Journal of Network Intelligence ©2026 ISSN 2414-8105
Taiwan Ubiquitous Information Volume **11**, Number **1**, February **2026**

Supervision of High Rise Building Construction Quality and Safety Status Based on Deep Learning

Jian-Cong Wu[1,*]

[1]Xiamen Haicang District Construction and Traffic Engineering Quality and Safety Station,
Xiamen 361026, P.R. China
cnwjcei@163.com

Xiang-Zhen Zhou[2]

[2]Faculty Information Science and Technology,
National University of Malaysia, Selangor 43600, Malaysia
619543699@qq.com

*Corresponding author: Jian-Cong Wu

Received October 7, 2024, revised February 22, 2025, accepted May 2, 2025.

Abstract. *This article proposes a method rooted in deep learning for monitoring the construction quality and safety status of high-rise buildings, aiming to solve the problems of quality control and safety management in complex construction environments. With the expansion of China's construction industry, quality problems and safety accidents occur frequently during the construction process, and traditional supervision methods are no longer sufficient to address the requirements of modern high-rise building construction. To this end, this article combines image recognition and deep learning techniques to build an automated supervision system to achieve monitoring and analyzing the construction process in real time. The system is based on convolutional neural networks and transfer learning to recognize personnel behavior and construction quality at the construction site. The system is capable of detecting whether construction personnel are operating in violation of regulations, while also identifying quality issues during the construction process, such as structural defects and non-standard processes. By automatically analyzing and processing on-site image data, the system effectively improves the accuracy of construction quality supervision and the efficiency of safety management. The research results indicate that construction supervision systems based on deep learning can significantly reduce risks caused by quality issues and unsafe behaviors, promote efficient management and quality improvement of construction projects.*

Keywords: deep learning; convolutional neural network; construction of buildings; quality and safety.

1. **Introduction.** Amid the swift expansion of China's economy and the ongoing progress of urbanization, The construction sector, being a foundational industry, is showing a trend of gradually expanding in scale. As an important component of urban development, high-rise buildings have complex construction processes and require extremely high quality control and safety guarantees. However, with the expansion of construction scale and the increase of construction personnel, safety issues at construction sites have become increasingly prominent, and production safety accidents have occurred frequently. According to statistics, most safety accidents during the construction process are caused by unsafe behavior of personnel, including failure to wear safety protective equipment, illegal operations, and entering hazardous areas. These behaviors not only pose a threat to the

construction personnel themselves, but also have adverse effects on the project's progress and quality [1, 2]. How to effectively identify and reduce unsafe behaviors during the construction process has become a significant concern which urgently needs to be addressed in the safety management of the construction industry.

At the same time, the control of building quality is also an important part of construction management. The development of tall structures involves multiple processes and complex technical links, and even a slight carelessness may lead to quality problems such as structural defects and non-standard construction. These quality issues not only affect the safety and service life of buildings, but may also bring huge economic losses. Therefore, ensuring the quality of the construction process is equally important as ensuring construction safety. However, traditional construction quality and safety supervision methods mainly rely on manual inspection and on-site patrols, which are inefficient and prone to problems such as missed inspections and misjudgments, making it challenging to fulfill the requirements of modern construction of tall buildings.

In this context, the evolution of AI technology, especially deep learning technology, provides new solutions for quality and safety management in the construction industry's building process. As an important branch of artificial intelligence, deep learning has demonstrated powerful capabilities in fields such as image recognition and object detection by learning large amounts of data and automatically extracting features. Combining image recognition technology, deep learning can provide real-time monitoring and analysis of personnel behavior and construction quality on construction sites, effectively improving the automation and intelligence level of construction supervision.

Recently, deep learning-based image recognition technology has been widely researched and applied in the construction industry. For example, deep learning network frameworks such as convolutional neural networks have been used for the recognition and monitoring of personnel safety protection equipment. These technologies can capture real-time images of construction sites through cameras and use deep learning algorithms to automatically detect whether construction workers are equipped with protective gear like helmets and seat belts, and promptly detect and correct unsafe behavior. In addition, object detection technology can also be used to identify other hazardous factors at the construction site, such as personnel entering hazardous areas. Through real-time tracking and analysis, the risk of accidents caused by violations during the construction process can be reduced [3, 4].

Deep learning can also play an important role in construction quality supervision. By analyzing the image data of the construction site, deep learning technology can automatically detect structural defects and identify quality issues such as non-standard operations during the construction process. This image recognition based quality supervision method not only reduces the subjectivity and uncertainty of manual detection, but also efficiently and accurately identifies problems in complex construction environments.

This article is based on deep learning technology, combined with text mining and image recognition, to supervise quality issues and safety hazards during the construction of tall buildings, with a focus on identifying and managing unsafe behaviors of construction personnel. Firstly, this article utilizes text mining techniques to extract feature words related to unsafe behavior from production safety accident reports in the construction industry, providing a theoretical basis for subsequent image recognition. Then, based on network frameworks such as convolutional neural networks, the actions of workers at the construction site are detected and classified to ensure that the construction process complies with safety regulations. At the same time, real-time identification and tracking of personnel in hazardous areas on construction sites can further enhance the comprehensiveness and accuracy of construction safety supervision.

