IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 2, 2020, accepted March 11, 2020, date of publication March 18, 2020, date of current version March 27, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2981561

Crack Detection of Concrete Pavement With Cross-Entropy Loss Function and Improved VGG16 Network Model

ZHONG QU[1,2], JING MEI[1], LING LIU[3], AND DONG-YANG ZHOU[3]
[1]College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
[2]School of Software Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
[3]College of Mobile Telecommunications, Chongqing University of Posts and Telecommunications, Chongqing 401520, China

Corresponding author: Zhong Qu (quzhong@cqupt.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61701060.

ABSTRACT Concrete pavement defects are an important indicator reflecting the safety status of pavement. However, it is difficult to accurately detect the concrete pavement cracks due to the complex concrete pavement environment, such as uneven illumination, deformation and potential shadows, etc. In order to solve these problems, we propose the crack detection algorithm of concrete pavement with convolutional neural network. Firstly, our method is used to classify cracks first and detect the classified crack images, different deep learning models are used in these two parts to achieve different functions. Secondly, in the crack classification section, in view of the low proportion of effective concrete pavement crack images in the mass images collected by crack detection vehicle, the output dimension of FC2 layer of LeNet-5 model is modified before crack detection. It can accurately identify the concrete pavement cracks from several types of disturbance characteristics by training the classification model. Finally, in order to improve the efficiency of crack detection, the algorithm scales the network model horizontally and accesses the convolution layer with the kernel size of 1×1, 3×3. Experiments show that the F_1 of our algorithm reaches to 0.896 in CFD dataset. Compared with VGG16, U-Net and Percolation, it is 25.2%, 2.8%, 39.1% improvement of F_1 respectively. For Cracktree200 dataset, the F_1 is 0.892. Compared with VGG16, U-Net and Percolation, it is 50.3%, 16.6%, 68.9% improvement of F_1 respectively. For DeepCrack dataset, the F_1 is 0.901. Compared with VGG16, U-Net and Percolation, it is 53%, 5.2%, 52.2% improvement of F_1 respectively.

INDEX TERMS Crack detection, cross-entropy loss function, VGG16 network, crack classification.

I. INTRODUCTION

Cracks are one of primary forms of early diseases on concrete pavement, which reduce the service life of pavement and endanger driving safe. Therefore, it is necessary to find and repair cracks early. The traditional detection methods based on artificial detection are labor-intensive and time-consuming which fail to detect concrete pavement cracks accurately.

With the development of digital image processing technology, many researchers have proposed a variety of detection methods for concrete pavement crack detection. In 2016, Shan *et al.* [1] proposed a stereo vision-based crack width detection method to evaluate the crack width of concrete structures' surfaces quantitatively. In 2018, Cho *et al.* [2] proposed an edge-based crack detection method, which used a width map to remove noise and residual noise can be removed by reclassification of crack regions. In 2019, Qu *et al.* [3], [4] proposed a genetic algorithm based on genetic programming (GP) and percolation model, which can detect cracks on the concrete pavement under different noises. In 2018, Su and Yang *et al.* [5] proposed an edge detection method based on the morphological segmentation to achieve the segmentation of concrete cracks. The segmentation results provide the basic information of cracks, which is helpful for the further inspection of concrete structures. In 2008, Shin *et al.* [6] proposed an assessment method of concrete crack depth based on PCA and natural networks. In this method, PCA is used to compress TRF and neural network is used to estimate fracture depth. In 2019, Hoang *et al.* [7] proposed a method for detection condition of

The associate editor coordinating the review of this manuscript and approving it for publication was Long Cheng.

concrete wall. This method uses image texture to evaluate the wall condition and uses logistic expression for pattern recognition. In 2017, Roberto *et al.* [8] proposed a method for the tunnel concrete crack detection, which uses the improved Gabor filter to process the data and can detect cracks in any direction. In 2016, Xie *et al.* [9] proposed a sensor based on ultrasonic concrete detection method. This method can detect concrete internal cracks and potential damage by analyzing time domain waveform and frequency domain spectra.

In recent years, the powerful learning capabilities of deep learning have performed well in image recognition [10], medical diagnosis [11] and path planning [12], [13].

At present, the scholars have proposed a series of crack detection and recognition algorithms based on neural networks for cracks in different situations. In 2018, Chen Jahanshahi [14] proposed a vision-based method for detecting concrete cracks. This method can effectively detect concrete cracks by combining the structural network with the sliding window. In 2018, Maeda *et al.* [15] obtained dataset of pavement cracks through smart phones and detected these cracks through convolution neural network. In 2017, Wang and Hu [16] applied CNN to pavement crack detection, which extracted skeleton of crack through grid, and applied PCA to classify the detected pavement cracks. In 2018, Fan *et al.* [17] proposed a supervised method based on deep learning. This method has a good detection effect on pavement cracks in different environments by modifying the proportion of positive to negative samples. In 2016, Zhang *et al*. [18] proposed a convolutional neural architecture for cracks automatic detection. His method has an outstanding recognition effect in complex noise environments. In 2017, Liu *et al.* [19] proposed an edge detection method. It obtains better performance by modifying the VGG16 model architecture. In 2018, Yang *et al.* [20] used a fully convolutional neural network to segment different types of cracks and pixel-wise to represent the predicted crack skeleton. In 2019, Li *et al.* [21] proposed pixel-level detection of concrete pavement, which obtains better performance by training Convolutional Network. In 2019, Ni *et al.* [22] proposed a new crack width detection method based on Zernike moment operator. It used dual-scale convolutional neural networks to detect cracks width. In 2019, Liu *et al.* [23] proposed a new deep convolutional neural architecture for cracks detection, it combined multi-scale and multi-level information of the target object. In 2019, Lee *et al.* [24] proposed a crack detection method based on image segmentation, the algorithm has a good robustness. In 2018, Dorafshan *et al.* [25] compared the performance of six common edge detectors and convolutional neural networks in crack detection of concrete images. In 2007, Gopalakrishnan *et al.* [26] proposed a pavement crack detection method based on neural network. This method trains convolutional neural network in the "big data" Imagenet database through migration learning. Experiments show that this method has good effect. In 2018, Zhang *et al.* [27] proposed a crack detection method based on the deep convolution neural network. The method first

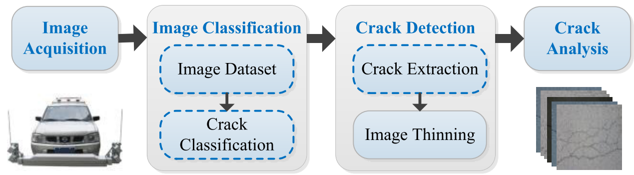


FIGURE 1. Related work.

classifies the road image, then effectively extracts the crack pixels through the block wise threshold method. Gopalakrishnan [28] gave a narrative review of the recently published research on pavement distress detection based on deep learning, comparing the overall framework, network architecture and crack detection performance of these papers.

