Advances in Research

Volume 25, Issue 5, Page 272-280, 2024; Article no.AIR.123837
ISSN: 2348-0394, NLM ID: 101666096

Identification of Concrete Crack Using Deep Learning Based Approach

Ziyan Bian [a*]

[a] *School of Civil and Transportation, North China University of Water Resources and Electric Power, Zhengzhou-450045, Henan, China.*

Author's contribution

The sole author designed, analysed, interpreted and prepared the manuscript.

Article Information

DOI: https://doi.org/10.9734/air/2024/v25i51160

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/123837

Opinion Article

Received: 24/07/2024
Accepted: 26/09/2024
Published: 01/10/2024

ABSTRACT

Cracks reflect the safety and durability of concrete structures, and the existing artificial crack detection has the disadvantages of low efficiency and large error. However, the use of deep learning of images to identify cracks has the advantages of high efficiency, small error and low cost. This paper systematically discusses the deep learning in the identification of concrete structure cracks, expounds the deep learning technology, studies the SqueezeNet network model and YOLO (You Only Look Once) network model in the field of concrete structure crack identification, and improves the model according to the test results. In this paper, the classification accuracy based on crack trend is improved by increasing the width and depth of the Fire module of the SqueezeNet network. Secondly, the target detection accuracy is improved by the transfer learning of the YOLO model. The results show that the classification accuracy of the improved SqueezeNet model is more than 82% and the recognition accuracy of YOLO model is more than 92%.

Keywords: Deep learning; crack recognition; SqueezeNet model; YOLO model.

**Corresponding author: E-mail: bianzy2022@163.com;*

***Cite as:** Bian, Ziyan. 2024. "Identification of Concrete Crack Using Deep Learning Based Approach". Advances in Research 25 (5):272-80. https://doi.org/10.9734/air/2024/v25i51160.*

1. INTRODUCTION

The emergence of cracks will seriously affect the strength and durability of the building structure, and the emergence and expansion of cracks are also the performance of structural safety issues [1-2]. It is the key to ensure the safety and reliability of concrete structure to carry out health monitoring of concrete structure performance, avoid the aggravation of structural damage and timely repair and maintenance.

Deep learning is an integral part of artificial intelligence. Using deep learning to identify apparent cracks in concrete structures can not only reduce labor costs, but also improve the accuracy of crack identification. Zhang [3] proposed a supervised crack classification algorithm based on CNN. This algorithm uses 500 pavement crack images for training and generates a CNN model. Yun [4] modified the G-set network by using the Global average pooling layer, the Batch Normalization layer and the Adjust Softmax layer to prevent the occurrence of over-fitting, and then adjusted the learning rate, activation function and convolution kernel size to complete the improvement of the model. The experimental results show that the improved accuracy reaches 0.9630 and the F1 score is 0.9623. Bibi [5] compared and analyzed ResNet18 and VGG11, and concluded that both models had good detection performance in target detection, and their prediction accuracy reached 100%. Alipour [6] proposed the identification of concrete cracks based on residual convolutional neural networks. At the same time, three methods suitable for this training were proposed and tested: joint training, sequence learning, and integrated learning. The training accuracy rate reached more than 97%. Adel [7] used U-Net network to detect cracks and pits on the surface of concrete structures. Through image enhancement, the training set was expanded from 1600 to 6400, and the results showed that the accuracy rate was 99.76 %. Yang [8] used FCN network to identify and segment cracks at different scales based on semantics, and then used single-pixel width skeleton to represent crack recognition and quantitatively measure the morphological characteristics of cracks. The prediction results show that the accuracy of crack segmentation is 97.96 %, and the recall rate is 81.73%. Cui [9] carried out real-time supervision of the bridge based on the combination of robot operating system and YOLOv3 algorithm, and the recognition accuracy of bridge cracks reached 82.3 %. Domestic and foreign scholars have not done much research on YOLOv2 model in building structure damage identification.

In addition, the target detection of concrete structure cracks based on YOLO model has not been studied much at home and abroad, and the research on target detection of cracks using YOLOv2 model is even rarer. In the application of practical engineering, In practical engineering applications, some special scenes have very high requirements for model accuracy and speed, so a detection algorithm that can achieve real-time detection while meeting the requirements of recognition accuracy is urgently needed.