Through the research in this article, unsafe behaviors on construction sites can be efficiently and accurately identified and managed, greatly reducing production safety accidents caused by human factors. In addition, the construction quality supervision method based on image recognition technology also provides strong technical support for quality control during the construction process. The findings of this article not only provide new ideas for the quality and safety supervision of high-rise building construction, but also lay the foundation for intelligent management within the construction sector.

1.1. **Related work.** Chi et al. [5] analyzed 9358 accidents that occurred within the U.S. construction industry from 2002 to 2011, identified key risk factors that have a significant impact between behavioral and situational factors, and concluded that the primary cause of construction accidents is the interaction between unsafe worker behavior and hazardous site conditions. Han et al. [6] found that unsafe behavior is the primary cause of construction accidents. To mitigate these accidents, the Importance Hazard Analysis (IDA) technique is utilized to evaluate the potential risks posed by unsafe behavior among workers in the construction industry. Research has found that unlike IPA, IDA identifies unsafe behaviors that may pose potential risks. Fang et al. [7] concentrated on the unsafe behavior of construction workers, drawing on cognitive and social psychology theories as well as existing accident causality models, adopted a five stage form to analyze and study the cognitive process of construction workers when they face potential hazards on construction sites, and developed the CM-CWUB system for monitoring. Asilian et al. [8] constructed a conceptual model to analyze the influencing factors by conducting field observations, in-depth interviews, and work group discussions on 113 workers in various large-scale projects. The research results indicate that the conceptual model consists of four categories - general management, organizational elements, safety supervision management, and personal characteristics. General management and organizational culture are regarded as important prerequisites and contributing factors that lead to human errors and unsafe behaviors among construction workers. Guo et al. [9] used case studies to explore the temporal patterns of unsafe behavior among workers, which can be divided into two aspects: (1) cross temporal allocation characteristics of different occupations, and (2) association rules. The results indicate that the time distribution of unsafe behaviors is a rectangular distribution, suggesting that workers' unsafe behaviors have explosive and memorable characteristics. Zhang et al. [10] used psychological methods, based on the five personality models and planned behavior theory, to hypothesize the relationship between risk propensity analysis and the intention of construction workers' unsafe behavior, and studied how personality traits affect unsafe behavior among construction workers. Han and Lee [11] found that approximately 80% to 90% of incidents in the construction sector are related to hazardous worker behavior [11]. Due to the difficulty of real-time observation of worker behavior on construction sites, a framework for visual detection of hazardous behavior has been proposed for behavior monitoring. Lee et al. [12] identified pedestrian behavior recognition as a key factor in early accident prevention for ADAS. Leveraging the strong recognition capabilities of CNNs, a CNN-based detection system was developed for real-time recognition of vehicle driver behavior. Furthermore, by integrating CNN with a random forest classifier, the CNN output is randomly connected to the random forest instead of being fully connected. Experimental results demonstrate that this method has been effectively implemented in the PUB dataset captured by night infrared cameras. with higher behavior recognition accuracy than algorithms that only use CNN correlation, and faster processing time. Yu et al. [13] found that existing unsafe behavior management methods, such as Behavior Based Safety (BBS), mainly rely on manual observation and recording, which not only consumes significant time and expense, but also cannot cover the entire construction site and all workers. To address this issue, a parameterized method based on image skeleton is proposed for identifying unsafe behaviors of construction workers. Ding et al. [14] found that computer vision-based pattern recognition techniques for identifying unsafe behavior at construction sites rely on manually computing intricate image features, requiring intricate parameters for feature extraction [14]. Considering the difficulties of temporary features, lighting, and the complexity of viewpoint arrays required for identifying unsafe operations, a new hybrid deep learning model has been developed, which integrates CNN and LSTM networks to automatically identify unsafe operations by workers. Yang et al. [15] applied supervised learning methods to assess the relative influence of various cognitive factors on safe behavior within the framework of the TRA. Data was gathered through a survey questionnaire based on TRA from 80 workers involved in a tunnel construction project, along with observational data from BBS. Using TRA cognitive factors as input variables, a prediction model for unsafe behavior was created using six commonly employed machine learning algorithms and logistic regression. Fang et al. [16] considered that structural supports (such as concrete and steel) provide stability to engineering structures by transferring loads, and during the construction process of engineering structures, workers often tend to use shortcuts to perform daily activities and save time by reducing the construction period [16]. Therefore, when workers engage in such unsafe behavior, the likelihood of injury or even suicide greatly increases. To address this issue, a computer vision method has been developed using a mask-based convolutional neural network (R-CNN) for detecting workers crossing structural supports during project construction. Ran et al. [17] found that vision-based techniques struggle to detect unsafe behavior from data obtained in digital images and videos. To solve this problem, a novel deep learning method was proposed, utilizing a spatiotemporal attention-based network to filter out redundant information in videos, thereby automatically determining the identity of workers. The results show that this method can accurately recognize the identity of workers from videos captured at construction sites. Fang et al. [18] found that advances in deep learning and computer vision methods may provide managers and engineers with more information-based security management technologies. However, in practical applications, the use of deep learning and computer vision is limited due to a series of technical (such as accuracy and reliability) and management challenges. These challenges are the result of the dynamism and complexity of architecture, as well as a series of difficulties related to obtaining video surveillance data. To address these challenges, a computer vision based construction safety framework has been designed and developed by integrating a range of digital technologies with data fusion techniques. Shahverdy et al. [19] found that current driver behavior monitoring systems have been widely used as intelligent transportation systems (ITS) to reduce the risk of traffic accidents. Most of the methods for monitoring driver behavior in the past relied on computer vision technology, which has the potential to invade privacy and cause false positives. Therefore, a novel and effective deep learning method was proposed, utilizing the advantages of deep neural networks to construct a two-dimensional convolutional neural network (CNN) based on repeated drawing techniques. The experimental results confirm that this method can effectively detect unsafe behaviors of drivers.