The actual concrete pavement image captured by the crack detection vehicle includes not only cracks and intact surface, but also crack-free images such as leaves, stains, and marking lines. Moreover, the proportion of concrete pavement crack images is deficient. It will waste much time if we first detect and analyze large-scale images. In order to resolve the problem, we can process a large number of unordered graphs to find the crack image of the concrete pavement effectively.

Based on the above problems, we modified the FC2 (Fully Connected) layer output dimension of the LeNet-5 [29] model. By training the classification model to classify original images, we can accurately identify concrete pavement crack images from several types of images with similar features. Then, we optimized the VGG16 [30] model to detect concrete pavement crack images automatically, and output grayscale images with black as background pixels and white as target pixels. The model can locate the crack edges accurately, thus improving the accuracy of crack detection. The blue dotted box showed in Fig. 1 is our main work.

The rest of this paper is organized as below. Section 2 introduces some of the network models referenced in this paper. Section 3 introduces the modified network model and the considerations when building a database. Section 4 demonstrates the effectiveness of the proposed method by experiments and comparing with other methods. Section 5 summarizes the work of this paper.

II. RELATED WORK

Deep learning-based neural network frameworks excel in image detection and recognition [31], [32]. Inspiration of our improved network structure mainly comes from LeNet-5, VGG16, InceptionNet [33], etc. These networks are introduced here. The inception module is shown in Fig. 2.

A. CRACK CLASSIFICATION

Table 1 shows the network structure of LeNet-5. The LeNet-5 network composes of two Conv layers and average pooling layers. A flattening convolutional layer is connected to the pooling layer, flowing two fully-connected layers. Finally, a Softmax classifier is used at the output layer. This network has a good effect on image classification [34]–[36] etc.

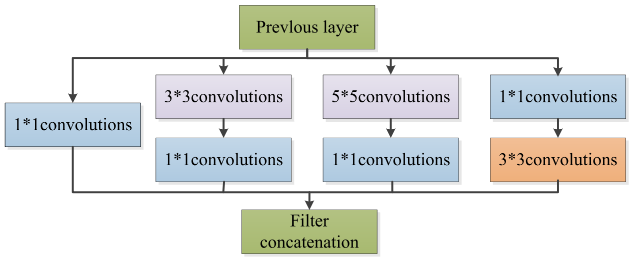


FIGURE 2. Inception module.

TABLE 1. Detailed output size, Stride sizes and a kernel size of standard LeNet-5.

Layer	Conv1	Subsampling	Conv3	Subsampling
Stride	1	2	1	2
Kernel Size	5×5	2×2	5×5	2×2

TABLE 2. Detailed channel number and stride sizes of standard VGG16.

Layer	Conv1_1	Conv1_2	Pool1
Channel number	64	64	-
Stride	1	1	2

Layer	Conv2_1	Conv2_2	Pool2
Channel number	128	128	-
Stride	2	2	4

Layer	Conv3_1	Conv3_2	Conv3_3	Pool3
Channel number	256	256	256	-
Stride	4	4	4	8

Layer	Conv4_1	Conv4_2	Conv4_3	Pool4
Channel number	512	512	512	-
Stride	8	8	8	16

Layer	Conv5_1	Conv5_2	Conv5_3	Pool5
Channel number	512	512	512	-
Stride	16	16	16	32

B. CRACK DETECTION

The VGG16 network consists of 13 Conv layers and three fully connected layers and is connected to the pooling layer after each stage. Table 2 shows the network structure of VGG16. It is known for many experiments that the target characteristics obtained by each stage of VGG16 which increases as the number of layers increases.

Fig. 3 (c)-(g) is the output of each layer of VGG16. It can be seen that the convergence image of the Conv1 and Conv2 layers in the VGG16 model has fine lines at the crack edge, but there are many non-edge pixel points; the crack edge of Conv5 layer is thick, but it has abundant edge features, so the crack edge position can be accurately located. The VGG16 model combines five layers of convolution image features in the fusion layer, which causes the thick edges of the crack. The crack image is not clear, the gray image of the crack has black spots and so on. In view of the existing problems, we modified the VGG16 model. The modified model mainly uses the VGG16 classic two-layer model and combines with the advantages of the Inception Module model. It accesses the Conv layer with the kernel size of 1×1 and resets the partial convolution and parameter values. In this paper, the model uses fewer parameters to extract effective features and fuse multiple convolutions to locate the crack edges, which improves the efficiency and accuracy of crack detection.

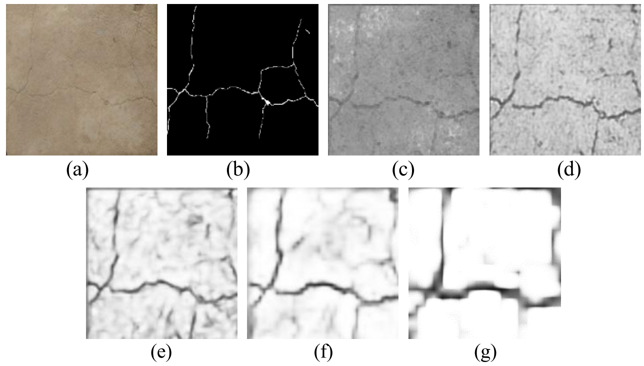


FIGURE 3. The output of vgg16 per layer, as the basis for the network modification: (a) original image; (b) ground truth; (c) Conv1; (d) Conv2; (e) Conv3; (f) Conv4; (g) Conv5.

III. CONCRETE PAVEMENT CRACK DETECTION MODEL

A. OVERALL FRAMEWORK OF THE PROPOSED METHOD

The overall method flow chart of the deep learning model (overall framework) is proposed in this paper, which consists of two parts: crack classification and crack extraction. Different deep learning models are used in these two parts to achieve different functions.

The main works in the crack classification part are as follows:

i. Build dataset CCD1500 (Crack Classification Dataset). CCD1500 is a classified dataset related to cracks in the concrete pavement.

ii. Data process (including labelling, random order, and preventing overfitting).

iii. A transfer learning (i.e., fine-tuning of the developed neural network model) of the LeNet-5 is used to develop the classifier.

