consumption rate.

However, statistics indicate that research on supply chain management for prefabricated buildings can mainly be divided into four parts: prefabricated component production, storage and inventory, delivery and transportation, and overall supply chain performance. According to statistical analysis, the proportions of each part are 47%, 11%, 17%, and 25% respectively [39], with the least research on storage and inventory. The current academic exploration mainly focuses on the planning of building material storage locations and dynamic inventory control in the vicinity of production parks, while systematic research on real-time inventory monitoring and dynamic control mechanisms for prefabricated building materials on construction sites is still insufficient [39].

This study utilizes information-based computer vision object detection and recognition technology, focusing on the inventory management of prefabricated components at the construction site of prefabricated buildings. Through real-time monitoring of the quantity of various prefabricated components in the material yard, the construction schedule and material supply chain ordering, transportation, and entry plans are adjusted in a timely manner according to the actual construction situation, enabling the time and cost advantages of the prefabricated construction mode to be realized.

3. Identification model for prefabricated components

3.1. Data Collection and Processing

The data collected by this research institute are images of prefabricated panels and prefabricated stairs, from the construction sites of 74 construction projects in Zhejiang Province (mainly Hangzhou, Huzhou, Taizhou, Quzhou, Wenzhou, Ningbo, Shaoxing, Jinhua, and Lishui). Mainly using imaging systems based on Power over Ethernet (PoE) cameras and mobile phone shooting to capture on-site images. On average, each construction site has 3 PoE security cameras. The total number of cameras in the imaging system based on PoE cameras is greater than 200. The collected images of prefabricated components cover various scenarios, such as (1) different positions of prefabricated components (mainly based on different stages of construction); (2) Different weather conditions; (3) Different times; (4) Different perspectives, distances, and shadows, as shown in Fig. 1.

Fig. 1: Various scenes in the collected images.

This study collected a dataset of 3000 images of prefabricated panels and prefabricated stairs with a resolution of 1920*1080 (some images have watermarks that have been removed or cropped) from actual construction processes. Use an annotation tool called Labelme [40] to label the collected image dataset, where prefabricated panels are labeled as PCSL (Precast Slab) and prefabricated stairs are labeled as PCST (Precast Stair), as shown in Fig. 2. More than 70000 target objects were annotated. The entire dataset is divided into training set, validation set, and testing set in a ratio of 7:1:2.

ISSN 2616-5767 Vol.8, Issue 3: 13-22, DOI: 10.25236/AJETS.2025.080303

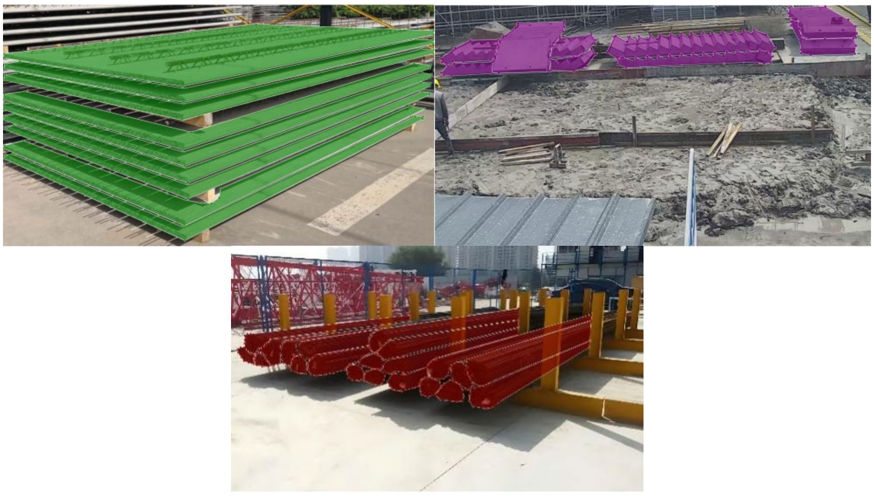

Fig. 2: Annotated schematic diagram.

