WARSAW UNIVERSITY OF TECHNOLOGY Index 351733
FACULTY OF CIVIL ENGINEERING
COMMITTEE FOR CIVIL AND WATER ENGINEERING
POLISH ACADEMY OF SCIENCES ISSN 1230-2945

DOI: 10.24425/ace.2025.154119

ARCHIVES OF CIVIL ENGINEERING

Vol. LXXI ISSUE 2 2025

pp. 243–256

Research paper

Intelligent building construction site safety inspection model based on YOLOX

Hairong Huang[1], Lian Yuan[2], Huiji Wang[3], Haoran Yuan[4]

Abstract: The construction industry is a high-risk and high accident rate industry, and it is crucial to conduct safety inspections on construction sites. Therefore, the study introduces an improved YOLOX algorithm and performs lightweight processing such as replacing the backbone network and pruning channels. At the same time, the optimized YOLOX algorithm will be applied to the construction of a model for safety detection in intelligent building construction sites. Results showed that the improved model proposed in the study had the best inference speed and average accuracy, with an average accuracy of 95.01%. In the experimental analysis under different detection categories, the model proposed in the study had the highest detection accuracy for whether to wear a safety helmet, with an accuracy rate of 96.39%, which was 10.05% higher than the YOLOX model. At the same time, the accuracy of the model in detecting whether to wear welding masks, masks, and welding gloves was as high as 92.37%, 94.49%, and 94.61%, respectively. In addition, the recall rate of the model proposed by the research institute in helmet wearing detection was as high as 95.48%. The improved model proposed by the research institute has performed well in safety inspection of construction sites, not only possessing high-speed processing capabilities but also high-precision detection performance, providing reliable technical support for real-time monitoring and early warning of intelligent building construction.

Keywords: YOLOX, architecture, construction, security testing, MobileNetv3, attention module

[1]MSc., School of Architecture and Engineering, Zhejiang Tongji Vocational College of Science and Technology, Hangzhou, 310014, China, e-mail: z20120100803@zjtongji.edu.cn, ORCID: 0009-0007-4344-5623

[2]MSc., Engineering Department, Tengda Construction Group Co., Ltd., Taizhou, 318050, China, e-mail: yuanlian0313@sina.com, ORCID: 0009-0005-1421-0345

[3]PhD., School of Humanities and Public Administration, Jiangxi Agricultural University, Nanchang, 330045, China, e-mail: wanghuiji@pku.edu.cn, ORCID: 0009-0007-8468-222X

[4]UGS., College of Engineering and Architecture, Wenzhou University, Wenzhou, 325035, China, e-mail: henny1106@sina.com, ORCID: 0009-0009-3417-2146

1. Introduction

The safety inspection of intelligent building construction sites has always been an important and challenging issue. As is well known, construction site workers have poor safety awareness, high mobility, and low overall quality [1]. In terms of construction safety, with the development of new technologies, it has become possible to automatically identify hidden dangers on construction sites. There are currently two main methods, one is recognition based on wearable technology. This method mainly adopts functions such as positioning, monitoring, alarm, and reminder. But this method requires construction workers to wear additional equipment on their own, which is somewhat cumbersome for construction workers themselves. The second method is based on computer vision technology, which does not have the aforementioned limitations. As computer vision and deep learning technology develop, intelligent security detection models using image recognition have become a new solution [2]. Traditional intelligent security detection models based on rules or features overly rely on rule and feature extraction, and may be affected by factors such as lighting, angle, occlusion, etc., resulting in a decrease in detection accuracy [3, 4]. In response to this type of problem, an improved YOLOX algorithm is introduced to design and implement an intelligent building construction site safety detection model to improve the safety and work efficiency of the construction site. The contribution of the research is the introduction of an improved YOLOX algorithm that has undergone lightweight processing. This model helps to accurately identify and locate various safety hazards in construction sites, and is expected to provide an efficient and accurate safety detection solution for the construction industry.

The research content includes four parts. The first part provides a review of the YOLOX algorithm and construction safety detection. The second part introduces the construction of an intelligent building construction site safety detection model using the improved YOLOX algorithm. The third part conducts experimental analysis on the application of the algorithm and model proposed in the research. The fourth part summarizes and discusses the experimental results, and proposes future prospects.

