

Research Article

Concrete Crack Detection Algorithm Based on Deep Residual Neural Networks

Xiuying Meng 

Henan Vocational College of Water Conservancy and Environment, Zhengzhou 450000, China

Correspondence should be addressed to Xiuying Meng; mengxiuyinghb@163.com

Received 29 April 2021; Revised 22 May 2021; Accepted 26 May 2021; Published 24 June 2021

Academic Editor: Shah Nazir

Copyright © 2021 Xiuying Meng. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Crack is the early expression form of the concrete pavement disease. Early discovery and treatment of it can play an important role in the maintenance of the pavement. With ongoing advancements in computer hardware technology, continual optimization of deep learning algorithms, as compared to standard digital image processing algorithms, utilizing automation of crack detection technology has a deep learning algorithm that is more exact. As a result of the benefits of greater robustness, the study of concrete pavement crack picture has become popular. In view of the poor effect and weak generalization ability of traditional image processing technology on image segmentation of concrete cracks, this paper studies the image segmentation algorithm of concrete cracks based on convolutional neural network and designs an end-to-end segmentation model based on ResNet101. It integrates more low-level features, which make the fracture segmentation results more refined and closer to the practical application scenarios. Compared with other methods, the algorithm in this paper has achieved higher detection accuracy and generalization ability.

1. Introduction

Concrete structure [1] is the basic element of constructing urban environment, which needs to be maintained systematically. Crack [2] is a common damage form of concrete structure, which is one of the important indexes to evaluate the safety of concrete structure [3, 4]. For bridges [5–7], roads [8, 9], and other engineering structures [10, 40], it is necessary to carry out regular management and maintenance; crack detection [11–14] is an important part of management and maintenance. At present, the crack detection is mainly artificial detection, which has a lot of problems, such as in the crack detection of the bridge, the staff need to observe the cracks on the bottom and side of the bridge by building scaffolding, which requires a lot of preparation work and low detection efficiency. The safety of staff working in high places is also difficult to be guaranteed. There are blind spots in the manual detection process, and the inspection effect is often affected by the professional accomplishment and experience level of the detection personnel, and the detection accuracy is low. Due to these

shortcomings of manual detection methods, a crack damage detection method based on crack image is developed [15–18].

The traditional detection of crack damage is based on pictures. In the early stage, the detection method based on morphology is often low and its detection accuracy is often low and is greatly affected by light and noise. In the later stage, methods such as seed points and tensor voting are used. Better detection accuracy however calculations are frequently complex, and computation efficiency is low. For crack pictures with complex backgrounds, false detections and missed detections may occur, so it is necessary to develop other concrete crack detection methods.

In recent years, with the continuous improvement of Internet technology [19] and the support of related hardware, artificial intelligence technology with deep learning [20] as the core has developed rapidly. Convolutional neural network is a feed-forward neural network that applies deep learning. It has significant advantages such as high parallelism, good robustness, and strong generalization ability. Due to the characteristics of weight sharing, the

convolutional neural network has a much smaller amount of calculation than the BP neural network, which reduces the burden on the computer and does not require manual feature extraction. These all make it have a wide range of applications in the field of computer vision, especially in image classification and detection. When there are enough samples, the classifier of the trained convolutional neural network can accurately identify concrete damage. Nowadays, there are many convolutional neural network models with more efficient training and more accurate classification every year. The detection of concrete cracks through convolutional neural networks helps to reduce the subjective influence of people and is more efficient and economical. It is an excellent solution to the problem of crack detection. Following are the main innovative points of this paper:

To propose an improved fully convolutional neural network model based on ResNet101, using ResNet101 as the backbone network for feature extraction and then fusing the extracted features of different scales with the up-sampling and restoring feature maps. Because it combines shallow position information and high-level semantic information, it makes the image edge detail segmentation more refined.

To use the concrete crack detection method based on deep residual neural network proposed in this paper is a nondestructive detection technology, which has urgent needs and extremely high application value in the field.

The rest of this paper is organized as follows: Section 2 shows the related work of the paper. Section 3 discusses the methodology of the paper. The experiments and results are shown in Section 4. Section 5 shows the conclusion of the paper.