Based on the above literature, scholars and experts have conducted relatively complete research on the detection and analysis of construction quality and unsafe behaviors, and have achieved a large number of results, providing this article with a lot of basic research and theoretical support. However, existing research still has the following two issues.

(1) Most scholars conduct research on the mechanisms of unsafe behavior, or on issues such as classification and control, while there is a lack of extensive in-depth research on the identification of unsafe behavior in the field of construction. In the era of informatization, most inspections and classifications that rely on human labor have been replaced by more information-based technologies, and the construction process is dynamic, making it difficult for supervisors to monitor unsafe behaviors of construction workers at all times. Therefore, using image information can comprehensively capture the operational behavior of construction workers for recognition. With the rise of image recognition technology, neural network algorithms have once again demonstrated their advantages in today's big data era. Therefore, this article. Using deep learning theory and convolutional neural network technology to detect and detect unsafe behaviors among construction workers.
(2) The use of convolutional neural networks in the recognition of unsafe behaviors involves the modification of algorithms and the study of accuracy. For the recognition and detection of safety protective equipment, most scholars have only studied the recognition of safety helmets unilaterally, and there is a limited amount of research on the in-depth analysis of unsafe practices in construction personnel and the integration with safety warning technology.

1.2. **Contribution.** The primary contribution of this article is to propose and implement a method for monitoring the construction quality and safety of high-rise buildings relying on deep learning technology, which is specifically evident in the following areas:

(1) **Recognizing hazardous actions of construction workers using deep learning techniques:** This paper proposes a deep learning based object detection algorithm, including network frameworks such as convolutional neural networks, to automatically identify and track unsafe actions of construction workers. By analyzing the images of the construction site, the system can accurately identify whether the construction personnel have engaged in illegal operations and other behaviors, significantly improving the intelligence and automation level of construction site safety management.
(2) **Dual functions of integrated construction quality supervision and safety management:** The contribution of this article is not limited to the supervision of safety status, but also applies deep learning technology to construction quality supervision. Through image recognition technology, the system is able to detect potential quality issues that may arise during the construction process, such as structural defects, non-standard processes, etc., thereby ensuring a balance between construction quality and safety, and improving the comprehensive supervision capability of the construction site.
(3) **Real time dynamic monitoring and analysis:** This article realizes the real-time dynamic monitoring function of the construction site, especially the real-time recognition and tracking of personnel behavior in dangerous areas. By using deep learning algorithms to process real-time image data, the system can instantly detect potential safety hazards and provide timely feedback, thereby reducing the risk of accidents caused by unsafe behavior during the construction process.
(4) **Application of deep learning in complex construction environments:** This article successfully overcomes the limitations of traditional supervision methods in terms of high labor costs, low efficiency, and insufficient accuracy through the application of deep learning technology. The system can adapt to complex construction

environments, automatically identify and classify personnel behavior and construction processes on site, effectively reducing errors in manual supervision and improving the accuracy of construction quality and safety management.

(5) **Provide a new path for intelligent management in the construction industry:** By introducing deep learning into construction quality and safety supervision, this article provides a new technological path for intelligent management in the construction industry. Automated supervision not only improves management efficiency, but also provides more reliable data support for safety management personnel, promoting the transformation of construction from traditional management mode to digitalization and intelligence.

2. **Theoretical analysis.**

2.1. **Principles of convolutional neural networks.** A convolutional neural network is a form of deep learning model similar to the multi-layer deep perceptron of artificial neural networks, and can perform representation learning. The CNN is capable of performing translation-invariant classification on input data due to its hierarchical structure, which is why it is referred to as a translation-invariant artificial neural network [20]. Figure 1 displays a schematic representation of a CNN, comprising an input layer, hidden layers, and an output layer.

