Application Research on Convolution Neural Network for Bridge Crack Detection

Jinghang Cen, Jiankang Zhao [a], Xuan Xia, Chuanqi Liu

School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

[a]zhaojiankang@sjtu.edu.cn

Keywords: image processing, bridge crack, detection, convolution neural network.

Abstract. The bridge crack detection still relies on human visual measurement nowadays, which means low efficiency and high cost. In view to this situation, the convolution neural network (CNN) was introduced into bridge crack detection to improve the efficiency and reduce the error caused by manual work. In this paper, we designed a recognition algorithm based on convolution neural network, which can directly input the crack images obtained by UAV, avoids the complicated feature extraction pre-processing used in the traditional image processing. Basis on that, this paper puts forward an algorithm of image filter window with variable size control, and also carries on the experiments to set the threshold value of the length of the connected domain. Finally, an intelligent recognition scheme is obtained, which is suitable for crack recognition, and the recognition rate is over 95%, means that it has a good application prospect.

1. Introduction

Bridge engineering plays an important role in China's transportation system. The safety issues of the bridge cracks relate to our property and personal safety directly. The current method used in crack detection is still mainly rely on human eye observation. This method is not only inefficient, but also has a potential threat on personal safety, and will introduce a larger of subjective factor interference [1].

In recent years, many domestic and foreign experts have been committed to the application of digital image processing methods in the detection of bridge cracks. In order to extract the characteristics of the cracks more accurately, researchers have proposed methods such as Minimum Path Selection (MPS)[2-4], Minimum Spanning Tree(MST)[5-6], Crack Fundamental Element (CFE) [7-8] and so on. These traditional image processing methods need to detect the characteristics of objectives at first, and then use a specific pattern recognition algorithm to recognize and classify. Therefore, the accuracy of the recognition and classification is not only relates on the pattern recognition algorithm, but also depends on the selection of the objectives' feature. So the traditional image processing methods have great limitation because the target's feature needed to be specified and designed by human.

With the development of machine learning, the computer began to have a ability to learn the potential of the laws and characteristics from a large number of existing data [9]. Especially, when the convolution neural network combine the feature extraction and target classification into one, the objectives feature extraction is no longer dependent on manual design, but on the automatic machine learning, excluding the artificial interference such as unreasonable feature design. At present, the machine learning algorithm has been introduced into the field of image processing, which may replace the combination of human design and machine learning, and become the mainstream image recognition technology [10]. For example, in [11], the convolution neural network LeNet-5 is introduced into the study of license plate character recognition. And then the detection don't need to denoise the images or select the characteristics from the character.

In this paper, a recognition algorithm based on convolution neural network is designed, which can directly input the crack images obtained by UAV, and avoid the complicated pre-processing, which is a necessary feature extraction method in traditional image processing. On this basis, this paper puts

forward an algorithm of input image filter window with variable size control, and carries on the experiments to set the threshold value of the connected domain length. The final experiment shows that this method can identify the cracks quickly and effectively, has a good application prospects.

2. Bridge Crack Recognition Algorithm

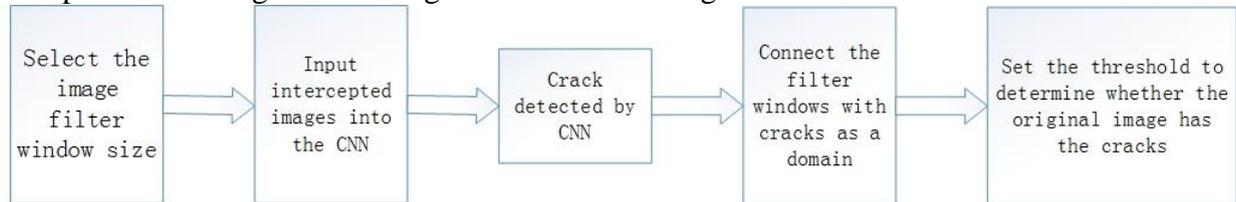The process of bridge crack recognition is shown in Fig.1:

Fig. 1 Bridge crack recognition algorithm

Select the size of the filter window to get the screenshots, then input the selection part into the CNN to identify. As the training images are 48×48, filter window length should set to be a multiple integers of 48. For example, if you set the length of the filter window of 96, so each time 96×96 size of the picture will be extracted, and compressed to 48×48, then input into the CNN.