2. Related Work

In recent years, convolutional neural networks have been widely used in the field of civil engineering. For cracks and other structural damage problems, scholars usually use classified convolutional neural networks to locate the damaged parts and use semantic segmentation networks to describe the damaged parts at pixel level to achieve the description of the damage.

Classical convolutional neural networks include AlexNet [21] and GoogLeNet [22] which have specific requirements on the size of the input image. When using these networks, the image to be tested needs to be divided into small samples of specific sizes. The network is trained to classify these samples, determine the position of the crack sample on the picture to be tested, and realize the location of the crack in the picture.

Dorafshan et al. [23] compared the detection effects of common edge detectors and deep convolutional neural networks on concrete cracks. In this article, six edge detection methods, Roberts, Prewitt, Sobel, Laplacian of Gaussian, Butterworth, and Gaussian, are used to detect concrete crack images. The final binary image produces residual noise, and the best method can only detect cracks with a width greater than 0.1 mm. When using the deep

convolutional neural network (DCNN) AlexNet, the correct rate of labeled samples reached 99%. The network pair in the migration learning mode can correctly detect 86% of the crack samples and can detect cracks with a width greater than 0.04 mm, while the network in the fully trained mode can also detect cracks greater than 0.08 mm. In terms of time, if the training process of the network is not considered, the calculation time of DCNN is also shorter than the most effective edge detection algorithm. These all show the superiority of using convolutional neural network to detect concrete cracks.

Kim and Cho [24] used concrete pictures on the Internet to construct a dataset, which included five categories, such as cracks and plants. AlexNet is trained by transfer learning, that is, the network parameters of the pretrained AlexNet are fine-tuned by using the existing dataset of concrete images. In the test process, two kinds of sliding windows with 50% overlap rate were used to scan the image so that the cracks appearing on the edge of window 1 appeared in the central area of window 2, so as to avoid the cracks being missed due to appearing on the edge of window, and the detection effect was enhanced by adjusting the threshold value of classification layer.

The classification network locates the cracks based on the crack samples and cannot accurately describe the shape of the cracks. Semantic segmentation is to classify each pixel in the image. Therefore, when the semantic segmentation network is applied to crack detection, the precise location and shape description of the crack can be realized at the same time.

Ye et al. [25] used full convolutional networks (FCN) to identify cracks in concrete bridge images at the pixel level and compared them with the results of edge detection methods. The results show that the edge detection method is greatly affected by noise, and when there is enough training data collected from the real situation, the full convolutional network (FCN) can eliminate a lot of noise interference, the performance is reliable, and the position of the crack in the image is displayed, and the ability of the path, its detection, has stronger robustness. Du et al. [26] proposed a new model based on the improvement of the original DeepLab v2 to adapt to the particularity of crack damage detection. Compared with the previous method, this method can obtain very high-precision output, and a new method of marking cracks is proposed, which is beneficial to the measurement of crack length and width. Hoskere et al. [27] proposed a structural damage detection method based on multiscale pixel-level deep convolutional neural network [28–31]. This method utilizes two networks [32], derived from VGG19 and ResNet45, for structural damage classification and semantic segmentation, respectively. By parallel, the two networks can simultaneously process multiple types of damage classification and pixel-level segmentation. Due to the high accuracy of damage detection in the classification network [33, 34], the segmentation network only needs to perform semantic segmentation [35] on the images judged as damage and does not need to identify the nondamage parts, which reduces the probability of false detection. Hoskere constructed a dataset of 1695 images from 250 different

structures, including six types of damage. On this dataset, the applicability of the method to civil infrastructure was verified, and the damage classification was achieved at the pixel level.

3. Methodology

Some scholars have proposed a variety of methods based on digital image processing to detect cracks on the surface of concrete structures, but there may be many problems such as uneven illumination intensity and distortion in the actual collected images, and the traditional image processing algorithms cannot solve these problems well. Due to its powerful feature extraction ability [36], deep learning technologies [37–39] can better overcome the interference brought by external environmental factors, and it has better detection accuracy compared with traditional methods. Therefore, this paper proposes a concrete crack detection algorithm based on deep residual neural network to achieve pixel-level segmentation detection of concrete crack images. Next in this section, the proposed network model for crack segmentation and the whole training process are described in detail.