2. Related works

YOLOX is an accurate and real-time object detection model, with improved algorithms widely used in computer vision. SONG et al. proposed the SEYOLOX-tiny model to accurately and robustly detect the male ears of corn in the field. This network model optimized the YOLOX model by embedding attention mechanisms. The results indicated that the model fully met the accuracy and robustness requirements of the visual system for detecting corn tassels [5]. Guo et al. proposed the YOLOX-SAR algorithm to address the scattering characteristics. The experimental results showed an excellent performance and explanatory power in SAR image target detection, with mAP reaching 89.56% and maintaining real-time detection speed [6]. Zhang et al. found that there were omissions and errors in the detection of prohibited items, so they introduced a detection framework based on the state-of-the-art YOLOX object detection network. The results showed that the average accuracy of this method on the public safety SIXray dataset was improved by 5.0% compared to the benchmark YOLOX-S model [7].

Construction safety inspection is significant for the safety of people's lives and property, and improving the quality of engineering. De Melo and Costa proposed a new method that utilizes elastic engineering concepts and drone system technology. The results showed that this method helped to improve the safety conditions of construction sites [8]. Li et al. proposed a phased inspection plan to explore the safety inspection of drones on construction sites. The results showed that drones had significant potential advantages in construction site safety inspections [9]. Akinolu et al. used a two-step literature selection method and bibliometric review to compile relevant publications from the Scopus database on safety issues related to construction sites. Analysis showed that the emerging trends in research on building health and safety technologies were concentrated in areas such as project health [10].

In summary, many researchers have conducted extensive research on the safety detection of construction sites, but have not applied advanced object detection technologies. Therefore, an improved YOLOX model is introduced in the study to achieve more efficient and accurate object detection in construction site safety inspection.

3. Intelligent building construction site safety detection model using YOLOX object detection algorithm

For the safety inspection of intelligent building construction sites, this chapter first introduces the construction process of a safety inspection model, and application investigation of safety hazards. Subsequently, for the selection of target detection algorithms in the third stage, an improved YOLOX algorithm is introduced, and the backbone network is replaced and channel pruning is performed to achieve the lightweight of the model.

3.1. Construction of a safety inspection model for intelligent building construction sites

The construction of an intelligent building construction site safety detection model is mainly divided into four stages, including safety hazard data collection, dataset production, model training and evaluation, and application investigation of safety hazards. In the stage of collecting safety hazard data, it covers three main steps, including creating a hazard list, arranging equipment, and collecting data [11]. The production stage of the dataset mainly consists of two stages: data annotation and construction. In the annotation process, research mainly uses manual methods to annotate the collected image data for security risks, and the quality of annotation will directly affect the recognition effect of subsequent models. In the process of constructing a dataset, the research focuses on the production of datasets for hidden dangers listed in the subsystem areas of each engineering system. The framework diagram of the data collection and production stage is shown in Fig. 1.

The third stage is the training and evaluation of safety detection models for intelligent building construction sites. The core tasks of this stage include selecting a suitable deep learning framework, building a model training environment, and selecting suitable object detection

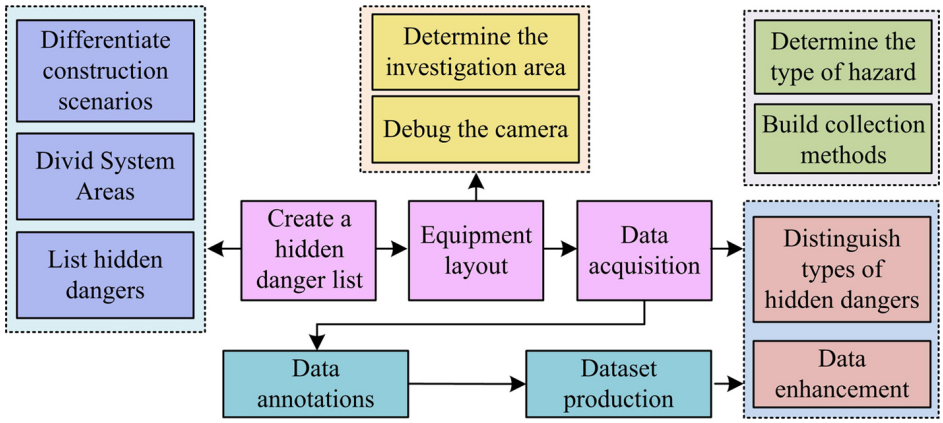

Fig. 1. Frame diagram of the data acquisition and production stage

algorithms. This stage belongs to the core link of the model and is also a prerequisite for achieving safety detection on construction sites. The fourth stage is the application investigation of safety hazards, which mainly includes three stages: debugging of equipment and operating environment, intelligent security detection model for video input, and intelligent detection of safety hazards. When safety hazards are discovered through the intelligent building construction site safety inspection model at the construction site, the model will immediately issue a reminder to the on-site workers, urging them to rectify in a timely manner. At the same time, a facial recognition system will be launched to identify and record the identities of workers at the construction site. Finally, the testing results will be included in the monthly safety assessment and safety education will be provided to relevant personnel. The framework diagram of the safety detection model for intelligent building construction sites, as well as the application and investigation of safety hazards, is shown in Fig. 2.

