Research Article

Intelligent Analysis of Hydraulic Concrete Vibration Time Based on Convolutional Neural Network

Hao Liu,[1,2] **Chengzhao Liu,**[1,2] **Jiake Fu,**[1,2] **Chenzhe Ma,**[1,2] **Ye Zhang,**[3] **and Yumeng Lei** [ID][4]

[1]*Zhongnan Engineering Corporation Limited, Power China, Changsha 410014, China*
[2]*Hunan Provincial Key Laboratory of Hydropower Development Key Technology, Changsha 410014, China*
[3]*State Key Laboratory of Eco-Hydraulics in Northwest Arid Region, Xi'an University of Technology, Xi'an 710000, China*
[4]*College of Civil Engineering and Architecture, Zhejiang University, Hangzhou 310058, China*

Correspondence should be addressed to Yumeng Lei; ymlei1017@163.com

The vibrating of concrete is one of the most important procedures that directly determines the quality of construction projects. The concrete vibration quality in field construction is mainly judged by the experience of workers, lacking quantitative indicators, and necessary supervision. However, the lack of research about concrete vibration quality led to these problems are still existing. There are some methods are proposed that are too difficult, or too expensive to use in field construction. Combined with the pouring project of Jianquan Pumped Storage Power Station in Yunyang, China, this research developed an intelligent detection system for concrete vibration time. The system took the convolutional neural network as the basic framework, and divided the concrete vibration process into three different states: vibrating, not vibrating, and no vibration tube, realized the concrete vibration time through the analysis of concrete vibration video data. The detection of concrete vibration process video with multiple stages in the actual project shows that the detection error of the system for each state is kept within 1 s, the accuracy is high, which can meet the quality management requirements of the construction process. The system can be quickly deployed to the construction site by using mobile phones, cameras, and other common equipment. It has the advantages of simple frame structure, low hardware requirements, and accurate detection results. In addition, the current system training process is only for the concrete pouring process of Jianquan Power Station, and the training sample can be further expanded in the future to enhance the applicability and accuracy of the system in other engineering applications in order to play a better role.

1. Introduction

Concrete is widely used in the field of hydropower engineering, civil engineering, and other construction engineering as a basic material, and its quality control has become one of the most important things that needs to paying attention in the construction process. Vibration is a key control process in the concrete pouring process. Suitable vibration can enhance the concrete compactness and improve durability. If the vibration is not sufficient, excessive internal bubbles will lead to the reduction of concrete strength, while excessive vibration will lead to the segregation of aggregates and cementitious materials, which will seriously affect the construction quality [1–3]. Generally, the concrete pouring quality is evaluated by the concrete compressive strength, which

is realized by drilling and sampling test of concrete after consolidation. This method needs to be carried out after construction, which is difficult to timely correct construction quality defects, poor timeliness, and cannot reflect the whole pouring process. In the pouring process, the vibration quality control is often based on the experience judgment of the onsite construction personnel. With strong subjectivity, experience judgment is difficult to achieve quantitative judgment, and it needs continuous long-term monitoring by manual. Therefore, it is necessary to develop a tool to realize rapid quantitative analysis of concrete vibration quality to meet the needs of mechanization and rapid construction.

Existing researches on automatic vibration quality monitoring mainly adopt positioning tracking, image analysis, and other methods, which are realized by global positioning

system (GPS) [4–8], radio frequency identification (RFID) [9, 10], ultra wide band (UWB) [11–13], computer vision [14–22], and so forth.

The concrete vibration positioning tracking method mainly judges the movement track of the vibrating tube through location-sensing technology to realize the monitoring of the overall pouring and vibration quality of the warehouse surface. Tian and Bian [7] installed the GPS antenna and inductive electrode on the vibrator, drew the vibration plan by computer and evaluated the quality of vibration construction combined with the depth of vibration. Denghua et al. [8] measured the insertion depth of the vibrating tube through the ultrasonic rangefinder, and combined the depth data with GPS positioning data to evaluate the concrete vibrating construction quality with information entropy and random algorithm. Su and Liu [10] judged the progress of the pouring process by placing RFID sensors at the fixed boundary. Gong et al. [11, 12] set UWB labels on the tip of the vibrating tube, tracking the track of the vibrator, and visualized the vibrating process by representing the energy accumulation process with the duration. Combining UWB, Quan and Wang [13] realized the real-time tracking of the concrete vibration process, and based on this, the hybrid neural network recognition model was used to realize the automatic derivation of relevant control indicators. The positioning tracking mainly focuses on the whole pouring quality of concrete, which is relatively lacking of the details of each vibration process.

Image analysis of concrete vibration quality based on digital image processing through monitoring videos, images, to achieve the vibration tube and concrete surface state judgment, which can reflect the local vibration details better. Hacıefendioğlu et al. [14] studied the relationship of masonry structures between the change of strength and surface properties characteristics after high temperature and used convolutional neural network (CNN) to carry out classification training on samples of different temperatures taken by portable microscopes, realized the prediction of structural strength loss under high-temperature environment. Similarly, considering the temperature characteristics of the concrete surface in the process of vibration, Burlingame [15] developed a concrete vibration quality monitoring system based on thermal imaging technology. Liu [16] located the position of the vibrator in the concrete vibrating video, based on YOLO v2 [17] image recognition algorithm, and judged the track and quality of the vibration process in combination with the vibration depth. Xu et al. [18] and Makantasis et al. [19] respectively considered the multiscale features of steel box girders of bridges crack images and the multidimensional features of tunnel defects in the actual engineering environment, and carried out bridge crack detection and tunnel defect analysis using CNN, which verified the feasibility of the application of the neural network model in the detection and analysis of material surface properties. Furthermore, Wang et al. [20] classified concrete surface states as unqualified, middle, qualified, combined with Internet of Things technology and neural networks, and evaluated the vibration quality with residual neural network (ResNet) [21].
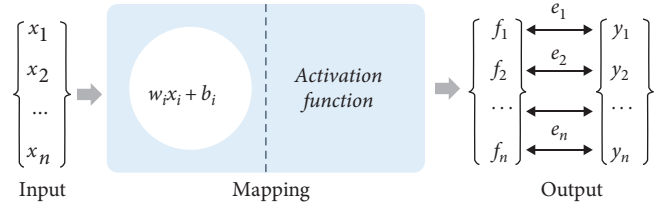

FIGURE 1: Mapping structure of simple neurons.

Ren et al. [22] installed an industrial camera on a vibrator, also used ResNet to process the video image of concrete vibrating, and combined with a semisupervised learning method (Co-MixMatch) to simplify the preprocessing process of the neural network, achieved a good accuracy.