2. BASIC INTRODUCTION OF DEEP LEARNING

Deep Learning (DL) is an overall architecture constructed by combining mathematical knowledge with computer algorithms, and then combines massive training data and computer computing power to adjust parameters. Compared with traditional machine learning, deep learning is more efficient and reasonable.

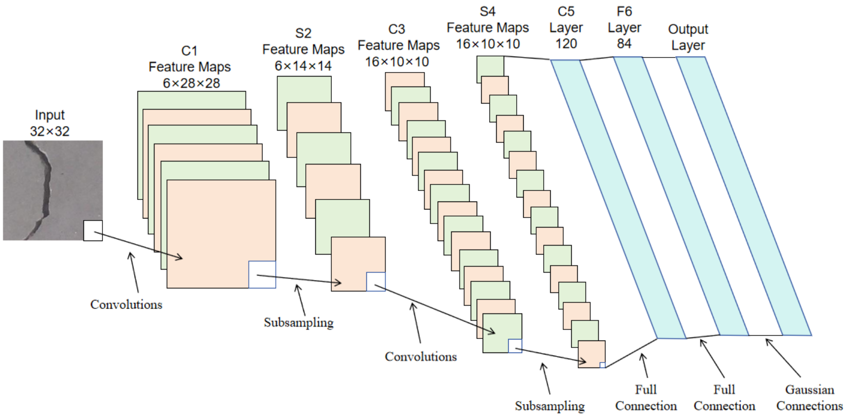


Fig. 1. Typical convolutional neural network

Deep learning is a machine learning method that builds and trains models based on neural networks. Neural network is one of the core components of deep learning. The neuron applies the activation function to the weighted sum and outputs the function value. It can be known that the artificial neuron is essentially a computing unit. Convolutional neural network is a multi-layer nonlinear feedforward neural network. It was originally used for feature recognition and extraction of complex images to solve the problem that traditional image recognition technology can only extract shallow features of images and realize automatic processing of complex images. Typical convolutional neural networks include input layer, convolution layer, activation layer, pooling layer, fully connected layer and output layer [10], and typical convolutional neural networks [11] are shown in Fig. 1.

3. CONCRETE CRACK RECOGNITION BASED ON IMAGE DEEP LEARNING

The identification of cracks mainly includes two major tasks: crack classification and crack target detection. This study is based on Matlab 's Deep Learning Designer function to form a deeper and more complex network model by building the basic units mentioned above, so as to identify cracks in deep learning.

3.1 Classification of the Crack Images

Image classification is the task of assigning corresponding labels to images in a given classification set. For the cracks of the building structure, due to their different forms, the structural damage states reflected are also different. The SqueezeNet structure is a lightweight and efficient CNN model proposed by Han [12] and others. Its parameter number is only 1/50 of Alex-Net, but the model performance such as accuracy is close to Alex-Net. The SqueezeNet is connected with 8 Fire modules, where the first convolutional layer, Fire4 module and Fire8 module are connected with max pooling layer after it, and the tenth convolutional layer is followed by an average pooling layer, which greatly reduces the number of parameters. In order to transform the output of the neural network into a probability distribution, a Softmax layer is added at the end of the model for classification. The core of SqueezeNet is the Fire module as shown in Fig. 2.

The SqueezeNet network is less used in the field of architecture, and the standard for relevant scholars to use the SqueezeNet network to classify cracks is not high enough. It is only judged whether there are cracks in the image as shown in Fig. 3. In practical engineering, concrete cracks can be divided into transverse cracks, longitudinal cracks, and oblique cracks according to their inclination angles, and the existence of cracks is no longer used as a classification standard. This paper will use Matlab software to construct the network model of crack classification based on deep learning, and then train and test the dataset.