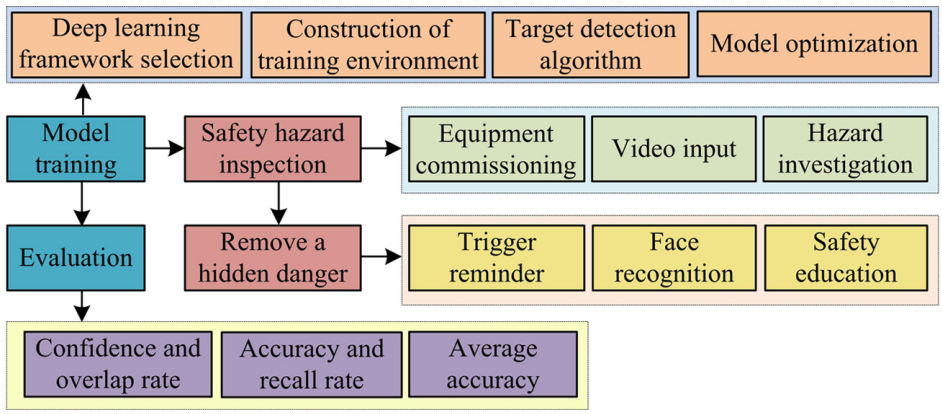

Fig. 2. Framework diagram for the training and evaluation phase and application troubleshooting phase

In the safety detection model of intelligent building construction sites, the performance and layout method of collection equipment have a significant impact on the quality of collected image data. As the basic equipment for image data collection, the performance and usage of cameras are particularly important in construction sites. Due to the complex and dusty environment of construction sites, the quality of images captured by cameras will directly affect the accuracy of the model in identifying safety hazards. Therefore, to ensure the accuracy and stability of safety inspections on construction sites, objective evaluation of camera image quality is studied to test its performance and various indicators. The specific evaluation content includes indicators such as camera resolution, color reproduction, dynamic range, distortion correction, and clarity, to ensure that the camera can collect high-quality image data in the construction site and provide a reliable basis for safety hazard inspection. When collecting data for training, research mainly places cameras in multiple positions to collect data from different angles, and simultaneously analyzes the occlusion situation at different positions. In addition, for the production of datasets, research is conducted on capturing image data from construction sites and capturing monitoring images of construction sites to obtain the required safety hazard image data.

3.2. Intelligent building construction site safety detection based on improved YOLOX model

In the training and evaluation of safety detection models for intelligent building construction sites, there are many difficulties in safety detection, including the complex construction environment and the presence of obstacles. In response to this type of problem, an improved YOLOX model is introduced, which combines the Bilateral Attention Module (BiCAM) and Residual Feature Pyramid Network (Res-FPN) to enhance its performance in safety detection tasks at construction sites. The traditional Convolutional Block Attention Module (CBAM), after learning, can obtain attention weights for channels and spaces. Although this method can obtain contextual information of local regions, it is too focused on feature information, resulting in the loss of edge information of densely detected targets [12]. To address this issue, BiCAM is introduced in the study, and the schematic diagram of this module is shown Fig. 3.

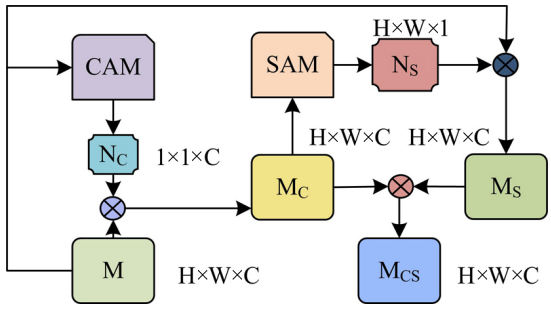

Fig. 3. BiCAM module schematic diagram

In BiCAM, the input feature map needs to be first fed into the Channel Attention Module (CAM) for processing, in order to obtain CAM weights. Subsequently, the input features are multiplied by this weight to obtain a feature map containing channel weights, which are calculated as shown in Equation (3.1).

$$M_C = M \otimes N_C \tag{3.1}$$

In Equation (3.1), M represents the input feature and N_C represents the weight of CAM. Subsequently, the study feeds the feature maps containing channel weights into the Spatial Attention Module (SAM) for processing to obtain the weights of the SAM. Simultaneously the original input feature map is multiplied with this weight to achieve the focus of the target position, and the calculation is shown in Equation (3.2).

