

Real-time monitoring for vibration quality of fresh concrete using convolutional neural networks and IoT technology

Dong Wang, Bingyu Ren, Bo Cui*, Jiajun Wang, Xiaoling Wang, Tao Guan

State Key Laboratory of Hydraulic Engineering Simulation and Safety, Tianjin University, Tianjin 300350, China



ARTICLE INFO

Keywords:

Vibration quality
Fresh concrete
Real-time monitoring
IoT technology
Convolutional neural network
Image classification

ABSTRACT

Vibration quality is critical to ensure the concrete strength, which directly affects the long-term safe operation of concrete structures. The vibration duration and vibration depth are key parameters to guarantee vibration quality. However, traditional manual inspection on concrete surface to judge the vibration duration and estimation of vibration depth is subjective and unreliable. Moreover, existing studies monitor the vibration duration based on the knowledge from prior experiments, ignoring the influence of concrete heterogeneity. Thus, a real-time monitoring method for vibration quality of fresh concrete based on ResNet with 50 layers (ResNet-50) and Internet of Things (IoT) technology is proposed. The IoT-based monitoring framework is proposed to measure vibration depth and capture concrete surface image (CSI). A three-category classification model of CSI is established based on fine-tuned ResNet-50 model using a self-constructed dataset with 15,006 images to determine proper vibration duration. A large-scale hydraulic engineering application verifies the performance of the proposed method.

1. Introduction

Concrete, the prevailing building material in infrastructure, is widely constructed in dams, bridges, foundations, etc. The consolidation of concrete directly affects the safe operation and service life of concrete structures. Vibration has been considered as a critical method to make concrete consolidation. Appropriate vibration can ensure the consolidation of concrete [1,2], and thus improve the long-term strength and durability of concrete. Under-consolidation can lead to honeycombing, cold joints, and surface pores, while over-consolidation can cause segregation and uneven distribution of aggregates [3]. Typical defects of concrete caused by improper vibration are shown in Fig. 1. These defects demand time and cost to repair, and may lead to the collapse of buildings [1,4]. For proper vibration, the vibration duration and vibration depth are the key indicators that need to be enforced in the specification [5]. For concrete consolidation, the vibration duration is judged on features of the concrete surface when laitance emerges on the concrete surface without big bubbles, and sufficient insertion depth is required to ensure the bond quality of placement between successive layers. Due to the complexity of the concrete consolidation mechanism and heterogeneity of concrete materials, the consolidation of concrete may be different under the same vibration duration. Yet, there is no exact theory

and reliable system to judge the rationality of vibration duration during fresh concrete placement. Currently, judging consolidation adequacy on-site strongly depends on the vibrator operator's experience as only they can see the concrete surface appearance during vibration. Furthermore, vision-based manual inspection of concrete surface image (CSI) to judge the vibration duration and estimation of vibration depth is subjective, labor-consuming, and crude management. Thus, a real-time monitoring method for vibration quality of fresh concrete is urgently needed that can contribute to a more objective quality assurance of concrete during on-site construction.

To assure concrete vibration quality, several attempts have been made to address this problem. A thermal imaging technology was developed to assess consolidation adequacy [6]. As sensor technology and internet of things (IoT) technology are widely applied in civil engineering, positioning sensors, such as ultra-wideband (UWB) and GPS, were utilized to monitor the location and depth of vibration [3,7] to achieve the quantitative analysis of the vibration process. To consolidate concrete, the appropriate vibration duration should be realized considering the features of CSI during concrete placement [5], which is a problem in qualitative image analysis. However, to the best of our knowledge, no relevant research of image processing methods for CSI analysis during vibration exists. Recently, the application of computer

* Corresponding author.

E-mail address: cuib@tju.edu.cn (B. Cui).

vision and image processing technology [8], especially the state-of-the-art deep learning method, has been successfully applied to concrete damage detection [9,10] and surface quality analysis of concrete [11], which brings the light to address this problem.

Therefore, to the best of our knowledge, it is necessary to simultaneously satisfy the qualitative and quantitative requirements of the specification [5] to ensure concrete vibration consolidation. In this paper, a real-time monitoring method for vibration quality of fresh concrete based on convolutional neural network (CNN) and IoT technology is proposed. Specifically, qualitative analysis on CSI to judge the appropriate vibration duration and quantitative analysis on vibration depth are carried out. The proposed real-time monitoring method mainly consists of three parts: (1) a monitoring framework based on IoT technology is proposed, in which an infrared sensor is utilized to measure the vibration depth and a camera sensor is adopted to capture the CSI; (2) the self-constructed surface image dataset of concrete vibration is established with 15,006 surface images covering different periods during vibration placement; Furthermore, the three-category classification model of CSI (unqualified, middle, and qualified) is established based on CNN model; (3) the CSI and vibration depth are monitored in real-time by applying the proposed method in a real-world construction project.

The rest of this paper is organized as follows. [Section 2](#) provides an overview of the related work. The gaps in knowledge and practice are elaborated in [Section 3](#). The overall research framework and methodology are presented in [Section 4](#). [Section 5](#) introduces the results of the case study and [Section 6](#) provides the discussion. Finally, the conclusion is provided in [Section 7](#).

2. Related works

2.1. Quality assurance method of concrete vibration

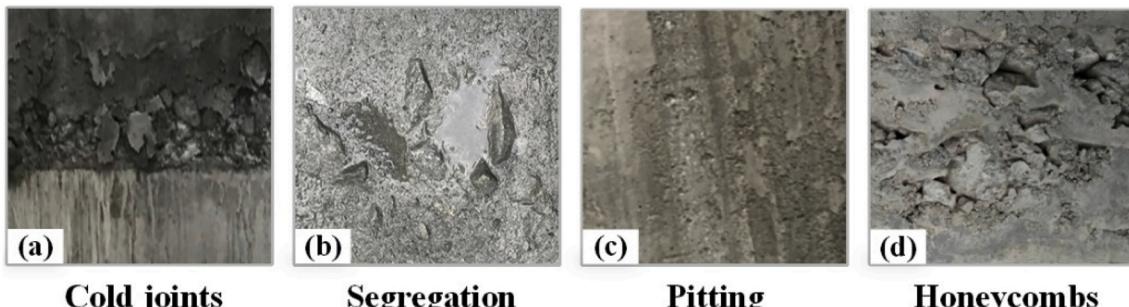
Currently, judging consolidation adequacy on-site mainly depends on the subjective experience of the vibrator operator. Several attempts have been made to inspect vibration more objectively to improve vibration quality assurance. For example, Burlingame [6] used thermal imaging technology to monitor the trajectory of vibrator's tips based on the principle that the temperature of the vibrator was significantly higher than that of the surrounding concrete during construction, and realized the preliminary exploration of vibration information visualization. This method realized the perception of vibration position. Positioning technologies, such as GPS [3] or UWB [7], were applied to track the position of the vibrator, and vibration parameters (vibration duration and depth) were monitored by two tags or antennas. Then, a visualization system was developed to improve the level of intelligent vibration. Additionally, an innovative construction method using computer vision and ultrasonic technology was proposed to monitor concrete placement and vibration for real-time quality control [1], which has provided a sufficient research foundation for the present study. With respect to certain automated concrete operations in heavy construction

activities (dam construction), heavy vibration equipment with multiple vibrators is commonly utilized to consolidate fresh concrete to meet the requirements of rapidly mechanized construction of dam concrete placement. Zhong et al. [12] proposed a real-time monitoring method for the vibration process (vibration position, depth, and duration) of concrete dams and established a dynamic evaluation model for predicting concrete strength based on random forest algorithm.

The above researches achieved real-time monitoring of the vibration process and quality evaluation. However, to realize the proper vibration duration, the above studies are based on prior experiment with quantitative vibration duration during concrete placement, which may not be the best way to ensure consolidation of fresh concrete [13]. Thus, ensuring suitable vibration quality for each vibration should be the qualitative analysis of CSI and quantitative analysis of a certain insertion depth to satisfy the specification [5]. Therefore, this study focuses on using IoT technology to perceive vibration depth and CSI. Additionally, image processing technology to classify CSI and mathematical analysis to assess the insertion depth are separately used to simultaneously analyze whether the vibration depth and CSI satisfy the requirements of the specification, so as to ensure the vibration quality of fresh concrete.

2.2. Image processing method of concrete

Image processing technology is an effective way to obtain valuable information through feature extraction and analysis of images. It is widely used to address real-world civil engineering problems in concrete defect and crack image analysis. For example, Hutchinson et al. [14] proposed a statistical-based method for image analysis to evaluate concrete damage, and the results of the image analysis illustrated its promise. Liu et al. [15] proposed a method to detect concrete surface bugholes using image processing technology, and presented the evaluation parameters of the surface bugholes to assess concrete surface quality; With the development of image processing technology, especially deep learning techniques, the traditional image processing technology has transited to the deep neural network technology, which is one of the major breakthroughs in computer vision domain. Hüthwohl et al. [16] proposed a three-staged concrete defect classifier to realize multi-classification and detailed assessments of bridge health conditions based on a deep CNN, and the classifier achieved an average mean score of 85% in terms of multiple defect type classification. Dung and Anh [10] proposed a semantic segmentation method of concrete crack images based on deep fully convolutional network, and a concrete specimen loading experiment validated the proposed method by achieving approximately 90% in average precision. Zhang et al. [17] established a bridge crack recognition model based on CNN and developed a bridge structure health monitoring system based on IoT technology. Cha et al. [9] proposed a vision-based method to detect concrete surface cracks using CNN, which achieved comparatively better performance than the traditional image processing technology. Additionally, deep CNNs have been widely applied in pavement crack identification [18,19], underground sewer pipe damage identification [20,21], etc.



[Fig. 1. Examples of quality defects of concrete.](#)

The above-mentioned studies illustrate the rapid development of deep learning technology and its widespread application in solving civil engineering problems. These researches focused on object detection or image classification in the operation and maintenance phase of infrastructure concrete, which tends to monitor and analyze the structure afterwards. Quality control during the construction period can effectively reduce the risk of damage and maintenance costs in the maintenance phase. However, in the concrete construction stage, to the best of our knowledge, there is no relevant research to detect whether the CSI during the construction stage is in conformity with the existing specification. In recent years, many vision-based methods have been proposed to address the image classification problems, namely, traditional machine learning (ML) methods and state-of-the-art deep learning methods. Many scholars have pointed out that machine learning methods have achieved remarkable results in solving image problems in civil engineering [22–24]. Driven by the innovation of new algorithm structure and hardware computing resources, the deep learning methods have achieved explosive development. Many researches have revealed that deep learning methods can achieve better performance than traditional ML methods in some domains in terms of accuracy performance [18,25,26]. Besides, the advantages and superiority of deep learning methods have been proved in concrete research [9,27,28]. Inspired by those research ideas, this study intends to exploit state-of-the-art deep learning technology to detect and analyze the CSI during the vibration process, which classifies the concrete image into various categories belonging to a classification problem. Among the state-of-the-art deep learning-based classification models, the residual neural network (ResNet) [29] achieved an error rate of 3.57% with the ImageNet test and won the first place in the ILSVRC 2015 classification task. Therefore, the ResNet model is utilized to classify the CSI in this study. The specific model structure is described in [Section 4.2.1](#).