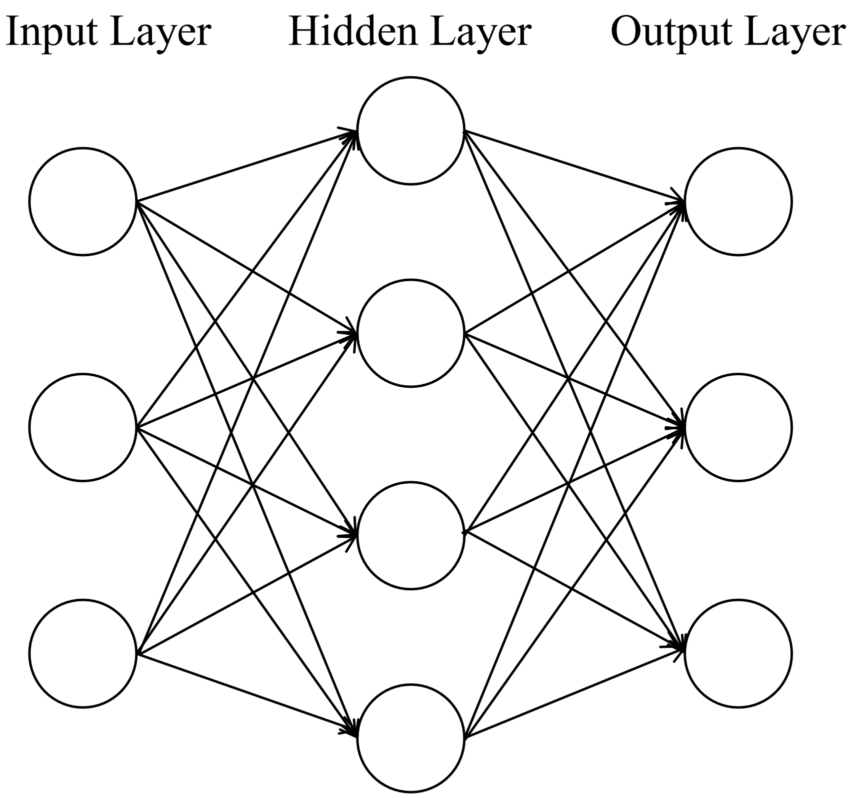

Figure 1. CNN schematic diagram

The activation values of each output layer are obtained through the forward propagation formula.

$$net_j = \sum_{i=0}^{m} V_{ji} X_i \tag{1}$$

$$H_j = f(net_j) \tag{2}$$

where, net_j represents the input; V_{ji} represents the weight; X_i is the output of the i-th input layer neuron; H_j is the output of the j-th neuron in the hidden layer.

$$net_k = \sum_{j=0}^{m} W_{kj} X_j \tag{3}$$

$$O_k = f(net_k) \tag{4}$$

where, net_k represents the input of the output layer neurons; W_{kj} represents the weight of the output layer neurons; O_k represents the output of the output layer neurons.

Backpropagation:

$$E = \frac{1}{2} \sum_{k=1}^{l} (Y_k - O_k)^2 \tag{5}$$

where, E represents the cost function; Y_k is the expected output of the k-th neuron in the output layer.

$$\delta_k^O = -(Y_k - O_k) \cdot f'(net_k) \tag{6}$$

$$\delta_j^H = \sum_{k=1}^{l} \delta_k^O \cdot W_{kj} f'(net_j) \tag{7}$$

where, δ_k^O represents the residual; δ_j^H represents the residual.

$$W_{kj}^1 = -\eta \frac{\partial E}{\partial W_{kj}} = \eta \delta_k^O H_j \tag{8}$$

$$V_{ji}^1 = -\eta \frac{\partial E}{\partial V_{ji}} = \eta \delta_j^H X_i \tag{9}$$

where, W_{kj}^1 is the weight increment; V_{ji}^1 is the weight increment; η is the step size for gradient descent.

$$W_{kj} = W_{kj} + W_{kj}^1 \tag{10}$$

$$V_{ji} = V_{ji} + V_{ji}^1 \tag{11}$$

In the forward propagation process, the values are calculated for each node in the network. The combination of backpropagation and dynamic programming avoids redundant operations, greatly improving the speed of gradient calculation and making it more conducive to random gradient descent to adjust neural network parameters and complete the training process.

2.2. **Principles of transfer learning.** The dataset of construction workers' operational behavior is limited, and relying solely on a limited amount of data for learning will result in poor model accuracy, overfitting, and poor generalization performance. Therefore, this article adopts transfer learning theory to transfer the parameters trained on a large dataset to a new model, enabling the new model to achieve faster training speed, higher accuracy, and accelerate model convergence [21, 22]. Figure 2 illustrates the transfer learning process.

Using the ImageNet dataset and transfer learning to train ResNet50, transfer learning can enable the model to achieve higher starting points, slopes, and asymptotic performance. On this basis, use the homework behavior dataset for training and adjust parameters.

