Deep Learning-Based Corroded Crack Identification in Reinforced Concrete Using SCNet Model

YING XU and SIWEI ZHOU

ABSTRACT

In order to improve the efficiency and accuracy of corroded cracks detection and classification in reinforced concrete, a corroded cracks identification model Steel Corrosion Net (SCNet), based on deep learning Convolutional Neural Network (CNN), is proposed. The SCNet combines massive initial data with a multi hidden layer neural network framework, and achieves feature learning and accurate classification through model training. The data set of 39000 crack figures is firstly built by original data collection and data enhancement. The training process of the SCNet consists of defining the loss function, the selecting back propagation optimization algorithm, continuously entering data to the network framework and running back the propagation algorithm until the error drops to a certain range. Afterwards, a SCNet three-classification neural network model is built and tested using TensorFlow learning framework and Python. According to the training and testing accuracies of the model, the structure and parameters of the SCNet network are optimized. The results of SCNet are compared with those obtained by two traditional testing methods. The results show that the proposed SCNet model achieves a classification accuracy of 96.8%. In other words, it can effectively identify and classify the corroded cracks in reinforced concrete, with high accuracy and measurability.

KEY WORDS

Concrete cracks; steel corrosion; Convolutional Neural Network; data enhancement; neural network optimization

Shenzhen Key Lab of Urban Civil Engineering Disaster Prevention & Reduction, Guangdong Provincial Key Laboratory of Intelligent and Resilient Structures for Civil Engineering, Harbin Institute of Technology, Shenzhen, Guangdong, 518055, China

1. INTRODUCTION

The corrosion of steel bars in reinforced concrete structure usually causes volume expansion of steel bars, leading to the development of concrete cracks until the final concrete protective layer spalls off. This results in a structural damage and then reduces the bearing capacity[1]. Traditional surface crack monitoring due to reinforcement corrosion including high-power magnifier, inclinometer and crack measurement card, is regularly done manually[2]. However, these methods have some disadvantages such as poor safety, high cost and low efficiency. New technologies such as optical fiber, SEM, thermal imaging and ultrasonic[3], are also used in crack detection. However, these technologies also have their limitations. More precisely, they are not suitable for surface crack detection, and they are susceptible to environmental interference and costly. At present, several sophisticated structural health monitoring systems deployed in large buildings[4-5] require a large number of sensors, data collection system, and certain environmental compensation. With the development of image processing, several damage detection methods, based on computer vision, are used to study the concrete surface[6]. However, shortcomings in these methods exist, such as the lack of further classification, the poor light and the presence of noise.

The deep learning algorithm, which is a type of machine learning, consists in constructing a neural network framework to perform learning on the initial data, training and learning the neural network framework through a large-scale data set, and updating the weights to extract features[7]. A Convolutional Neural Network (CNN) is a neural network based on deep learning, inspired by the visual cortex of animals[8]. It can efficiently capture the grid topology of the image as the basis for classification[9], which allows to efficiently perform image identification.

In the past decades, several researches on Artificial Neural Network (ANN) have been carried out. However, they have recently been applied in image processing, such as crack identification in civil engineering. For instance, Cai[10] proposed a bridge crack video detection system using video image processing to make up for the shortcomings of the bridge crack detection technology. Landstrom et al.[11] proposed a method of automatic on-line detection of steel plate surface cracks based on three-dimensional profile data, using morphological image processing and logical reasoning. Moon et al.[12] developed an automatic detection system which can analyze concrete surface and effectively identify visual cracks. In this method, cracks are identified and distinguished from the background image by filtering, improving subtraction and using morphological operations.

The efficiency of cracks feature extraction was also improved using the Back Propagation (BP) neural network. Zhang[13] was the first to use deep learning for the crack detection of roads. Cha et al.[14] combined the trained CNN model with a sliding window technology to perform the detection and recognition of concrete cracks.

However, the current deep learning CNN models can only identify cracks on the concrete surface, while they are not able to judge whether the cracks are belonging to the corrosion cracks. In this paper, the CNN is applied to identify and classify reinforced concrete corrosion cracks. A reinforced concrete corrosion crack identification model (SCNet), based on deep learning CNN, is proposed. The results verify the accuracy and measurability of the proposed model, which is able to identify the corrosion cracks of reinforced concrete in practical engineering applications.