The main works in the crack detection part of concrete pavement are as follows:

i. The model needs to be optimized. The VGG16 model has some problems when extracting crack positions. Due to this shortcoming, the VGG16 model is modified according to the characteristics of the crack image.

ii. The optimized model needs to be trained to let itself learn the crack features automatically and to find the best parameters and optimization algorithms of the model through the continuous calculation of errors and update learning.

iii. Using the concrete pavement crack images obtained by LeNet-5, the model can detect the crack features automatically and output the crack grayscale images instead.

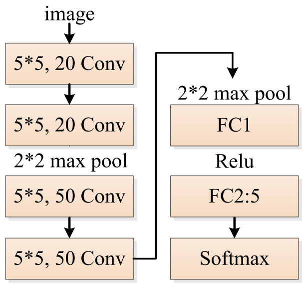


FIGURE 4. Fine-tuned LeNet-5 model diagram for classification.

B. CRACK CLASSIFICATION

At first, it is essential to train the classifier and obtain the performance of the classifier depending on the diversity of the training samples. Therefore, we choose images taken under various conditions from the big data of the Internet as a classification dataset. A simple data collection software is used to obtain all kinds of images by searching for keywords through search engine websites (such as Google), and to establish a dataset CCD1500 of the deep learning classification model, such as cracks, complete concrete pavement and non-cracked objects. It is noteworthy that some objects have significant similarities compared to cracks and are easily identified as cracks by mistake. There is no strict regulation in the size of images during the process of images capture. The image needs to be normalized before input into the neural networks and to be trimmed by the preprocessing software with the standard of image size as 256×256 pixels.

In this paper, the output dimension of the FC2 layer of the LeNet-5 model is modified, as shown in Fig. 4. By training the classification model, it can accurately identify the concrete pavement cracks from several types of interference features. Meanwhile, by fine-tuning more layers with a smaller learning rate, the time to train a new model is significantly saved.

There are some features of concrete pavement crack images. For example, the macro direction is consistent, the shape is tree-like growth, and it has growth and aggregation characteristics, etc. Usually, cracks and fake cracks (such as the lining joints produced by tunnel construction, the pavement joints caused by concrete pavement construction, and the expansion joints and settlement joints that prevent deformation of concrete pavements) have similar features, and they are all elongated. Since the deep learning model can automatically learn the target features during the training process, if only two types are used in the training process (the crack class and the non-crack class), the fake cracks might be seen as cracks because of similar features, causing classify these cracks incorrectly. Therefore, we proposed five types of representative images (the examples might be more, but this paper only represents some classic cases). The figures of five classified images are shown in Fig. 5(a)-(e). The joints between concrete pavement are classified as fake crack class (Fig. 5(a)). Various types of cracks, such as transverse cracks, longitudinal cracks, and reticular cracks, are classified as crack class (Fig. 5(b)). Scratches caused by human activities on different surfaces are classified as artificial scratch class (Fig. 5(c)). Images of real concrete pavement taken at a different time and in different lighting conditions are classified as complete surface class (Fig. 5(d)). Plants such as fallen leaves and vines covered on the concrete pavement are classified as plant class (Fig. 5(e)). This specific crack classification method can improve the accuracy of the classification model when detecting concrete cracks.

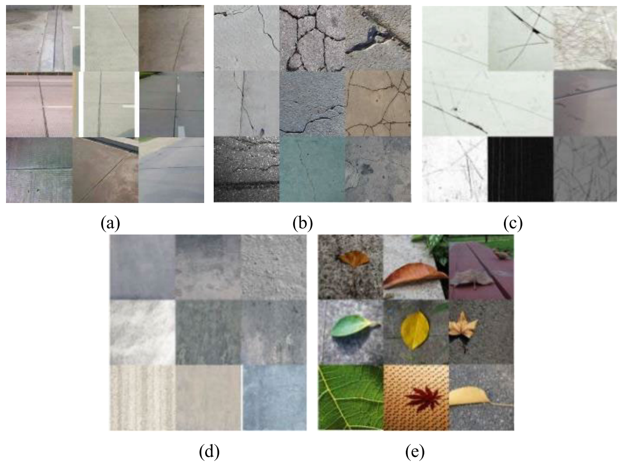


FIGURE 5. More detailed crack classification examples: (a) fake crack; (b) crack; (c) artificial scratch; (d) intact surface; (e) plant.

C. CRACK EXTRACTION

Neural network usually consists of multiple layers, such as input, convolution, pooling, activation, and output layers.

This section introduces the improved VGG16 network model and the specific crack detection process of this paper. We refer to different network structures and analyze a lot of experimental data which are the basis for modifying the network structures. The novel network proposed in this paper is shown in Fig. 6.

Compared with the VGG16 model, the changes made in this paper are as follows:

i. Ignoring Conv1, Conv2 and Conv4 layers of the VGG16 model, the model parameters are significantly reduced because the model layers are reduced at the same time. Thus, the efficiency of crack detection is improved by extracting useful features with fewer parameters.

ii. Use the horizontal expansion method [37]. And insert $1 \times 1 - 1$ Conv layer followed by an Eltwise layer. Then the Deconv layer is used for up-sampling convolution features.

iii. Convolution kernels are used. It inserts a Conv layer with a kernel size of 1×1 and a channel depth of 64 in the Conv3 layer of the model and a Conv layer with a kernel size of 3×3 and a channel depth of 64 in the Conv5 layer of the model. The goal is to enhance the expressive ability of the convolutional layer.

iv. Connect two layers of up-sampling layer and insert $1 \times 1 - 1$ Conv layer in the FC2 layer to fuse the feature map of the two-layer model.

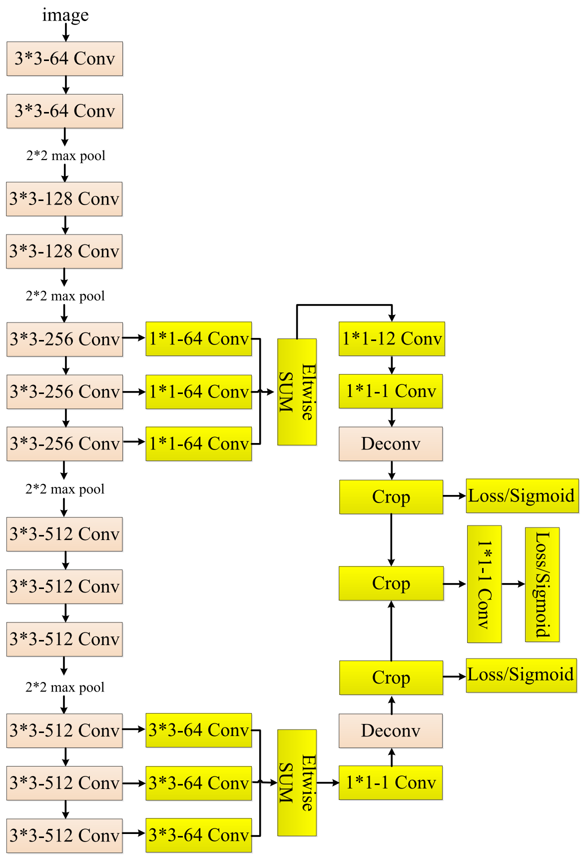


FIGURE 6. **The training model structure of concrete pavement crack detection model.**