2.3. IoT technology for concrete construction

With the popularization of internet and rapid development of sensor technology, the emergence of IoT technology came into being at a historic moment, which has escalated the network ubiquitous capabilities by connecting every object for interaction via embedded systems, thereby constructing a new physical network world. IoT has been widely immersed into various domains, such as civil engineering [30,31], smart cities or building [32,33], health monitoring [34], etc. For example, Zhong et al. [35] developed a real-time construction quality monitoring system to monitor the construction process of RCC dams using IoT technology. This system not only realizes the real-time, on-line, all-weather, and remote monitoring of the compaction process, but also makes the construction parameters (i.e., rolling passes, speed, track, etc.) digital, which effectively improves the compaction quality of RCC dams. Relevant technologies are also widely applied in road rolling construction [36,37]. Real-time monitoring of construction processes based on IoT technology not only improves construction quality, but also makes the construction process traceable. Similarly, for the vibration construction process, vibration depth and CSI are two critical parameters that need to be accurately perceived and analyzed whether they satisfy the requirements of the vibration specification. Therefore, based on the laboratory research regarding the application of IoT technology in construction process monitoring [35,38,39], this study concentrates on applying an infrared sensor to measure vibration depth and a camera sensor to capture CSI during the concrete vibration process using IoT technology.

3. Research gap

Although existing techniques were used to monitor the vibration quality of fresh concrete, there is still an obvious limitation that the CSI could not be analyzed intuitively during the construction phase. This research attempts to use interdisciplinary knowledge to bridge the gaps

from the domains of concrete vibration quality analysis, deep learning application, and IoT-based real-time monitoring application. The specific research gaps addressed by this study are illustrated below:

1. Deep learning techniques have hitherto not been adopted to analyze the surface images of fresh concrete during vibration due to the lack of effective data and professional technical personnel. A solution for this limitation will be stated in this study by using a self-constructed dataset and deep learning technique.
2. Heterogeneous data that include insertion depth (i.e., text-data) and CSI (i.e., image data) have not been simultaneously monitored during the fresh concrete placement phase. This study will address this limitation by utilizing multiple sensors.
3. The real-time analysis of vibration quality, including insertion depth and CSI, is not realized due to the lack of integration between IoT technology and image processing technology. In this paper, IoT technology and deep learning method are integrated to address this limitation in practice.

In response to the above mentioned specific gaps, a real-time monitoring method of concrete vibration quality based on IoT and deep learning technology is proposed. Three specific research objectives that need to be achieved in this paper are as follows:

1. To propose a classification model for CSI based on deep learning using the self-constructed dataset.
2. To develop a monitoring framework based on IoT technology to perceive CSI and insertion depth using multiple sensors. Besides, its performance should be evaluated by comparison with the manual method.
3. To verify the effectiveness and reliability of the proposed method to monitor concrete vibration quality in real-time based on IoT technology and deep learning in a real-world case study.

4. Framework and methodology

4.1. Research framework

The overall research framework of this paper is illustrated in [Fig. 2](#). The main objective of this paper is to construct a real-time monitoring method to assure the vibration quality of fresh concrete using CNN and IoT technology. The contents include IoT-based real-time monitoring for vibration quality and the establishment of CNN-based classification model. In the first part, the IoT-based real-time monitoring framework, including four layers (i.e., sensing layer, network layer, server layer, and application layer), is constructed. The sensing layer includes the infrared sensor to measure the insertion depth and the camera sensor to capture the CSI (For the depiction of other layers, refer to [Section 4.2.2](#)). In the second part, a CNN-based classification model is used to accurately analyze the CSI. The image data are derived from a camera and smartphone. Next, self-constructed datasets are utilized to train the classification model using a fine-tuned ResNet with 50 layers (ResNet-50), and a testing dataset is used to validate the constructed model. Finally, the proposed classification model is utilized to analyze the CSI captured by the camera sensor during the real-time monitoring process.

4.2. Methodology

4.2.1. Deep convolutional neural network for concrete image classification

Deep convolutional neural network [40] is one of the major breakthroughs in computer vision domain. The general CNN architecture includes input, convolution, pooling, activation, and output layers. Compared to the traditional neural network model, the most prominent structural change is the addition of the convolutional and pooling layers, which improves the ability of abstract feature extraction. The convolutional layers are stacked to extract feature maps from the raw image and

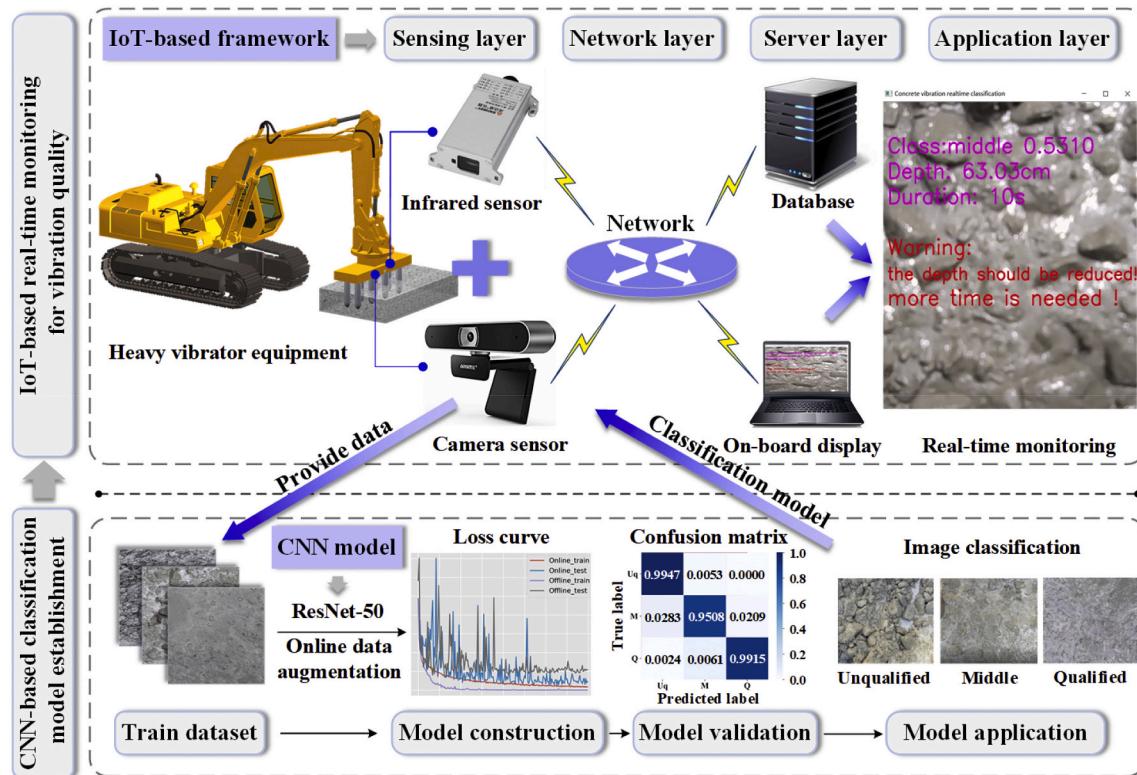


Fig. 2. Research framework.

pooling layers choose abstract features that represent the image well. These changes in the structure promote the CNN model to make a great contribution in the computer vision domain, such as object detection, image classification, etc. In recent years, with the rapid development of CNN model, many outstanding models, such as AlexNet [41], GoogLeNet [42], VggNet [43], and ResNet [29], all of which have won the ImageNet Large Scale Visual Recognition Challenge, promote the development of computer vision [44], and have been successfully applied in solving civil engineering problems [8,9].

Among the aforementioned state-of-the-art models, the ResNet model [29] addressed the degradation problem caused by the increase in

network depth by changing the structure of the network model using a residual network. The two core components of ResNet are identity mapping and residual mapping. Identity mapping is added to the proposed residual mapping structure, which changes the optimization target from $H(x)$ to $H(x)-x$. Hence, the optimization process for residuals mapping can be easily achieved to 0. Thus, the degradation problem can be addressed and the difficulty of optimizing the network parameters is also greatly reduced, theoretically. Two common residual structures of ResNet model are shown in Fig. 3 [45]. As the structure of the deep network model, such as the network depth and residual structure, has a great impact on the performance, this paper adopts ResNet with the 50

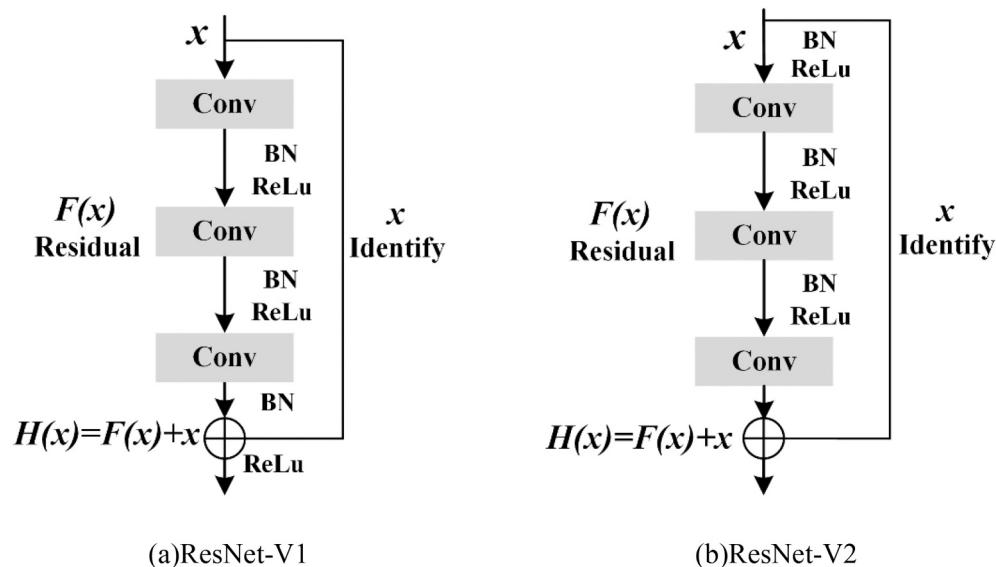


Fig. 3. Schematic of different residual structures.

layers depth and leverages the residual structure shown in Fig. 3(b) to trade-off the computation cost and classification accuracy. Besides, the network model is fine-tuned as the output layer is set to three categories instead of the original 1000 categories to fit this research. The schematic of ResNet-50 used in this paper for image classification is shown in Fig. 4. A detailed structure description can be found in the literature [44,46].

4.2.2. Vibration monitoring using IoT technology

- Framework of IoT-based vibration monitoring

This paper proposes an IoT-based vibration process monitoring framework to sense and analyze the vibration depth and CSI. The monitoring framework is mainly composed of four tiers: sensing layer, network layer, server layer, and application layer, as shown in Fig. 5(a). The sensing layer is the foundation of the framework to sense the vibration depth and CSI using various sensors. The network layer is responsible for the transmission of data and connects various layers, including the wired network, wireless network, and radio transmission. The server layer works on data decoding, data processing, and then stores the data in the server database. The application layer is the head of the framework that analyzes the vibration depth and CSI, and provides early warning and feedback on whether the vibration process satisfies the specification. The detailed research and application refer to other studies of our research team [35,38,39]. As the transmission of live video consumes tremendous amount of resources, only the CSI that satisfies the specification during the vibration process is stored and the live video stream captured by the camera is not stored in the database in this study. The above treatment not only depresses the requirement of the network bandwidth and storage space, but also makes each vibration process traceable. Additionally, it reduces the research cost too.