3.2. Developing the prefabricated component recognition model

This study focuses on the recognition and counting of prefabricated components on construction sites. Considering the complex environment of the construction site, where production activities such as material stacking, worker construction, and construction machinery operation are intertwined, in order to further improve the accuracy and reliability of intelligent means, this study proposes an improved YOLOv8 object detection method based on a multi-head self-attention mechanism (MSM) to enhance the speed and accuracy of model object detection and counting.

Typically, a typical self attention mechanism consists of three matrix operations: Q, K, and V, which are essentially self operations. MSM has made certain improvements to the typical self attention mechanism, allowing each attention operation to extract effective feature information from multiple dimensions through grouping. The specific structure is shown in Fig. 3.

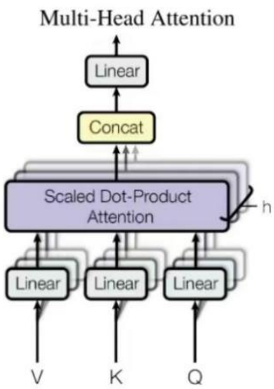

Fig. 3: Schematic diagram of multi-head self-attention mechanism.

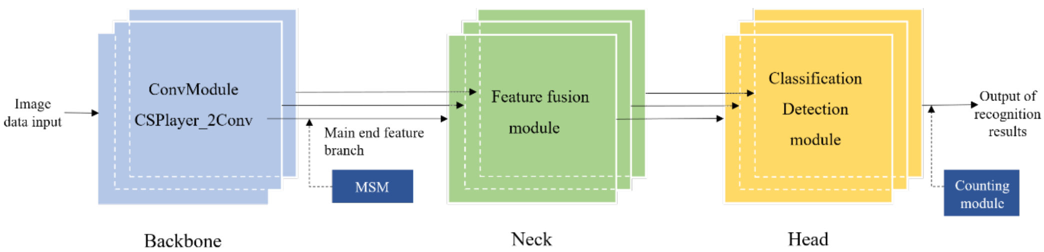

Fig. 4: MSM-YOLOv8 network structure diagram.

The schematic diagram of the improved YOLOv8 network structure is shown in Fig. 4. The black arrows represent the data flow during network operation, and different colors represent different network modules. For example, blue represents ConvModule and CSPlayer_2Conv, while green represents feature fusion module.

ISSN 2616-5767 Vol.8, Issue 3: 13-22, DOI: 10.25236/AJETS.2025.080303

The network structure mainly consists of three parts: backbone, neck, and head. Before inputting image data into Backbone, basic data preprocessing operations including data augmentation will be performed. The main function of Backbone is to extract feature information of the target area from the input image. When image data is input into Backbone, the target region features are extracted sequentially through convolution module, C2f module, and SPPF module. Then, the obtained features are further processed through MSM multi head self attention mechanism module to increase the feature weight of the target region and extract more effective feature information. The main function of the Neck section is to perform feature fusion. From the figure, it can be seen that there are three different scale network branches in the main input of the neck, including the main feature branch after feature enhancement using MSM. The three feature branches fused with Neck features are fed back into the Head section for classification and detection of target area features. The main output information here includes the position information and classification information of prefabricated components, that is, the position of the component in the picture and the category to which the component belongs. On this basis, a statistical quantity module is introduced to achieve recognition and counting functions, ensuring the output of quantity information for prefabricated components.

3.3. Model training and evaluation

The training process of the building material recognition model relies on the RTX4090 graphics card, with the specific software utilized and the versions of the virtual environments created omitted for brevity. The Epoch is set to 200, Batch Size to 8, Optimizer to Stochastic Gradient Descent (SGD), Learning Rate to 0.01, Weight Decay to 0.0005, and Momentum to 0.9.