D. MODEL TRAINING

The optimized model is used for crack detection in this paper, and the specific process of model training is shown in Fig. 7.

i. Input the original image for parameter initialization.

ii. Convolve and pool the target object to obtain image features. The result of pooling can reduce the number of features and parameters. Through repeated operations of convolution and pooling, the parameter quantity of the network does not increase much, but the error rate drops considerably, thus improving the efficiency of crack detection.

iii. Perform edge feature extraction on the target objects. In the training process, treat pixel points with cracks as real sample label T, and pixel points without cracks as the real sample label N. It is evident that the crack detection problem can change into the one that is to determine whether a given pixel point is a crack pixel point. Moreover, the larger the predicted value probability is, the higher the possibility will be.

iv. By calculating deviations between the target value and the actual output value, and fine-tuning the learning rate continuously, then fusing the convolution features of each layer at the connected layer, the crack grayscale image with more accurate is output through the loss layer.

E. CROSS-ENTROPY LOSS FUNCTION

The Cross-Entropy loss function [38] is used to calculate deviations between the target value and the actual output value in this paper. The optimal value interval used by positive and negative samples is obtained through a large number of experiments. This function turns the output of the neural network into a probability distribution. So that the Cross-Entropy can calculate the distance between the predicted probability distribution and the probability distribution of actual output. Meanwhile, the prediction accuracy can be improved by adding the balance parameter $\theta = 1.1$.

Algorithm 1 The Algorithm of Cross-Entropy Loss

Input: Data, GroundTtuth
Output: TotalLoos
/* Casually select sample pixels X and implement data = sigmoid(x) function operation */
/* Cross-Entropy: groundtruth = sigmoid(groundtruth) */
1: **For** i in W do{
/*W = h *w indicates the amount of pix, w, h indicate the height and width of the pixels, respectively */
2: **If** 0< groundtruth (i) < 0.3 then
3: {Num_pos++;
4: loss_pos + = L(data(i),groundtruth(i));}
5: **else if** 0.7<groundtruth < 1 then
6: {Num_neg++;
7: Loss_neg + = L(data(i),groundtruth(i));}
8: Weight_loss_pos = loss_pos*num_poss/(num_pos+ num_neg)
9: Weight_loss_neg = loss_pos*num_poss/(num_pos+ num_pos) }
10:Retuen Weight_loss_pos+ Weight_loss_neg

$W = h \times w$ indicates the amount of pix; w, h indicate the height and width of the pixels, x is input value. TotalLoos represents the cross-loss value of all pixels in an image, Loss_pos and Loss_neg represent the cross-loss value of the positive and negative samples of each layer, Weight_loss_pos and Weight_loss_neg represent the cross-loss values of the positive and negative samples, respectively.

IV. EXPERIMENTS RESULTS

The first part of the experiment is to fine-tune the LeNet-5 structure for classification. The second part is to detect cracks by using modified VGG16 model. This section shows the results of these two parts. All experiments in this paper based on hardware environment Intel Core i7-8700k, GPU GTX 1070TI 8G, and software environment unbuntu18.04 system, caffe2.0.

A. DATASETS

i. CCD1500: We collected about 1500 images (i.e., 1150 images for training, 350 images for testing images), which are divided into five categories, include fake crack, crack, artificial scratch, intact surface and plant. Each category image is fixed at 300 frames, through balancing the data and reducing bias to get the best classification

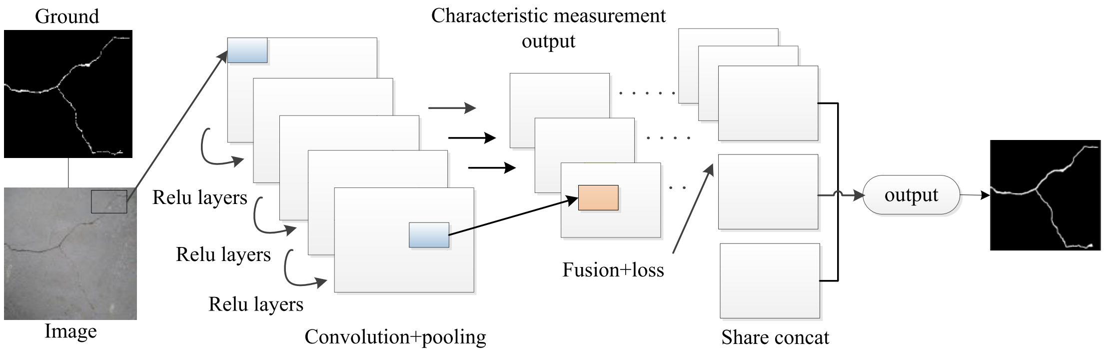


FIGURE 7. The specific process of crack detection model training.

effect. There is no strict regulation in the size of images. Before input into the neural networks, the image needs to be trimmed by the preprocessing software with the standard of image size as 256 × 256 pixels. The crack images are from SDNET2018 [39]. We named the dataset as Crack Classification Dataset 1500 (CCD1500). CCD1500 is composed of our laboratory dataset and available on the Internet datasets of SDNET2018.

ii. SDNET2018: It contains over 56,000 images of cracks (We randomly selected 300 crack images and added them to the CCD1500) and non-crack concrete. It is an annotated image dataset for training, validation, and benchmarking of artificial intelligence neural networks-based crack detection algorithms for concrete.

iii. We collected about 861 images of cracks. All the image size is 448 × 448 pixels. We named the dataset as Crack detection Dataset 861 (CDD861), CDD861 is composed of available datasets from the Internet datasets of CFD, Cracktree200, DeepCrack etc. To equalize the samples, we randomly selected a part from the CDD861 dataset as the training set, the rest was testing set (i.e., 768 images for training, 93 images for testing images).

iv. CFD [40]: It contains 118 annotated concrete surface road crack images, which have problems of noise, cracks that are not clear, etc, All the image size is 448 × 448 pixels.

v. DeepCrack [23]: It contains 537 concrete surface crack images, which have complex background, and cracks that are not clear. All the image size is 448 × 448 pixels.

vi. Cracktree200 [41]: It contains 206 concrete pavement crack images, which have shadows and uneven lighting. All the image size is 448 × 448 pixels.