In order to reduce the workload of the computer, the crack training set selects 120 high-quality and representative crack images, including 40 transverse cracks, 40 longitudinal cracks, and 40 oblique cracks. After the crack classification is completed, the folder is named as the corresponding crack type name, and then imported into the data of the convolutional network. Because the deep network requires a large amount of training data to better summarize and achieve good accuracy, in order to improve the accuracy, the image enhancement operation is performed on the data set. Image enhancement is the process of generating new images for training deep learning models. The imported data is randomly rotated, shifted, cut, and flipped to process the original training set to achieve the purpose of expanding the number of training sets.

Through 30 iterations of the fracture data set, the accuracy function of the training set and the verification set of the initial SqueezeNet model are shown in Fig. 4.

The final accuracy of the loss function of the training set is 33.33 % after 30 iterations. It can be seen from the Fig. 4 that the function value tends to be gentle in the later stage of training, and the little change proves that there is no phenomenon of gradient disappearance, and the predicted value is close to the target value. The module can be fused according to the feature maps of different sizes and the training effect is good. Therefore, inspired by the Inception module, the SqueezeNet is improved.

In order to further expand the data, a column of multi-layer networks is connected in parallel in the Expand layer of the Fire module. Firstly, the maximum pooling layer is added to reduce the dimension of the data, and then the convolution layer with the convolution kernel of 1 × 1 is connected to amplify the data. Finally, the ReLU and BN layers are connected to linearly correct and normalize the data. Different convolution kernel sizes provide different receptive fields, so that the model can extract features from different levels. The result matrices processed by different convolution layers are integrated in the deep dimension to form a deeper matrix., which not only improves the accuracy of the improved convolutional neural network. And prevent the occurrence of over-fit. The purpose of adding the maximum pooling layer is to reduce the dimension of the larger matrix to ensure that the visual information of the multi-size feature map is aggregated and extracted while reducing the dimension of the large-size matrix.

The improved SqueezeNet network is trained after setting the same hyperparameters as the unimproved SqueezeNet. The progress map after training is shown in Fig. 5, and the final verification accuracy is 83.33 %.

3.2 Target Detection of Cracks

The YOLO model comes from 2016.Joseph, Santosh Divvala, Ross Girshick [13] proposed a single-stage target detection network YOLO. The model deals with the object detection problem as a regression problem. A convolutional neural network is used to predict the bounding box and category probability of the input image, and the regression method is used to detect the target. The execution speed is fast, and the characteristics of the target object can be extracted faster and more effectively, so as to improve the accuracy of the target detection.In order to better understand the YOLO model, this paper first introduces the concepts of bounding box, anchor box and confidence.