$$M_S = M \otimes N_S \tag{3.2}$$

In Equation (3.2), N_S represents the weight of SAM. Next, the study will add and operate feature maps containing channels and spatial weights to further integrate features from both spatial and channel dimensions. Finally, the sigmoid function is used to calculate the matrix based on channel space attention weights, as shown in Equation (3.3).

$$M_{CS} = \sigma(M_C \oplus N_S) \tag{3.3}$$

In Equation (3.3), $\sigma(\cdot)$ represents the sigmoid function. The weight information represents the contribution of channels or spaces in the feature map to the final output, which helps the model obtain richer semantic information in specific regions, thereby improving detection performance. Additionally, in the feature extraction module, to overcome the network degradation and gradient vanishing, the research draws on the idea of residual connection and proposed the Res-FPN structure [13]. The improved YOLOX model has improved the recognition accuracy of dense and occluded targets to some extent, but it has problems such as large scale, too many parameters, and insufficient inference speed. Therefore, a series of lightweight optimization measures are proposed in the study, resulting in the YOLOX-M3 model. Firstly, the study replaces the backbone network of the YOLOX model with MobileNetv3, which is a lightweight neural network that can perform fast inference on mobile and embedded devices without the support of GPU [14]. In this network, the computational cost of the convolutional kernel on a certain fully convolutional layer is calculated as shown in Equation (3.4).

$$\mathrm{Cos}_{\mathrm{full}} = D_F \times D_F \times A \times B \tag{3.4}$$

In Equation (3.4), $\mathrm{Cos}_{\mathrm{full}}$ represents the computational cost of the convolutional kernel on the entire convolutional layer. $D_F \times D_F$ represents the size of the input feature. A and B represents the number of input and output channels, respectively. The calculation cost of point by point convolution is shown in Equation (3.5).

$$\mathrm{Cos}_{1\times1} = D_F \times D_F \times A \times B \tag{3.5}$$

In Equation (3.5), $\mathrm{Cos}_{1\times1}$ represents the computational cost of point by point convolution. The cost of deep separable convolution calculation is shown in Equation (3.6).

$$\mathrm{Cos}_{\mathrm{all}} = D_F \times D_F \times A \times D_K \times D_K + D_F \times D_F \times A \times B \tag{3.6}$$

In Equation (3.6), $\mathrm{Cos}_{\mathrm{all}}$ represents the computational cost of depthwise separable convolution. $D_K \times D_K$ represents the size of the convolution kernel. The ratio of the calculation cost of deep separable convolution to standard convolution is calculated in Equation (3.7).

$$\frac{\mathrm{Cos}_{\mathrm{all}}}{\mathrm{Cos}_{\mathrm{full}}} = \frac{1}{D_K^2} + \frac{1}{B} \tag{3.7}$$

To achieve compression and acceleration of the model, a channel pruning strategy is further introduced in the study. At the same time, approve the data and use the scaling factor of the batch standardization layer as the pruning parameter [15]. The specific steps are to first set up a dataset $\{X_1, X_2, \ldots, X_m\}$, and the mean calculation of this dataset is shown in Equation (3.8).

$$\mu_m = \frac{\sum_{i=1}^{m} X_m}{m} \tag{3.8}$$

The variance calculation of this data is shown in Equation (3.9).

$$\sigma_m^2 = \frac{\sum_{i=1}^{m} (X_i - \mu_m)^2}{m} \tag{3.9}$$

Subsequently, the mean and variance are normalized, and the calculation is shown in Equation (3.10).

$$\hat{X}_i = \frac{X_i - \mu_m}{\sqrt{\sigma_m^2 + \varepsilon}} \tag{3.10}$$

In Equation (3.10), ε is a parameter close to 0, with the aim of avoiding a denominator of 0. According to $\hat{X}_i$, the output result can be obtained, and its calculation is shown in Equation (3.11).

$$Y_i = \beta \hat{X}_i + \gamma \tag{3.11}$$

In Equation (3.11), β and γ respectively represent the translation factor and scaling factor. Next, γ will be applied to each channel and sorted by size, compared to the preset threshold. Crop the corresponding channels below the threshold and perform repeated fine-tuning training. The calculation of training results is shown in Equation (3.12).

$$L = \sum_{(X,Y)} l(f(X, W), Y) + \lambda \sum_{\gamma \in \Gamma} g(\gamma) \tag{3.12}$$

In Equation (3.12), $l(f(X,W),Y)$ represents the original loss function of the network, $g(\gamma)$ represents the penalty term of γ, λ represents the sparse factor, and Γ represents the set of scaling factors. The study removes the connection between the front and back of the channel through channel pruning, and selects the channel through scaling factors. At the same time, the regularization term of the scaling factor is added to the original loss function for joint training to optimize the network. Assuming the channel pruning threshold is set to 0.5, 50% of the channels will be pruned and the complete model structure will be ensured, without affecting the dimensionality matching of the backbone network. Fig. 4 shows the schematic diagram of channel pruning.