The model evaluation method adopts the following formula:

$$Precision = \frac{TP}{TP + FP} \times 100\% \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \tag{2}$$

$$F1 - score = \frac{2 \times Recall \times Precision}{Recall + Precision} \tag{3}$$

$$AP = \int_0^1 (Precision \times Recall) dx \times 100\% \tag{4}$$

$$mAP = \frac{\sum_{i=1}^{C} AP_i}{C} \times 100\% \tag{5}$$

Among them, TP, FP, FN, and TN represent true positive, false positive, false negative, and true negative, respectively. AP (Average Precision) represents the area under the precision recall curve enclosed by the curve and coordinate axis, mAP (Mean Average Precision) represents the average precision across multiple categories, mAP50 represents the average precision at 50% IoU threshold, mAP50-95 represents the average mAP value within the range of 50-95% IoU threshold, and C represents the total number of categories. The larger the indicators, the better the model performance.

Table 1: Model operation results.

(a) Identification of inventory quantity of PCSL.

Model	Precision (%)	Recall(%)	F1-score	mAP50(%)	mAP50-95(%)	MB
YOLOv8	85.2	78.2	0.816	80.2	49.1	14.1
MSM-YOLOv8	89.6	83.5	0.864	83.1	50.7	15.0

(b) Identification of inventory quantity of PCST.

Model	Precision (%)	Recall(%)	F1-score	mAP50(%)	mAP50-95(%)	MB
YOLOv8	84.8	77.5	0.810	79.4	50.2	14.1
MSM-YOLOv8	87.5	83.8	0.856	82.4	53.7	15.0

According to the results in Table 1, overall, the MSM-YOLOv8 inventory quantity recognition model has shown some improvement in the recognition reliability of each category. Taking the identification of

ISSN 2616-5767 Vol.8, Issue 3: 13-22, DOI: 10.25236/AJETS.2025.080303

the number of PCSL as an example, compared with the original YOLOv8, the MSM-YOLOv8 model has improved accuracy, recall, F1 score, mAP50 and mAP50-95 by 4.4%, 5.3%, 0.048, 2.9%, and 1.6%, respectively. Due to the introduction of MSM, the size of the model has also increased by 0.9MB. Similarly, the recognition effect of the PST has also been improved to some extent, fully demonstrating that the introduction of MSM module helps to enhance the recognition ability of the model.

4. Inventory management model

4.1. Developing the inventory management model

Based on the prefabricated component recognition model established above, we can obtain the actual number of components on the construction site. Combined with inventory theory, considering that the supply of prefabricated components is divided into four stages: production, transportation, on-site storage, and hoisting construction, this article establishes a corresponding dynamic inventory management model for construction sites for prefabricated components.

Basic assumptions of the model:

(1) When transporting prefabricated components to the construction site, they are first stacked in the material area and then lifted for construction;

(2) Consider the time delay that may occur during the production phase after ordering prefabricated components;

(3) Consider the possible transportation delay of prefabricated components when transported to the construction site;

(4) Consider the amount of prefabricated components in transit during transportation.

The parameter settings are as follows:

(1) The inventory I_t for the t-th cycle and the arrival quantity A_t for that cycle can be monitored through a visual recognition model.

(2) Time delay L, transportation delay TD_{t-L}, component in transit IT_t, demand D_t, inventory adjustment IS_t, inventory adjustment coefficient α_s, and component in transit adjustment ITS_t are model related parameters that can be adjusted according to actual engineering needs.

(3) The order quantity Q_t and the order time S for the $(t+1)$-th cycle are the variables we need to find the optimal solution for.