- Hardware scheme

The hardware of the IoT-based monitoring framework mainly includes a camera sensor, an infrared sensor, a data transmission module, and an on-board display. The detailed parameters of the two sensors are shown in Table 1. A USB port is used to connect the camera with the on-board display, which displays the video stream of the CSI captured by the camera. The infrared sensor includes the transmitter and receiver. The infrared beams from the transmitter are reflected by the concrete surface and then captured by the receiver, thereby measuring the distance between the infrared sensor and the concrete surface. The RS232 serial communication component is utilized to connect the infrared sensor with the data transmission module, which is responsible for packaging and transmitting the data sensed by the infrared sensor to the remote server software through radio transmission and wireless network. Then, the data are processed and stored in the database. The data transmission module is placed inside the vibrator equipment and powered by the vibrator equipment using a stable working voltage under 12 to 24 V. The frequency of the data transmission module is adjustable and is initially set to 10 Hz considering the impact of the

complex construction environment.

- Analysis principle of the vibration process

The analysis and calculation principle of vibration depth and duration are shown in Fig. 5(b). The primary vibration process can be divided into three stages: insertion, vibration, and pulling-out. Generally, the insertion process is relatively fast. The vibration process lasts for a certain time to consolidate the concrete, and finally the vibrator is pulled out slowly to avoid forming the holes left by the vibrator. During the vibration process, the vibrator is required to insert a certain depth to ensure the bond quality of placement between successive layers, of which a certain depth is 5 to 10 cm deeper than the thickness of the current pouring layer [5]. The distance d between the top of the vibrator and concrete surface is measured by the infrared sensor mounted on the vibrator equipment. The vibration depth can be calculated, as shown in Eq. (1). During the pulling-out process, only the CSI (Fig. 5(c)) is classified into qualified category that the vibrators could be pulled out. Thus, the vibration duration can be calculated that it starts when the vibration depth is near the maximum depth until the insertion depth is reduced and finally pulled out. The calculation formula of the vibration duration is shown in Eq. (2).

$$L_{vd} = l_0 - d \quad (1)$$

where l_0 is the length of the vibrator, and d is the real-time data of the infrared sensor.

$$T_{vd} = t_3 - t_1 \quad (2)$$

where t_1 is the time when the vibration depth is near the maximum depth, and t_3 is the time when the vibrator is pulled out.

4.2.3. Real-time monitoring process of vibration quality

The real-time monitoring procedure of concrete vibration quality is illustrated in Fig. 6. Firstly, the classification model of CSI is established based on the deep CNN. The main process includes the acquisition of image data (see Section 5.1), processing of image data and division of samples, and establishment of classification model by training of CNN model. Secondly, the vibration depth and CSI are sensed using various sensors based on IoT technology, which includes the sensing layer, network layer, server layer, and application layer (see Section 4.2.2). Thus, the digital sensing framework of vibration placement process is established. Finally, the CNN-based image classification and IoT technology are applied to real-time monitoring concrete vibration process in a real-world case, where the evolution of the CSI is classified using the proposed CNN-based classification model and the insertion process is monitored using an infrared sensor.

During the application stage, one specific problem that the integration of text and image data needs to be addressed to realize the comprehensive monitoring of vibration depth and CSI. The specific information fusions are as follows. Firstly, the live video captured by the camera is loaded using the OpenCV package in python [47,48]. Secondly, the frames from the live video are parsed using the OpenCV

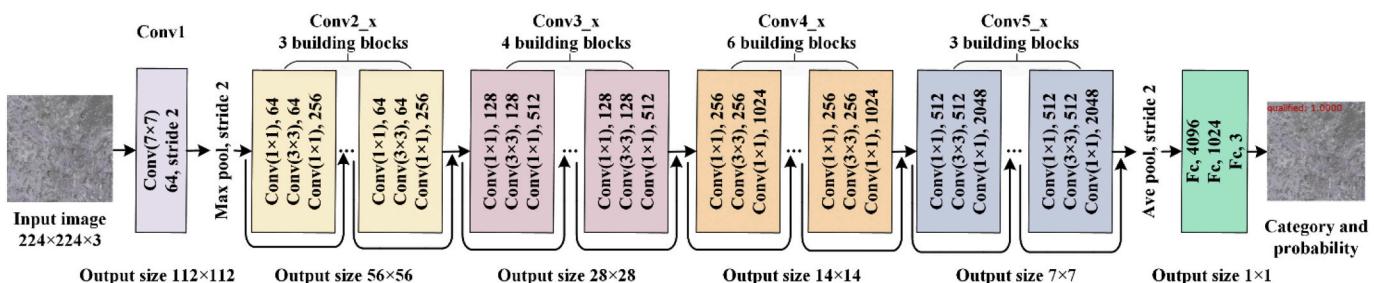


Fig. 4. Schematic of ResNet with 50 layers.

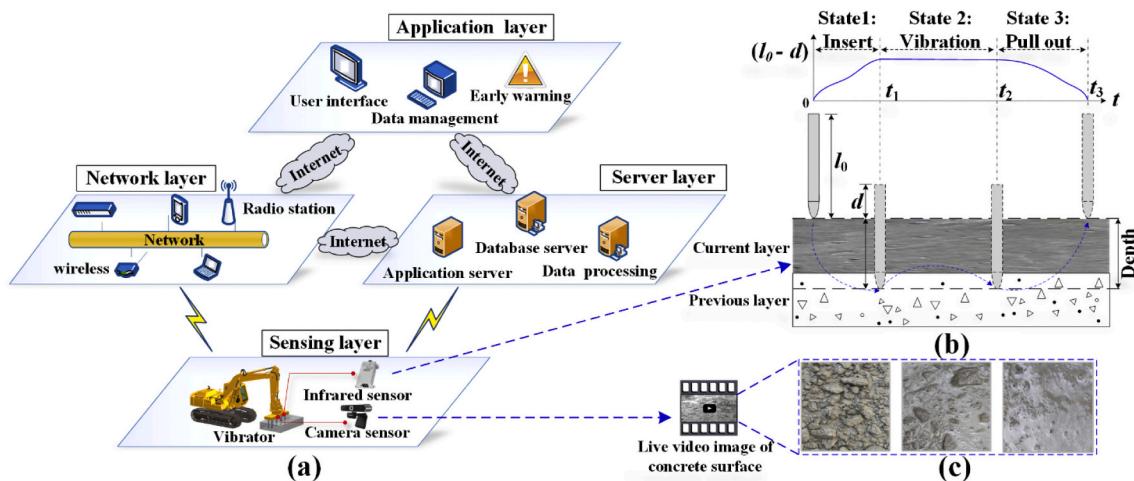


Fig. 5. Vibration monitoring using IoT technology: (a) Architecture of vibration process monitoring; (b) Calculation principles of vibration depth and duration; (c) Perception of CSI.

Table 1
Description of sensors used in the vibration process monitoring.

Commercial name	Technical features	Part number	Applications
Infrared sensor	-Accuracy: $\pm(2 \text{ mm} + 0.05\%)$ -Range: $0 \text{ }^{\circ}\text{C} \sim +40 \text{ }^{\circ}\text{C}$ -Resolution: 1 mm	SW-LDS50A	-Mounted in heavy vibrator equipment -Measure vibration depth
Camera sensor	-Pixel: 20,000 thousand -Resolution: 1920×1080 -Feature: Auto-focus -Max FPS: 30	AONI-A35	-Mounted in heavy vibrator equipment -Capture concrete surface image

package and then transferred to the CSI classification model. Thus, the CSI is classified into a specific category. Thirdly, the vibration depth measured by the infrared sensor is processed and stored in the SQL database, and read using python programming language by connecting the fixed IP address. Finally, the classification results, vibration depth results, and warning information are printed on the live video frame, which realizes the comprehensive information fusion from the camera and infrared sensor. Thus, the real-time monitoring method for vibration quality of fresh concrete can not only assist the operators to timely operate the equipment according to the analysis results, but also reduce the subjectivity of surveillance personnel for judging the CSI and estimation of vibration depth.

5. Case study

In this study, a high arch dam in southwest China is taken as the application example (Fig. 7(a)). The arch dam has a height of 155 m and consists of 17 dam monoliths. The concrete placement of the dam is mainly carried out using heavy vibrator equipment and assisting with manual vibrators. The sensors mounted on the heavy vibration equipment are shown in Fig. 7(b).

5.1. Proposed dataset

5.1.1. Data collection

In this paper, image data are obtained from the recording of the vibration process using the camera and different vibration state images using the smartphone. The camera is mounted on heavy vibration equipment (Fig. 7(b)) and slightly inclined to the concrete surface. To

preprocess the image data, one image data is extracted every 30 frames from the recording video with a pixel resolution of 1920×1080 as same as the camera resolution. To capture the CSI where the camera could not shoot, the mobile smartphone (Xiaomi 5 Pro, Xiaomi Technology Co., Ltd.), which has a resolution of 3000×4000 pixels, is leveraged to capture the static CSI with orthophotos. Considering that the image resolution from the mobile smartphone is larger than that of the camera, to facilitate the uniform pixel resolution of the training samples, the photos taken by the mobile smartphone are divided into 1300×1000 . On one hand, it is beneficial to keep close to the pixels taken by the camera; on the other hand, it can make the most use of the information captured by the mobile smartphone to enlarge the sample size. Thus, the total image data collected in this paper is 7503, of which 3520 images are from the video stream (see Table 2).

In this paper, we divide the CSI into three types, namely, unqualified, middle, and qualified. The example of different categories is depicted in Fig. 8(a). All the images are manually labeled by the author. The typical characteristics are different among the three categories. For instance, in the unqualified category, there are more coarse aggregates with large particle sizes, and almost no laitance suspended on the concrete surface. This is the typical image appearance with a short vibration duration. In the middle category, the size of aggregate particle is smaller than that of the unqualified type with the presence of laitance on the concrete surface. Besides, another critical feature is that the fluctuation level of the surface texture is extreme irregularity; In the qualified category, there are lots of laitance and a few small bubbles, and the surface texture is evolved into relatively flat with a probably small number of micro exposed aggregate.