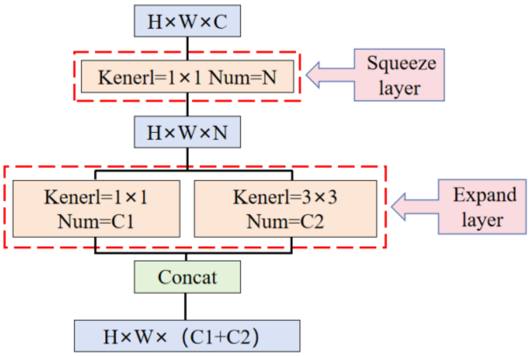


Fig. 2. Fire module

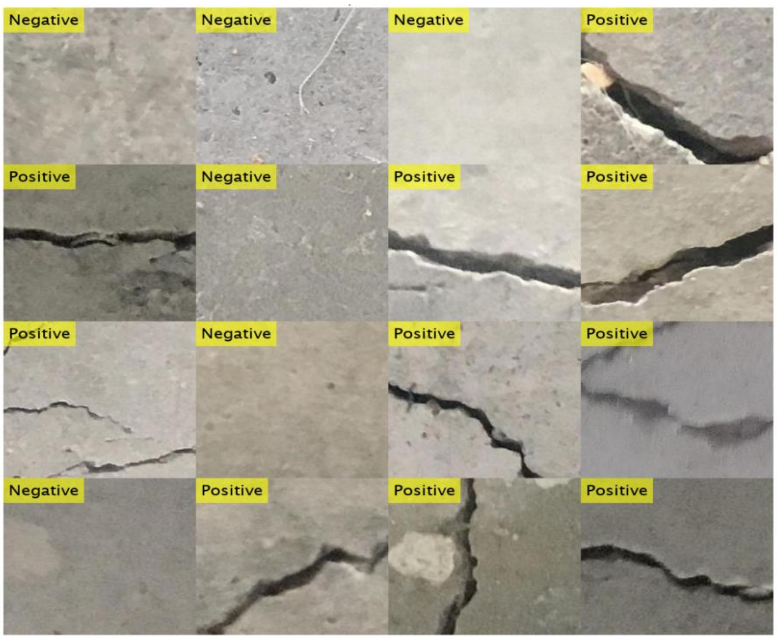


Fig. 3. Classification of concrete crack images

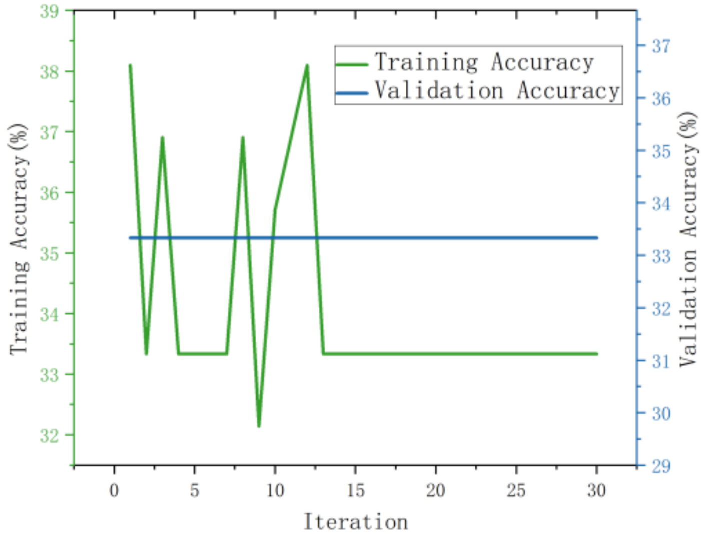


Fig. 4. Accuracy function of the original SqueezeNet

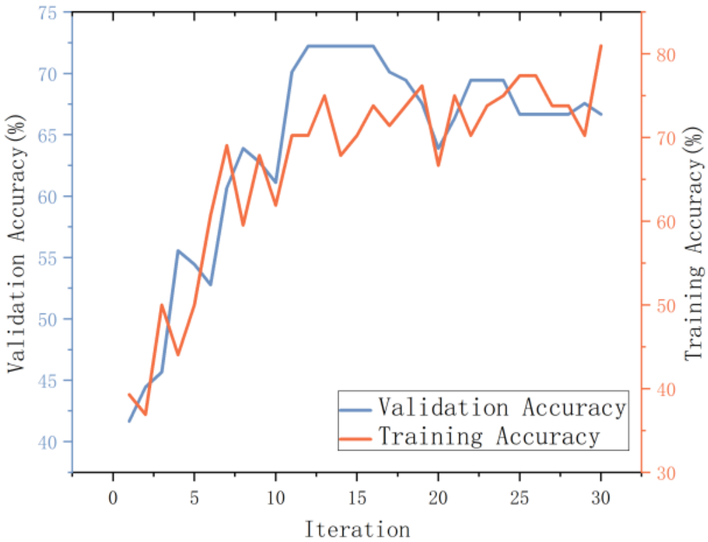


Fig. 5. Accuracy function of the improved SqueezeNet

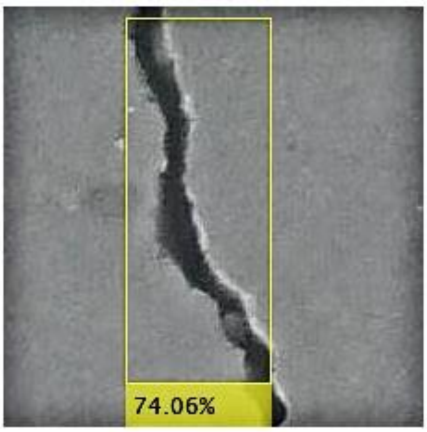


Fig. 6. The experiment result of YOLO model

The bounding box which can contain the rectangular box of the object. When constructing the data set, Image Labeler is used to manually label the training set and select the target object, so that the neural network can learn the characteristics of the target. In the later test, the model will identify the target and automatically select the target detection object.