In addition, experts and scholars are also trying to explore some new detection methods, such as 3D laser scanning, 3D modeling, new detection equipment, and algorithms [23–26]. However, the analysis of the existing research of automatic vibration quality monitoring shows that the current research in this field is relatively lacking, most of them need to upgrade the existing tools at a high cost, or add highly refined professional instruments which increases the difficulty of construction. It is too difficult to promote those methods in practical projects.

This research develops a concrete vibration time detection system based on image analysis. With CNN as the main framework of this system, the rapid analysis of concrete vibration quality can be achieved through the concrete vibration video collected by the camera. The detection and reminder of the concrete vibration time length can be realized only by adding cameras, mobile phones, or other recording devices in the construction process without additional auxiliary facilities, and the system could provide suggests for construction management personnel. Through the analysis of video from the pouring project of Jianquan Pumped Storage Power Station in Yunyang, China, the system realized the vibration time accurately, proved that the system is useful for the pouring construction process.

2. Methodology of Concrete Vibration Time Intelligent Detection System

2.1. Convolutional Neural Network. CNN can extract feature information from images through convolution structure, which is widely used in image data classification processing [27, 28]. As shown in Figure 1, a neuron is the basic structure of a neural network, simple neuron structure can form a corresponding result sequence $\{f_1, f_2, \cdots, f_n\}$, by nonlinear mapping of the sample sequence $\{x_1, x_2, \cdots, x_n\}$ in the form shown in Equation (1), then feedback the error sequence $\{e_1, e_2, \cdots, e_n\}$ to the neuron in the form of gradient by comparing the result sequence with the real value result sequence $\{y_1, y_2, \cdots, y_n\}$ and adjust the weight in the neuron, realize the mapping between the input result and the target result finally.

$$f(x_i) = \text{Activation function}\, (w_i x_i + b_i). \tag{1}$$

$$\begin{bmatrix} u_{11}^{l-1} & u_{12}^{l-1} & \cdots & u_{1m}^{l-1} \\ u_{21}^{l-1} & u_{21}^{l-1} & \cdots & u_{2m}^{l-1} \\ \cdots & \cdots & \ddots & \cdots \\ u_{n1}^{l-1} & u_{n2}^{l-1} & \cdots & u_{nm}^{l-1} \end{bmatrix}$$

$$\sum_{i \in n, j \in m} \left(k_{ij}^{l-1} u_{ij}^{l-1} + b_{ij}^{l-1} \right)$$

Activation function

$$\{u_{pq}^{l}\}$$

Input: U^{l-1} Mapping: *convolution*   Output : u_{pq}^{l}
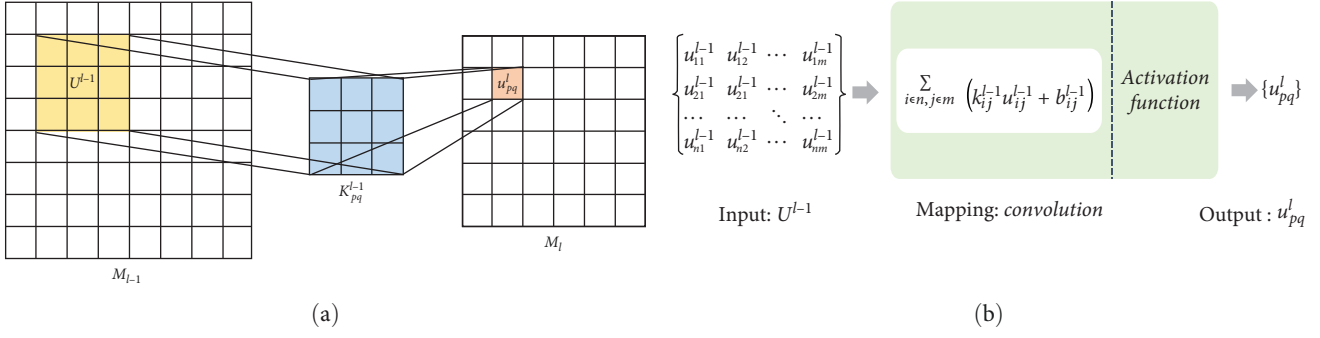
(a)   (b)

FIGURE 2: Convolution operation and mapping structure of convolutional neural network. (a) Convolution operation of 2D matrix and (b) mapping structure of convolutional neurons. M_{l-1} is the characteristic matrix of last convolution output, that is, the input matrix of this convolution process, M_l is the output matrix of the layer convolution $l-1$, U^{l-1} is the submatrix of M_{l-1}, which has the same size as the convolution kernel K_{pq}^{l-1}, and the remaining symbols are the same as before.

During image processing, it is necessary to establish the mapping relationship between the 2D data and the target results, which requires considering the influence of surrounding data in a certain range on the basis of the scale of the unit data itself. Convolution is a mathematical method to sum the product of 2D data in the form of translation and inversion, which meet the needs of 2D image data processing [29]. The neural network formed by convolution is called CNN. As shown in Figure 2, when image convolution analysis is processing, the 2D data matrix of the image M_{l-1} is formed into a submatrix U^{l-1}, according to the range of the convolution window, which has the same size with the convolution kernel K_{pq}^{l-1}. Summing the product of U^{l-1} and K_{pq}^{l-1}, according to Equation (2) to form a feature result u_{pq}^{l}. By sliding the convolution window, multiple feature results are formed and their combination is the output of the CNN unit.

$$u_{pq}^{l} = \text{Activation function} \left\{ \sum_{i \in n, j \in m} \left(k_{ij}^{l-1} u_{ij}^{l-1} + b_{ij}^{l-1} \right) \right\}, \tag{2}$$

where u_{pq}^{l} means output of the convolution process at row p and column q in the convolution process at layer $l-1$, k_{ij}^{l-1} is the size of convolution kernel K_{pq}^{l-1} at row i and column j in the convolution layer $l-1$, b_{ij}^{l-1} is the offset at row i and column j in the convolution layer $l-1$, and u_{ij}^{l-1} is the size of submatrix U^{l-1} at row i and column j in the convolution layer $l-1$ [30].

CNN formed a neural network hidden layer through multiple convolutional neurons, and established the corresponding relationship between input and output data through a large number of nonlinear mapping relationships. Similar to a simple neuron, CNN feedback the error between the output result and the target value result into the neuron in the form of gradients, and finally realized the mapping between the input result and the target result by adjusting the size of the convolution kernel k_{ij} in each convolutional neuron.