In the maintenance of the surface of concrete structures, cracks are one of the most common and serious diseases and are the early features of many other diseases, as shown in Figure 1.

3.1. Segmentation Detection Algorithm Based on Deep Residual Neural Network. At present, most of the semantic segmentation network models based on deep learning adopt encoder-decoder network architecture, which is the mainstream structure of semantic segmentation network models due to its simple structure, strong scalability, and good segmentation effect. In this paper, the encode-decoder architecture is still adopted, and by combining the advantages of FCN network model and U-Net network model, an improved full convolutional neural network (CrackNet model for short) is proposed. CrackNet also contains the narrowing path for collecting image features and the expanding path for locating and restoring features. The two are symmetrical structures and constitute a U-shaped network. The deeper and more generalizable ResNet101 network is used in the CrackNet encoder. Compared with VGG16 network, ResNet101 network has fewer parameters and can extract deeper features. In the decoder, the advanced semantic features encoded are decoded and restored to the size of the original image by bilinear interpolation and convolution operation. Another important feature of CrackNet is that it splices and integrates the features of different scales extracted in the process of down-sampling with the feature images restored from up-sampling, which further enhances the segmentation effect of crack edge details. Figure 2 is the network structure diagram of this paper.

3.2. Model Training. The training process of our model mainly includes the selection of parameters such as fracture dataset annotation, loss function, optimization algorithm,

and learning rate. Finally, this experiment completed the network model training under the Keras deep learning framework.

3.2.1. Data Labeling and Data Augmentation. Application scenarios of crack detection in this paper are mainly surface structures composed of concrete such as bridges and roads, but currently there is no standard crack dataset specifically used for semantic segmentation. In order to obtain enough concrete crack images to train the model, part of crack datasets was found in some papers, and a large number of crack images were supplemented by manual photography. Before training the network model, these collected crack images need to be annotated first. In this paper, the professional semantic segmentation and annotation software Labelme is used to annotate the crack images. There is a significant difference in the target detection task dataset of annotation and semantic segmentation, as in the mark of target detection. You only need to paint the rectangular box in the target and the appropriate categories, whereas annotation and semantic segmentation are more difficult. In particular, this slender aim requires cracking along the margins of each point to indicate until the contour map is complete. It forms an enclosed space. Figure 3 shows the process of labeling crack images with Labelme software.

For semantic segmentation tasks, a large number of training samples is the key factor to improve the segmentation effect of network models, but sometimes it is not easy to obtain a large number of datasets. An effective way to improve performance is to enlarge the existing data and expand the diversity of the existing data. This article uses Augmentor, a very useful Python data enhancement tool, to augment datasets. Its operation is very simple. According to the original image and label, the program performs rotation, stretch, zoom-in, and zoom-out of the labeled image with a certain probability to complete the expansion of the dataset. After augmenting the data by Augmentor library, this paper finally obtained 3000 fracture datasets, among which 2500 data were selected as the training set, and the remaining 500 were used as the test set.

3.2.2. Loss Function. The cross-direction loss function is one of the most widely used loss functions in deep learning. This article adopts the cross-direction loss function, but it is different from the cross-direction loss of the traditional image classification network in that the loss function in this paper needs to predict and evaluate each pixel and then calculate the pixel average. If the crossover loss is greater, the corresponding gradient will be greater and the training speed will be faster. Since this crack image segmentation task belongs to a two-class classification problem, the calculation equation of the cross-function is as follows:

$$L = y \log y' + (1 - y) \log (1 - y'), \quad (1)$$

where y is the truly marked crack value, y' is the predicted value of the model, L is the cross-child loss calculated according



FIGURE 1: Cracks on the concrete surface.