B. INTRODUCTION OF EXISTING METHODS

VGG16: VGG16 is a very classic convolutional neural network model, and its network structure is very simple. We train CrackForest on CDD861.

U-net: U-net [42] is a point-to-point network. The structure includes a contraction path and a symmetric expansion path (include contracting path and expansive path.). It improves the FCN and improves the expansion path. We train CrackForest on CDD861.

TABLE 3. Correct rate to classified cracks.

	Total Images	Correct image	Correct rate
fake crack	70	64	**0.914**
Crack	70	67	**0.957**
artificial scratch	70	62	**0.886**
intact surface	70	65	**0.928**
Plant	70	63	**0.900**

Percolation: Percolation model [43] is a crack detection method based on image processing. It improves the accuracy by performing a second seepage treatment on the pixels around the crack that has not been percolated.

C. CRACK CLASSIFICATION EXPERIMENTAL RESULTS

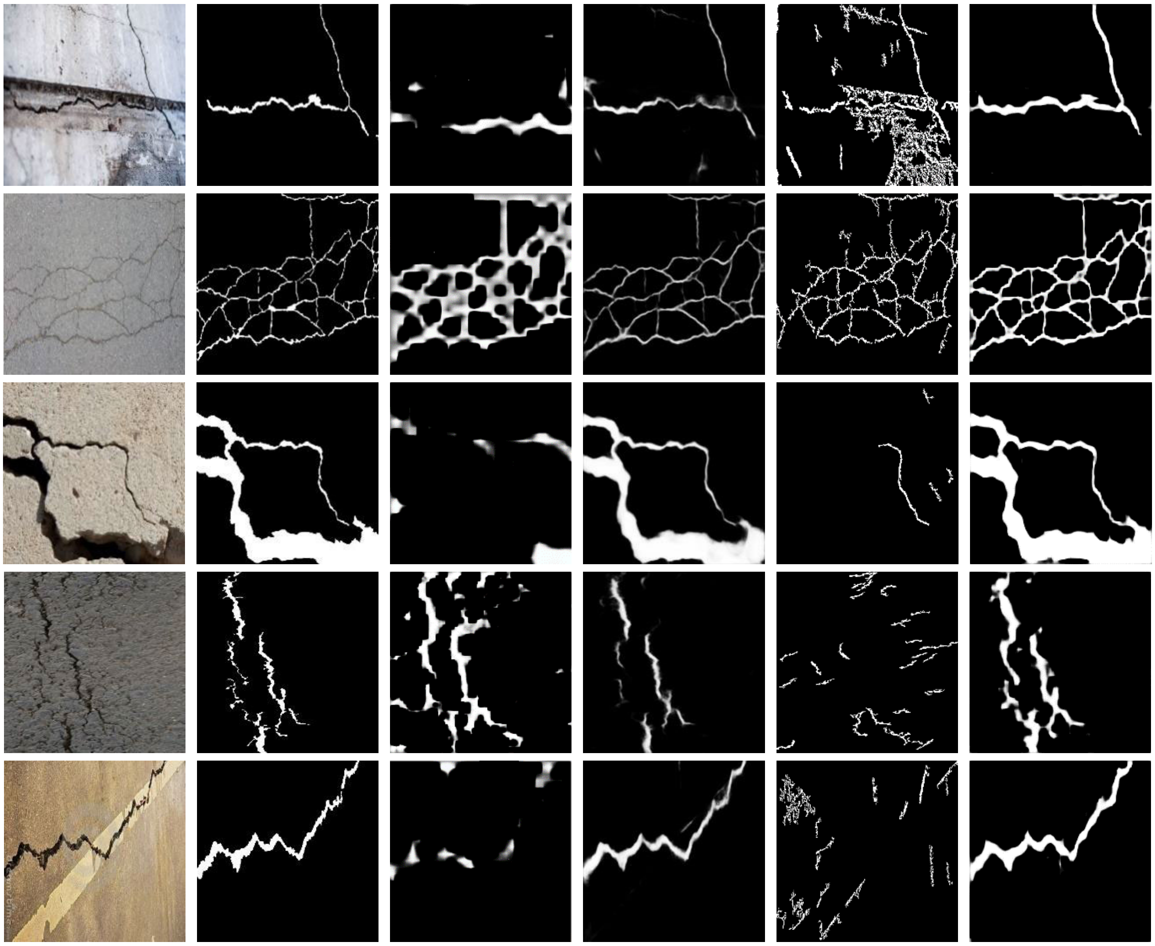
We train and test our classification model on CCD1500, Table 3 gives the number and accuracy of each type of classified images. As observed in Table 3, the recognition rate of crack is above 95%, 350 images are shared for 1.1s, and an average image is less than 0.003s. Experiments show that our algorithm can quickly classify crack images.

D. PERFORMANCE EVALUATION CRITERIA

We implement our network on Caffe framework. And our network is improved on deep learning model of VGG16. For our proposed network. We train our network on CDD861. The Caffe model is used to initialize the test model parameters. As is shown in Fig. 8, we show experimental results plots for four different methods. Then we present and analyze the experimental results on different test set in details.

In order to verify the effectiveness of the proposed crack detection algorithm, we used three datasets as experimental analysis objects. The cracks appeared in the text have various shapes and various structure, such as transverse, longitudinal, oblique and reticular cracks, and they have a certain width. Some cracks are faulty, small and unclear.

In order to further evaluate the performance of the algorithm in this paper, we compared the algorithm of this paper with the algorithm based on digital images and VGG16 algorithm and quantified the experimental results



(a) (b) (c) (d) (e) (f)

FIGURE 8. Crack detection result: (a) Original image; (b) Target map; (c)VGG16; (d) U-Net; (e) Percolation; (f) Our algorithm.

through comprehensive evaluation. Indexes $F1$, *Precision* and *Recall*, which are defined as follows:

$$Precision = \frac{ture\ positive}{ture\ positive + false\ positive} \quad (1)$$

$$Recall = \frac{ture\ positive}{ture\ positive + false\ negative} \quad (2)$$

$$F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (3)$$

E. CRACK DETECTION EXPERIMENTAL RESULTS

i) Results on CFD: It can be seen from Table 4 that our algorithm of crack detection can accurately locate the edge of the crack. As a result, the crack edges of result images are fine, and it can achieve the best F_1 on the CFD, with F_1 value of 0.896. Compared to VGG16, U-Net and Percolation, it is 25.2%, 2.8%, 39.1% improvement for F_1 respectively. Percolation method achieves the lowest F_1 and Recall, and deep learning method achieves better results.

ii) Results on Cracktree200: It can be seen from Table 5 that our algorithm has the highest F_1 value on Cracktree200, the F_1 value is 0.892. The VGG16, U-Net and Percolation are 50.3%, 16.6%, 68.9% less than the results of proposed method respectively. The performance of Percolation has the worst performance, which only holds F_1 value of 0.203.