2. CNN THEORY

CNN consists of structural layers such as the input layer, the hidden layers, the output layer, etc. The hidden layers include convolution, pooling, activation, full connection, Softmax and Dropout layers. Figure 1 shows the basic CNN framework diagram used for classification[15].

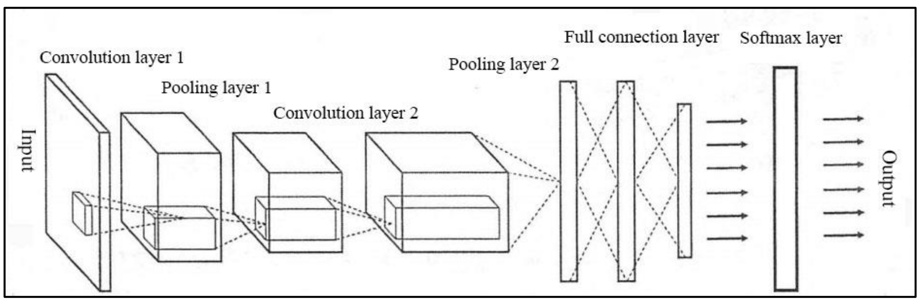


Figure 1. CNN framework of image classification

This paper consists in building the SCNet model based on CNN theory. The basic concept consists in first summarizing the basic design rules of typical classification neural networks. That is, the order of input-convolution-pooling-full connection-output. The number of layers is then optimized using three convolutional layers to connect, and small convolution kernels to increase the receptive field. Finally, ReLU is used as activation function. It significantly improves the training efficiency and learning rate, and uses the dropout layer to deal with the overfitting problem, which makes the network have better generalization capacity.

3. BUILDING DATA SETS

In deep learning, several networks with high recognition rate benefit from their huge and real sample sets. Simultaneously, the accuracy and quality of image features directly affect the training and detection of the subsequent models. The construction of neural network data set mainly includes image data acquisition and data enhancement. The data set of the collected images is divided into three categories; reinforced concrete corrosion cracks, other causes of concrete cracks and complete concrete. Other causes of concrete cracks are mainly divided into plastic shrinkage cracks, structural cracks, temperature cracks and cracks caused by alkali aggregate reaction. Plastic shrinkage cracks refer to the irregular crack or parallel crack of concrete perpendicular to the longitudinal reinforcement. Structural cracks occur when the external load exceeds the bearing capacity of the concrete, and when no color change on the surface of the cracks exists. Alkaline aggregate reaction crack is caused by expansion and compression around the aggregate, which is characterized by surface concrete mesh cracking, besides transparent or yellowish gel precipitation.

The main difference between the corrosion cracks of reinforced concrete and other causes of concrete cracks, is that the corrosion cracks have a unique color characteristic of reddish brown rust marks around the cracks and a full-length crack shape which is basically consistent with the direction of the longitudinal steel bar. On the other hand, other causes of concrete cracks are characterized by irregular cracking or mesh cracking besides transparent or yellowish gel precipitation. The data set includes a training set and a test set, with a data ratio of 4:1.

3.1 Raw Image Data Acquisition

Considering the accuracy and diversity of the sample set, three methods of network search, self-shooting and corrosion test are used to collect the original image data. In this paper, 1530 pictures are used, where 210 pictures are available online, 540 pictures were taken by self-shooting, and 780 pictures were taken by corrosion test. In order to clarify the features of the picture data, this paper unifies the shooting standard of the picture data and sets the image acquisition standard. Firstly, the camera is taken perpendicular to the concrete surface in order to avoid geometric distortion. This also ensures that all the image data features are obtained in the same shooting environment. The shooting distance is also a crucial factor affecting the image quality. It determines the range of vision in the actual shooting. In order to clearly show the crack features in the image, a shooting distance of 20 cm is used. Illumination also has a large impact on the image features. The background of the image will highly change under natural light leading to uneven background brightness. In addition, self-shooting and network-searching methods are used. According to the shooting specifications, the self-shooting method can be taken in accordance with the shooting specifications, and is directly used to construct the data set. The images searched on the network may not meet the collection specifications. Therefore, they require scaling and brightness adjustments to meet the collection standards. The methods used to collect the original image ensured the diversity of the data.