2.2. Activation Function, Softmax, and Loss Function. The neural network establishes the nonlinear mapping relationship between samples and targets through an activation function, which could project values of any size into a fixed range [31]. As shown in Equations (3)–(5), the commonly used activation functions mainly include Sigmoid, Tanh, Relu, and so forth [32, 33]. Sigmoid function maps data to the range of (0, 1), Tanh function maps data to the range of (−1, 1), and Relu function maps data to the range of (0, +∞).

$$\text{Sigmoid}(x) = \frac{1}{1 + e^{-x}}, \tag{3}$$

$$\text{Tanh}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \tag{4}$$

$$\text{Relu}(x) = \max(0, x). \tag{5}$$

Softmax method is used to determine the classification probability of CNN output. As shown in Equation (6), the softmax function can be used to convert the feature results of the neural network into a probability distribution within the range of [0, 1], and the sum of probabilities is 1, which determines the probability that the feature results belong to a certain class [28].

$$S_i = \text{Softmax}(x_i) = \frac{e^{x_i}}{\sum_{i=0}^{n} e^{x_i}}, \tag{6}$$

where S_i represents the possible probability of the ith result, n represents the number of output nodes by the neural network, which means the number of categories classified by the neural network, and x_i represents the output of the ith node.

The result of softmax means the probability of the category of the image, and the difference between softmax result and training target is the accuracy of the classification prediction of CNN. Usually, the difference between the predicted result and the real value is judged in the form of loss function in neural networks. In this study, the cross entropy loss function is used as the basis for the evaluation of classification accuracy, as shown in Equation (7), which represents the information entropy difference between the probability distribution of the neural network predicted and the
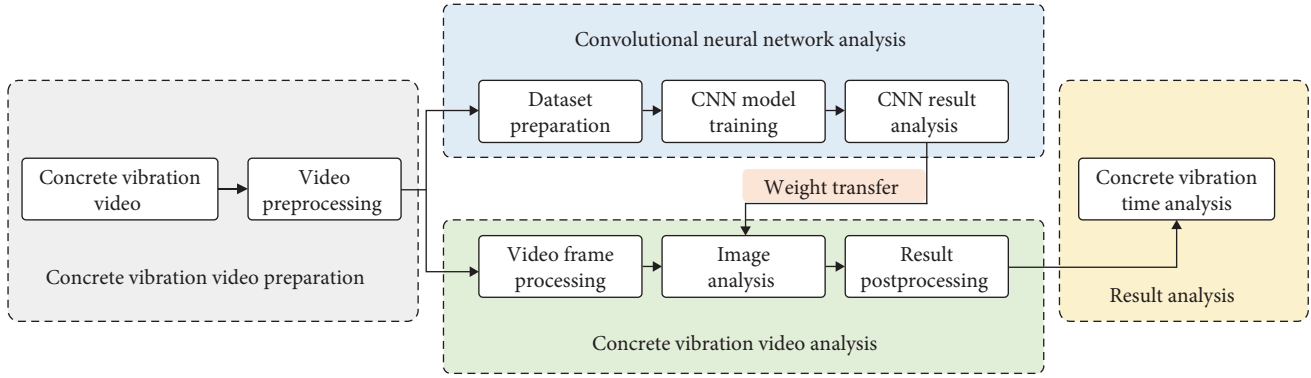
FIGURE 3: Concrete vibration time analysis system framework.

probability distribution of the true value of the sample. Feeding back the cross entropy loss to the hidden layer of the neural network, adjusting the size of the convolution kernel in each neuron, uniting the prediction result and the training objective, and then the image classification process is realized [34].

$$L = \text{Cross entropy loss}(S) = -\sum_{i=0}^{n} y_i \log(S_i), \qquad (7)$$

where L is the cross entropy loss, y_i is the true probability of the class i appearing, and S_i is the output of softmax, which represents the probability of picture belongs to class i.

3. Framework of Concrete Vibration Time Intelligent Detection System

3.1. System Framework. The concrete vibration time intelligent detection system based on CNN includes four parts: concrete vibration video preprocessing, CNN analysis, vibration video analysis, and concrete vibration time analysis, as shown in Figure 3.

(1) Concrete vibration video preprocessing part is mainly used for video reading and editing, combining multiple independent video files into one video file, or editing a long single video into multiple independent files of a certain length of time, and then according to the time sequence, slicing the file into image for neural network training;

(2) The CNN analysis part is used for the system preparation process before concrete vibration time detection, by the training of CNN, the system can realize the ability of concrete vibration image detection and analysis. In this part, the system first organizes the number of concrete vibration images from the last part into neural network training data by means of gray-scale, image scaling, image enhancement, and data augmentation, and then through the marking, training, and verification process, the system is equipped with classification ability of concrete vibration state;

(3) The vibration video analysis part can analyze and sort the concrete vibration video frame image based

on the training of the neural network. In the field detection process, the concrete vibration video data from the detector is transmitted to this part, similar to the first part, which is edited and sliced in the order of time to forming the concrete vibration image data that could be analyzed by the system, and then the results are passed to the next part after the CNN classification;

(4) The concrete vibration time analysis part can establish the concrete vibration timeline, realize the detection of concrete vibration time, and give the construction prompt and early warning. Since the concrete vibration state detected by CNN cannot be completely accurate, in this part, the system corrects the output results in chronological order to ensure the accuracy of the concrete vibration time detection. The specific and detailed correction principles will be described below.

3.2. CNN in Concrete Vibration Time Intelligent Detection System. In the CNN analysis part, the detection system established a five-layers deep learning network in the form of supervised deep learning to train and learn the concrete vibration video frame image data after slice processing, so that the system could have the ability to classify the vibration image.

Before the neural network training, it is necessary to slice the concrete vibrating video according to the time frame to form the vibrating image and establish the original dataset. As shown in Figure 4, according to the working state of the vibrating tube, the concrete vibrating images can be divided into three different categories: vibrating, not vibrating, and no vibration tube, which can be labeled to [1, 0, 0], [0, 1, 0], [0, 0, 1] in the form of one-hot coding. Taking each concrete vibration image as the training sample and the corresponding one-hot label as the training target, the training dataset of the neural network is formed.