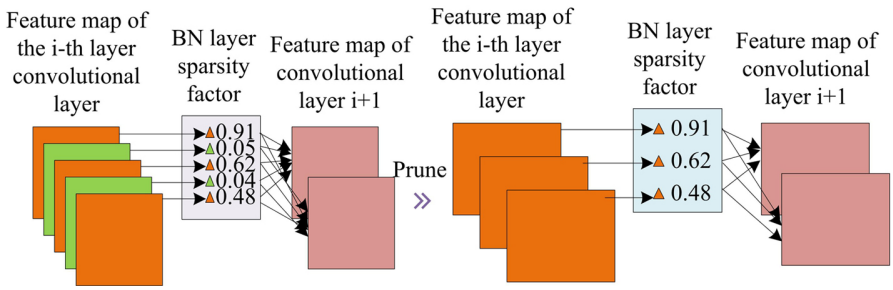

Fig. 4. Schematic diagram of channel pruning

4. Safety inspection model analysis for intelligent building construction site based on YOLOX

This chapter first analyzed the performance of the YOLOX-M3 model. Subsequently, the practical application effect of the YOLOX-M3 based intelligent building construction site safety detection model was verified, and the detection performance of different models in four safety hazards of not wearing welding masks, masks, welding gloves, and safety helmets correctly was compared.

4.1. Performance analysis of YOLOX-M3 model

The study first verified the lightweight effect of different pruning rates on the YOLOX-M3 model, setting pruning rates of 30%, 50%, and 70%, respectively. Among them, a pruning rate of 30% indicates that 30% of the original model parameters were retained during the pruning process, and the remaining 70% of the parameters were pruned. The dataset used in the experiment is the open-source dataset Safety Helmet Wearing Dataset (SHWD) for helmet wearing detection. The images in SHWD are mainly searched through web crawlers and networks, covering different scenes, visual ranges, lighting, personal poses, and occlusion situations. There are a total of 7581 images, including 9044 positive samples with helmets and 111514 negative samples without helmets. When training the YOLOX-M3 model, research pruning the model parameters based on the set pruning rate. Then, use the pruned model to train on the training set to adjust the pruned model parameters to adapt to the pruned structure.

Finally, use the test set to evaluate the detection performance and lightweighting effect of the model under different pruning rates, in order to select the optimal pruning rate. At the same time, inference speed and Average Precision (AP) were selected as performance evaluation indicators. To ensure the results effectiveness, 10 tests were conducted and the average value was taken as the final result. The performance evaluation results under different pruning rates are shown in Fig. 5. The inference speed of the model increased as pruning rate increased. When the pruning rate is 70%, the average inference speed of the YOLOX-M3 model reaches 89 FPS. At the same time, it can be seen that the AP value of the model decreases with the increase of pruning rate. Among them, when the pruning rate is 30%, the average AP value of the YOLOX-M3 model is as high as 96.13%. Taking into account various indicators, the study ultimately selected 50% as the pruning rate of the mode.

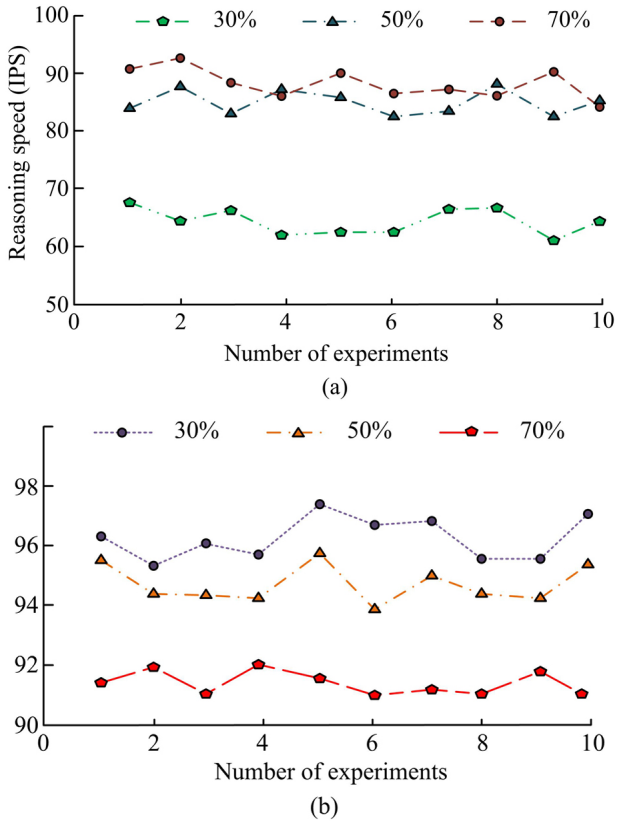

Fig. 5. Performance evaluation results at different pruning rates; (a) Reasoning speed under different pruning rates, (b) AP values under different pruning rates