In this model, the optimal goal of inventory management is to minimize the sum of the delay cost c_p^i caused by insufficient inventory and the inventory holding cost c_h^i within t ordering cycles.

$$I_t = I_{t-1} + A_t \tag{6}$$

$$A_t = TD_{t-L} \tag{7}$$

$$IT_t = IT_{t-1} + Q_{t-1} - A_t \tag{8}$$

$$IS_t = \alpha_s (D_t - I_t) \tag{9}$$

$$Q_t = D_t + IS_t + ITS_t \tag{10}$$

$$ITS_t = L \times D_t - IT_t \tag{11}$$

$$Q_t = D_t + \alpha_s (D_t - I_t) + (L \times D_t - IT_t) \tag{12}$$

$$C = min\left(\sum_{i=1}^{t} c_p^i t_{I=0} + c_h^i IL_t\right) \tag{13}$$

This is a dynamic inventory management model for prefabricated components that minimizes the total inventory cost. Due to the complexity and nonlinear constraints of the objective function, traditional

ISSN 2616-5767 Vol.8, Issue 3: 13-22, DOI: 10.25236/AJETS.2025.080303

optimization methods are difficult to effectively solve. In this paper, the particle swarm intelligence (PSO) algorithm is used to solve it.

4.2. Particle swarm optimization

According to the PSO principle, in a D-dimensional search space, there is a population $X = (X_1, X_2, \cdots, X_n)$ with n individuals. Each particle i in the group is represented by its position vector and velocity vector, namely vectors $X_i = [X_{i1}, X_{i2}, \cdots, X_{iD}]^T$, $V_i = [V_{i1}, V_{i2}, \cdots, V_{iD}]^T$. The memory of particles reflects their own experience, namely$P_i = [P_{i1}, P_{i2}, \cdots, P_{iD}]^T$, which represents the self or local best position (P_{best}) found by the particle; On the other hand, the vector $P_g = [P_{g1}, P_{g2}, \cdots, P_{gD}]^T$ represents the global best position (G_{best}) of the population, reflecting the population's experience.

Based on the objective function of prefabricated component inventory management mentioned above, the two variables that need to be solved are order quantity and order time, and a two-dimensional search space is constructed: the first dimensional vector X_{i1} represents Q_t (which needs to be an integer, using integer encoding), and the second dimensional vector X_{i2} represents S. Each particle represents the order quantity and order time situation within a certain period. In addition, since the objective function is to minimize the total inventory cost, the value of the objective function is the fitness value that measures the position of particles. The formula for updating the velocity and position of particles is as follows.

$$v_i(m+1) = \omega \cdot v_i(m) + c_1 \cdot r_1[Pbest_i(m) - x_i(m)] + c_2 \cdot r_2[Gbest_i(m) - x_i(m)] \tag{14}$$

$$x_i(m+1) = v_i(m) + x_i(m) \tag{15}$$

In order to enable the PSO to have high global search capability in the early stage of iteration and high local search capability in the later stage, the inertia weight ω and learning factor c adopt dynamic inertia weight and learning factor, as shown in the following equation.

$$\omega = \begin{cases} \omega_{min} - \dfrac{(\omega_{max} - \omega_{min}) \times (f_i - f_{min})}{f_a - f_{min}}, f_i \leq f_a \\ \omega_{max}, f_i \geq f_a \end{cases} \tag{16}$$

$$\begin{cases} c_1(m) = c_{1,s} + \left(c_{1,f} - c_{1,s}\right) \times \ln\left[\dfrac{(e-1) \times m}{M} + 1\right] \\ c_2(m) = c_{2,s} + \left(c_{2,f} - c_{2,s}\right) \times \ln\left[\dfrac{(e-1) \times m}{M} + 1\right] \end{cases} \tag{17}$$

In the formula, ω_{max} and ω_{min} are the maximum and minimum values of the inertia weight; f_{min} is the minimum fitness value in the current particle swarm; $c_1(m)$ and $c_2(m)$ are the individual learning factor and group learning factor of the algorithm in the mth iteration, respectively; $c_{1,s}$ and $c_{1,f}$ are the starting and ending values of individual learning factors; $c_{2,s}$ and $c_{2,f}$ are the starting and ending values of the group learning factor; m is the current iteration count; M is the maximum number of iterations; e is a natural constant.

By using formulas (14) and (15), we can iterate step by step and find the optimal solutions Q_t and S with the minimum value of C. The specific process will be demonstrated in the next section.