5.1.2. Data augmentation and statistics processing

Extensive researches have demonstrated that the performance of deep CNN models strongly depends on the availability of large training data [21,49]. Data augmentation [50,51] is an effective method to enlarge the original dataset. Generally, the mode of data augmentation can be categorized into offline and online modes [52,53], both of which may prevent overfitting problem. Offline data augmentation transforms the original image data and stores it on the disk before inputting image into the network model for training. However, online data augmentation transforms the original image data randomly before inputting image into the network model and the processed images are not stored on the disk. Thus, the difference between the two data augmentation modes mainly lies in whether the processed images are stored on the disk and the images need to be processed before inputting to network model for training. To obtain a superior classification model, the offline data

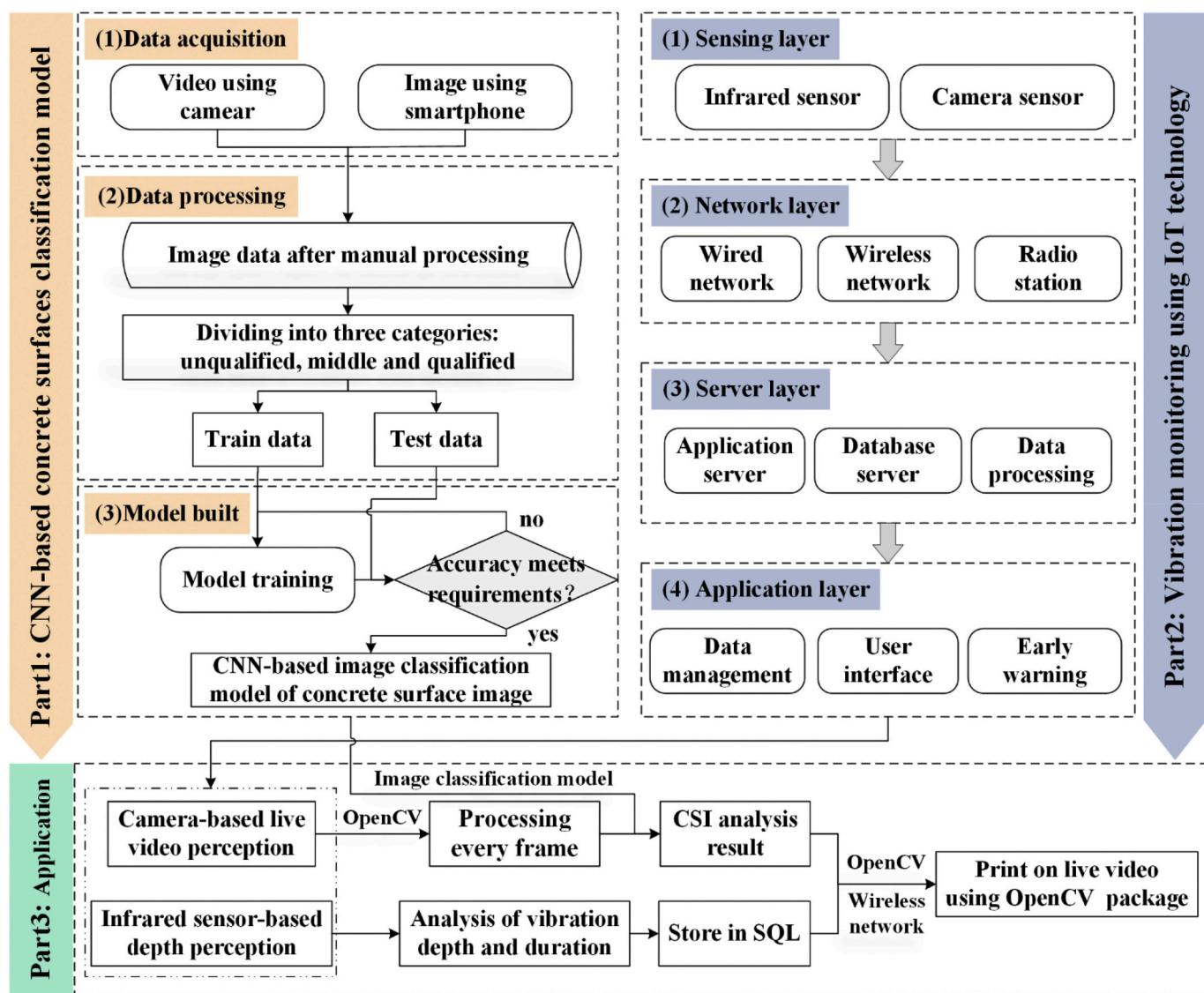


Fig. 6. Real-time monitoring procedure of concrete vibration quality.

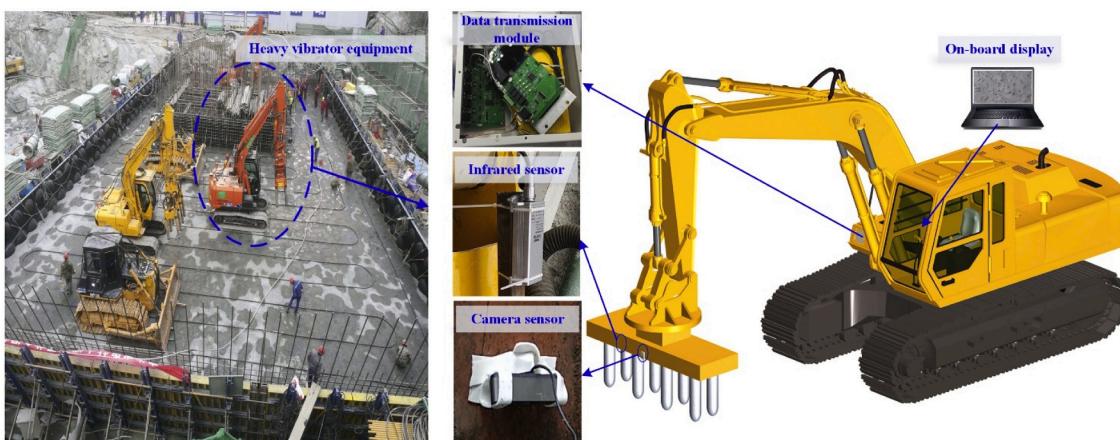
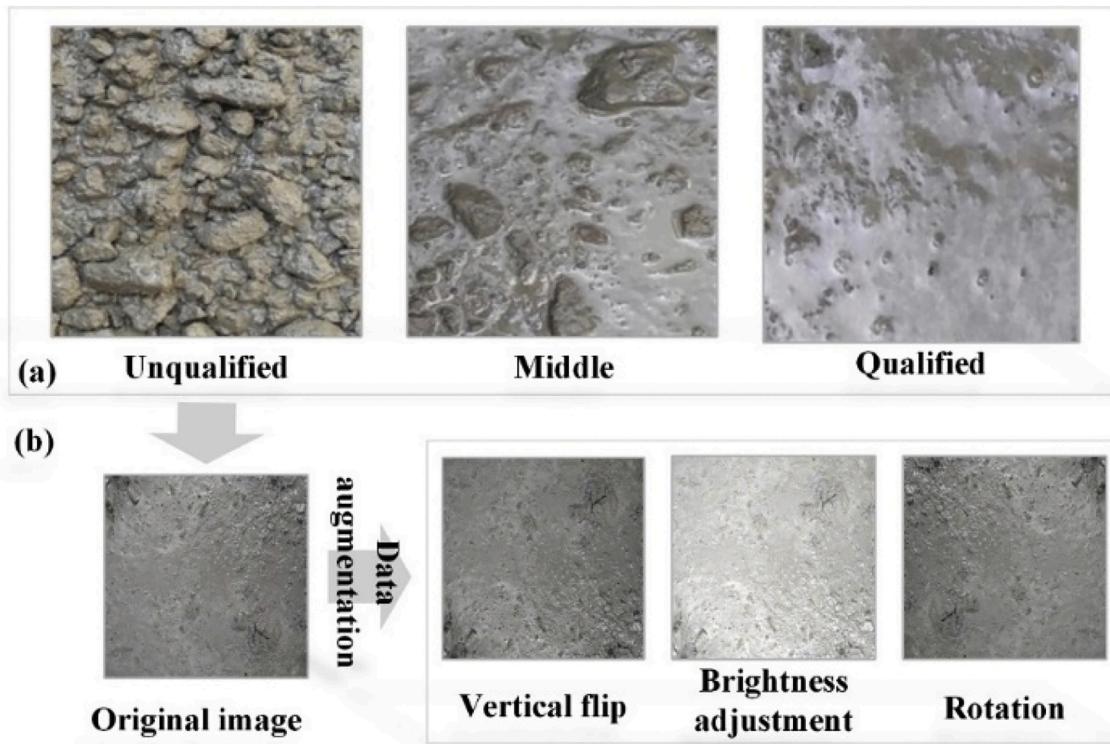


Fig. 7. Experiment and devices: (a) On-site testing scenario; (b) Equipment demonstration, including data transmission module, infrared sensor, camera sensor, and on-board display.

Table 2

Detailed description of the CSI dataset.

Categories	Camera-based method		Smartphone-based method		Total images after augmentation
	No. of images	After augmentation	No. of images	After augmentation	
Unqualified	1255	2510	1273	2546	5056
Middle	1038	2076	1200	2400	4476
Qualified	1227	2454	1510	3020	5474
Total	3520	7040	3983	7966	15,006

**Fig. 8.** Examples of data image: (a) Three categories of CSI; and (b) Data augmentation.

augmentation method is utilized to expand the original samples to 15,006 images and store them on a disk, which is the first stage to process the image data. Then, the 15,006 images are utilized to train the classification model of CSI using online data augmentation before being given to the network, which is the second stage to process the image data. Methods used for data augmentation in this paper include vertical flip, horizontal flip, rotation, and brightness adjustment, as shown in Fig. 8(b). The dataset statistics before and after data augmentation are shown in Table 2. After enlarging the samples, all the sample images are scaled down to the same size of 224 × 224 pixels and divided into three folders annotated in correspond to their categories.

5.2. CNN-based classification model for concrete surface

5.2.1. Model training and implementation

In this paper, a total of 15,006 images are leveraged to establish the classification model of CSI (see Table 2). Meanwhile, the impact between offline mode and online mode on the model performance is explored. Specifically, whether the image data are randomly transformed before they are inputted to the network in the second stage. In the following training process, the training mode that adopts online data augmentation to randomly transform the input images before starting the training network is called the “Online” mode. On the contrary, the training mode that does not randomly transform before starting the training network is called the “Offline” mode. The proposed dataset is

divided into training samples and test samples, with a proportion of 85% (12,756 images) and 15% (2250 images), respectively. The training samples are utilized for the model training process, and the test samples are utilized for model evaluation, which are not used in the model training process. During the model training, the cross-entropy function is adopted, which is a common loss function with a satisfactory performance in deep learning [54]. Additionally, it can quantify the error by comparing the predicted logits of the model and the ground truth during the training process. The stochastic gradient descent [9] optimizer is used for every parameter update in each iteration. The initial learning rate is set to 0.001 [46]. The weight decay and momentum parameters are assigned as 5×10^{-6} and 0.9, respectively. The ResNet-50 model is trained with 200 epochs and the batch size is assigned as 32. To prevent overfitting in the iterative process, early stopping of the deep learning model is applied in the training process [55]. The common evaluation protocols (i.e., precision, recall, and F₁-score) are utilized to evaluate the classification performance of the CSI model. The specific formulas for these protocols can refer to [56].

In our experiment, training is performed on a workstation running Ubuntu 16.04 OS with an NVIDIA TITAN XP consumer GPU and 128 GB RAM using Keras. The testing and real-time monitoring process on the construction site are performed on a PC running Windows 10 OS with an Intel(R) Core(TM) i7-6820HQ CPU @ 2.70 GHz and 32 GB RAM using Keras.