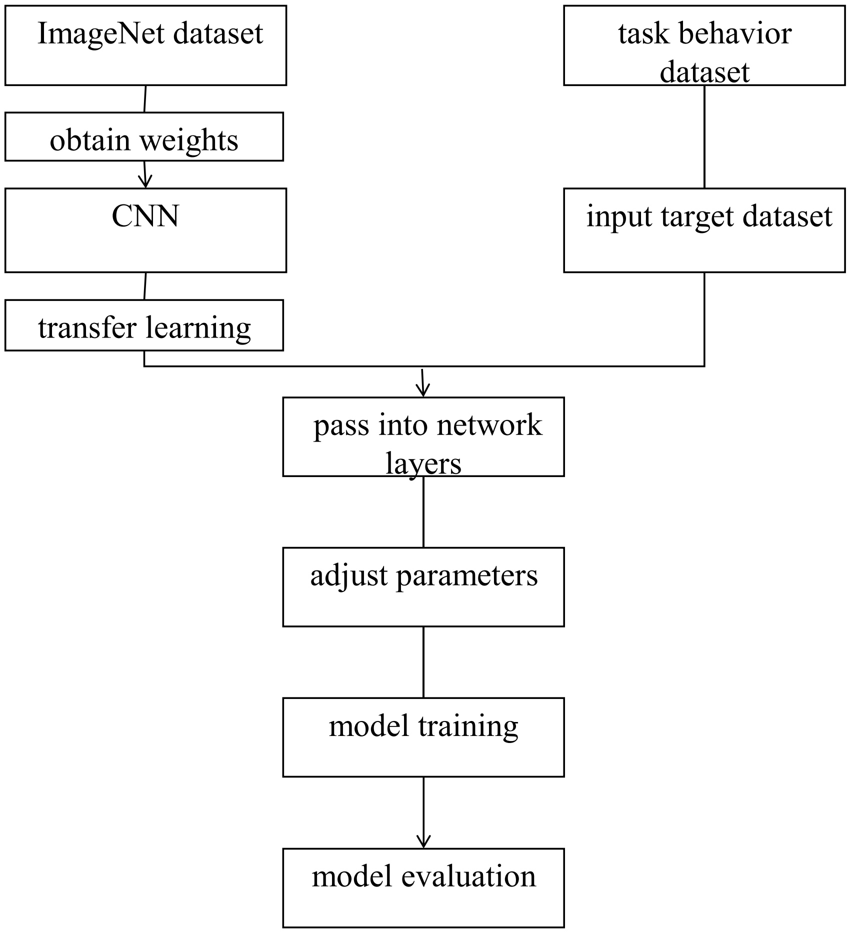

Figure 2. Transfer learning process

3. **High rise building construction quality and safety status supervision model based on deep learning.** High rise construction sites have a large number of personnel and diverse job types, making construction quality and safety issues particularly important. This article takes tower crane operators as an example to explore the operational behavior of construction personnel and the influencing factors of construction quality. Tower cranes are the most commonly used lifting equipment in construction sites. As of the end of 2017, the number of tower cranes in China was 394,000 to 427,000, with a median of 411,000. In 2018 and 2019, the domestic sales volume of tower cranes was around 15,000 and 28,000 respectively. By late 2019, the estimated number of tower cranes in China was around 440,000. The widespread use of tower crane equipment has effectively accelerated engineering efficiency, but it also comes with the risk of production safety accidents and construction quality issues.

Statistical analysis of tower crane accidents indicates that the primary direct cause is improper operation, with 21 accidents occurring. These factors not only affect the safety of construction, but also directly affect the quality of construction. The improper operation of tower crane drivers may not only cause safety accidents, but also lead to a decrease in project quality, such as component damage and inaccurate installation caused by improper lifting.

Therefore, in response to accidents and construction quality issues caused by unsafe behavior of tower crane drivers, this article identifies and analyzes the operational behavior of tower crane drivers, which is of great significance for reducing accidents and improving construction quality.

This article uses ResNet50 network to classify and predict the behavior of tower crane drivers. Compared to traditional deep learning networks, ResNet networks effectively reduce the number of parameters and address the issues of gradient vanishing and network degradation through the introduction of residual connections. While reducing the number of parameters, ResNet50 deepens the network layers, thereby improving the classification and prediction performance.

In traditional deep learning networks, the forward and backward propagation processes of the model often result in the gradual vanishing or explosion of gradients, especially when the network layers are deep, the optimization process is difficult to converge. Given a deep network with input x and output y after several layers of mapping, its expression is:

$$y = f(x, W) = W_n \sigma(W_{n-1} \sigma(\ldots \sigma(W_1 x))) \tag{12}$$

where x is the input data, $W_1, W_2, \ldots, W_n$ represents the weight of each layer, and σ represents the activation function.

In the process of backpropagation, the chain rule calculates the gradient of each layer:

$$\frac{\partial L}{\partial W_i} = \frac{\partial L}{\partial W_n} \cdot \frac{\partial W_n}{\partial W_{n-1}} \cdots \frac{\partial W_{i+1}}{\partial W_i} \tag{13}$$

when n is large, gradient vanishing or exploding problems are prone to occur, making it difficult to train the network.

The core of ResNet is the introduction of residual blocks, and its basic formula is:

$$y = F(x, \{W_i\}) + x \tag{14}$$

where x is the input, $F(x, \{W_i\})$ is a nonlinear transformation through several convolutional layers and activation functions, and y is the output.

By introducing skip connections, the output not only relies on the result of convolutional transformation, but also incorporates the input x itself, thereby alleviating the problem of gradient vanishing.

The specific residual block usually includes two convolutional layers and their corresponding batch normalization and ReLU activation functions, as follows:

$$F(x, \{W_i\}) = W_2 \sigma(BN(W_1 \sigma(BN(x)))) \tag{15}$$

where W_1 and W_2 denote the weights of the convolutional layer, σ is the ReLU activation function, and BN is the batch normalization operation.

Therefore, the final output of the residual block is:

$$y = W_2 \sigma(BN(W_1 \sigma(BN(x)))) + x \tag{16}$$

ResNet50 consists of 50 layers of network, mainly composed of multiple stacked residual blocks. Assuming the input to the l-th layer is $x^{(l)}$, the output after passing through the residual block is:

$$x^{(l+1)} = F(x^{(l)}, W^{(l)}) + x^{(l)} \tag{17}$$

where $W^{(l)}$ is the convolution kernel parameter of the l-th layer. The final network output obtained by stacking multiple layers of residual blocks is:

$$y = F(x^{(L)}, W^{(L)}) \tag{18}$$

where L represents the total number of layers in the network, and F is the classification or prediction function of the last layer.