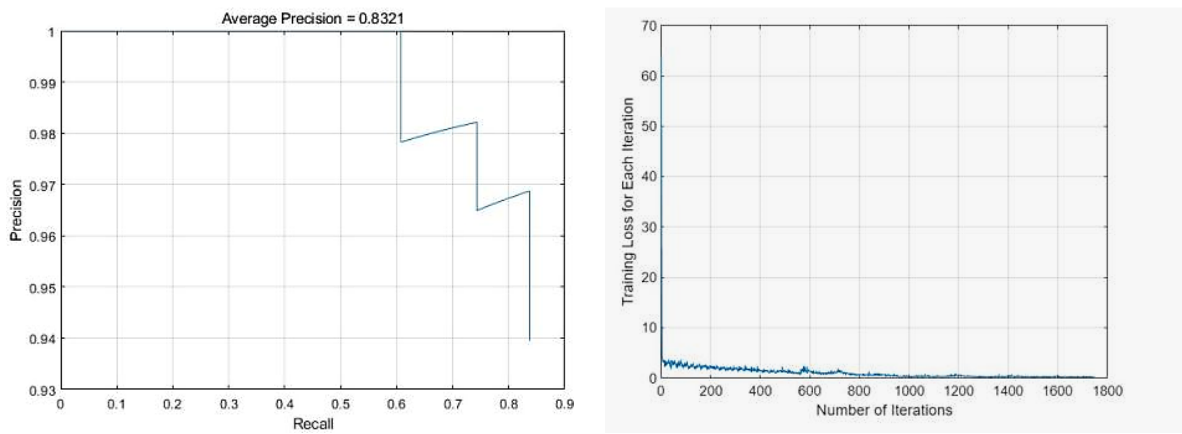
The anchor box is a rectangular box generated by setting the size, number, shape and then taking a point of the image as the center. The center is the anchor point. In target detection, a series of anchor boxes are generated on the picture with some rules. The anchor boxes are used as possible candidate regions, and the anchor boxes are used to make a preliminary prediction box. The model predicts whether these candidate regions contain objects. If the object is included, the model predicts the object category.

The advantage of YOLO is that YOLO uses a CNN network to achieve target detection, and its training is end-to-end. Therefore, the algorithm of YOLO model is simple and fast. The speed of the algorithm is 100 times that of fast-RCNN and 1000 times that of RCNN. In addition, compared with YOLO, other traditional methods can only predict based on local information of pictures.

In the experiment result (as shown in Fig. 6), it is shown that the mean average accuracy of the YOLO model is low. The mAP of the improved YOLOv2 model increases with the increase of the image size under different image input sizes, but the calculation speed decreases significantly. Due to the inconsistent size, the number of information carried by the image is also different, and the calculation parameters are also inconsistent, so the calculation speed will be different. The larger the parameter, the slower the calculation speed.

The YOLOv2 model deletes the fully connected layer and the Dropout layer but adds the BN(Batch Normalization) layer. The Dropout layer randomly discards data according to its own set parameters. If the Dropout layer is connected before the BN(Batch Normalization) layer, the neurons in the Dropout layer will generate Variance Shift during training. YOLOv2 uses Anchor, and the model is predicted by the priori box obtained by the K-means clustering algorithm. Compared with the initial YOLO, the model predicts the coordinates based on the offset scaling of the anchor point. The skeleton network of the model is DarkNet19, which well balances the contradiction between accuracy and calculation speed. Through the Image Labeler function, it uses the bounding box to mark it and generate a mat file. In this study, YOLOv2 is used for migration training, and the network used to train and identify cars is used to train and identify cracks. The pre-marked crack image mat file is imported for training. After 100 rounds of training, a total of 1750 iterations are obtained. The average accuracy of the final training is 83.21 % and the Loss function converges as shown in Fig. 7.