The filter window's step length of movement is half of its side length. For example, if the filter window is 96×96, then the move step is 48.

Input the selection part into the CNN to identify.

The filter window portion which is judged with crack is marked to form a connected domain. Note the diagonal of the smallest circumscribed rectangle of the connected domain as the length of the connected domain. Select the appropriate threshold, if the length of the connected domain exceeds the threshold, it is determined that there is a crack in the original image. If the threshold is too large, the small cracks will be ignored, resulting in missing, if it's too small, will be some of the interference items identified as cracks, resulting in error.

3. CNN parameters design and pre-processing

3.1 The structure of the CNN

Convolution neural network (CNN) is a two-dimensional network model in deep learning network. Its weight-sharing network structure makes it more similar to biological neural network, reduces the complexity of network model and reduces the number of weights, so that the image can be input directly, and the complex feature extraction and data reconstruction process in the traditional identification algorithms can be avoided [12].

Convolution neural networks recognize the displacement, scaling, and twist invariance by combining three methods: local receptive field, weight-sharing and sub-sampling [13]. The local receptive field refers to that the neurons of each layer only connect with the neural cells in a small neighborhood of the upper layer. Through the local receptive field, each neuron can extract the primary visual features, such as the direction of the line, Endpoints, corners, etc. Weight-sharing makes the CNN fewer parameters and little training data. And sub-sampling can reduce the resolution of the feature, to achieve the displacement, scaling and other forms of distortion of the invariance.

In general, after the convolution layer, there will be a sub-sampling layer to reduce the computational time and to establish the controls and structural invariance [11].
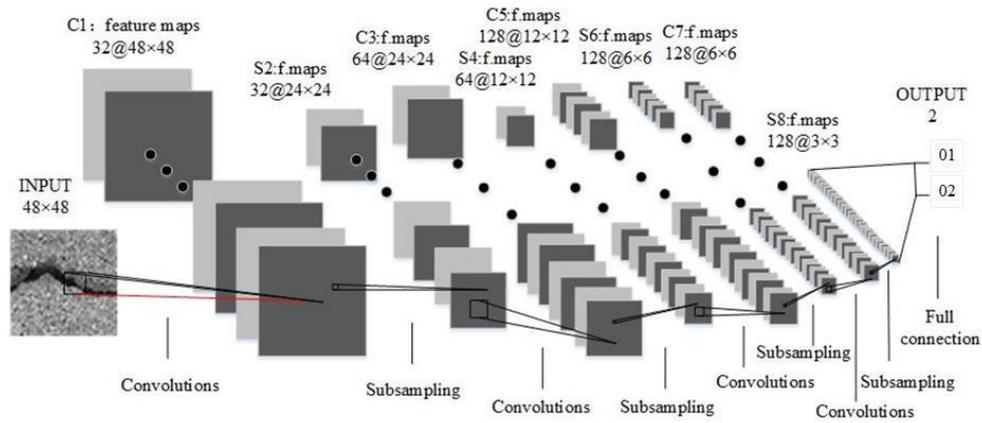
Fig. 2 Convolution neural network model

The structure of the convolution neural network is shown in Fig.2. The input of each neuron comes from a local neighborhood of the previous layer and is weighted by a set of weights. These features are extracted to form a higher level feature in the next layer. The neurons with the same feature share the same set of weights.

In consideration of the contradiction between that the neural network input layer should not be too large, and the crack information should be retained as much as possible, the final size of the input layer training image is decided to 48×48. In order to get the 48×48 images after the convolution, and to keep the border information, we fill the input image with 0 and make its size to be 58×58. The first convolution layer (C1) contains 32 convolution feature images, convoluted by the 11×11 size convolution filter window to extract the intrinsic characteristics. And after that, each image size is still 48×48. Then, the 32 feature images of C1 layer are sampled by 2×2 maximum pool sampling to obtain 32 new feature images with size of 24×24.