In the task of classifying and predicting the behavior of tower crane drivers, the commonly used loss function is cross entropy loss, and its formula is:

$$L = -\sum_{i=1}^{N}[y_i \log(p_i) + (1 - y_i)\log(1 - p_i)] \tag{19}$$

where N denotes the number of samples, y_i is the actual label of the i-th sample, and p_i is the model's predicted probability for the i-th sample.

By minimizing the loss function, the network can adjust the weight parameter W, thereby improving classification accuracy. During the optimization process, stochastic gradient descent or its variants (such as Adam optimizer) are usually used to adjust the network's weight parameters. The formula for gradient update is:

$$W^{(t+1)} = W^{(t)} - \eta\frac{\partial L}{\partial W} \tag{20}$$

where $W^{(t)}$ is the weight at the t-th iteration, η is the learning rate, and $\frac{\partial L}{\partial W}$ is the gradient of the loss function relative to the weight.

By using the ResNet50 network, this article achieves efficient classification and prediction of tower crane driver behavior. The residual block structure of ResNet50 addresses the issue of gradient vanishing and degradation in deep networks, and reduces the number of parameters by introducing skip connections, improving the classification accuracy of the model.

4. **Experimental results and analysis.**

4.1. **Dataset construction.** If behavior classification relies on accident causation theory and other criteria, it may result in fewer classification categories, ultimately leading to problems such as false detection. Therefore, this article classifies the collected 22352 behaviors of various tower crane drivers into 8 categories according to recognition targets, namely normal operation, looking at mobile phones (right hand), looking at mobile phones (left hand), making phone calls (right hand), making phone calls (left hand), adjusting instruments, holding things in the back seat, and holding glasses.

4.2. **Model performance evaluation indicators.** The evaluation of the recognition effect of tower crane driver behavior using ResNet50 network can be evaluated from the perspective of accuracy. Accuracy includes four indicators: accuracy, recall, F1 score, and confusion matrix.

$$precision = \frac{TP}{TP + FP} \tag{21}$$

where TP represents the number of positive samples, and FP denotes the number of negative samples.

$$recall = \frac{TP}{TP + FN} \tag{22}$$

$$F1 = 2 \times \frac{precision \times recall}{precision + recall} \tag{23}$$

The Confusion Matrix is a measure of classification accuracy for all categories.