The CNN architecture in the detection system is shown in Figure 5, which mainly includes three parts: input layer, middle hidden layer, and output layer. In the input layer, the system feeds the concrete vibration image into the neural network. In the middle hidden layer, a convolution operation
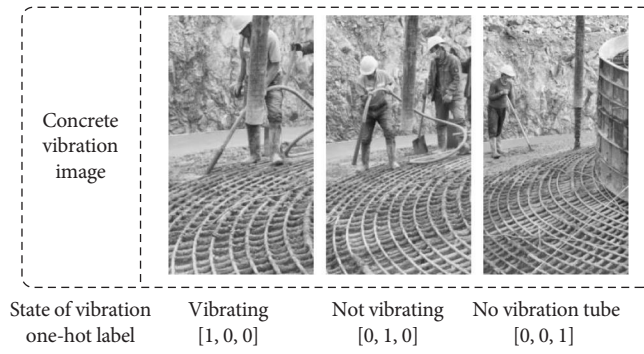
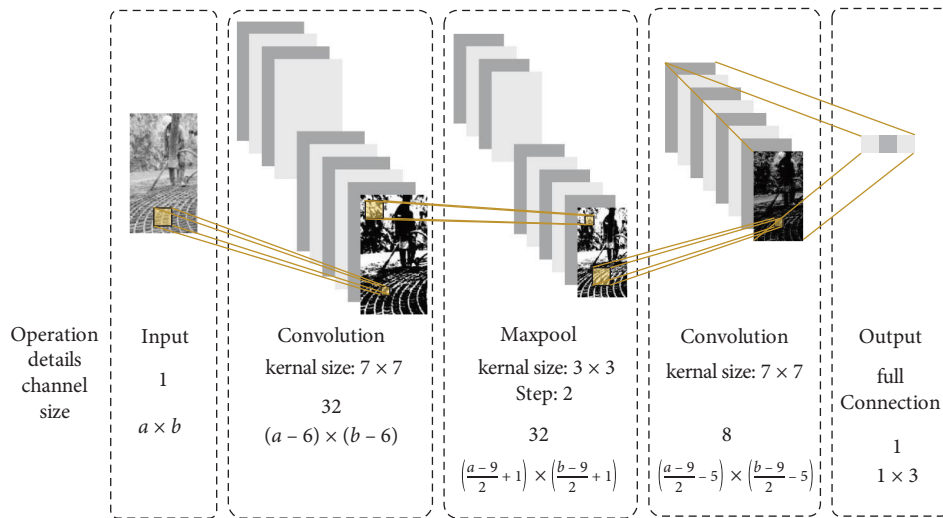FIGURE 4: Concrete vibration image with one-hot label.

FIGURE 5: Convolutional neural network in concrete vibration time analysis system.

is carried out on the image with 7×7 size convolution kernel according to step 1 to form a 32-channel feature image, then, the feature matrix is pooled according to the scale of 3×3 size of pooled kernel and step 2 in the way of Maxpool to extract the main features in the feature image. After Maxpool, the feature image is also convolved with the convolution kernel size of 7×7 and step size of 1 to form an 8-channel feature image. In the output layer, 8-channel feature images are extracted in a fully connected layer and the 1×3 size predictive coding is output. The classification probability of output results is analyzed through Softmax, which achieves the feature extraction and classification of concrete vibration images by CNN.

4. Concrete Vibration Time Intelligent Detection System Test

4.1. Preparation of Detection System. Because of the supervised deep learning method, it is necessary to collect certain video data of concrete vibration as the samples of system training and learning which make the system have the ability to detect concrete vibration state before using the detection system. This research relies on the pouring project of

Jianquan Pumped Storage Power Station to carry out the test of the concrete vibration time detection system.

Jianquan Pumped Storage Power Station is located in Yunyang, Chongqing, China, where is a tributary of the Yangtze River. The station consists of upper reservoir, lower reservoir, water conveyance system, and switching station. The installed capacity of the project is 1,200 MW. Both upper reservoir and lower reservoir are concrete-faced rockfill dams with a rated head of 332 m, the maximum dam height of 98 and 78 m, respectively, and the design strength of concrete is 35 MPa. The power station is built for improving the peak shaving and valley filling, frequency and phase modulation, and emergency standby capacity of Chongqing power grid.

The video samples of concrete vibration are collected in the pouring construction site of the power station, and the vibration process from the insertion of the concrete vibrator tube to the complete pulling out is taken as one complete vibration process. The sample dataset includes 13 complete vibration processes, three incomplete vibration processes (only including the vibrating with the vibrator, no pulling process) and three concrete pouring processes without vibration. By slicing video data in the unit of frame, a total of 2,477 video slicing image data of concrete vibration were obtained.
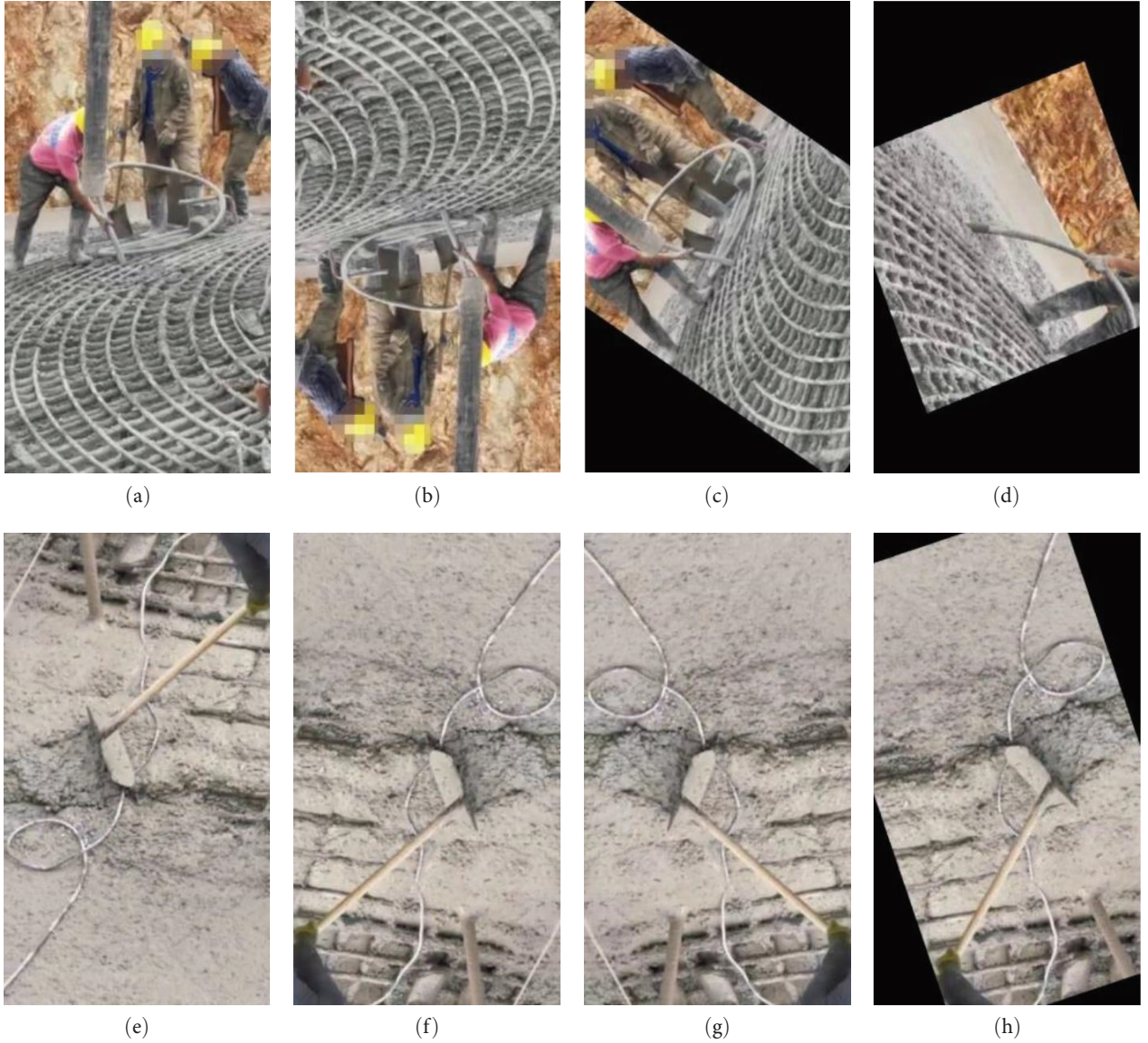
FIGURE 6: Image data augmentation in concrete vibration time analysis system. (a) Concrete vibration image-1; (b) image flip; (c) image random rotation; (d) image random clipping; (e) concrete vibration image-2; (f) image flip; (g) image mirroring; and (h) image random rotation.