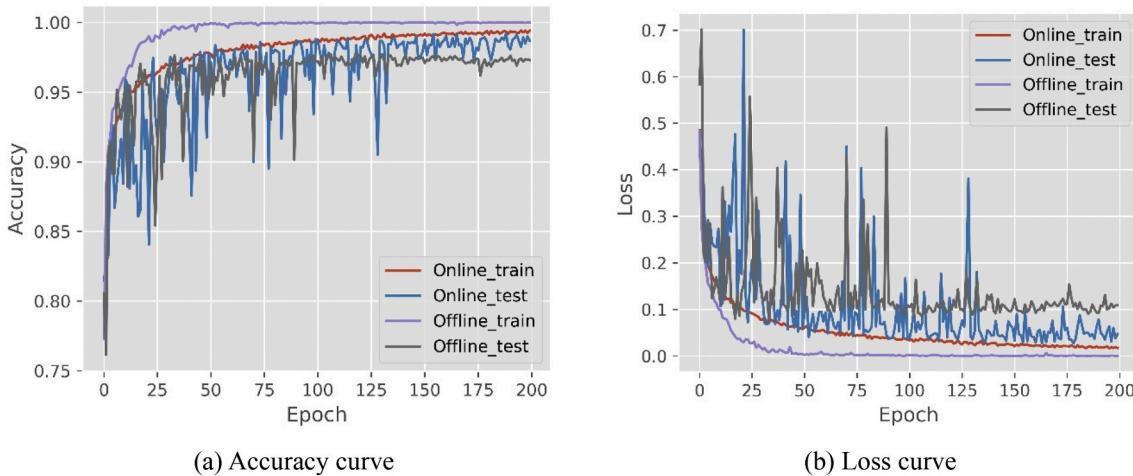


Fig. 9. Accuracy and loss curves of each epoch in training process.

5.2.2. Model testing and classification result

The accuracy and loss curves of the “Online” and “Offline” modes are shown in Fig. 9(a) and (b), respectively, which illustrate the difference in the training process between two data augmentation mode. In the “Online” mode, the training loss curve has a downward trend throughout the training process corresponding to the slow upward trend of its accuracy curve. However, in the “Offline” mode, the training loss curve tends to be zero and the testing loss curve fluctuates around 0.1 without an obvious downward trend after the training epoch is 100. This indicates that the performance of the model tends to be optimal. The same rules can be revealed from its training and testing accuracy curves that also tend to be stable with a little fluctuation. After 200 epochs of training, the values of the loss and accuracy curves of the “Online” mode are within that of the “Offline” mode, which indicates that the model with the “Online” mode achieves a good balance between the training and test samples without overfitting, and has a better generalization performance than the model with “Offline” mode.

The confusion matrix results of the “Online” and “Offline” modes using the test samples are shown in Fig. 10(a) and (b), respectively. In the “Online” mode, the proportion of predicted labels that are consistent with the ground truth is significantly higher than that of the “Offline” mode, where the proportion of the unqualified, middle, and qualified categories is as high as 99.47%, 95.08%, and 99.15%, respectively. The main reason is that the original images are randomly transformed before transferring to the network for training, which enables the CNN model to learn more features, and thus improves the performance of the model.

compared to the “Offline” mode. The classification accuracy of the qualified and unqualified categories of the model is more than 96.95%, while the accuracy of the middle category is relatively low. The main reason may be that the middle category is between them and has the risk of misclassification to the other categories. Besides, the imbalance of the dataset may be another reason leading to the sub-optimal classification performance of the middle category.

The results of classification performance using evaluation protocols are listed in Table 3. For the “Online” mode, the average performance of precision, recall, and F₁-score are 98.07%, 97.90%, 97.97%, respectively, which are better than that of the “Offline” mode with the same average performance of 95.83%. In terms of classification performance of the three categories, the unqualified category has the highest classification accuracy, followed by the qualified category and middle

Table 3
Classification results of CSI.

Metrics	Mode	Concrete image category			Average
		Unqualified	Middle	Qualified	
Precision	Offline	97.36%	93.30%	96.84%	95.83%
	Online	97.29%	98.61%	98.31%	98.07%
Recall	Offline	97.10%	93.44%	96.95%	95.83%
	Online	99.47%	95.08%	99.15%	97.90%
F1-score	Offline	97.23%	93.37%	96.90%	95.83%
	Online	98.37%	96.81%	98.73%	97.97%

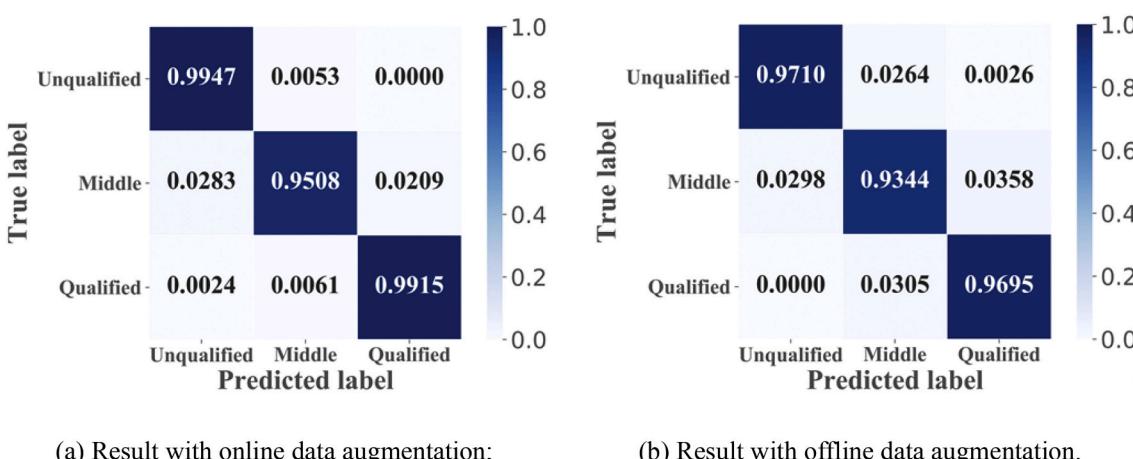


Fig. 10. Confusion matrix of CSI classification.

category. The main reason may be that the concrete surface texture of unqualified class is relatively dry, and the aggregate gaps and particles are relatively large during the initial state of vibration, of which the features are easy to extract and learn using CNN method. For the qualified category, the concrete surface is relatively flat, and small aggregates may interfere with each other. However, the middle category is between them, and has the risk of misclassification to the other categories. The results shown in Fig. 10 coincide with the above explanation. Next, the classification model of CSI using “Online” data augmentation mode will be utilized to classify the CSI on the construction site.

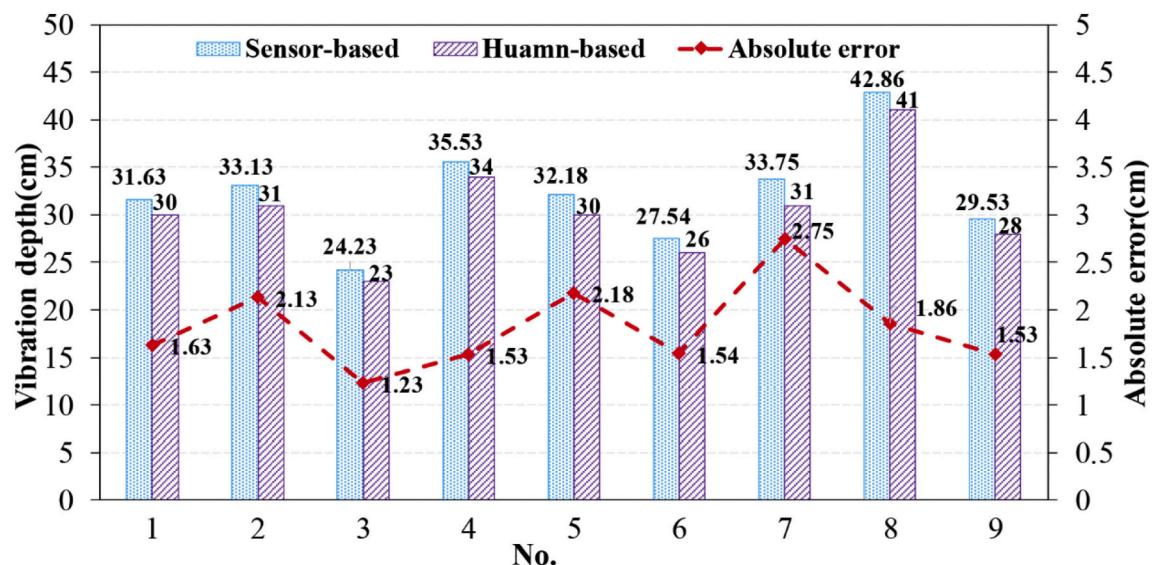
5.3. Assessment of IoT-based performance

To assess the performance of the IoT-based vibration depth monitoring and the classification performance of the proposed model in this paper, nine individual comparison experiments are conducted on-site between the proposed method and professionals. The objective tools,

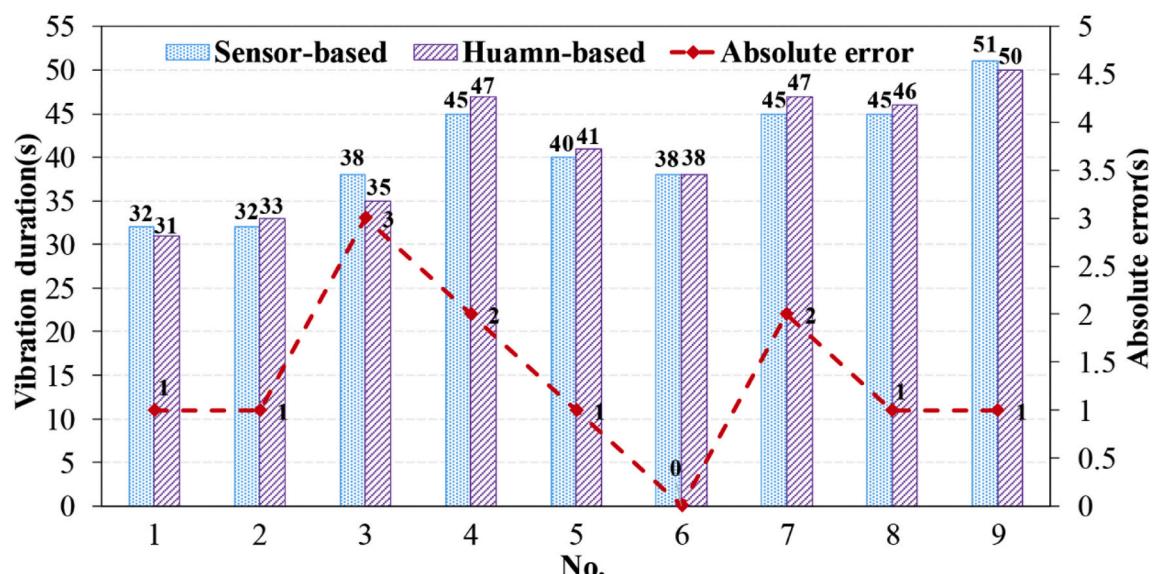
such as measuring tape and stopwatch, are separately used by professionals to verify the vibration depth and duration perceived by the IoT technology. The bias analysis is shown in Fig. 11. Similarly, the contrast results of CSI selected stochastically during different vibration stages are shown in Fig. 12, which are classified into different categories by the proposed classification model and manual inspectors with professional experience. The maximum absolute error of vibration depth (Fig. 11(a)) and duration (Fig. 11(b)) are distinct and within 2.75 cm and 3 s, respectively. The classification results of CSI are consistent with that of the experienced professionals, which verify the high accuracy and reliability of the proposed method.

5.4. On-site application

Field experiments were conducted to verify the real-time monitoring performance of the proposed method. The real-time monitoring results of a complete vibration process are shown in Fig. 13, which are



(a) Bias analysis of vibration depth.