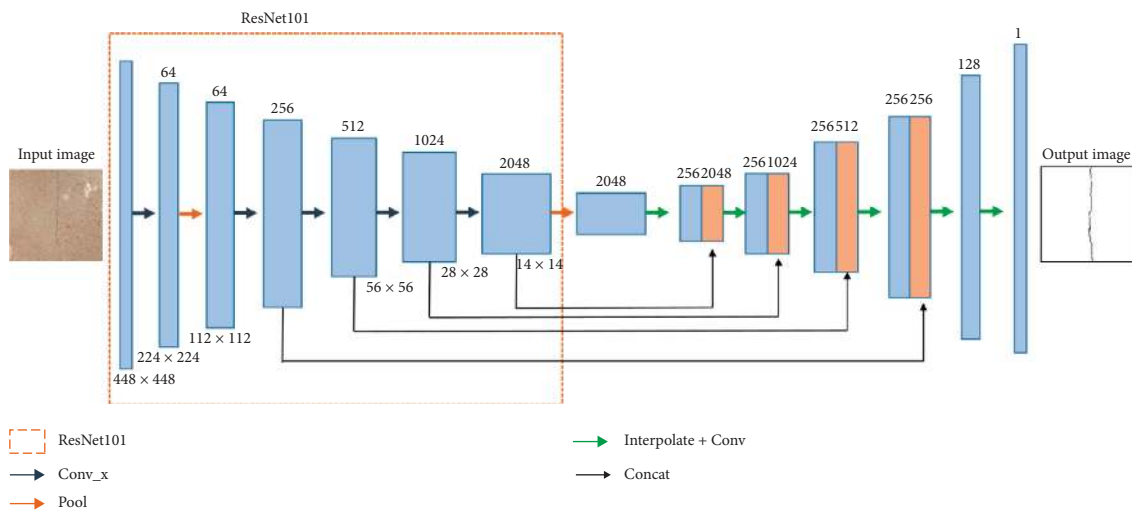


FIGURE 2: Network structure of the proposed model.

to the cross-child loss function, and the overall loss of the image is equal to the average loss value of all pixel points.

4. Experiments and Results

4.1. Experimental Setup. All the experiments in this section were performed on a Lenovo workstation with Nvidia GeForce GTX 1080 Ti and PyCharm installed. In this paper, the previously produced fracture dataset is divided into 2500 training sets and 500 test sets. Since the color information is of little significance to the concrete image and will take up too much storage space, in order to reduce this redundant information, this paper converts the image of three-channel concrete cracks captured into single-channel grayscale image. In order to maximize the use of computer memory

without causing memory overflow, the batch-size of this experiment is 4, the size of the input image is 448×448 , the loss function is the cross-saturated loss function, the optimization algorithm is the gradient descent method of driving quantity, the learning rate is 0.001, and the momentum coefficient γ is 0.9. The number of iterations is 100 epochs.

4.2. Evaluation Method. Precision and recall are two basic quantitative evaluation indexes in semantic segmentation. FP represents the total number of pixels that were wrongly judged as cracks in the background, and TP represents the total number of pixels that were correctly extracted from the cracks. FN represents the total number of pixels that belong