The study continued to validate the lightweight effect of YOLOX-M3, and two mainstream models, YOLOv4 and YOLOv5s, were selected for comparative analysis. The performance indicators of different models are shown in Fig. 6. The YOLOX-M3 model had significantly

better indicators than other models. Among them, the YOLOX-M3 model had a reasoning speed of up to 83 FPS, which was improved by 35 FPS and 23 FPS compared to the other two models, respectively. Meanwhile, the AP value of the YOLOX-M3 model was as high as 95.01%, while the AP values of the YOLOv4 and YOLOv5s models were only 88.14% and 92.39%. The YOLOX-M3 model maintained high detection accuracy while also having faster inference speed.

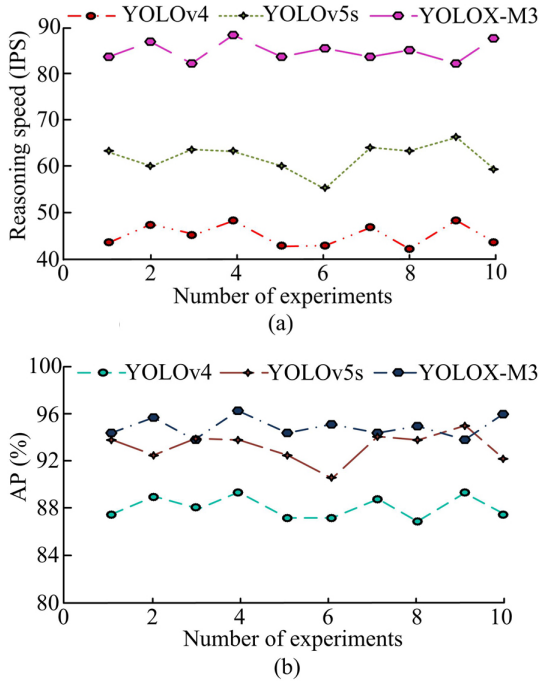

Fig. 6. Results of the performance indicators for the different models; (a) Reasoning speed of different models, (b) AP values for different models

4.2. Application analysis of safety inspection model for intelligent building construction site

To verify the practical application effect of the safety detection model for intelligent building construction sites, this study took electric welding operations in building construction as an example to detect four safety hazards of construction workers who did not wear welding masks, masks, welding gloves, and safety helmets correctly. The intelligent identification of hidden dangers consists of two stages. The first stage is the training and testing of the intelligent identification model for hidden dangers. The second stage is to use the model trained in the first stage to conduct intelligent inspection experiments on hidden dangers in steel structure engineering scenarios. The collected data comes from smartphones, online channels, and on-site cameras. After completing data collection, the study divided the experimental data

into 5618 training sets, 701 testing sets, and 724 validation sets. The data collection for the study belongs to descriptive statistics, which considers the diversity of data sources, real-time characteristics, diversity conditions, as well as the partitioning and utilization of the dataset.

Table 1. Experimental environment and parameter setting

Entry name	Parameter
CPU	AMD Ryzen 7 5800X
Memory	FURY32G
Video card	GeForce RTX 2080Ti
Solid State Disk	1T
Operating system	windows10
Programming language	python3.7
Training Platform	PyTorch

The study first validated the performance advantages of the YOLOX-M3 model under different Intersection over Union (IoU) conditions. Faster Region based Convolutional Networks (Faster R-CNN) model, Single Shot MultiBox Detector (SSD) model, and YOLOX model were selected for performance comparison. The evaluation indicators include accuracy and recall. The test samples are 701 test sets from the welding operation. The results of various indicators under different categories are shown in Fig. 7. From Fig. 7(a), it can be seen that the YOLOX-M3 model has significantly better security detection accuracy for various categories than the YOLOX model, Faster R-CNN model, and SSD model. Among them, the YOLOX-M3 model has the highest detection accuracy for whether to wear a safety helmet, with an accuracy rate of 96.39%. Compared with the YOLOX model, R-CNN model, and Faster SSD model, it has improved by 10.05%, 18.95%, and 14.58%, respectively. Meanwhile, the YOLOX-M3 model has detection accuracies of 92.37%, 94.49%, and 94.61% for whether to wear welding masks, masks, and welding gloves, respectively. As shown in Fig. 7(b), the recall rate of the YOLOX-M3 model in helmet wearing detection is as high as 95.48%. The YOLOX-M3 model has better security detection performance than other models in various detection categories, and it has significant detection performance advantages.