Due to the similarity of images in the vibration process, although thousands of sample images have been formed, the scene feature information is relatively simple, which is insufficient to extract the unified feature information in the concrete vibration image. Therefore, in the dataset preparation part before CNN training, random methods were used to augment the image data, which included image flipping, random angle rotation, random size clipping, and mirror image. As shown in Figure 6, a total of 8,276 concrete vibration images were formed to meet the requirements of neural network training.

After the image data augmenting, the size of the dataset is shown in Table 1. The trainset data contained 7,366 images, which were used for the learning and training of neural network. The validation dataset contains 910 images to verify the accuracy of the current neural network for image recognition, which does not participate in model training. The

TABLE 1: The augmentation dataset in concrete vibration time analysis system training.

Dataset	Vibrating	Not vibrating	No vibration tube	Total
Train	3,631	539	3,196	7,366
Validation	390	130	390	910
Total	4,021	669	3,586	8,276

accuracy of CNN is judged in the way of Equation (8), which represents the percentage of correct image data of a certain scale verification set.

$$\text{Accuracy} = \frac{\text{Correct classification data size}}{\text{Validation data size}} \times 100\%, \qquad (8)$$

TABLE 2: Concrete vibration time analysis system training parameter and result.

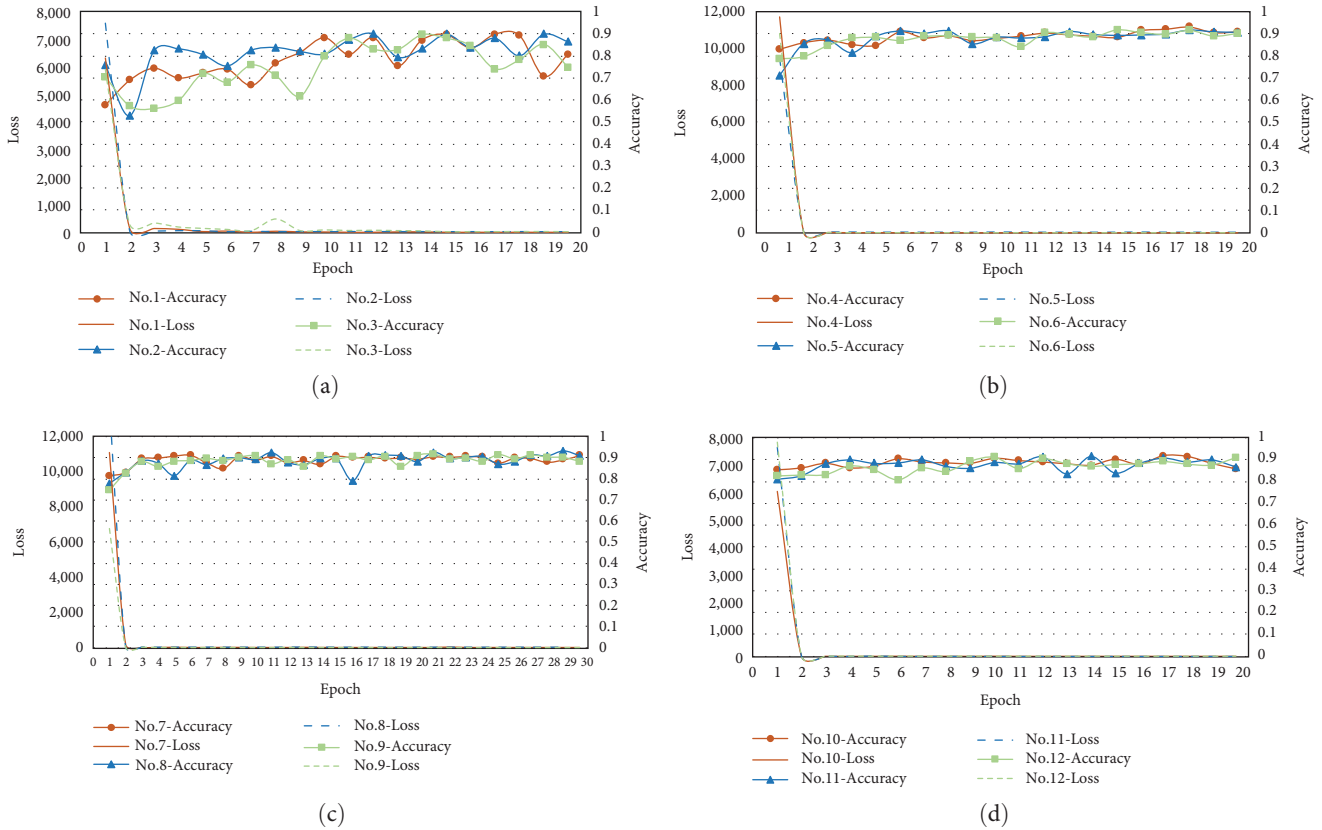| No	Convolutional kernel	Batch size	Epoch	Test accuracy
1	3×3	16	20	0.7975
2	3×3	32	20	0.85625
3	3×3	64	20	0.75375
4	7×7	16	20	0.915
5	7×7	32	20	0.91875
6	7×7	64	20	0.90625
7	7×7	16	30	0.9025
8	7×7	32	30	0.91375
9	7×7	64	30	0.875
10	11×11	16	20	0.87875
11	11×11	32	20	0.89375
12	11×11	64	20	0.905