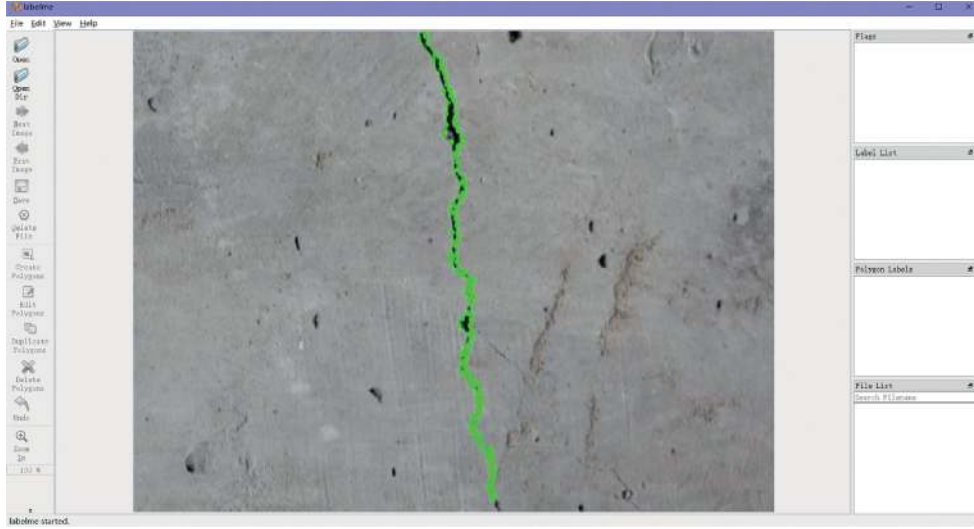


FIGURE 3: Data labeling.

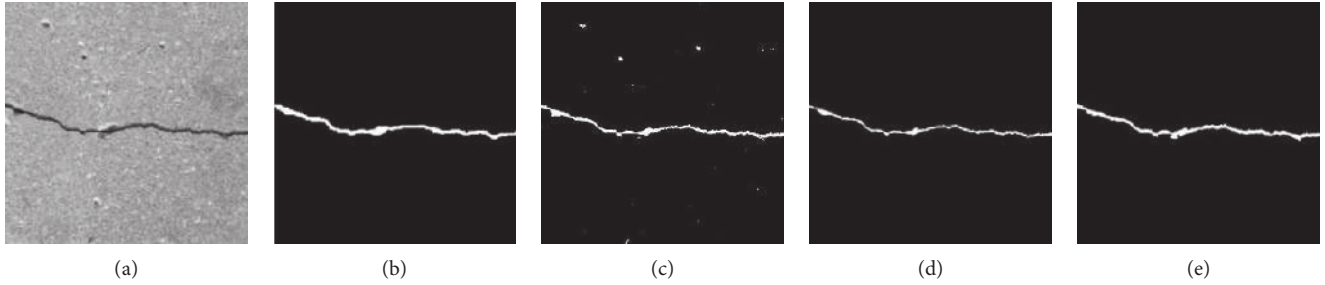


FIGURE 4: Visualization results of segmentation detection. (a) Input image. (b) Ground truth. (c) FCN-8s. (d) FCN-32s. (e) Ours.

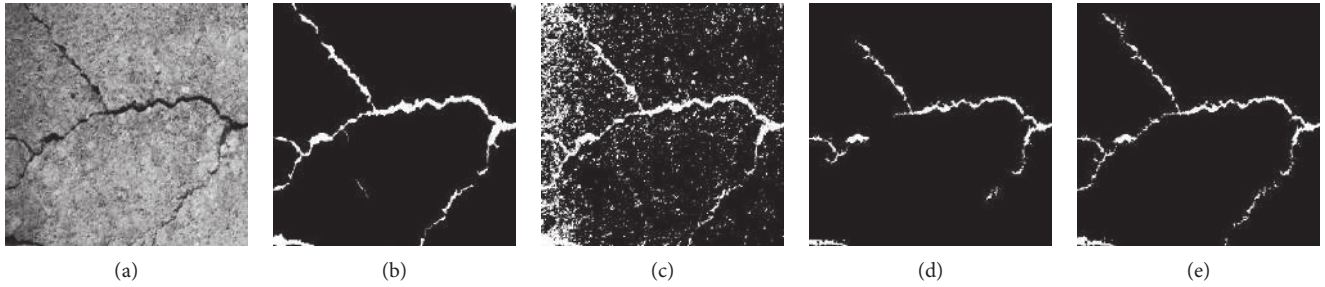


FIGURE 5: Visualization results of segmentation detection. (a) Input image. (b) Ground truth. (c) FCN-8s. (d) FCN-32s. (e) Ours.

to the crack area but are misjudged as the background. The calculation equations of fracture accuracy index P and recall rate index R are as follows:

$$P = \frac{TP}{TP + FP}, \quad (2)$$

$$R = \frac{TP}{TP + FN}. \quad (3)$$

4.3. Comparative Experiments. This paper conducts a comparative experiment with the classic semantic segmentation model FCN. Figures 4 and 5 show the visualization results of the comparative experiment.