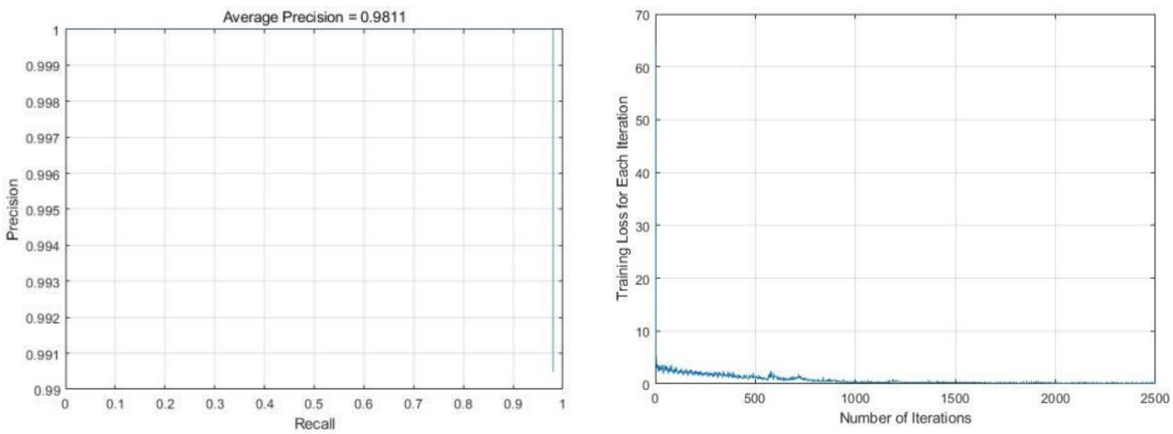
In order to further improve the average accuracy of YOLOv2, the network model is increased by training times for a total of 2500 iterations. The average accuracy of the model is 98.11 % and the Loss function converges as shown in Fig. 8.



a. P-R curve

b. Loss function

Fig. 7. The result of training



a. P-R curve b. Loss function

Fig. 8. The result of training

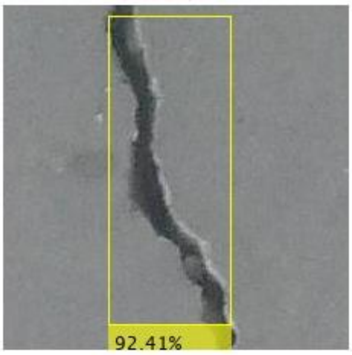


Fig. 9. Test result

Through the target detection of cracks, it is found that the increase of iterative rounds obviously makes the recognition accuracy higher and the crack location correct. This shows that the migration learning of the crack detection model based on YOLOv2 is more successful.

After the establishment of the network model, the model evaluation index should be used for evaluation. The evaluation indexes selected in this paper are Precision, Recall and F1-score.

Before introducing the evaluation index, the four indexes of TP, TN, FP and FN are introduced, which are the parts of the accuracy rate, recall rate and F1 score. P and N represent the detected state, and T and N represent whether the detection is wrong. The positive case is the crack, and the negative case is the non-crack. The following is an explanation of the four indicators.

Precision represents the proportion of the actual positive samples in the samples predicted by the model. For the crack detection of concrete structures, it is the proportion of the actual cracks selected by the bounding box. Recall represents the proportion of samples predicted by the model as positive samples in the actual positive samples, that is, the proportion of all crack areas in the picture selected by the bounding box. The precision and recall rate are contradictory, and neither of them can achieve the maximum at the same time. The F_β scoring rule can comprehensively consider the accuracy and recall rate.

$$\Pr ecision = \frac{TP}{TP + FP}$$

$$\mathrm{Re}\, call = \frac{TP}{TP + FN}$$

$$F_\beta = (1+\beta^2)\frac{\Pr ecision \times \mathrm{Re}\, call}{\beta^2 \times \Pr ecision + \mathrm{Re}\, call}$$

where β presents the importance ratio of precision rate to recall rate is generally considered to be consistent with the importance of precision rate and recall rate.

In order to verify the crack target detection after YOLOv2 transfer learning, this study randomly selected a crack image from the number of test sets to verify the accuracy of the verification results as shown in Fig. 9.