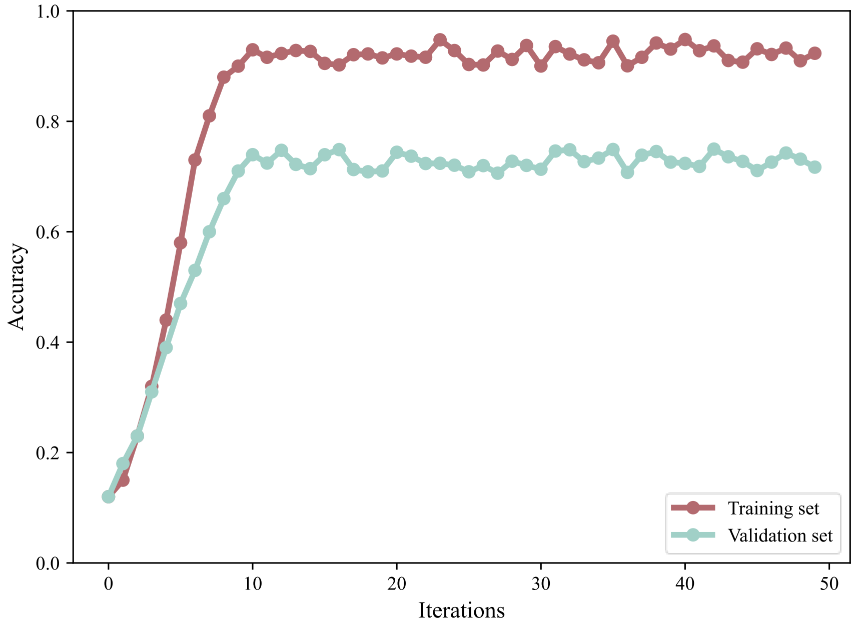

Figure 3. Accuracy without transfer learning

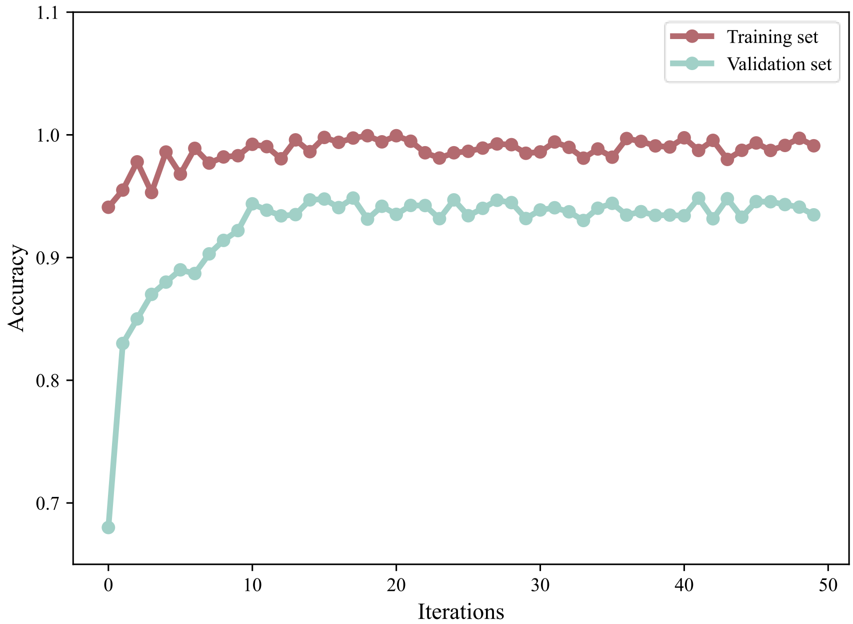

Figure 4. Accuracy after using transfer learning

4.3. **Comparison of transfer learning effects.** The model's training iteration curve is displayed in Figure 3.

Figure 3 illustrates that the ResNet50 network model, without the use of transfer learning, exhibits a slow increase in accuracy on the training set during the training process. On the validation set, accuracy change slowly. From Figure 4, it can be seen that after transfer learning, the network can accurately fit the samples, quickly improve accuracy. By comparison, it can be seen that using transfer learning significantly improves the network.

4.4. **Comparison of the effectiveness of various network models.** In order to avoid overfitting caused by too few samples, the collected 22352 image information were used as the training and testing sets, and data augmentation was performed before training to prevent overfitting. The resulting evaluation metrics are shown in Table 1.

Table 1. ResNet model test results

Evaluation indicators	Precision	Recall	F1
Normal operation	1.00	0.99	0.99
Looking at the phone (right hand)	1.00	1.00	1.00
Use your phone (right hand)	1.00	1.00	1.00
Looking at the phone (left hand)	1.00	1.00	1.00
Make a phone call (left hand)	1.00	1.00	1.00
Adjusting instruments	1.00	1.00	1.00
Carrying things in the back seat	1.00	1.00	1.00
Helping glasses	0.99	0.98	0.99
mean value	0.99	0.99	0.99

According to Table 1, based on the ResNet50 model for detection, the accuracy of identifying 8 common behaviors of tower crane drivers is 99%, The recall rate reaches 98%, and the F1 score is also 98%. This indicates that the model effectively recognizes common behaviors of tower crane drivers and can accurately identify unsafe actions.

To highlight the advantages of the ResNet50 model utilizing transfer learning, comparative experiments were conducted using directional gradient histograms, random forest classifiers, and a lightweight convolutional neural network structure. The comparison results are presented in Table 2.

Table 2. Comparison of ResNet50 with HOG-RF and lightweight CNN networks

Network type	lightweight CNN networks	HOG-RF	ResNet50
Normal operation	0.96	0.87	1.00
Looking at the phone (right hand)	0.96	0.92	1.00
Use your phone (right hand)	0.98	0.94	1.00
Looking at the phone (left hand)	0.98	0.93	1.00
Make a phone call (left hand)	0.97	0.95	1.00
Adjusting instruments	0.98	0.98	1.00
Carrying things in the back seat	0.97	0.96	1.00
Helping glasses	0.93	0.93	0.99
mean value	0.96	0.93	0.99

According to Table 2, the recognition rate of the ResNet50 network with transfer learning exceeds that of both the lightweight CNN structure and the HOG-RF algorithm. Notably, the recognition rate for glasses support showed the greatest improvement at 6%, whereas the smallest increases, 2%, were observed for using the phone with the right hand and looking at the phone with the left hand. These results highlight a significant enhancement in accuracy with the ResNet50 network using transfer learning compared to other methods.

5. **Conclusion.** This study introduces a novel approach to monitoring the construction quality and safety status of high-rise buildings based on deep learning, aiming to address the increasingly complex construction environment in the modern construction industry and the needs that traditional supervision methods cannot meet. With the continuous expansion of the construction sector in China, frequent quality problems and safety accidents during the construction process have brought significant challenges to the smooth implementation of projects. To this end, this article combines image recognition technology and deep learning algorithms to develop an automated supervision system that achieves precise supervision of the construction process through real-time image monitoring and data analysis. This system is based on convolutional neural networks and transfer learning, and can intelligently recognize personnel behavior and construction quality on construction sites. The system can not only detect and warn personnel of illegal operations, but also identify potential quality issues that may arise during the construction process, such as defects in building structures and non-standard manufacturing processes. These are achieved through automated analysis and processing of image data on construction sites, greatly improving the accuracy of construction quality supervision and reducing errors caused by manual supervision. The research results show that the deep learning based supervision system can effectively reduce the risks caused by quality issues and unsafe behaviors, significantly improving the safety management efficiency during the construction process. This system has contributed significantly to enhancing the efficiency of construction project management and ensuring project quality, providing strong support for intelligent supervision of construction sites. Overall, this study provides an innovative and efficient solution for quality control and safety management in modern construction engineering, which is expected to further promote the digitalization and intelligent management process of the construction industry.