3.2 Data Enhancement

In this paper, the data enhancement is performed using OpenCV random clipping, rotation and random color transformation. Note that the various functions provided by OpenCV can modify the picture to get more data. The effect of the rotation operation is shown in Figure 2

Using random clipping, rotation and random change of the picture saturation and brightness, the effect of expanding the amount of data is achieved. In other words, one origin image can produce 30 image data. In the process of image selection, the final data set is obtained by eliminating some noise images with insufficient cracks and/or unsatisfactory brightness, as well as unifying the image size. Note that 80% (31200) off the data was used as training set, while 20 % (7800) was used as testing samples.

4. IDENTIFICATION AND CLASSIFICATION OF CORROSION CRACKS

In this paper, CNN is used to classify images with different concrete conditions. The data set is divided into three categories according to the characteristics of obvious rust marks around reinforced concrete corrosion cracks and long cracks along steel bars. The specific classification method consists in first adding the corresponding labels (0, 1, 2) to the training set samples according to three different classifications. Afterwards, through supervised-learning, these known data and their corresponding output label are trained, the parameters are adjusted, and an optimal model is obtained. This model is then used to map all the inputs of the unknown data to the corresponding output, output the corresponding label, and obtain the classification results. Finally, we get the network input, define the neural network framework structure, repeatedly modify the parameter training model, gradually reduce the result error, and test the performance of the model on unknown data.

This paper splits the deep learning CNN into four independent programs: file input, forward propagation process, training part and testing part. This decoupling allows to modify each part independently, without affecting subsequent procedures. It also makes the whole process more flexible

Figure 2 Rendering of rotation operation

4.1 Building the Network Framework

After optimization and adjustment, a SCNet model based on the classic classification neural network VGG16 is designed. VGG16 is used in several studies[15-16]. The model is composed of five large convolution layers and three fully connected layers. The large convolution layer includes small convolution layers and a pooling layer. The VGG16 model improves the network performance by increasing the neural network depth. The main characteristic of the model is that it uses a 3×3 small convolution kernel and a 2×2 pooling window. It also increases the network depth to 16-19 layers. The network framework of the SCNet model is shown in Figure 3. The model consists of an input layer, six large convolutional layers, three fully connected layers, a softmax classification layer and an output layer. The large convolution layer includes a convolution layer and a pool layer. The 1st, 2nd, 3rd and 4th large convolution layer consist of three convolution operations with 32, 64, 128 and 256 convolution kernels, for each convolution followed by one maximum pooling, respectively. The operation of the fourth and fifth convolution layers is the same. The sixth large convolution layer includes only one convolution layer. The size of all the convolution kernels is 3×3, the step size is 1, the pooling window size of the pooling operation is 2×2, and the step size is 2.

4.2 Training and Testing of the Neural Networks

The training process of the neural network consists of the content expression form of defining the loss function, the selecting back propagation optimization algorithm, continuously entering data to the network framework and running back the propagation algorithm until the error drops to a certain range. Figure 4 shows the