(b) Bias analysis of vibration duration.

Fig. 11. Bias analysis of vibration parameters.

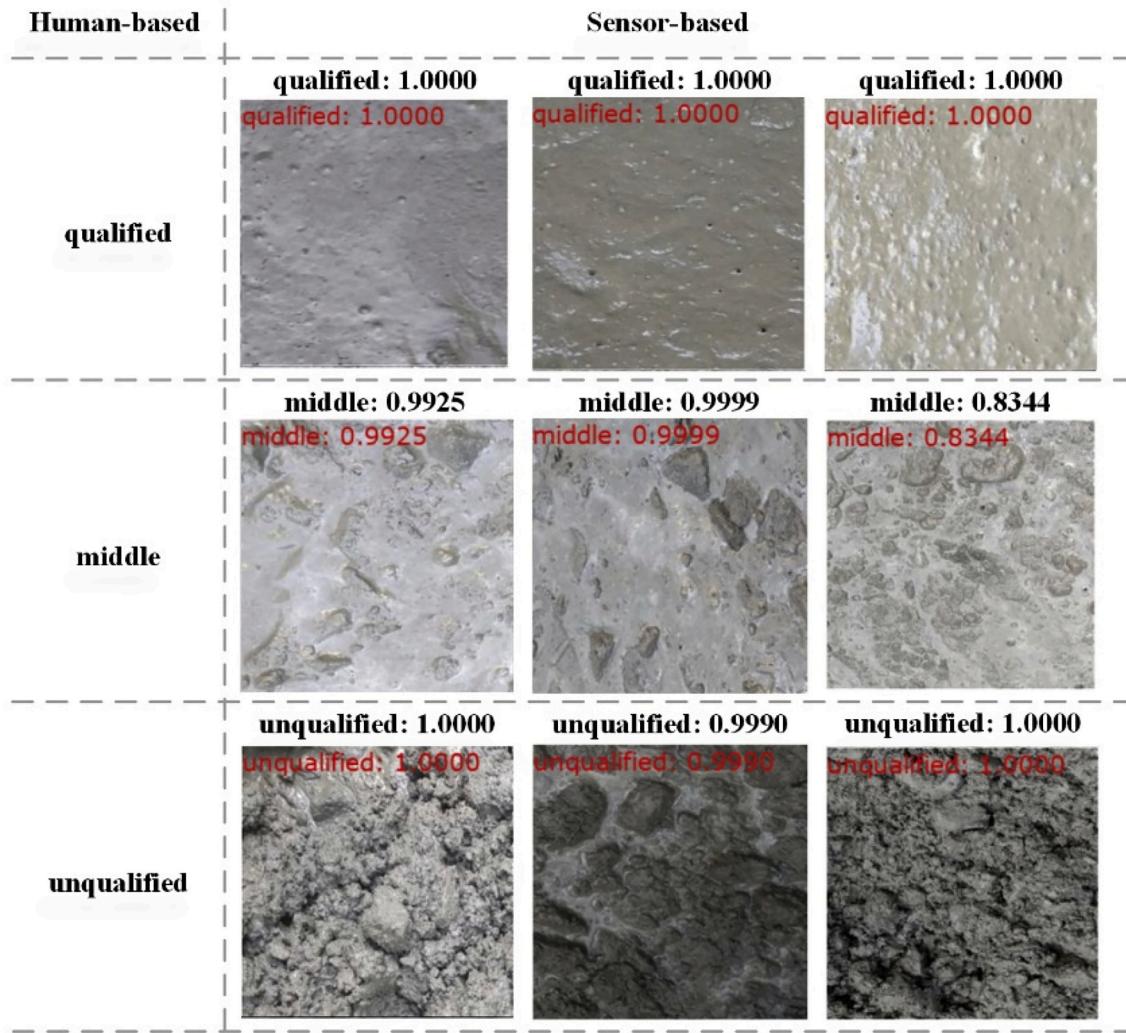


Fig. 12. Comparison between sensor-based and human-based results.

presented with a 5 s interval. The real-time monitoring contents include the category classification of CSI, vibration depth, vibration duration, and early warning information, of which the early warning information is mainly the results comparing the vibration process with the specification. To facilitate the presentation of results, the figure before insertion is shown in Fig. 13(a), the insertion process (stage 1) in Fig. 13(b) to (d), the vibration process (stage 2) in Fig. 13(d) to (l), the process of pulling-out (stage 3) in Fig. 13(l) to (n), and the state of complete pulling-out in Fig. 13(o).

The classification results of CSI show that it slowly transits from the initial unqualified category to the middle category, and finally turns to the qualified category. The colors of the three categories are yellow, pink and green, respectively. The insertion depth gradually increases from zero, and then maintains a certain value with some fluctuation, and slowly decreases until it becomes zero. The time starts from the vibration process and ends when the vibrators are pulled out. The real-time monitoring results of the vibration process (i.e., CSI, insertion depth, and vibration duration) are consistent with the real-world construction process, which verifies the effectiveness and reliability of the proposed monitoring method. It should be noted that during the classification of CSI, the probability of classification gradually reduces from the unqualified category to the middle category, and then gradually increases and tends to be stable after the classification category changes. The same rule is also applicable to the evolution process of images from the middle category to the qualified category, which can be verified from Fig. 13(c)

to (e) (the classification probability changes from 1.0000 to 0.5310, and then increases back to 1.0000) and Fig. 13(f) to (j) (the classification probability changes from 1.0000 to 0.6193 and then increases back to 1.0000), indicating that the classification model used in this paper has good classification performance and reliability.

6. Discussion

As discussed in the research gaps section, the experiment results respond to three research objectives. The first objective, the performance of the proposed classification model for CSI based on deep learning using the collected datasets achieves an average F₁-score of 97.97% (Table 3). The second goal is the performance of IoT-based monitoring framework. Compared to manual recording, the absolute errors of vibration depth and duration acquired from the proposed IoT-based monitoring framework are within 2.75 cm and 3 s (Fig. 11), respectively. Besides, the classification results of CSI using the proposed classification model are consistent with that of the experienced staff (Fig. 12). For the third objective, the real-time monitoring results (CSI, vibration depth, vibration duration) in the field construction process are consistent with the real-world construction process (Fig. 13), which proves the reliability and accuracy of the proposed method for real-time monitoring of the vibration process in a real-world scenario.

To compare the classification performance of different methods, firstly, the traditional classification methods, such as the classic BPNN

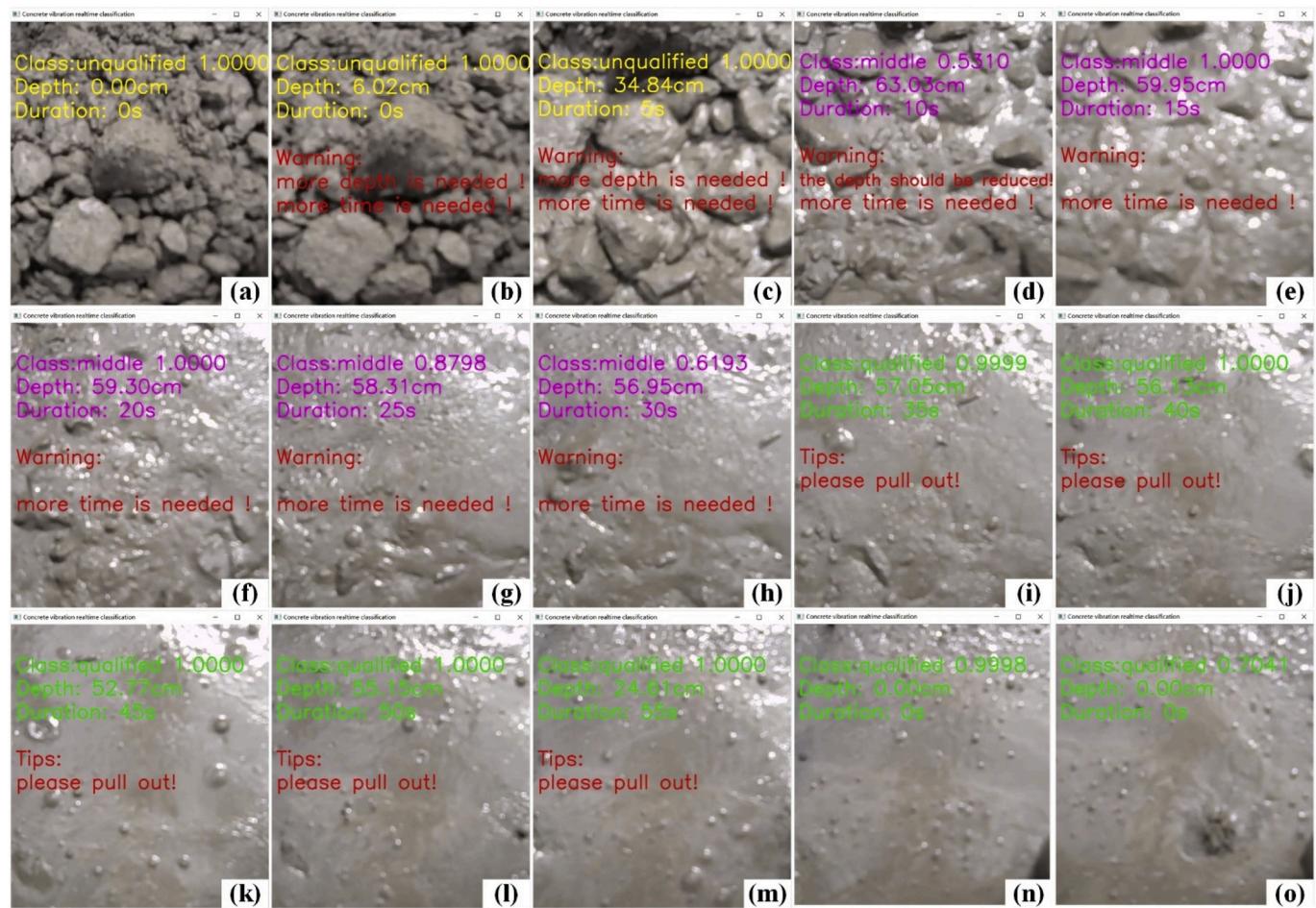


Fig. 13. Real-time monitoring results of live video.

[57] method and traditional machine learning SVM [24] method, are compared with the state-of-the-art CNN method to verify the advantage of CNN method. Secondly, the different CNN network structure models, such as AlexNet [21] and VggNet-16 [56], are compared with the ResNet-50. Finally, the performance of ResNet variant models is discussed in terms of different network depth (i.e., ResNet-18 and ResNet-50) and residual structure(i.e., ResNet-V1 and ResNet-V2), of which the residual structure refer to Fig. 3. It should be noted that the model parameters of the comparison methods are referred from the corresponding literature and then the methods are utilized for CSI classification in this study. The comparative analysis of various classification methods of CSI is shown in Table 4, which reveals the advantages of CNN models over traditional methods in terms of evaluation protocols. Besides, the network depth and residual structure also have a significant impact on the performance of the model.

In this paper, two data augmentation modes are compared to illustrate their impact on the construction of the classification model based on the CNN approach. The performance of the classification model using the “Online” mode during the training process is significantly better than that of the “Offline” mode, which can be used as a trick to improve the model performance, especially when the amount of image data is small.