5. Case presentation

Apply the identification model and inventory management model to a prefabricated construction project in Zhejiang Province. Initially, within the project's material yard, the effectiveness of the visual model in monitoring prefabricated components is illustrated in Fig. 5.

ISSN 2616-5767 Vol.8, Issue 3: 13-22, DOI: 10.25236/AJETS.2025.080303

Fig. 5: Visual model recognition effect.

In actual prefabricated construction projects, the building material identification and counting model proposed in this study can identify PCSL and PCST. The identification results of each prefabricated component are automatically marked with a bounding box, and the predicted probability can be indicated in the upper left corner of the bounding box. Similarly, the material quantity monitoring model has a counting module that can automatically count the total quantity of various materials in the image and automatically label them in the lower right corner of the image.

According to the ordering plan of the project, a single standard floor construction requires 23 PCSLs and 6 PCSTs, with an ordering cycle of 4 days. The inventory holding cost of each prefabricated component is 200 CNY per day, and the cost of project delay is 10000 CNY per day. Assuming that the lifting construction is delayed for 24 hours due to weather conditions, the total inventory cost generated is 15800 CNY. Based on this, the identification model and inventory management model are used to optimize and adjust the order quantity and order cycle.

Set the standard using PSO: particle number is 30; Iteration times N=200; ω_{max}=0.95, ω_{min}=0.5; The values of $c_{1,s}$ and $c_{2,s}$ are 2.5 and 0.5; The values of $c_{2,s}$ and $c_{2,f}$ are 0.6 and 2.4, respectively. Perform optimization and solution.

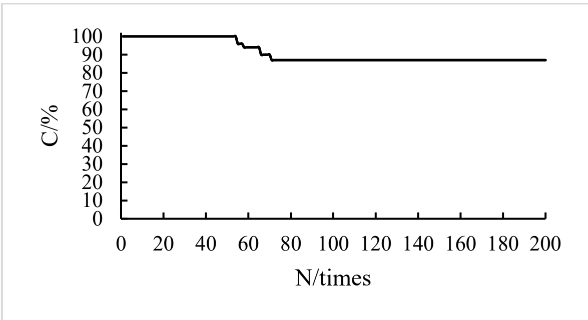

Fig. 6: PSO operation results.

Repeat the program 200 times to obtain the iterative convergence curve of the objective function (as shown in Fig. 6). It can be seen that the PSO established in this article gradually reduces the optimal solution obtained in each round of 200 iterations until the optimal solution is obtained. At the 71st iteration, the particle swarm algorithm found the optimal solution, which resulted in the lowest total inventory cost of 87% of the original plan, a reduction of 13%. At this point, the corresponding next cycle order quantity is 20 prefabricated panels and 5 prefabricated stairs, and the order time needs to be postponed by 16.3 hours. During the iteration process, the algorithm converges quickly and can quickly find the approximate optimal value, avoiding the randomness of the results and providing a new method for solving practical engineering case related problems.

ISSN 2616-5767 Vol.8, Issue 3: 13-22, DOI: 10.25236/AJETS.2025.080303

6. Conclusion

In the context of engineering intelligence and informatization, this article starts from the inventory management of prefabricated components and combines information technologies such as computer vision and PSO algorithm to construct an automatic and efficient dynamic inventory management model. The main contributions of this article are as follows: Firstly, image data collection was mainly carried out for different engineering projects in various regions of Zhejiang Province, and a dataset containing 3000 prefabricated components was obtained and processed; In addition, by improving the YOLOv8 model, a more accurate identification model for prefabricated components was obtained; Finally, by combining the identified data with PSO, automatic identification and dynamic inventory optimization of prefabricated components in prefabricated buildings were achieved.

However, this article also has some shortcomings. Due to limitations in engineering data collection, the proposed model is mainly aimed at prefabricated building projects in Zhejiang Province, which mostly use prefabricated shear wall structures, with only slabs and stairs as prefabricated components, and lack attention to other types of building materials. Future work will expand the dataset and focus on dynamic inventory management of more types of building materials, thereby further improving the practicality of the model.