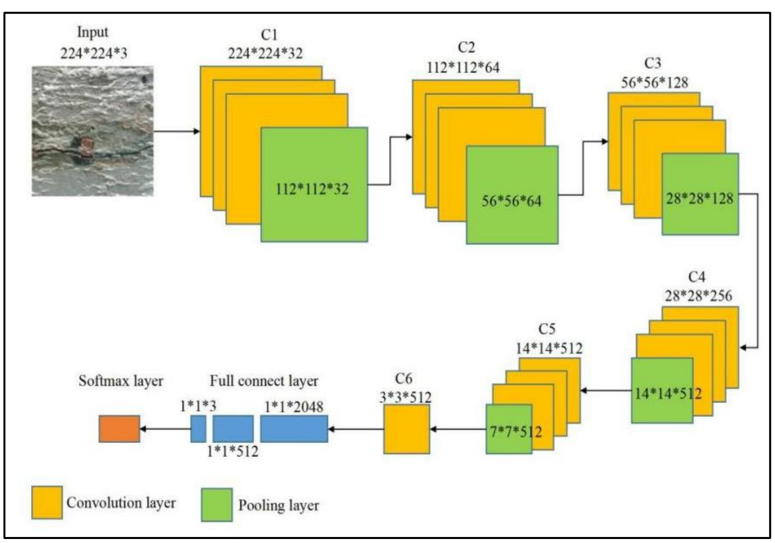


Figure 3. SCNet network model sketch map

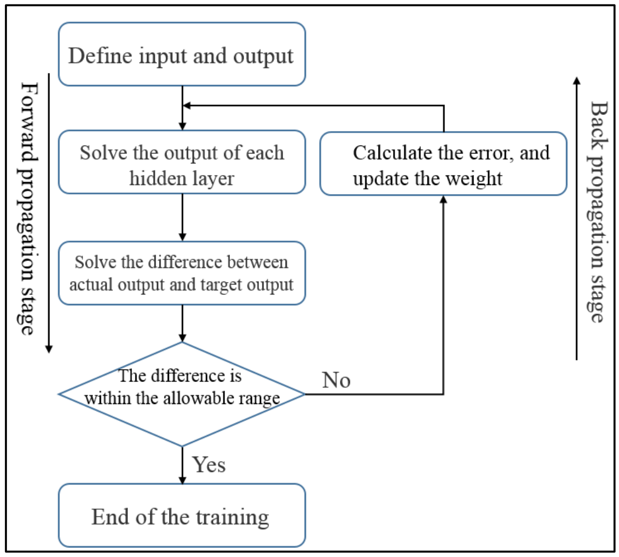


Figure 4 CNN training process

CNN training process. The first stage is the forward propagation which consists in calculating the input data and the final statistics error. The second stage is the back propagation which consists in propagating the errors of the actual output and the theoretical output, from high to low level.

In order to improve the model training accuracy and optimize the model, the activation function and the dropout layer are added when constructing the neural network framework. The sliding average, the regularization loss, the back propagation optimization algorithm and the exponential attenuation learning rate, are also added to the training process.

The testing process of the neural network model consists in letting the model determine which classification label (0, 1 and 2) the image data in the data set belongs to. It then compares the judged label with the real classification label, using the SCNet model. Finally, the accuracy of the SCNet model is obtained by counting the number of images with correct labels. Note that the training accuracy is the correct rate of the training set, while the test accuracy is the correct rate of the test set. This paper separates the test program from the previous programs so that the test program can be called as a subroutine during the training process. After each training epoch, the test program and the latest model are utilized to test the training set and test set, respectively. The training accuracy and test accuracy of the updated model are thus obtained. Finally, the next epoch is trained on the original model.

4.3 Results Analysis

Training the neural network is a very complex procedure, and it usually takes a long time (it may reach few days or weeks). In order to facilitate the debugging and the optimization of the neural network, a TensorFlow visualization tool (the TensorBoard) is used. This latter efficiently shows the composition of the calculation graph and the trend of some indexes over time.

The executive process of the SCNet model is similar to the training process（Figure 4）. Firstly, an unknown image with a size of $224\times224\times3$ is used as input to the model. The image size changes to $112\times112\times32$, $56\times56\times64$, $28\times28\times18$, $14\times14\times256$, $7\times7\times512$ and $3\times3\times512$, after the 1st, 2nd, 3rd, 4th, 5th and 6th large convolution layer, respectively. The size of the image then changes to $1\times1\times3$ after three full-connection layers leading to three classifications. Finally, the image enters the softmax layer, and the probability for each corresponding classification is obtained. The state of the concrete surface from this image can then be identified.

The efficiency of a neural network is determined by its performance on the test data, because its ultimate purpose is to judge the classification results of the unknown data. In this paper, the updated model is saved at each training epoch of the SCNet model. The model accuracies on the training set samples and the test set samples are verified and recorded.

4.3.1 COMPARISION OF THE NETWORK MODEL RESULTS WITH DIFFERENT STRUCTURAL FRAMEWORKS

A network structural framework is crucial for deep learning CNN. The number of hidden layers, the location of the convolutional pooling layer and the size of the fully connected layer affect the training accuracy and the test accuracy. Figure 5 compares the SCNet model with classic classification neural network model VGG16, when all the network parameters and training hyperparameters are consistent.