TABLE 1: Comparison results of crack segmentation algorithms.

	FCN		Ours	
	P	R	P	R
Sample 1	0.9261	0.9366	0.9452	0.9525
Sample 2	0.9312	0.9425	0.9569	0.9655

It can be seen from Figures 4 and 5 that although the FCN segmentation method can overcome isolated noise points, the effect of wider crack width is acceptable, but because of the low-level features of its fusion, some cracks are not as detailed as others, and the noise extract is not as effectively split. The segmentation algorithm in this article can still maintain a relatively stable and accurate extraction

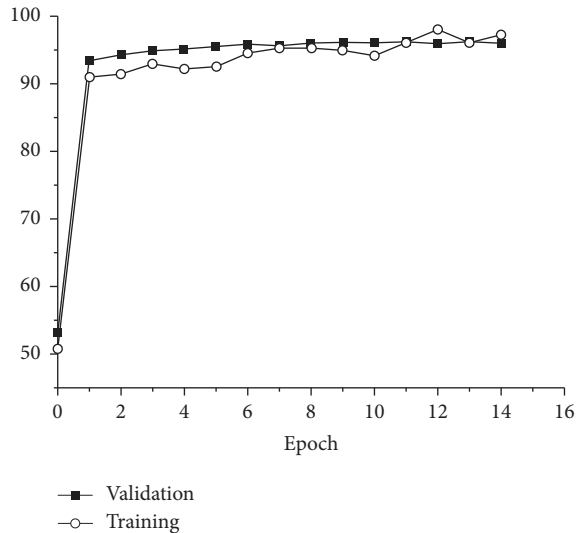


FIGURE 6: P curve during training.

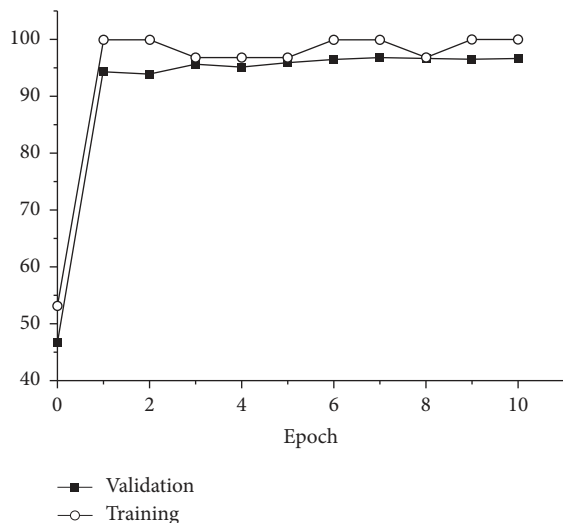


FIGURE 7: R curve during training.

for different types of cracks. It can be clearly seen from Table 1 that the algorithm in this paper is superior to the other two methods in all indicators. In addition, Figures 6 and 7 also show the P and R curves of the algorithm in the training process.

5. Conclusion

In this paper, an image segmentation algorithm for concrete cracks based on convolutional neural network is studied, and an end-to-end segmentation model based on ResNet101 is designed. In order to better train the network model, Labelme software was used to manually annotate the collected real scene fracture dataset and expand the data. Then, Keras deep learning framework was used to train the model, and the extracted feature map was visually displayed. Finally, the proposed algorithm is compared with the FCN method, and the experimental results show that the proposed

segmentation model has better performance in segmentation accuracy and generalization ability.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The author does not have any possible conflicts of interest.