(a)

(b)

(c)

(d)

FIGURE 7: The system training process in 12 kinds of different training parameters. (a) No. 1–3 system training process; (b) no. 4–6 system training process; (c) no. 7–9 system training process; and (d) no. 10–12 system training process.

where correct classification data size is the number of verification sets for which the model classification is correct. Validation data size is the total size of the sample data.

4.2. Training and Learning of Detection System. On the basis of image augmented dataset, the system learns and trains the vibration image through deep learning, and obtains the features of targets through the analysis and summary of different images. For the concrete vibrating detection, the target characteristics are the working state of the concrete vibrating

tube. The main parameters of neural network that affect the accuracy of system detection include: convolution kernel, batch size, and epochs. The size of the convolution kernel represents the range of local feature extraction of CNN in the process of image feature extraction, which changes with the scale of the input image. Batch size represents the number of concrete vibration images put into each neural network training process. Epochs represent the training times of all sample data in the neural network. During the system training, all three parameters have a certain impact on the final

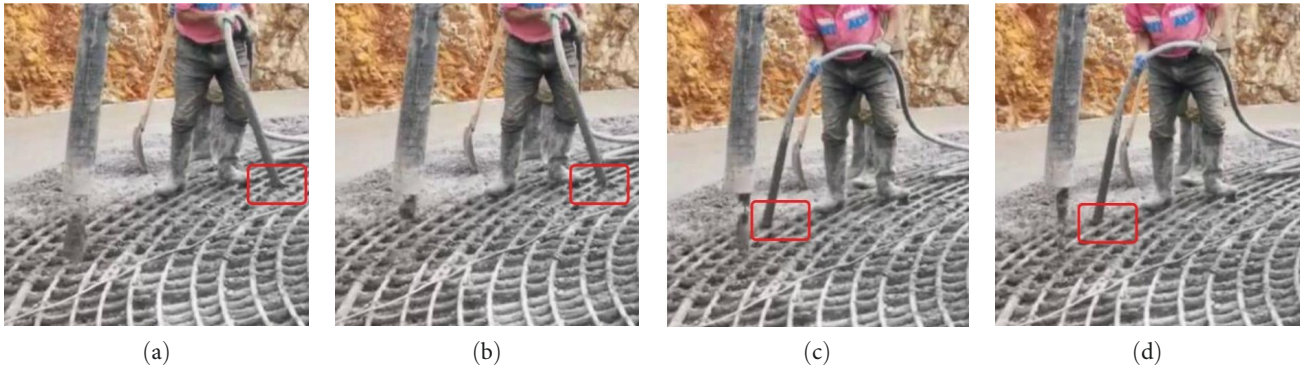| (a) | (b) | (c) | (d) |

FIGURE 8: The different state of vibration tube in pulling out and inserting process. (a) State of vibrating in pulling out process; (b) state of not vibrating in pulling out process; (c) state of not vibrating in inserting process; and (d) state of vibrating in inserting process.
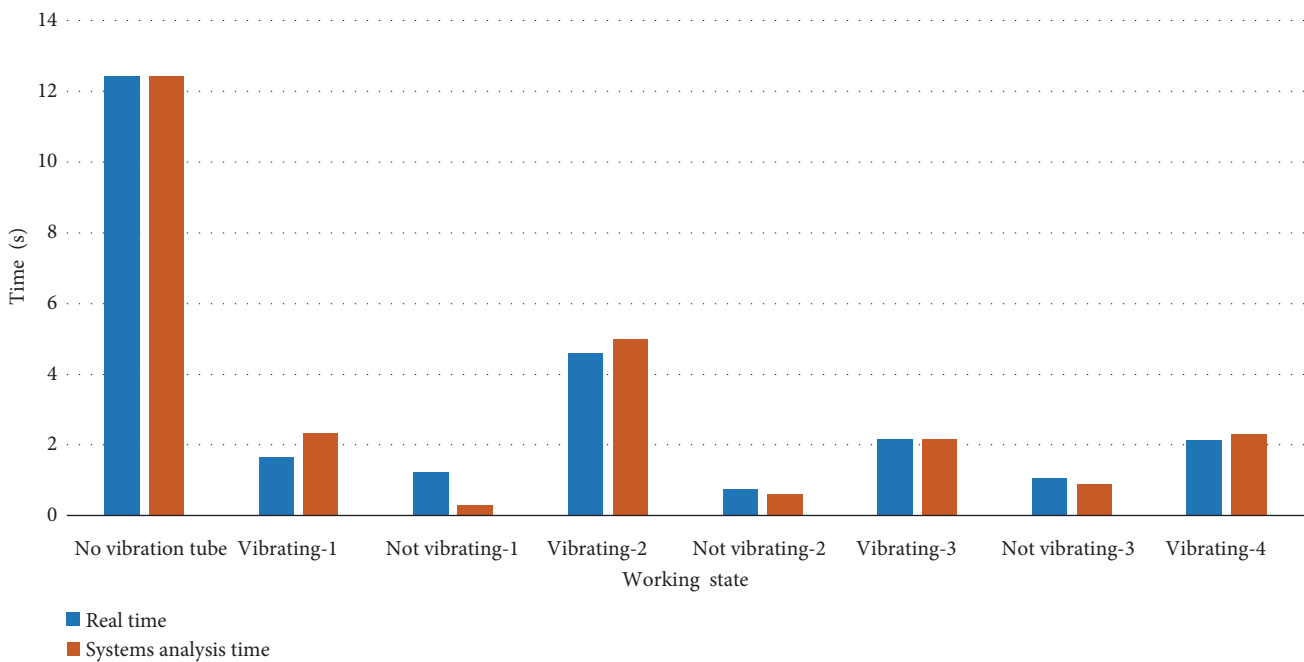

FIGURE 9: The concrete vibration time analysis system test results.

accuracy. Different parameter combinations are set in this research, and the corresponding results are shown in Table 2. The learning rate of the CNN is set at 0.00001. During the training, 500 concrete vibration images are randomly taken from 910 of the verification dataset each time as the judgment basis for the accuracy of the network in each epoch, and the training process is shown in Figure 7. In the accuracy testing part, 800 concrete vibrating images that did not participate in the training were used as the evaluation basis, including 350 pictures of vibrating, 100 pictures of not vibrating, and 350 pictures of no vibration tube.