REFERENCES

[1] J. Wanberg, C. Harper, M. R. Hallowell, and S. Rajendran, "Relationship between construction safety and quality performance," *Journal of Construction Engineering And Management*, vol. 139, no. 10, 04013003, 2013.

[2] Y. Ma, Y. Peng, and T.-Y. Wu, "Transfer learning model for false positive reduction in lymph node detection via sparse coding and deep learning," *Journal of Intelligent & Fuzzy Systems*, vol. 43, no. 2, pp. 2121-2133, 2022.

[3] J. Lee, and S. Lee, "Construction site safety management: a computer vision and deep learning approach," *Sensors*, vol. 23, no. 2, 944, 2023.

[4] X. Wang, and N. El-Gohary, "Deep learning-based relation extraction and knowledge graph-based representation of construction safety requirements," *Automation in Construction*, vol. 147, 104696, 2023.

[5] S. Chi, S. Han, and D. Y. Kim, "Relationship between unsafe working conditions and workers' behavior and impact of working conditions on injury severity in US construction industry," *Journal of Construction Engineering And Management*, vol. 139, no. 7, pp. 826-838, 2013.

[6] I.-H. Han, G.-N. Yang, S. Kim, G.-H. Kim, and Y. Shin, "Decision Making Method Based on Importance-Dangerousness Analysis for the Potential Risk Behavior of Construction Laborers," *Mathematical Problems in Engineering*, vol. 2015, no. 1, 502121, 2015.

[7] D. Fang, C. Zhao, and M. Zhang, "A cognitive model of construction workers' unsafe behaviors," *Journal of Construction Engineering and Management*, vol. 142, no. 9, 04016039, 2016.

[8] H. Asilian-Mahabadi, Y. Khosravi, N. Hassanzadeh-Rangi, E. Hajizadeh, and A. H. Behzadan, "A qualitative investigation of factors influencing unsafe work behaviors on construction projects," *Work*, vol. 61, no. 2, pp. 281-293, 2018.

[9] S. Guo, P. Zhang, and L. Ding, "Time-statistical laws of workers' unsafe behavior in the construction industry: A case study," *Physica A: Statistical Mechanics and Its Applications*, vol. 515, pp. 419-429, 2019.

[10] J. Zhang, P. Xiang, R. Zhang, D. Chen, and Y. Ren, "Mediating effect of risk propensity between personality traits and unsafe behavioral intention of construction workers," *Journal of Construction Engineering And Management*, vol. 146, no. 4, pp. 04020023, 2020.

[11] S. Han, and S. Lee, "A vision-based motion capture and recognition framework for behavior-based safety management," *Automation in Construction*, vol. 35, pp. 131-141, 2013.

[12] E. J. Lee, B. C. Ko, and J.-Y. Nam, "Recognizing pedestrian's unsafe behaviors in far-infrared imagery at night," *Infrared Physics & Technology*, vol. 76, pp. 261-270, 2016.

[13] Y. Yu, H. Guo, Q. Ding, H. Li, and M. Skitmore, "An experimental study of real-time identification of construction workers' unsafe behaviors," *Automation in Construction*, vol. 82, pp. 193-206, 2017.

[14] L. Ding, W. Fang, H. Luo, P. E. Love, B. Zhong, and X. Ouyang, "A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory," *Automation in Construction*, vol. 86, pp. 118-124, 2018.

[15] Y. M. Goh, C. U. Ubeynarayana, K. L. X. Wong, and B. H. Guo, "Factors influencing unsafe behaviors: A supervised learning approach," *Accident Analysis & Prevention*, vol. 118, pp. 77-85, 2018.

[16] W. Fang, B. Zhong, N. Zhao, P. E. Love, H. Luo, J. Xue, and S. Xu, "A deep learning-based approach for mitigating falls from height with computer vision: Convolutional neural network," *Advanced Engineering Informatics*, vol. 39, pp. 170-177, 2019.

[17] R. Wei, P. E. Love, W. Fang, H. Luo, and S. Xu, "Recognizing people's identity in construction sites with computer vision: A spatial and temporal attention pooling network," *Advanced Engineering Informatics*, vol. 42, 100981, 2019.

[18] W. Fang, L. Ding, P. E. Love, H. Luo, H. Li, F. Peña-Mora, B. Zhong, and C. Zhou, "Computer vision applications in construction safety assurance," *Automation in Construction*, vol. 110, 103013, 2020.

[19] M. Shahverdy, M. Fathy, R. Berangi, and M. Sabokrou, "Driver behavior detection and classification using deep convolutional neural networks," *Expert Systems with Applications*, vol. 149, 113240, 2020.

[20] N. Ketkar, J. Moolayil, N. Ketkar, and J. Moolayil, "Convolutional neural networks," *Deep Learning with Python: Learn Best Practices of Deep Learning Models with PyTorch*, pp. 197-242, 2021.

[21] N. Agarwal, A. Sondhi, K. Chopra, and G. Singh, "Transfer learning: Survey and classification," *Smart Innovations in Communication and Computational Sciences: Proceedings of ICSICCS 2020*, pp. 145-155, 2021.

[22] M. Iman, H. R. Arabnia, and K. Rasheed, "A review of deep transfer learning and recent advancements," *Technologies*, vol. 11, no. 2, 40, 2023.