References

- [1] S. W. Tang, Y. Yao, C. Andrade, and Z. J. Li, "Recent durability studies on concrete structure," *Cement and Concrete Research*, vol. 78, pp. 143–154, 2015.
- [2] S. Kabir, "Imaging-based detection of AAR induced map-crack damage in concrete structure," *NDT & E International*, vol. 43, no. 6, pp. 461–469, 2010.
- [3] G. Peiying, D. Chang, and T. Lei, "Determination of local damage probability in concrete structure," *Procedia Engineering*, vol. 28, pp. 489–493, 2012.
- [4] L. Yang and H. Jianguo, "Comparison of safety level of Chinese and foreign crack control of concrete members," *Building Structure*, vol. 2, 2011.
- [5] P. L. Fuhr and D. R. Huston, "Corrosion detection in reinforced concrete roadways and bridges via embedded fiber optic sensors," *Smart Materials and Structures*, vol. 7, no. 2, pp. 217–228, 1998.
- [6] P. L. Fuhr and D. R. Huston, "Corrosion detection in reinforced concrete roadways and bridges via embedded fiber optic sensors," *Smart Materials and Structures*, vol. 7, no. 2, pp. 217–228, 1998.
- [7] S. Yehia, O. Abudayyeh, S. Nabulsi, and I. Abdelqader, "Detection of common defects in concrete bridge decks using nondestructive evaluation techniques," *Journal of Bridge Engineering*, vol. 12, no. 2, pp. 215–225, 2007.
- [8] H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama, and H. Omata, "Road damage detection and classification using deep neural networks with smartphone images," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 12, pp. 1127–1141, 2018.
- [9] M. Foedisch and A. Takeuchi, "Adaptive real-time road detection using neural networks," in *Proceedings. The 7th International IEEE Conference on Intelligent Transportation Systems (IEEE Cat. No. 04TH8749)*, pp. 167–172, IEEE, Washington, WA, USA, October 2004.
- [10] M. Z. Naser, R. A. Hawileh, and J. A. Abdalla, "Fiber-reinforced polymer composites in strengthening reinforced concrete structures: a critical review," *Engineering Structures*, vol. 198, Article ID 109542, 2019.
- [11] C. V. Dung and L. D. Anh, "Autonomous concrete crack detection using deep fully convolutional neural network," *Automation in Construction*, vol. 99, pp. 52–58, 2019.
- [12] A. Mohan and S. Poobal, "Crack detection using image processing: a critical review and analysis," *Alexandria Engineering Journal*, vol. 57, no. 2, pp. 787–798, 2018.
- [13] Q. Zou, Z. Zhang, Q. Li, X. Qi, Q. Wang, and S. Wang, "Deepcrack: learning hierarchical convolutional features for crack detection," *IEEE Transactions on Image Processing*, vol. 28, no. 3, pp. 1498–1512, 2018.
- [14] X. Yang, H. Li, Y. Yu, X. Luo, T. Huang, and X. Yang, "Automatic pixel-level crack detection and measurement using