The C3 layer is expanded the image size to 30×30 at first, and then convolved by 64 $7 \times 7 \times 32$ convolution filter windows, and obtains 64 24×24 images. After the same sampling, 64 12×12 images are obtained.

Similarly, the C5 is expanded to 16×16, through 128 $5 \times 5 \times 64$ convolution window, we can get 128 12×12 feature images, after the same pooling operation, 128 6×6 features images are obtained.

Finally, the C7 layer is expanded to 8×8, through 128 $3 \times 3 \times 128$ size of convolution window, we can get 128 6×6 feature images, after the pooling operation, 128 3×3 feature images are obtained. And then, through two fully connected layer with size of $3 \times 3 \times 128$, we can get two classification mark, 01 or 02, the former means having cracks, the latter means absence of cracks.

In addition, the CNN in this paper uses the ReLu function and the maximum pool sampling method.

3.2 Neural network related parameter

The experimental environment is Matlab R2014b, Inter Core CPU 3.30GHz, memory 4GB, and 64-bit operating system.

Add GPU support the MatConvNet, learning rate is logspace(-2,-9,1000), momentum is 0.9, numEpochs is 50, and the batchSize is 200.

3.3 Cross-validation

The purpose of cross validation is to obtain a reliable and stable model, the specific approach is to take out most of the data modeling, leave a small part of the sample to be predicted by the newly established model, and find out the prediction error of the small part, record their sum of squares. This process will be carried out until all the samples are predicted once. Cross-validation is very effective in overcoming the over-fitting.

In this paper, we use the 5 fold cross-validation, that is, the sample data is divided into 5 copies, take turns using four of them as training data, the remaining one as testing data. The progress repeats five times till each subsample is validated once. The mean of the five results is the accuracy of the algorithm.

3.4 Training images preprocessing

(a)Comparison of transverse crack pretreatment

(b)Comparison of longitudinal crack pretreatment

(c)Comparison of oblique crack pretreatment
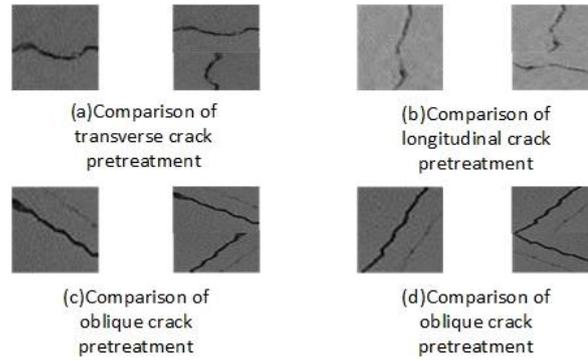
(d)Comparison of oblique crack pretreatment

Fig. 3 Pre-progress of the training images

The process is shown in Fig.3.

In order to highlight the spatial feature information of the crack, the input image is processed as follows: the original image of 48×48 size is compressed to 48×24, meanwhile, rotate original image 90 °and also compressed to 48×24, combine two of them into a new 48×48 ones as the final training picture.

4. Training results of the CNN

4.1 Training progress of the CNN

The training images are self-collected images, the size is 48×48, a total of two categories, respectively, 2304 crack samples marked as 01, and 5368 non-crack samples marked as 02.