The system training in 12 parameter combinations showed that the loss of the system became smaller at epoch 2–3, and oscillates slightly within the range of 100, then it was closed to 0 at epoch 9–11. The accuracy at epoch 0–10 was relatively large, and it was basically stable at epoch 10–20 with a stability near 0.9 and a maximum value of 0.91875. The accuracy of the system fluctuated greatly in small

convolution kernel (3 × 3), and it was better and the stability after the convolution kernel being modified to a larger value. According to the system testing result, the system parameters were finally set as No. 5 parameter combination, convolution kernel size was 7 × 7, batch size was 32 and epoch was 20.

4.3. Concrete Vibration Time Detection and Results Analysis. With the training and learning process of CNN, the intelligent detection system of concrete vibration time has the ability to extract and classify the features of concrete vibration image. The weight, which was obtained from network training, was transferred to the vibration video image analysis part of the detection system, and then the concrete vibration video analysis can be started.

In the vibration video image analysis part, the system first slices the video data according to the time axis which is similar to the vibration video preprocessing part, and the slice interval can be set freely according to the frame rate of

TABLE 3: The test result of concrete vibration time analysis system.

Working states	Real time (s)	Systems analysis time (s)	Error (s)
No vibration tube	12.43	12.43	0
Vibrating-1	1.63	2.30	0.67
Not vibrating-1	1.2	0.30	0.90
Vibrating-2	4.61	5.00	0.39
Not vibrating-2	0.76	0.60	−0.16
Vibrating-3	2.17	2.17	0
Not vibrating-3	1.06	0.90	0.16
Vibrating-4	2.14	2.30	0.16

the video. The interval of the slices represents the allowable time error range of the system detection. For example, an image data is taken every 6 frames in a 30 frames per second video, the allowable error range of the system detection time is 0.2 s. The state of concrete vibration can be judged by analyzing the state of each vibration image on the time axis.

As shown in Table 2, the accuracy of the neural network is stable at 90% for the dataset of Jianquan power station concrete pouring process, there are still some errors in the classification results of concrete vibration image. The analysis of results shows that errors are mainly found at the moment when the vibration tube is pulled out or inserted into the concrete, which is the beginning or end of a vibration process. Figure 8 shows the pulling out and inserting process of the vibration tube, which clearly shows that due to the influence of resolution, distance, and other factors, sometimes it is difficult to divide the boundary of the vibration state. Therefore, it is necessary to correct the error in the part of vibration time analysis to improve the accuracy of detection.

Compared with the vibrating process, the duration of vibration state changing is extremely short, which means the main purpose of result correction is to determine the boundary point of the vibrating state, avoiding the result of consecutive and overlapping occurrence of "Vibrating" and "Not Vibrating" states. Based on the purpose, the principles of vibration state vibration state correction are:

(1) If the state of the current frame is inconsistent with the state of the previous time frame, check the next five frame images states since this time frame. If the state is consistent with three or more of the five vibration image states, it is considered that the vibration state has changed. If not, it is considered that the state of vibration has not changed and corrected the vibration state of this frame to the vibration state of the previous frame.

(2) Analyzing the duration of each vibrating state. If the duration is less than 0.3 s, adjust the part of the vibrating state into the previous vibrating state.

Taking a 26 s concrete vibration video in the project site as an example to test the accuracy of the concrete vibrating time detection system. The sample video includes three complete vibration processes, one incomplete vibration process, and one concrete pouring without vibration. The test results are shown in Figure 9 and Table 3. As Table 3 shows, the detection system has realized accurate identification of random working states in the concrete vibration, and the maximum time error of each state duration is 0.9 s, which can be used in construction quality control of concrete pouring engineering, and the system has realized automatic intelligent identification of concrete vibration time.

5. Conclusion and Discussion

Based on the CNN, this research established a five layers deep learning network in the form of supervised deep learning and divided the concrete vibration process into three categories according to the working state of the vibrator: vibrating, not vibrating, and no vibration tube. With the different states of the vibration tube, the concrete vibration time detection system could analyze and divide vibrating time automatically. Combined with the pouring project of Jianquan Pumped Storage Power Station, the system effectiveness was verified through field sampling test, and the error of the detection of concrete vibration time was within 1 s.

The advantages of the detection system are the simple hardware requirements and frame structure, it can be quickly deployed to the actual construction site through cameras, mobile phones, and other devices, without adding redundant processes which might reduce the construction efficiency. It can provide a reliable reference for the quality management process of concrete vibration site construction

However, there are still some defects that need to be further improved in subsequent studies, including: the system adopts the deep learning model of supervised neural network, before application and deployment, a certain amount of the samples need to be collected for model training and system preparation before the detection process can be carried out. In the actual construction process, the environment of video, light, angle, and other factors will affect the quality of video imaging, but the main object "Vibrating Tube" detected by this system has not changed, therefore, 2–3 min of vibrating videos for several different typical working environments, such as daytime, night, and distant construction sites, are necessary to collect, then the state of the concrete vibrating tube in different time periods in each video is input for system preprocessing, which can ensure the accuracy of system detection. Similarly, this research focuses on the development of the concrete pouring process

of Jianquan power station. Before deployment, it experienced the typical environmental data collection process in advance, and achieved good results after system training. But the current sample dataset is still relatively small, which means that the accuracy may decline due to environmental changes in the deployment process of other projects. In the future, the sample data of the system can be expanded for specific projects to enhance the environmental adaptability of the system, and play a better role in project construction.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

References

[1] S. Lee and M. J. Skibniewski, "Automated monitoring and warning solution for concrete placement and vibration workmanship quality issues," *AI in Civil Engineering*, vol. 1, no. 1, pp. 1–19, 2022.

[2] Z. Tian, X. Sun, W. Su et al., "Development of real-time visual monitoring system for vibration effects on fresh concrete," *Automation in Construction*, vol. 98, pp. 61–71, 2019.