4. CONCLUDING REMARKS

In this paper, a method of crack recognition of concrete structure based on image deep learning is proposed. The inclination degree of cracks is classified by SqueezeNet network, and the target detection of cracks is carried out by YOLO network. The effectiveness of the method is verified by computer analysis. The main conclusions are as follows:

1) After the improvement of the Fire module of the SqueezeNet, the classification efficiency of the model can be improved. The classification model has good effect and high accuracy.
2) In this paper, the YOLO model is used to detect the cracks of concrete structures. Through transfer learning training, the accuracy of crack target detection is more than 90 %. Increasing the number of training iterations of the data set and repeatedly training the same model can improve the recognition accuracy of the model.

In recent years, the research and application of digital image recognition of cracks in reinforced concrete structures have been developed rapidly. However, in the complex environment of practical engineering, the shortcomings of crack recognition are also exposed. For example, the depth of cracks cannot be detected by digital images. Therefore, in the future research, the detection ability of crack depth can be supplemented by means of concrete radar, ultrasound, laser and so on. At the same time, it can also be combined with UAV and other equipment to detect and identify cracks in high-rise structures, which also lays a foundation for the repair and reinforcement of concrete structures in the future.

DISCLAIMER (ARTIFICIAL INTELLIGENCE)

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc) and text-to-image generators have been used during writing or editing of this manuscript.

COMPETING INTERESTS

Author has declared that no competing interests exist.

REFERENCES

1. Shu J, Honfi D, Plos M, et al. Assessment of a cantilever bridge deck slab using multi-level assessment strategy and decision support framework [J]. Engineering Structures. 2019;200:109666
2. Shu J, Zhang Z, Gonzalez I, et al. The application of a damage detection method using artificial neural network and train-induced vibrations on a simplified railway bridge model . Engineering Structures. 2013;52:408-421.
3. Zhang L, Yang F, Zhang Y D, et al. Road crack detection using deep convolutional neural network[C]//2016 IEEE international conference on image processing (ICIP). IEEE. 2016:3708-3712.
4. Que Y, Dai Y, Ji X, et al. Automatic classification of asphalt pavement cracks using a novel integrated generative adversarial networks and improved VGG model [J]. Engineering Structures. 2023;277:115406.
5. Bibi R, Saeed Y, Zeb A, et al. Edge AI-based automated detection and classification of road anomalies in VANET using deep learning[J]. Computational intelligence and neuroscience, 2021;2021: 1-16.
6. Alipour M, Harris D K. Increasing the robustness of material-specific deep learning models for crack detection across different materials[J]. Engineering Structures. 2020;206:110157.
7. Adel M, Yokoyama H, Tatsuta H, et al. Early damage detection of fatigue failure for RC deck slabs under wheel load moving test using image analysis with artificial intelligence[J]. Engineering Structures. 2021;246:113050.
8. Yang X, Li H, Yu Y, et al. Automatic pixel-level crack detection and measurement using fully convolutional network[J]. Computer-Aided Civil and Infrastructure Engineering. 2018;33(12): 1090-1109.

9. Cui Mida, Wang Chao, Chen Jinqiao, Dai Jiao, & Wu Gang. Real-time detection system of concrete bridge cracks based on ROS and YOLOv3. Journal of Southeast University/Dongnan Daxue Xuebao. 2023,53(1).
10. Yang Shuo, Xu Qingfeng, & Wang Zhuolin. Research progress on structural damage identification based on convolutional neural network. Journal of Architecture & Civil Engineering. 2022;39(4).
11. Lecun, L. Bottou, Y. Bengio, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE. 1998;86(11):2278-2324.
12. Iandola NF, Han S,Moskewicz WM, Ashraf K, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size.[J]. CoRR; 2016; abs/ 1602.07360.
13. Joseph Redmon, Santosh Divvala, Ross Girshick, et al. You Only Look Once: Unified, Real-Time Object Detection[C]. //2016 IEEE Conference on Computer Vision and Pattern Recognition: CVPR 2016, Las Vegas, Nevada, USA, 27-30 June 2016;761-1515, [v.2].: Institute of Electrical and Electronics Engineers. 2016;779-788.

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/123837