As the textures of CSI are continuously evolving during the vibration process, whether the proposed model can accurately extract the feature of images and classify the images is a question to be verified, especially the image in adjacent categories. The results in Section 5.4 show that the probability of classification decreases firstly and then gradually increases when the evolution of CSI is from the former category to the next

Table 4
Comparative analysis of different methods for CSI classification research.

Metrics	Method	Category			Average
		Unqualified	Middle	Qualified	
Precision	SVM [24]	79.07%	61.32%	88.64%	76.34%
	BPNN [57]	81.57%	63.04%	89.80%	78.14%
	AlexNet [21]	85.03%	84.93%	85.66%	85.21%
	VggNet-16 [56]	95.86%	80.96%	99.56%	92.13%
	ResNet-18-V2	98.07%	86.83%	96.86%	93.92%
	ResNet-50-V1	99.83%	79.28%	95.64%	91.58%
Recall	ResNet-50-V2	97.29%	98.61%	98.31%	98.07%
	SVM [24]	78.68%	69.00%	79.29%	75.66%
	BPNN [57]	68.45%	76.75%	88.14%	77.78%
	AlexNet [21]	93.42%	64.68%	94.59%	84.23%
	VggNet-16 [56]	97.63%	85.68%	83.07%	88.79%
	ResNet-18-V2	93.94%	94.34%	93.79%	94.02%
F ₁ -score	ResNet-50-V1	77.87%	95.83%	98.78%	90.83%
	ResNet-50-V2	99.47%	95.08%	99.15%	97.90%
	SVM [24]	78.88%	64.94%	83.70%	75.84%
	BPNN [57]	74.44%	69.22%	88.96%	77.54%
	AlexNet [21]	89.03%	73.43%	89.91%	84.12%
	VggNet-16 [56]	96.74%	87.70%	90.57%	91.67%
ResNet-18-V2	ResNet-18-V2	95.96%	90.43%	95.30%	93.90%
	ResNet-50-V1	87.49%	86.77%	97.18%	90.48%
	ResNet-50-V2	98.37%	96.81%	98.73%	97.97%

category. This situation happens not only from unqualified category to middle category, but also occurs from middle category to qualified category, which demonstrates the superior performance and reliability of proposed model to classify the images in adjacent categories.

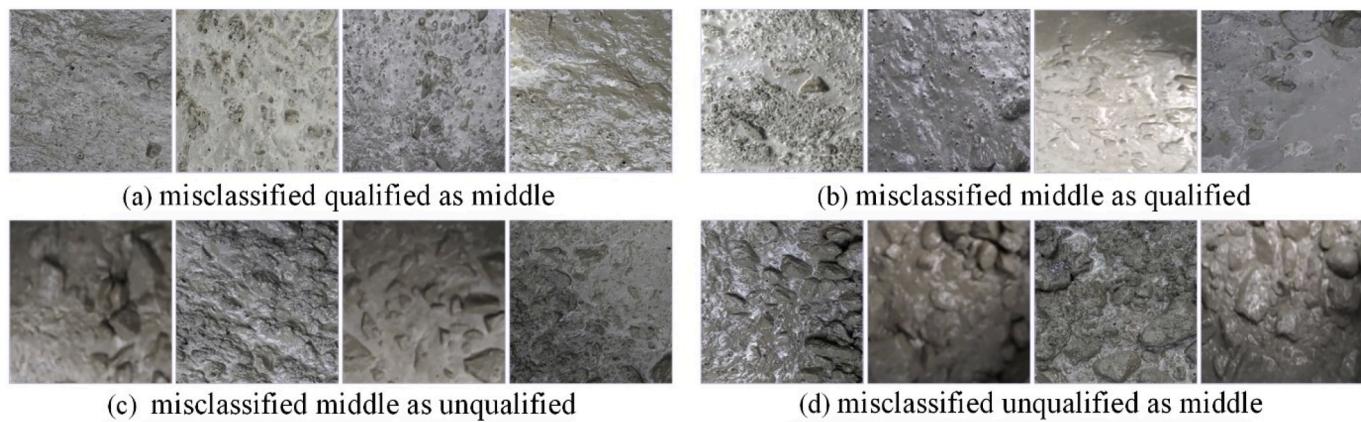


Fig. 14. Examples of misclassified images.

Typical examples of false positives are shown in Fig. 14. The image of the qualified category is wrongly classified into the middle category (Fig. 14(a)) and the middle category is wrongly classified into the qualified category (Fig. 14(b)). On one hand, this can be explained by the fact that the concrete image includes lots of small bubbles and some aggregates simultaneously (the third image in Fig. 14(a) and fourth image in Fig. 14(b)), which is difficult to distinguish artificially in a real-world situation. On the other hand, the broken bubbles on the concrete surface are prone to be identified as aggregate (second image in Fig. 14(a) and (b)), which leads to the misclassification between them. However, a possible way to mitigate the influence of misclassification is by continuously judging the category of CSI and quantitatively judging the probability of recognition. Besides, the middle category and unqualified category are misclassified with each other, as shown in Fig. 14(c) and (d), respectively. One possible explanation is that the boundary between them is difficult to elaborate, and the water retentiveness and fluidity of concrete also affect the classification results. The typical cases are that the small water retentiveness leads to the misclassification as unqualified (third image in Fig. 14(c)), and the large liquidity (second image in Fig. 14(d)) is misclassified as middle. This can be improved with more training samples. Overall, the classification performances of the three categories have achieved a satisfactory result with the lowest F₁-score of 96.81% (Table 3).

7. Conclusion

Vibration is one of the most critical processes in fresh concrete construction, which has a major influence on the long-term safe operation and service life of concrete structures. Proper vibration duration and insertion depth can contribute to concrete consolidation and ensure vibration quality. However, traditional manual inspection on CSI to judge the vibration duration and estimation of vibration depth is subjective. The existing studies on vibration quality assurance monitor vibration duration determined by prior experiment, which not only ignores the influence of concrete heterogeneity, but also cannot satisfy the specification to achieve proper vibration duration directly according to the image characteristics of concrete surface. Thus, a real-time monitoring framework is proposed to ensure vibration quality of fresh concrete based on fine-tuned ResNet-50 model and IoT technology. As the first study applying deep learning and IoT techniques to monitor both the CSI and insertion depth during concrete vibration construction, the major conclusions of this paper are as follows:

- (1) The IoT-based real-time monitoring framework for concrete vibration is proposed, in which an infrared sensor is utilized to measure the vibration depth and a camera sensor is adopted to capture the CSI.

- (2) A new, self-constructed surface image dataset of concrete vibration is established using a total of 15,006 images covering different periods during vibration construction. Then, the three-category classification model of CSI (i.e., unqualified, middle, and qualified) based on fine-tuned ResNet-50 model is constructed, which achieves a remarkable classification performance with an average F₁-score of 97.97%.
- (3) Compared to the manual method, the biases of insertion depth and vibration duration are separately within 2.75 cm and 3 s. The classification results of CSI are in conformance with experienced professionals. These verify the high accuracy and reliability of monitoring results.
- (4) A large-scale and real-life hydraulic engineering application example demonstrates that the real-time monitoring results are in compliance with the vibration process, which verifies the effectiveness and reliability of the proposed real-time monitoring method.

This paper verifies the feasibility and accuracy of the real-time monitoring method for concrete vibration construction. In the future, a larger dataset should be collected and utilized to train model so that it can still achieve appreciable classification performance under complex construction circumstances, such as weak light and drizzle. Besides, due to similar features between the bubbles and aggregates in the CSI, a high-precision fine-grained recognition model should be established to avoid misclassification of images, which has been a hot and difficult issue in classification research. Furthermore, embedding the live video recognition model into the existing real-time monitoring system using faster transmission technology to form a comprehensive system would be more instructive to ensure vibration quality.

Declaration of Competing Interest

The authors declare no conflict of interest.

Acknowledgement

This research was funded by the National Key R&D Program of China (No. 2018YFC0406704), the National Natural Science Foundation of China (No. 51839007), and the Natural Science Foundation of Tianjin (No. 17JCQNJC07100).

References

- [1] S.G. Lee, M.J. Skibniewski, Monitoring of concrete placement and vibration for real-time quality control, Proc. Creat. Construct. Conf. 011 (2019) 67–76, <https://doi.org/10.3311/CC2019-011>.
- [2] S. Popovics, A review of the concrete consolidation by vibration, Mater. Constr. 6 (1973) 453–463, <https://doi.org/10.1007/BF02473784>.