To ensure the objectivity of the experimental results, the study further collected safety inspection data from other construction sites online, including 50 sets of test data. This includes inspection data for the welding area and assembly area. Among them, the welding area is specifically used for welding operations, usually including welding equipment, workbenches, and related materials. Construction personnel conducting welding operations here need to wear welding masks, masks, welding gloves, and safety helmets to protect themselves from sparks and smoke. The assembly area is a place for assembling building materials and components, where construction personnel carry out assembly and installation work. In this area, construction workers also need to wear safety helmets and gloves to protect themselves from accidental injuries. The study used on-site safety inspection and video surveillance security inspection and research method for comparison. The study summarizes the response speed measured by each method and takes the average value for comparison. The experimental results are

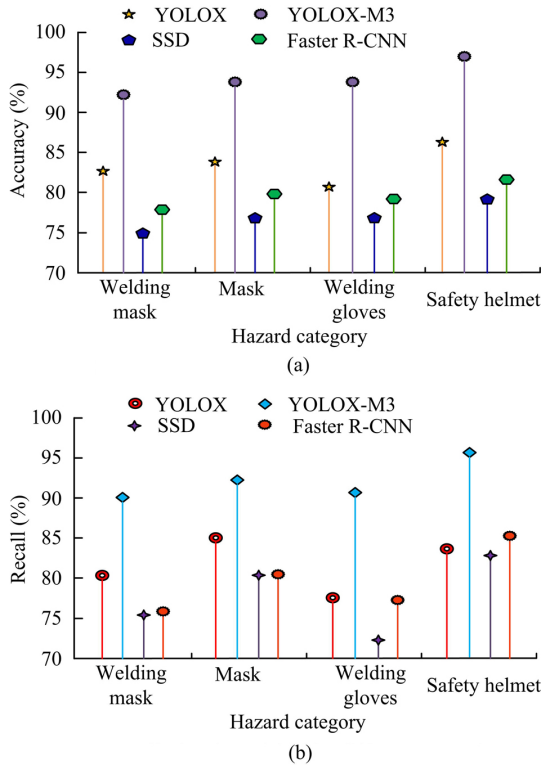

Fig. 7. Results of various indicators under different categories; (a) Accuracy of each model under different categories, (b) Results of various indicators under different categories

shown in Table 2. From Table 2, it can be seen that the YOLOX-M3 model has a significantly faster response speed in construction safety inspection compared to on-site safety inspections and video surveillance inspections. Among them, the average inspection time for the welding area during on-site safety inspections is as high as 29.68 minutes, which is an increase of 20.76 minutes compared to the YOLOX-M3 model. Meanwhile, in the safety inspection of the assembly area, the average detection time of the YOLOX-M3 model is only 9.25 minutes. The YOLOX-M3 model has significant advantages.

Table 2. Average response times for the different methods

Inspection methods	**Mean response time/minute**	
	Welding area	**Assembly area**
On site safety inspection	29.68	31.84
Video monitoring	19.82	22.58
Model inspection	8.92	9.25

5. Conclusions

The safety inspection of intelligent building construction sites is significant for improving construction safety, achieving informationization and modernization of building construction, and other aspects. The research mainly builds a safety detection model for intelligent building construction sites based on the improved YOLOX-M3 algorithm. Results showed that the inference speed of the model increases with the increase of pruning rate. When the pruning rate is 70%, the average inference speed of the YOLOX-M3 model reaches 89FPS. When the pruning rate is 30%, the average AP value of the YOLOX-M3 model is as high as 96.13%. The inference speed of the model increased as pruning rate increased, but the AP value decreased as pruning rate increased. The study ultimately selected 50% as the pruning rate of the model, with a reasoning speed and AP value of 85FPS and 94.86%, respectively. In addition, the average inspection time for the welding area during on-site safety inspections is as high as 29.68 minutes, which is an increase of 20.76 minutes compared to the YOLOX-M3 model. Meanwhile, in the safety inspection of the assembly area, the average detection time of the YOLOX-M3 model is only 9.25 minutes. The intelligent building construction site safety detection model based on YOLOX-M3 algorithm performed well in terms of inference speed, target detection accuracy, and safety evaluation accuracy. However, in the lightweight processing of YOLOX algorithm, pruning technology may slightly reduce the detection accuracy of the model. Subsequent research can consider introducing optimization processes such as knowledge distillation and quantization to further improve the inference speed of the model while maintaining high detection accuracy.