- fully convolutional network,” *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 12, pp. 1090–1109, 2018.
- [15] T. Nishikawa, J. Yoshida, T. Sugiyama, and Y. Fujino, “Concrete crack detection by multiple sequential image filtering,” *Computer-Aided Civil and Infrastructure Engineering*, vol. 27, no. 1, pp. 29–47, 2012.
- [16] H. G. Moon and J. H. Kim, “Intelligent crack detecting algorithm on the concrete crack image using neural network,” in *Proceedings of the 28th ISARC*, pp. 1461–1467, Seoul, South Korea, June 2011.
- [17] B. Kim and S. Cho, “Image-based concrete crack assessment using mask and region-based convolutional neural network,” *Structural Control and Health Monitoring*, vol. 26, no. 8, p. e2381, 2019.
- [18] Y. Ren, J. Huang, Z. Hong et al., “Image-based concrete crack detection in tunnels using deep fully convolutional networks,” *Construction and Building Materials*, vol. 234, Article ID 117367, 2020.
- [19] J. Chen, C. Du, Y. Zhang, P. Han, and W. Wei, “A clustering-based coverage path planning method for autonomous heterogeneous UAVs,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 99, pp. 1–11, 2021.
- [20] X. Zhang, Y. Yang, Z. Li, X. Ning, Y. Qin, and W. Cai, “An improved encoder-decoder network based on strip pool method applied to segmentation of farmland vacancy field,” *Entropy*, vol. 23, no. 4, p. 435, 2021.
- [21] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097–1105, 2012.
- [22] C. Szegedy, W. Liu, Y. Jia et al., “Going deeper with convolutions,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9, Golden, CO, USA, June 2015.
- [23] S. Dorafshan, R. J. Thomas, and M. Maguire, “Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete,” *Construction and Building Materials*, vol. 186, pp. 1031–1045, 2018.
- [24] B. Kim and S. Cho, “Automated crack detection from large volume of concrete images using deep learning,” in *Proceedings of the 7th World Conference on Structural Control and Monitoring*, pp. 22–25, Qingdao, China, July 2018.
- [25] X.-W. Ye, T. Jin, and P.-Y. Chen, “Structural crack detection using deep learning-based fully convolutional networks,” *Advances in Structural Engineering*, vol. 22, no. 16, pp. 3412–3419, 2019.
- [26] Z. Du, J. Yang, C. Ou, and T. Zhang, “Smallholder crop area mapped with a semantic segmentation deep learning method,” *Remote Sensing*, vol. 11, no. 7, p. 888, 2019.
- [27] V. Hoskere, Y. Narazaki, T. Hoang, and B. Spencer, “Vision-based structural inspection using multiscale deep convolutional neural networks,” 2018, <https://arxiv.org/abs/1805.01055>.
- [28] L. Nkenyereye, B. A. Tama, and S. Lim, “A stacking-based deep neural network approach for effective network anomaly detection,” *CMC-Computers Materials & Continua*, vol. 66, no. 2, pp. 2217–2227, 2021.
- [29] L. Zhang, X. Wang, X. Dong, L. Sun, W. Cai, and X. Ning, “Finger vein image enhancement based on guided tri-Gaussian filters,” *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 17–23, 2021.
- [30] K. Khan, J. Ali, K. Ahmad et al., “3D head pose estimation through facial features and deep convolutional neural networks,” *Computers, Materials & Continua*, vol. 66, no. 2, pp. 1757–1770, 2021.
- [31] G. Krishna Durbhaka, B. Selvaraj, M. Mittal, T. Saba, A. Rehman, and L. Mohan Goyal, “Swarm-LSTM: condition monitoring of gearbox fault diagnosis based on hybrid LSTM deep neural network optimized by swarm intelligence algorithms,” *Computers, Materials & Continua*, vol. 66, no. 2, pp. 2041–2059, 2021.
- [32] Q. Liu, L. Cheng, A. L. Jia, and C. Liu, “Deep reinforcement learning for communication flow control in wireless mesh networks,” *IEEE Network*, vol. 35, no. 2, pp. 112–119, 2021.
- [33] W. Cai and Z. Wei, “Remote sensing image classification based on a cross-attention mechanism and graph convolution,” *IEEE Geoscience and Remote Sensing Letters*, In Press, 2020.
- [34] R. Liu, “Multiscale dense cross-attention mechanism with covariance pooling for hyperspectral image scene classification,” *Mobile Information Systems*, vol. 2021, Article ID 9962057, 15 pages, 2021.
- [35] X. Zhang, Y. Yang, Z. Li, X. Ning, Y. Qin, and W. Cai, “An improved encoder-decoder network based on strip pool method applied to segmentation of farmland vacancy field,” *Entropy*, vol. 23, no. 4, p. 435, 2021.
- [36] X. Ning, W. Li, B. Tang, and H. He, “BULDP: biomimetic uncorrelated locality discriminant projection for feature extraction in face recognition,” *IEEE Transactions on Image Processing*, vol. 27, no. 5, pp. 2575–2586, 2018.
- [37] Y. Tong, L. Yu, S. Li, J. Liu, H. Qin, and W. Li, “Polynomial fitting algorithm based on neural network,” *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 32–39, 2021.
- [38] J. Zhang, J. Sun, J. Wang, and X.-G. Yue, “Visual object tracking based on residual network and cascaded correlation filters,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 20, 2020.
- [39] X. Ning, Y. Wang, W. Tian, L. Liu, and W. Cai, “A biomimetic covering learning method based on principle of homology continuity,” *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 9–16, 2021.
- [40] Z. Luo, “Application and development of electronic computers in aero engine design and manufacture,” *ASP Transactions on Computers*, vol. 1, no. 1, pp. 6–11, 2021.