- [3] Z. Tian, X. Sun, W. Su, D. Li, B. Yang, C. Bian, J. Wu, Development of real-time visual monitoring system for vibration effects on fresh concrete, *Autom. Constr.* 98 (2019) 61–71, <https://doi.org/10.1016/j.autcon.2018.11.025>.
- [4] I.C. Ezema, O. Olatunji, Building collapse in Lagos State, Nigeria : towards quality control in materials, batching and placement of concrete, *Coven. J. Res. Built Environ.* 6 (2018) 25–39, <http://journals.covenantuniversity.edu.ng/index.php/cjrbe/article/view/941/615>.
- [5] Ministry of Water Resources of the People's Republic of China, SL677-2014 Specifications for Hydraulic Concrete Construction, China Water&Power Press, Beijing, China, 2014 (in Chinese), <https://www.docin.com/p-2149559997.html> (Jul. 10, 2019).
- [6] S.E. Burlingame, Application of Infrared Imaging to Fresh Concrete: Monitoring Internal Vibration, Cornell University, May 2020. Available from, <http://www.concretepipe.org/secure/newscast/0608nc/0608nc.php> (Accessed date: 25 May 2020).
- [7] J. Gong, Y. Yu, R. Krishnamoorthy, A. Roda, Real-time tracking of concrete vibration effort for intelligent concrete consolidation, *Autom. Constr.* 54 (2015) 12–24, <https://doi.org/10.1016/j.autcon.2015.03.017>.
- [8] B. Zhong, H. Wu, L. Ding, P.E.D. Love, H. Li, H. Luo, et al., Mapping computer vision research in construction: developments, knowledge gaps and implications for research, *Autom. Constr.* 107 (2019) 102919, <https://doi.org/10.1016/j.autcon.2019.102919>.
- [9] Y.J. Cha, W. Choi, O. Büyüköztürk, Deep learning-based crack damage detection using convolutional neural networks, *Comput. Aid. Civil Infrastruct. Eng.* 32 (5) (2017) 361–378, <https://doi.org/10.1111/mice.12263>.
- [10] C.V. Dung, L.D. Anh, Autonomous concrete crack detection using deep fully convolutional neural network, *Autom. Constr.* 99 (2019) 52–58, <https://doi.org/10.1016/j.autcon.2018.11.028>.
- [11] B. Liu, T. Yang, Y. Xie, Factors influencing bugholes on concrete surface analyzed by image processing technology, *Constr. Build. Mater.* 153 (2017) 897–907, <https://doi.org/10.1016/j.conbuildmat.2017.07.156>.
- [12] D. Zhong, Z. Shen, J. Wang, B. Cui, B. Ren, D. Wang, Study on dynamic evaluation of vibration quality of concrete dam based on real-time monitoring, *J. Hydraul. Eng.* 49 (07) (2018) 775–786, <https://doi.org/10.13243/j.cnki.sxb.20180122>.
- [13] X. Gao, J. Zhang, Y. Su, Influence of vibration-induced segregation on mechanical property and chloride ion permeability of concrete with variable rheological performance, *Constr. Build. Mater.* 194 (2019) 32–41, <https://doi.org/10.1016/j.conbuildmat.2018.11.019>.
- [14] T.C. Hutchinson, Z. Chen, Improved image analysis for evaluating concrete damage, *J. Comput. Civ. Eng.* 20 (3) (2006) 210–216, [https://doi.org/10.1061/\(ASCE\)0887-3801\(2006\)20:3\(210\)](https://doi.org/10.1061/(ASCE)0887-3801(2006)20:3(210)).
- [15] B. Liu, T. Yang, Image analysis for detection of bugholes on concrete surface, *Constr. Build. Mater.* 137 (2017) 432–440, <https://doi.org/10.1016/j.conbuildmat.2017.01.098>.
- [16] P. Hüthwohl, R. Lu, I. Brilakis, Multi-classifier for reinforced concrete bridge defects, *Autom. Constr.* 105 (2019) 102824, <https://doi.org/10.1016/j.autcon.2019.04.019>.
- [17] L. Zhang, G. Zhou, Y. Han, H. Lin, Y. Wu, Application of internet of things technology and convolutional neural network model in bridge crack detection, *IEEE Access.* 6 (2018) 39442–39451, <https://doi.org/10.1109/ACCESS.2018.2855144>.
- [18] H. Nhat-Duc, Q.L. Nguyen, V.D. Tran, Automatic recognition of asphalt pavement cracks using metaheuristic optimized edge detection algorithms and convolution neural network, *Autom. Constr.* 94 (2018) 203–213, <https://doi.org/10.1016/j.autcon.2018.07.008>.
- [19] S. Li, Y. Cao, H. Cai, Automatic pavement-crack detection and segmentation based on steerable matched filtering and an active contour model, *J. Comput. Civ. Eng.* 31 (5) (2017) 1–9, [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000695](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000695).
- [20] S.S. Kumar, D.M. Abraham, M.R. Jahanshahi, T. Isley, J. Starr, Automated defect classification in sewer closed circuit television inspections using deep convolutional neural networks, *Autom. Constr.* 91 (2018) 273–283, <https://doi.org/10.1016/j.autcon.2018.03.028>.
- [21] S.I. Hassan, L.M. Dang, I. Mehmood, S. Im, C. Choi, J. Kang, et al., Underground sewer pipe condition assessment based on convolutional neural networks, *Autom. Constr.* 106 (2019) 102849, <https://doi.org/10.1016/j.autcon.2019.102849>.
- [22] C.F.J. Kuo, C.Y. Lai, C.H. Kao, C.H. Chiu, Integrating image processing and classification technology into automated polarizing film defect inspection, *Opt. Lasers Eng.* 104 (2018) 204–219, <https://doi.org/10.1016/j.optlaseng.2017.09.017>.
- [23] H. Seong, H. Son, C. Kim, A comparative study of machine learning classification for color-based safety vest detection on construction-site images, *KSCE J. Civ. Eng.* 22 (2018) 4254–4262, <https://doi.org/10.1007/s12205-017-1730-3>.
- [24] M.R. Halfawy, J. Hengmeechai, Automated defect detection in sewer closed circuit television images using histograms of oriented gradients and support vector machine, *Autom. Constr.* 38 (2014) 1–13, <https://doi.org/10.1016/j.autcon.2013.10.012>.
- [25] A. Zhang, K.C.P. Wang, B. Li, E. Yang, X. Dai, Y. Peng, et al., Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network, *Comput. Aid. Civil Infrastruct. Eng.* 32 (10) (2017) 805–819, <https://doi.org/10.1111/mice.12297>.
- [26] Y. Lecun, Y. Bengio, G. Hinton, Deep learning, *Nature* 521 (2015) 436–444, <https://doi.org/10.1038/nature14539>.
- [27] Y.J. Cha, W. Choi, G. Suh, S. Mahmoudkhani, O. Büyüköztürk, Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types, *Comput. Aid. Civil Infrastruct. Eng.* 33 (9) (2018) 731–747, <https://doi.org/10.1111/mice.12334>.
- [28] P. Hüthwohl, I. Brilakis, Detecting healthy concrete surfaces, *Adv. Eng. Inform.* 37 (2018) 150–162, <https://doi.org/10.1016/j.aei.2018.05.004>.
- [29] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778, <https://doi.org/10.1109/CVPR.2016.90>.
- [30] D.H. Liu, J.J. Chen, S. Li, Collaborative operation and real-time control of roller fleet for asphalt pavement compaction, *Autom. Constr.* 98 (2019) 16–29, <https://doi.org/10.1016/j.autcon.2018.11.005>.
- [31] C. Zhou, L.Y. Ding, Safety barrier warning system for underground construction sites using internet-of-things technologies, *Autom. Constr.* 83 (2017) 372–389, <https://doi.org/10.1016/j.autcon.2017.07.005>.
- [32] A.H. Alavi, P. Jiao, W.G. Buttler, N. Lajnef, Internet of things-enabled smart cities: state-of-the-art and future trends, *Measurement* 129 (2018) 589–606, <https://doi.org/10.1016/j.measurement.2018.07.067>.
- [33] M. Jia, A. Komeily, Y. Wang, R.S. Srinivasan, Adopting internet of things for the development of smart buildings: a review of enabling technologies and applications, *Autom. Constr.* 101 (2019) 111–126, <https://doi.org/10.1016/j.autcon.2019.01.023>.
- [34] P. Verma, S.K. Sood, Fog assisted-IoT enabled patient health monitoring in smart homes, *IEEE Internet Things J.* 5 (2018) 1789–1796, <https://doi.org/10.1109/JIOT.2018.2803201>.
- [35] Y.X. Liu, D.H. Zhong, B. Cui, G.L. Zhong, Y.X. Wei, Study on real-time construction quality monitoring of storehouse surfaces for RCC dams, *Autom. Constr.* 49 (2015) 100–112, <https://doi.org/10.1016/j.autcon.2014.10.003>.
- [36] Q. Xu, G.K. Chang, Adaptive quality control and acceptance of pavement material density for intelligent road construction, *Autom. Constr.* 62 (2016) 78–88, <https://doi.org/10.1016/j.autcon.2015.11.004>.
- [37] W. Hu, X. Shu, X. Jia, B. Huang, Geostatistical analysis of intelligent compaction measurements for asphalt pavement compaction, *Autom. Constr.* 89 (2018) 162–169, <https://doi.org/10.1016/j.autcon.2018.01.012>.
- [38] D.H. Zhong, D.H. Liu, B. Cui, Real-time compaction quality monitoring of high core rockfill dam, *SCIENCE CHINA Technol. Sci.* 54 (2011) 1906–1913, <https://doi.org/10.1007/s11431-011-4429-6>.
- [39] J.J. Wang, D.H. Zhong, B.P. Wu, M.N. Shi, Evaluation of compaction quality based on SVR with CFA: case study on compaction quality of earth-rock dam, *J. Comput. Civ. Eng.* 32 (3) (2018), 05018001, [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000742](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000742).
- [40] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, *Proc. IEEE* 86 (1998) 2278–2324, <https://doi.org/10.1109/5.726791>.
- [41] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, in: Proceedings of the 25th International Conference on Neural Information Processing Systems, 2012, pp. 1097–1105, <https://doi.org/10.1145/3065386>.
- [42] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, et al., Going deeper with convolutions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9, <https://doi.org/10.1109/CVPR.2015.7298594>.
- [43] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, in: International Conference on Learning Representations, ICLR 2015: International Conference on Learning Representations, 2015. <https://arxiv.org/abs/1409.1556>.
- [44] Y. Jang, Y. Ahn, H.Y. Kim, Estimating compressive strength of concrete using deep convolutional neural networks with digital microscope images, *J. Comput. Civ. Eng.* 33 (3) (2019) 1–11, [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000837](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000837).
- [45] K. He, X. Zhang, S. Ren, J. Sun, Identity mappings in deep residual networks, in: European Conference on Computer Vision, 2016, pp. 630–645, https://doi.org/10.1007/978-3-319-46493-0_38.
- [46] N. Wang, X. Zhao, P. Zhao, Y. Zhang, Z. Zou, J. Ou, Automatic damage detection of historic masonry buildings based on mobile deep learning, *Autom. Constr.* 103 (2019) 53–66, <https://doi.org/10.1016/j.autcon.2019.03.003>.
- [47] A. Kazemian, X. Yuan, O. Davtalab, B. Khoshnevis, Computer vision for real-time extrusion quality monitoring and control in robotic construction, *Autom. Constr.* 101 (2019) 92–98, <https://doi.org/10.1016/j.autcon.2019.01.022>.
- [48] M. Kouzehgar, Y. Krishnasamy Tamilselvam, M. Vega Heredia, M. Rajesh Elara, Self-reconfigurable façade-cleaning robot equipped with deep-learning-based crack detection based on convolutional neural networks, *Autom. Constr.* 108 (2019) 102959, <https://doi.org/10.1016/j.autcon.2019.102959>.
- [49] S. Karimpouli, P. Tahmasebi, Image-based velocity estimation of rock using convolutional neural networks, *Neural Netw.* 111 (2019) 89–97, <https://doi.org/10.1016/j.neunet.2018.12.006>.
- [50] C. Shorten, T.M. Khoshgoftaar, A survey on image data augmentation for deep learning, *J. Big Data* 60 (2019), <https://doi.org/10.1186/s40537-019-0197-0>.
- [51] K.M. Rashid, J. Louis, Times-series data augmentation and deep learning for construction equipment activity recognition, *Adv. Eng. Inform.* 42 (2019) 100944, <https://doi.org/10.1016/j.aei.2019.100944>.
- [52] P. Molchanov, S. Gupta, K. Kim, J. Kautz, Hand gesture recognition with 3D convolutional neural networks, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2015, pp. 1–7, <https://doi.org/10.1109/CVPRW.2015.7301342>.
- [53] H. Luo, M. Wang, P.K.Y. Wong, J.C.P. Cheng, Full body pose estimation of construction equipment using computer vision and deep learning techniques, *Autom. Constr.* 110 (2020) 103016, <https://doi.org/10.1016/j.autcon.2019.103016>.

- [54] M. Wang, J.C.P. Cheng, A unified convolutional neural network integrated with conditional random field for pipe defect segmentation, *Comput. Aid. Civil Infrastruct. Eng.* 35 (2) (2020) 162–167, <https://doi.org/10.1111/mice.12481>.
- [55] L. Prechelt, Early Stopping - but when?, *Neural Networks: Tricks of the Trade*. Lecture Notes in Computer Science, 2012, pp. 53–67, https://doi.org/10.1007/978-3-642-35289-8_5.
- [56] C.V. Dung, H. Sekiya, S. Hirano, T. Okatani, C. Miki, A vision-based method for crack detection in gusset plate welded joints of steel bridges using deep convolutional neural networks, *Autom. Constr.* 102 (2019) 217–229, <https://doi.org/10.1016/j.autcon.2019.02.013>.
- [57] Y.X. Fan, Q.L. Zhao, S.D. Ni, T. Rui, S. Ma, N. Pang, Crack detection based on the mesoscale geometric features for visual concrete bridge inspection, *J. Electron. Imag.* 27 (5) (2018) 1–11, <https://doi.org/10.1117/1.JEI.27.5.053011>.