References

[1] I. Jeelani, K. Han, and A. Albert, "Development of virtual reality and stereo-panoramic environments for construction safety training", *Engineering, Construction and Architectural Management*, vol. 27, no. 8, pp. 1853–1876, 2020, doi: 10.1108/ECAM-07-2019-0391.

[2] M.M. Hossain and S. Ahmed, "Developing an automated safety checking system using BIM: A case study in the Bangladeshi construction industry", *International Journal of Construction Management*, vol. 22, no. 7, pp. 1206–1224, 2022, doi: 10.1080/15623599.2019.1686833.

[3] M. Parsamehr, U.S. Perera, T.C. Dodanwala, P. Perera, and R. Ruparathna, "A review of construction management challenges and BIM-based solutions: perspectives from the schedule, cost, quality, and safety management", *Asian Journal of Civil Engineering*, vol. 24, no. 1, pp. 353–389, 2023, doi: 10.1007/s42107-022-00501-4.

[4] Y. Min, J. Guo, and K. Yang, "Research on real-time detection algorithm of rail-surface defects based on improved YOLOX", *Journal of Applied Science and Engineering*, vol. 26, no. 6, pp. 799–810, 2023, doi: 10.6180/jase.202306_26(6).0006.

[5] C. Song, F. Zhang, J.S. Li, J.Y. Xie, C. Yang, H. Zhou, and J. X. Zhang, "Detection of maize tassels for UAV remote sensing image with an improved YOLOX model", *Journal of Integrative Agriculture*, vol. 22, no. 6, pp. 1671–1683, 2023, doi: 10.1016/j.jia.2022.09.021.

[6] Q. Guo, J. Liu, and M. Kaliuzhnyi, "YOLOX-SAR: High-precision object detection system based on visible and infrared sensors for SAR remote sensing", *IEEE Sensors Journal*, vol. 22, no. 17, pp. 17243–17253, 2022, doi: 10.1109/JSEN.2022.3186889.

[7] Y. Zhang, W. Xu, S. Yang, Y. Xu, and X. Yu, "Improved YOLOX detection algorithm for contraband in X-ray images", *Applied Optics*, vol. 61, no. 21, pp. 6297–6310, 2022, doi: 10.1364/AO.461627.

[8] R.R.S. de Melo and D.B. Costa, "Integrating resilience engineering and UAS technology into construction safety planning and control", *Engineering, Construction and Architectural Management*, vol. 26, no. 11, pp. 2705–2722, 2019, doi: 10.1108/ECAM-12-2018-0541.

[9] J. Li, G. Zhou, D. Li, M. Zhang, and X. Zhao, "Recognizing workers' construction activities on a reinforcement processing area through the position relationship of objects detected by faster R-CNN", *Engineering, Construction and Architectural Management*, vol. 30, no. 4, pp. 1657–1678, 2023, doi: 10.1108/ECAM-04-2021-0312.

[10] M. Akinlolu, T.C. Haupt, D.J. Edwards, and F. Simpeh, "A bibliometric review of the status and emerging research trends in construction safety management technologies", *International Journal of Construction Management*, vol. 22, no. 14, pp. 2699–2711, 2022, doi: 10.1080/15623599.2020.1819584.

[11] D. Kim, J. Kong, J. Lim, and B. Sho, "A study on data collection and object detection using faster R-CNN for application to construction site safety", *Journal of the Korean Society of Hazard Mitigation*, vol. 20, no. 1, pp. 119–126, 2020, doi: 10.9798/KOSHAM.2020.20.1.119.

[12] P. Lin, P.C. Wei, Q.X. Fan, and W.Q. Chen, "CNN model for mining safety hazard data from a construction site", *Journal of Tsinghua University (Science and Technology)*, vol. 59, no. 8, pp. 628–634, 2019, doi: 10.16511/j.cnki.qhdxxb.2019.26.008.

[13] C.X. Zhu, J.C. Qian, and B.R. Wang, "YOLOX on embedded device with CCTV&TensorRT for intelligent multicategories garbage identification and classification", *IEEE Sensors Journal*, vol. 22, no. 16, pp. 16522–16532, 2022, doi: 10.1109/JSEN.2022.3181794.

[14] T. Umar, "Applications of drones for safety inspection in the Gulf Cooperation Council construction", *Engineering, Construction and Architectural Management*, vol. 28, no. 9, pp. 2337–2360, 2021, doi: 10.1108/ECAM-05-2020-0369.

[15] J. Bęc, "Influence of anchoring and bracing system on dynamic characteristics of façade scaffolding", *Archives of Civil Engineering*, vol. 69, no. 4, pp. 493–506, 2023, doi: 10.24425/ace.2023.147672

Received: 2024-04-09, Revised: 2024-06-19