TABLE 4. The precision, recall, F_1 of compared method test on CFD.

Methods	*Precision*	*Recall*	F_1	*FPS*
VGG16	0.570	0.740	0.644	30
U-Net	0.855	0.882	0.868	6
Percolation	0.582	0.447	0.505	0.05
Ours	0.889	0.903	0.896	30

TABLE 5. The precision, recall, F_1 of compared method test on DeepCrack.

Methods	*Precision*	*Recall*	F_1	*FPS*
VGG16	0.305	0.538	0.389	30
U-Net	0.760	0.694	0.726	6
Percolation	0.121	0.631	0.203	0.05
Ours	0.829	0.966	0.892	30

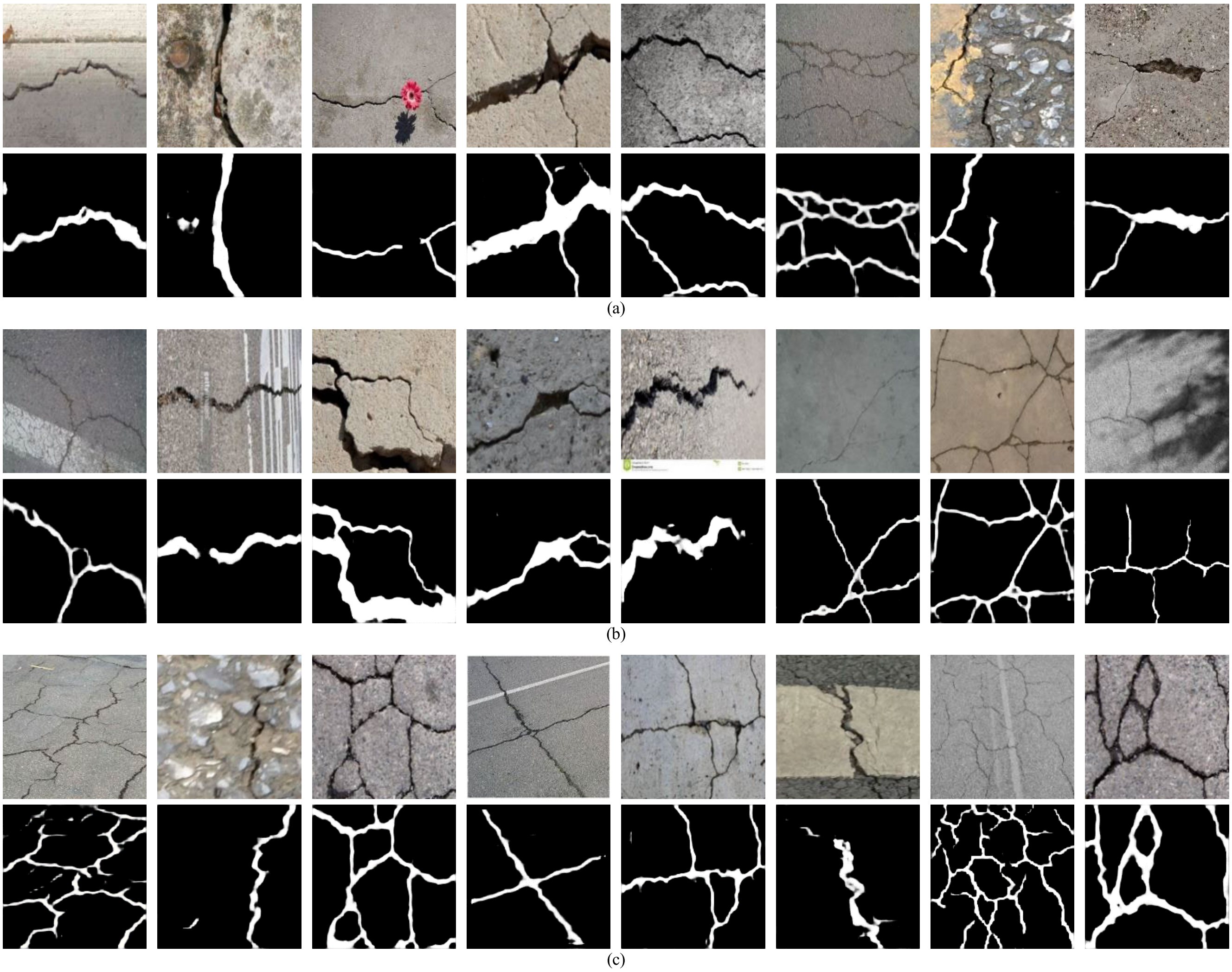


FIGURE 9. More results of our proposed algorithms on the test dataset. (a) Original images and results of the first group; (b) Original images and results of the second group; (c) Original images and results of the third group.

TABLE 6. The precision, recall, F1 of compared method test on Cracktree200.

Methods	*Precision*	*Recall*	F_1	*FPS*
VGG16	0.407	0.341	0.371	30
U-Net	0.848	0.851	0.849	6
Percolation	0.453	0.326	0.379	0.05
Ours	0.912	0.891	0.901	30

iii) Results on DeepCrack: This dataset contains images of road cracks on different surfaces, which have many complex characteristics, such as road traffic signs, pollutants, etc. It can be seen from Table 6, the method we proposed gets the highest F_1 on DeepCrack dataset, which is 0.901 Compared to VGG16, U-Net and Percolation, it is 53%, 5.2%, 52.2% improvement for F_1 respectively.

iv) More results: for the reason of simple network structure, we found that the method we proposed and VGG16 have the ability to extract the crack features faster. As we can see from Table 4, our method processes the image at 30 *FPS*, which are faster than the U-Net and Percolation methods. Due to the complexity of the U-Net network, the U-Net method runs at a slow speed of 6 *FPS*. For non-deep learning methods, Percolation takes about 20 seconds to process the image. In addition, for different datasets, the F_1 of our method has been kept 90%, which is less affected by image noise. However, The VGG16, U-Net and Percolation method are more efficient for the images with simple cracks and single background, and the image effects is worse which is affected by the shadow and unclear cracks, as well as the Precision is greatly influenced by image noise.