References

[1] Zhai P, Wang J, Zhang L, Extracting worker unsafe behaviors from construction images using image captioning with deep learning–based attention mechanism. Journal Of Construction Engineering And Management. 2023:04022164.
[2] Yu Y, Li H, Cao J, Luo X, Three-dimensional working pose estimation in industrial scenarios with monocular camera. IEEE international conference on computer vision. 2021, Things 8 (3):1740–1748.
[3]Chen C, Xiao B, Zhang Y, Zhu Z, Automatic vision-based calculation of excavator earthmoving productivity using zero-shot learning activity recognition. Automation in Construction. 2023, 146:104702.
[4] Liu H, Wang D, Xu K, Zhou P, Zhou D, Lightweight convolutional neural network for counting densely piled steel bars. Automation in Construction.2023, 146:104692.
[5] Yan X, Zhang H, Gao H, Mutually coupled detection and tracking of trucks for monitoring construction material arrival delays. Automation in Construction. 2022,142:104491.
[6] Yan X, Zhang H, Computer vision-based disruption management for prefabricated building construction schedule. Journal Of Computing In Civil Engineering. 2021,35(6):04021027.
[7] Yan X, Zhang H, Zhang W, Intelligent monitoring and evaluation for the prefabricated construction schedule. Computer-aided Civil And Infrastructure Engineering. 2022, 38 (3):1–17.
[8] Lu L, Dai F, Automated visual surveying of vehicle heights to help measure the risk of over height collisions using deep learning and view geometry. Computer-aided Civil And Infrastructure Engineering. 2022, 38 (2):194–210.
[9] Zhang E, Shao L, Wang Y, Unifying transformer and convolution for dam crack detection. Automation in Construction. 2023, 147 :104712.
[10] Qiu Q, Lau D, Real-time detection of cracks in tiled sidewalks using YOLO-based method applied to unmanned aerial vehicle (UAV) images. Automation in Construction. 2023, 147:104745.
[11] Yang X, Guo R, Li H, Comparison of multimodal RGB-thermal fusion techniques for exterior wall multi-defect detection. Journal of Infrastructure Intelligence and Resilience. 2023:100029.
[12] Zhang W, Zhang H, Yu L. Collaborative Planning for Stacking and Installation of Prefabricated Building Components Regarding Crane-Collision Avoidance. Journal of Construction Engineering and Management. 2023,149(6).
[13] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection. Proceedings of the IEEE conference on computer vision and pattern recognition. 2016:779-788.
[14] Redmon J, Farhadi A. YOLO9000: better, faster, stronger. Proceedings of the IEEE conference on computer vision and pattern recognition.2017:7263-7271.
[15] Redmon J, Farhadi A. Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767,2018.
[16] Bochkovskiy A, Wang C Y, Liao H Y M. Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934,2020.
[17] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector. European conference on computer vision.Springer,Cham,2016:21-37.
[18] Jeong J, Park H, Kwak N. Enhancement of SSD by concatenating feature maps for object detection. arXiv preprint arXiv:1705.09587,2017.