[3] Z. Li and G. Cao, "Rheological behaviors and model of fresh concrete in vibrated state," *Cement and Concrete Research*, vol. 120, pp. 217–226, 2019.

[4] N. Pradhananga and J. Teizer, "GPS-based framework towards more realistic and real-time construction equipment operation simulation," in *Proceedings of the 2012 Winter Simulation Conference (WSC)*, pp. 1–12, IEEE, 2012.

[5] R. Navon and Y. Shpatnitsky, "Field experiments in automated monitoring of road construction," *Journal of Construction Engineering and Management*, vol. 131, no. 4, pp. 487–493, 2005.

[6] R. Anderegg, D. A. Von Felten, and K. Kaufmann, "Compaction monitoring using intelligent soil compactors," in *Proceedings of Geotechnical Engineering in the Information Technology Age*, pp. 1–6, Geo-Congress 2006, 2006.

[7] Z. Tian and C. Bian, "Visual monitoring method on fresh concrete vibration," *KSCE Journal of Civil Engineering*, vol. 18, no. 2, pp. 398–408, 2014.

[8] Z. Denghua, S. Ziyang, and W. Jiajun, "Study on dynamic evaluation of vibration quality of concrete dam based on real-time monitoring," *Journal of Hydraulic Engineering*, vol. 49, no. 7, pp. 775–786, 2018.

[9] S. Moon and B. Yang, "Effective monitoring of the concrete pouring operation in an RFID-based environment," *Journal of Computing in Civil Engineering*, vol. 24, no. 1, pp. 108–116, 2010.

[10] Y. Y. Su and L. Y. Liu, "Real-time construction operation tracking from resource positions," *Computing in Civil Engineering*, pp. 200–207, 2007.

[11] J. Gong, Y. Yu, R. Krishnamoorthy, and A. Roda, "Real-time tracking of concrete vibration effort for intelligent concrete consolidation," *Automation in Construction*, vol. 54, no. 1, pp. 12–24, 2015.

[12] J. Gong, Y. Yu, and T. P. Williams, "Real-time 3D concrete vibration effort tracking and visualization with ultra-wide-band technologies," in *Proceeding of the Transportation Research Board 94th Annual Meeting*, Washington DC, USA, Article ID 01555776, 2015.

[13] Y. Quan and F. Wang, "Machine learning-based real-time tracking for concrete vibration," *Automation in Construction*, vol. 140, Article ID 104343, 2022.

[14] K. Hacıefendioğlu, A. F. Genç, S. Nayır, S. Ayas, and A. C. Altunışık, "Automatic estimation of post-fire compressive strength reduction of masonry structures using deep convolutional neural network," *Fire Technology*, vol. 58, no. 5, pp. 2779–2809, 2022.

[15] S. E. Burlingame, "*Application of infrared imaging to fresh concrete: monitoring internal vibration*," MA thesis, Cornell University, United States, 2004.

[16] Y. Liu, "*Develop of concrete vibration monitoring system based on stereo vision*," MA thesis, Harbin Institute of Technology, China, 2018.

[17] J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7263–7271, IEEE, 2017.

[18] Y. Xu, Y. Bao, J. Chen, and Z. Wangmeng, "Surface fatigue crack identification in steel box girder of bridges by a deep fusion convolutional neural network based on consumer-grade camera images," *Structural Health Monitoring*, vol. 8, no. 3, pp. 653–674, 2019.

[19] K. Makantasis, E. Protopapadakis, A. Doulamis, K. Loupos, and N. Doulamis, "Deep convolutional neural networks for efficient vision based tunnel inspection," in *2015 IEEE International Conference on Intelligent Computer Communication and Processing (ICCP)*, pp. 335–342, IEEE, 2015.

[20] D. Wang, B. Ren, B. Cui, J. Wang, X. Wang, and T. Guan, "Real-time monitoring for vibration quality of fresh concrete using convolutional neural networks and IoT technology," *Automation in Construction*, vol. 123, Article ID 103510, 2021.

[21] K. He, X. Zhang, and S. Ren, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, IEEE, 2016.

[22] B. Ren, H. Wang, D. Wang, T. Guan, and X. Zheng, "Vision method based on deep learning for detecting concrete vibration quality," *Case Studies in Construction Materials*, vol. 18, Article ID e02132, 2023.

[23] J. Teizer, C. H. Caldas, and C. T. Haas, "Real-time three-dimensional occupancy grid modeling for the detection and tracking of construction resources," *Journal of Construction Engineering and Management*, vol. 133, no. 11, pp. 880–888, 2007.

[24] M.-K. Kim, J. C. P. Cheng, H. Sohn, and C.-C. Chang, "A framework for dimensional and surface quality assessment of precast concrete elements using BIM and 3D laser scanning," *Automation in Construction*, vol. 49, pp. 225–238, 2015.

[25] C.-S. Park, D.-Y. Lee, O.-S. Kwon, and X. Wang, "A framework for proactive construction defect management using BIM, augmented reality and ontology-based data collection template," *Automation in Construction*, vol. 33, pp. 61–71, 2013.

[26] J. Li, Z. Tian, X. Sun, Y. Ma, and H. Liu, "Working state determination for concrete internal vibrator using genetic simulated annealing clustering method," *Case Studies in Construction Materials*, vol. 17, Article ID e01163, 2022.

[27] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[28] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[29] P. Li, J. Zhang, and P. Krebs, "Prediction of flow based on a CNN–LSTM combined deep learning approach," *Water*, vol. 14, no. 6, pp. 993–1006, 2022.

[30] Y. M. Lei, Z. Y. Chen, and S. Yu, "Intelligent detection of gradation for earth-rockfill materials base on deep Otsu convolutional neural network," *Journal of Hydraulic Engineering*, vol. 52, no. 3, pp. 369–380, 2021.

[31] I. Goodfellow, B. Yoshua, and C. Aaron, *Deep Learning*, MIT Press, USA, 2016.

[32] M. Cococcioni, F. Rossi, and E. Ruffaldi, "A fast approximation of the hyperbolic tangent when using posit numbers and its application to deep neural networks," in *Applications in Electronics Pervading Industry, Environment and Society*, pp. 213–221, Springer, 2020.

[33] S. Dimitriadou and K. G. Nikolakopoulos, "Artificial neural networks for the prediction of the reference evapotranspiration of the Peloponnese Peninsula, Greece," *Water*, vol. 14, no. 13, pp. 2027–2042, 2022.

[34] J. Goodman, "Classes for fast maximum entropy training," in *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 1, pp. 561–564, IEEE, 2001.