In order to verify the superiority of our method, we give more results in Fig. 9. More results of our proposed method on the CDD861 dataset are present. Experiments have shown that our method can accurately locate the edge of the crack, and has a good detection effect on short and unclear cracks, and there is no blurring or fracture phenomenon. Before the crack detection, the original crack image is not denoised, but the F_1 value of the method is less affected, proving that the method has a better effect.

V. CONCLUSION

We propose the crack detection algorithm of concrete pavement with convolutional neural networks. It uses the fine-tuned LeNet-5 network to classify the original image, which saves much time for the disordered image detection. In the crack detection part, we used the optimized VGG16 model to extract the concrete crack characteristics. For performance evaluation, two crack datasets were constructed. in which CCD1500 is used as the crack classification dataset and CDD861 is used as the crack detection dataset. As is shown in the experiments, compared with VGG16, U-Net and Percolation algorithms, our algorithm has the highest F_1 value on CFD dataset, Cracktree200 dataset, DeepCrack dataset.

These datasets contain images of road cracks on different background, which have many complex background characteristics, such as road traffic signs, pollutants, plants, etc. The cracks in the image have different shapes, including longitudinal cracks such as transverse cracks, linear cracks, and network cracks such as block cracks, fracturing cracks, and some of the images have short and unclear cracks. From the experiment result, Percolation algorithms based on image processing have better detection results for the background of cleaner linear crack. But for the background of interference noise, the effect of image with more complex cracks is poor, for example of the uneven sand surface crack. Because the neural network model has a strong ability to learn, can automatically extract the features of cracks, so the neural network also has a good effect on cracks in complex backgrounds. Therefore, in the aspect of crack detection, the algorithm is based on neural network in performance is superior to traditional methods of image processing. Our algorithm retains the high PFS value of VGG16, improves the efficiency of VGG16 model for crack detection. Our algorithm has relatively high accuracy as well as recall rate for all kinds of cracks in different environments, which can achieve rapid and accurate detection of crack images.

Because the original image was not pre-processed, we will add image preprocessing methods and reduce background noise to improve the accuracy and the efficiency of crack detection in the future. At present, this algorithm is mainly used in concrete crack detection. In the future, we will study how to apply this algorithm to a wider range of fracture environments, such as glass surface cracks, metal surface cracks, etc. As we all know, the detection speed of YOLO series (such as YOLO v3) network is fast. In the future, we will study YOLO series network and improve the detection rate further.

REFERENCES

[1] B. Shan, S. Zheng, and J. Ou, "A stereovision-based crack width detection approach for concrete surface assessment," *KSCE J. Civil Eng.*, vol. 20, no. 2, pp. 803–812, Mar. 2016.

[2] H. Cho, H.-J. Yoon, and J.-Y. Jung, "Image-based crack detection using crack width transform (CWT) algorithm," *IEEE Access*, vol. 6, pp. 60100–60114, 2018.

[3] Z. Qu, Y.-X. Chen, L. Liu, Y. Xie, and Q. Zhou, "The algorithm of concrete surface crack detection based on the genetic programming and percolation model," *IEEE Access*, vol. 7, pp. 57592–57603, 2019.

[4] Z. Qu, S.-Q. Chen, Y.-Q. Liu, and L. Liu, "Eliminating lining seams in tunnel concrete crack images via line Segments' translation and expansion," *IEEE Access*, vol. 7, pp. 30615–30627, 2019.

[5] T. C. Su and M. D. Yang, "Morphological segmentation based on edge detection-II for automatic concrete crack measurement," *Comput. Concrete*, vol. 21, no. 6, pp. 727–739, Jun. 2018.

[6] S. W. Shin, C. B. Yun, H. Futura, and J. S. Popovics, "Nondestructive evaluation of crack depth in concrete using PCA-compressed wave transmission function and neural networks," *Experim. Mech.*, vol. 48, no. 2, pp. 225–231, Apr. 2008.

[7] N.-D. Hoang, Q.-L. Nguyen, and X.-L. Tran, "Automatic detection of concrete spalling using piecewise linear stochastic gradient descent logistic regression and image texture analysis," *Complexity*, vol. 2019, pp. 1–14, Jul. 2019, doi: 10.1155/2019/5910625.

[8] R. Medina, J. Llamas, J. Gómez-García-Bermejo, E. Zalama, and M. Segarra, "Crack detection in concrete tunnels using a Gabor filter invariant to rotation," *Sensors*, vol. 17, no. 7, p. 1670, Jul. 2017.

[9] X. Xie, D. Xu, X. Guo, F. Sha, and S. Huang, "Nonlinear ultrasonic nondestructive evaluation of damaged concrete based on embedded piezoelectric sensors," *Res. Nondestruct. Eval.*, vol. 27, no. 3, pp. 125–136, Jul. 2016.

[10] L. Cheng, Y. Liu, M. Tan, D. Du, M. Fei, and Z.-G. Hou, "A rapid spiking neural network approach with an application on hand gesture recognition," *IEEE Trans. Cogn. Devel. Syst.*, to be published, doi: 10.1109/TCDS.2019.2918228.

[11] Z. Liu, C. Cao, S. Ding, Z. Liu, T. Han, and S. Liu, "Towards clinical diagnosis: Automated stroke lesion segmentation on multi-spectral MR image using convolutional neural network," *IEEE Access*, vol. 6, pp. 57006–57016, 2018, doi: 10.1109/ACCESS.2018.2872939.

[12] C. Yang, X. Wang, L. Cheng, and H. Ma, "Neural-learning-based telerobot control with guaranteed performance," *IEEE Trans. Cybern.*, vol. 47, no. 10, pp. 3148–3159, Oct. 2017.

[13] C. Yang, G. Peng, Y. Li, R. Cui, L. Cheng, and Z. Li, "Neural networks enhanced adaptive admittance control of optimized robot–environment interaction," *IEEE Trans. Cybern.*, vol. 49, no. 7, pp. 2568–2579, Jul. 2019.

[14] F.-C. Chen and M. R. Jahanshahi, "NB-CNN: Deep learning-based crack detection using convolutional neural network and Naïve bayes data fusion," *IEEE Trans. Ind. Electron.*, vol. 65, no. 5, pp. 4392–4400, May 2018.

[15] H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama, and H. Omata, "Road damage detection and classification using deep neural networks with smartphone images," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 33, no. 12, pp. 1127–1141, Dec. 2018.

[16] X. Wang and Z. Hu, "Grid-based pavement crack analysis using deep learning," in *Proc. 4th Int. Conf. Transp. Inf. Saf. (ICTIS)*, Banff, AB, Canada, Aug. 2017, pp. 917–924.