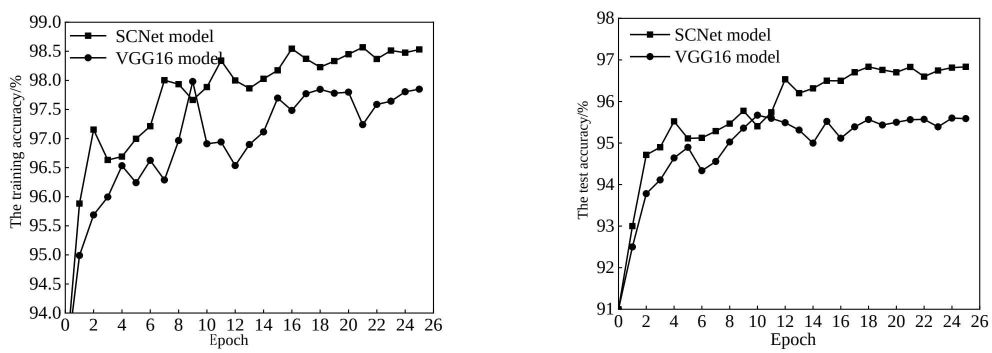
After comparing the performance of the two models on the data set, it can be seen that although the behavior is basically the same, the proposed SCNet model has a higher accuracy in both the training set and testing set. The proposed model reaches a training accuracy of 98.5% and a test accuracy of 96.8%.

Moreover, this paper uses optimization methods, such as the moving average, activation function, exponential decay learning rate and regularization loss, in the training process. In order to investigate the efficiency of these optimization methods, we use the control variable method which consists in comparing the model in the case where all the optimization methods are used, and the case where one of them is excluded.

4.3.2 COMPARISION OF THE NETWORK MODEL RESULTS WITH DIFFERENT TRAINING HYPERPARAMETERS

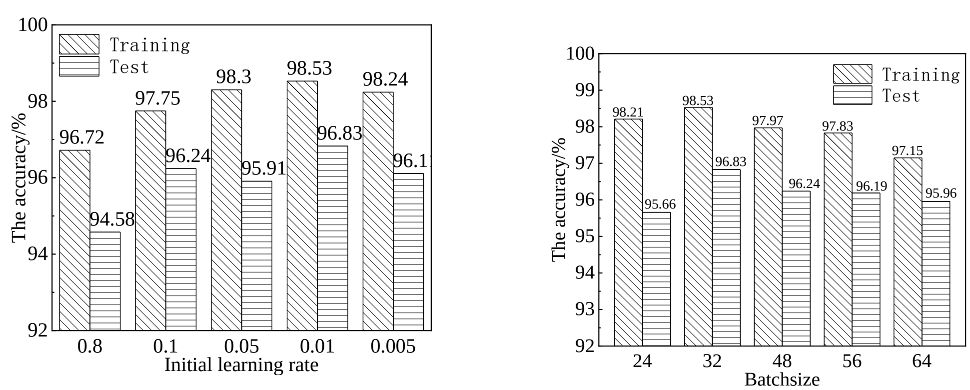
The model hyperparameters cannot be directly estimated from the internal data and should be manually specified. There is no theoretical basis leading to the optimal value of the model hyperparameters. However, the optimal value of the model hyperparameters can be adjusted by the accurate results of the model's classification with empirical judgments, based on a large number of experiments. The hyperparameters of the proposed SCNet model include the initial learning rate, exponential attenuation learning rate, batch size, back propagation optimizer and number of epochs.

Figure 6 shows the training accuracy and testing accuracy of several neural network models after 25 epochs of training on crack data set, when only one hyperparameter is changed. It can be seen that the initial model learning rate of 0.01 has the highest accuracy when the initial learning rate is 0.8, 0.1, 0.05, 0.01 and 0.005, respectively. The model of batch size 32 has the highest accuracy when the batch size of the models is 24, 32, 48, 56 and 64, respectively.



a) b)

Figure 5. Comparison of the accuracy of two different structural frame models. a) Comparison of the training accuracy; b) Comparison of the test accuracy.



a) Initial learning rate b) Batchsize

Figure 6. Comparison of the training accuracy and test accuracy for different neural network models with different training hyperparameters