ISSN 2616-5767 Vol.8, Issue 3: 13-22, DOI: 10.25236/AJETS.2025.080303

[19] Shen Z, Liu Z, Li J, et al. Dsod: Learning deeply supervised object detectors from scratch. Proceedings of the IEEE international conference on computer vision.2017:1919-1927.
[20] Li Z, Zhou F. FSSD: feature fusion single shot multibox detector. arXiv preprint arXiv:1712.00960,2017.
[21] Luo, X., Li, H., Yang, X., et al. Capturing and Understanding Workers' Activities in Far Field Surveillance Videos with Deep Action Recognition and Bayesian Nonparametric Learning. Computer-Aided Civil and Infrastructure Engineering, 2019, 34 (4): 333-351.
[22] Yan X, Zhang H, Zhang W, Intelligent monitoring and evaluation for the prefabricated construction schedule. Computer-aided Civil And Infrastructure Engineering. 2022, 38 (3):1–17.
[23] Yan, X. Z., Li, H., Li, A. R., et al. Wearable IMU-Based Real-Time Motion Warning System for Construction Workers' Musculoskeletal Disorders Prevention . Automation in Construction. 2017, 74: 2-11.
[24] Zhang, H., Yan, X. Z. & Li, H. Ergonomic Posture Recognition Using 3D View-Invariant Features from Single Ordinary Camera. Automation in Construction. 2018, 94: 1-10.
[25] Fang, Q., Li, H., Luo, X., et al. Detecting Non-Hardhat-Use by a Deep Learning Method from Far-Field Surveillance Videos. Automation in Construction. 2018, 85: 1-9.
[26] Cha, Y.-J., Choi, W. & Büyüköztürk, O. Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks. Computer-Aided Civil and Infrastructure Engineering. 2017, 32 (5): 361-378.
[27] Li, S., Zhao, X. & Zhou, G. Automatic Pixel-Level Multiple Damage Detection of Concrete Structure Using Fully Convolutional Network. Computer-Aided Civil and Infrastructure Engineering. 2019, 34 (7): 616-634.
[28] Yang, X., Li, H., Yu, Y., et al. Automatic Pixel-Pevel Crack Detection and Measurement Using Fully Convolutional Network . Computer-Aided Civil and Infrastructure Engineering. 2018, 33 (12): 1090-1109.
[29] Xue, Y. & Li, Y. A Fast Detection Method Via Region-Based Fully Convolutional Neural Networks for Shield Tunnel Lining Defects . Computer-Aided Civil and Infrastructure Engineering. 2018, 33 (8): 638-654
[30] Bang, S., Park, S., Kim, H., et al. Encoder-Decoder Network for Pixel-Level Road Crack Detection in Black-Box Images . Computer-Aided Civil and Infrastructure Engineering. 2019, 34 (8): 713-727.
[31] Maeda, H., Sekimoto, Y., Seto, T., et al. Road Damage Detection and Classification Using Deep Neural Networks with Smartphone Images . Computer-Aided Civil and Infrastructure Engineering. 2018, 33 (12): 1127-1141.
[32] Fang, Q., Li, H., Luo, X., et al. Detecting Non-Hardhat-Use by a Deep Learning Method from Far-Field Surveillance Videos. Automation in Construction. 2018, 85: 1-9.
[33] Kim, H., Ham, Y., Kim, W., et al. Vision-Based Nonintrusive Context Documentation for Earthmoving Productivity Simulation. Automation in Construction. 2019, 102: 135-147.
[34] Kim, H., Kim, H., Hong, Y. W., et al. Detecting Construction Equipment Using a Region Based Fully Convolutional Network and Transfer Learning. Journal of Computing in Civil Engineering. 2018, 32 (2): 04017082.
[35] Kolar Z, Chen H, Luo X. Transfer learning and deep convolutional neural networks for safety guardrail detection in 2D images. Automation in Construction. 2018, 89:58-70.
[36] H. Liu, D. Wang, K. Xu, P. Zhou, D. Zhou, Lightweight convolutional neural network for counting densely piled steel bars. Automation in Construction. 2023, 146:104692.
[37] X. Yan, H. Zhang, H. Gao, Mutually coupled detection and tracking of trucks for monitoring construction material arrival delays. Automation in Construction. 2022, 142:104491.
[38] Li H, Love P. Site-Level Facilities Layout Using Genetic Algorithms. Journal of Computing in Civil Engineering. 1998, 12(4): 227-31.
[39] Alanjari P, Razavialavi S, Abourizk S M. Hybrid Genetic Algorithm-Simulation Optimization Method for Proactively Planning Layout of Material Yard Laydown. Journal of Construction Engineering & Management. 2015, 141(10): 06015001.
[40] K. Wada, Labelme. https://github.com/wkentaro/labelme, 2016.