[17] Z. Fan, Y. Wu, J. Lu, and W. Li, "Automatic pavement crack detection based on structured prediction with the convolutional neural network," 2018, *arXiv:1802.02208*. [Online]. Available: http://arxiv.org/abs/1802.02208

[18] L. Zhang, F. Yang, Y. Daniel Zhang, and Y. J. Zhu, "Road crack detection using deep convolutional neural network," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 3708–3712.

[19] Y. Liu, M.-M. Cheng, X. Hu, J.-W. Bian, L. Zhang, X. Bai, and J. Tang, "Richer convolutional features for edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 8, pp. 1939–1946, Aug. 2019.

[20] X. Yang, H. Li, Y. Yu, X. Luo, T. Huang, and X. Yang, "Automatic pixel-level crack detection and measurement using fully convolutional network," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 33, no. 12, pp. 1090–1109, Dec. 2018.

[21] S. Li, X. Zhao, and G. Zhou, "Automatic pixel-level multiple damage detection of concrete structure using fully convolutional network," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 34, no. 7, pp. 616–634, Jul. 2019.

[22] F. T. Ni, J. Zhang, and Z. Q. Chen, "Zernike-moment measurement of thin-crack width in images enabled by dual-scale deep learning," *Comput.-Aided Civ. Inf.*, vol. 34, no. 5, pp. 367–384, May 2019.

[23] Y. Liu, J. Yao, X. Lu, R. Xie, and L. Li, "DeepCrack: A deep hierarchical feature learning architecture for crack segmentation," *Neurocomputing*, vol. 338, pp. 139–153, Apr. 2019.

[24] D. Lee, J. Kim, and D. Lee, "Robust concrete crack detection using deep learning-based semantic segmentation," *Int. J. Aeronaut. Space Sci.*, vol. 20, no. 1, pp. 287–299, Mar. 2019.

[25] S. Dorafshan, R. J. Thomas, and M. Maguire, "Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete," *Construct. Building Mater.*, vol. 186, pp. 1031–1045, Oct. 2018.

[26] K. Gopalakrishnan, S. K. Khaitan, A. Choudhary, and A. Agrawal, "Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection," *Construct. Building Mater.*, vol. 157, pp. 322–330, Dec. 2017.

[27] K. Zhang, H. D. Cheng, and B. Zhang, "Unified approach to pavement crack and sealed crack detection using preclassification based on transfer learning," *J. Comput. Civil Eng.*, vol. 32, no. 2, Mar. 2018, Art. no. 04018001.

[28] K. Gopalakrishnan, "Deep learning in data-driven pavement image analysis and automated distress detection: A review," *Data*, vol. 3, no. 3, p. 28, Jul. 2018.

[29] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998, doi: 10.1109/5.726791.

[30] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Int. Conf. Learn. Represent.*, San Diego, CA, USA, 2014, pp. 1409–1556.

[31] L. Cheng and W. Z. Liu, "An effective microscopic detection method for automated silicon-substrate ultra-microtome (ASUM)," *Neural Process Lett.*, pp. 1–8, Nov. 2019, doi: 10.1007/s11063-019-10134-5.

[32] L. Cheng, W. Liu, C. Yang, T. Huang, Z.-G. Hou, and M. Tan, "A neural-network-based controller for piezoelectric-actuated stick–slip devices," *IEEE Trans. Ind. Electron.*, vol. 65, no. 3, pp. 2598–2607, Mar. 2018.

[33] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Boston, MA, USA, Jun. 2015, pp. 1–9.

[34] X. Cao, P. Wang, C. Meng, X. Bai, G. Gong, M. Liu, and J. Qi, "Region based CNN for foreign object debris detection on airfield pavement," *Sensors*, vol. 18, no. 3, p. 737, Mar. 2018.

[35] K.-J. Hsu, Y.-Y. Lin, and Y.-Y. Chuang, "Weakly supervised salient object detection by learning a classifier-driven map generator," *IEEE Trans. Image Process.*, vol. 28, no. 11, pp. 5435–5449, Nov. 2019.

[36] G. Cheng, J. Han, P. Zhou, and D. Xu, "Learning rotation-invariant and Fisher discriminative convolutional neural networks for object detection," *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 265–278, Jan. 2019.

[37] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 5987–5995, doi: 10.1109/CVPR.2017.634.

[38] Y. Wei, Z. Wang, and M. Xu, "Road structure refined CNN for road extraction in aerial image," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 5, pp. 709–713, May 2017.

[39] S. Dorafshan, R. J. Thomas, and M. Maguire, "SDNET2018: An annotated image dataset for non-contact concrete crack detection using deep convolutional neural networks," *Data Brief*, vol. 21, pp. 1664–1668, Dec. 2018.

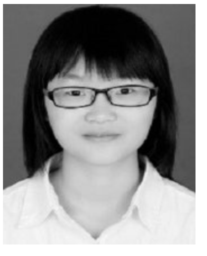
[40] Y. Shi, L. Cui, Z. Qi, F. Meng, and Z. Chen, "Automatic road crack detection using random structured forests," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 12, pp. 3434–3445, Dec. 2016.

[41] Q. Zou, Y. Cao, Q. Li, Q. Mao, and S. Wang, "CrackTree: Automatic crack detection from pavement images," *Pattern Recognit. Lett.*, vol. 33, no. 3, pp. 227–238, Feb. 2012.

[42] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, vol. 9351, Oct. 2015, pp. 234–241.

[43] Z. Qu, Y. Guo, F.-R. Ju, L. Liu, and L.-D. Lin, "The algorithm of accelerated cracks detection and extracting skeleton by direction chain code in concrete surface image," *Imag. Sci. J.*, vol. 64, no. 3, pp. 119–130, Apr. 2016.

ZHONG QU received the M.S degree in computer architecture and the Ph.D. degree in computer application technology from Chongqing University, in 2003 and 2009, respectively. He is currently a Professor with the Chongqing University of Posts and Telecommunications. His research interests include digital image processing, digital media technology, and cloud computing.



JING MEI is currently pursuing the M.S degree with the College of Computer Science and Technology, Chongqing University of Posts and Telecommunications. Her main research interest includes digital image processing.

LING LIU received the M.S degree in computer technology from the Chongqing University of Posts and Telecommunications, in 2014. Her main research interest includes digital image processing.

DONG-YANG ZHOU received the M.S degree in computer technology from the Chongqing University of Posts and Telecommunications, in 2016. His main research interest includes digital image processing.