5. CONCLUSTION

This paper proposed a SCNet three-classification neural network model for the classification of reinforced concrete corrosion cracks based on deep learning CNN. Compared with the traditional manual inspection methods, SCNet significantly improves safety by eliminating the need for inspectors to take field photos on high-rise buildings or bridges. Moreover, the SCNet method does not require visual judgment or installation of several sensors. Therefore, it reduces cost, saves manpower and highly increases the detection efficiency and accuracy. The main conclusions of the paper are summarized as follows:

1) Unified image data acquisition processes and shooting standards are set to make the features of the image data clear and explicit. This is convenient for subsequent neural network training and test classification.
2) A concrete corrosion image recognition and classification model (SCNet) is built using TensorFlow and Python. By adjusting the neural network structure, using optimization methods and training hyperparameters, the SCNet model can distinguish the corrosion cracks of steel bars with an accuracy of 96.8%.
3) The existing applications for concrete crack recognition based on conventional CNN, only perform the binary classification of the concrete conditions. On the contrary, the proposed model can identify intact concrete, concrete corrosion cracks and concrete cracks caused by any other reasons. Therefore, it can perform the triple classification of reinforced concrete cracks.

ACKNOWLEDGMENTS：

This research is supported by the National Natural Science Foundation of China (Grant No. 52078173)

REFERENCES

1．Lu CH, Zhao YX, Jin WL. Modeling of time to corrosion-induced cover cracking in reinforced concrete structures. Journal of Building Structures 2010, 31(02):85-92. [In Chinese].

2．Anitha MJ，Hemalatha R，Radha S. A Survey on Crack Detection Algorithms for Concrete Structures. Part of the Advances in Intelligent Systems and Computing book series (AISC, volume 1163), DOI:10.1007/978-981-15-5029-4_53, 2021:639-654.

3．Xu XD. Application study of ultrasonic nondestructive testing technology in bridge health condition assessment. Jilin University 2008; 8-13. [In Chinese].

4．Kurata M, Kim J, Lynch JP, et al. Internet-enabled wireless structural monitoring systems:development and permanent depolyment at the New Carquinez Suspension Bridge. Journal

of Structural Engineering 2012; 139(10): 688-702.
5．Jang S, Jo H, Cho S, et al. Structural health monitoring of a cable-stayed bridge using smart sensor technology: deployment and evaluation. Smart Structures and Systems 2010; 6(5-6): 439-459.
6．Cha YJ, Chen JG, Buyukozturk O. Output only computer vision based damage detection using phasebased optical flow and unscented Kalman filters. Engineering Structures 2017; 132: 300-313.
7．Lecun Y, Bengio Y, Hinton G. Deep learning. Nature 2015; 521(7553): 436-444.
8．Ciresan DC, Meier U, Masci J. Flexible, high performance convolutional neural networks for image Classification. In Proceedings of International Joint Conference on Artificial Intelligence 2011; 15(6): 1234-1242.
9．Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems 2017; 60(6): 84-90.
10．Cai GM. Study on the Video Detection system of Crack on the bottom of high-speed rail bridge. Beijing Jiaotong University 2011; 8-10. [In Chinese].
11．Landstorm A, Thurley MJ. Morphology-based crack detection for steel slabs. IEEE Journal of selected topics in signal processing 2012; 6(7): 866-875.
12．Moon H, Kim J. Intelligent crack detecting algorithm on the concrete crack image using neural network. In Proceedings of the 28th ISARC 2012; 22(6): 1461–1467.
13．Zhang L, Yang F, Zhang DY, et al. Road Crack detection using deep convolutional neural network. 2016 IEEE International Conference on Image Processing (ICIP) 2016; 3708-3712.
14．Cha YJ, Choi W, Buyukozturk O. Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks. Computer-aided Civil and Infrastructure Engineering 2017; 32(5): 361-378.
15．Simonyan K, Zisserman A. Very deep convolutional networks for Large-Scale image recognition[J]. Computer Science，2014，53（07）：88-95.
16．Sukegawa S, Yoshii K, Hara T, et al. Deep Neural Networks for Dental Implant System Classification [J]. Biomolecules, 2020, 10(7): 13.