Example of crack samples as shown in Fig.4, example of non-crack samples is shown in Fig.5
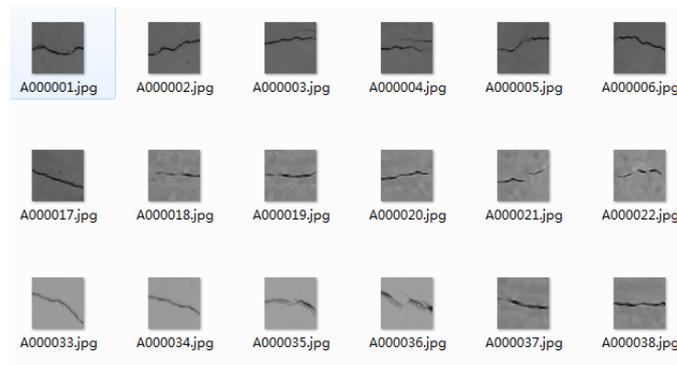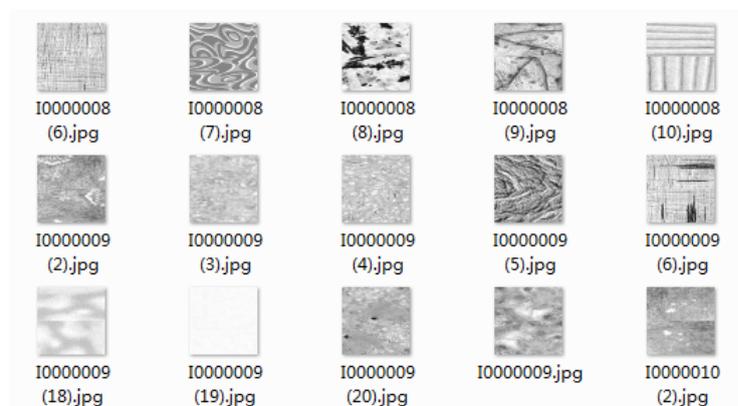

Fig. 4 Example of crack samples

Fig. 5 Example of non-crack samples

The neural network training process is shown in Fig.6, (a) shows the trend of the loss function of the training group and cross-verification group as the iterations increased. It can be seen from (a), the loss function approaches 0 in the iteration 24. Figure (b) shows the relationship between the top1 error and iterations. Top1 is the final prediction result, since our result only has two predict categories, so there

is a mistake unless the predictor is correct. So the top1 error is the probability of making a mistake. It can be seen from Fig (b), the top1 error of both training group and the cross-validation group reach the minimum value 0 when iteration is 24, which means that the neural network has a high degree of recognition.

So far, the neural network training is completed, we can use this CNN to test the test set images.
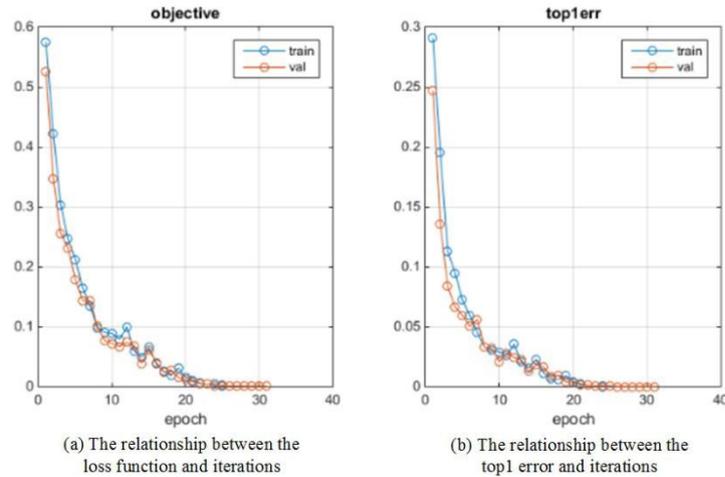

(a) The relationship between the loss function and iterations

(b) The relationship between the top1 error and iterations

Fig. 6 Training process of CNN

4.2 Experiment of test set images

Test images and training images were taken from a same batch of self-collected images, the size is also 48×48, with 131crack images and 314 non-crack images. So, a total of 445 images were identified by CNN and results were shown in table 1.

Table 1 Test results of 445 test set images

The number of images with cracks	The number of images without cracks	Recognition accuracy	Time of test(s)
125	320	93.13%	4.7513

The experimental results show that there is an unrecognized ratio of 6.87% for too small cracks. However, from the engineering point of view, too small cracks have not yet reached the point of threatening the security of the bridge, and the existence of cracks in the whole picture also need to consider the length of the connected domain. So the test results can be accepted.

5. Experimental results

5.1 Experiment of filter window size selection

Select 100 actual images with sizes of 1000×1666 or 1666×1000, of which 60 have cracks, 40 no cracks. When the determination threshold is 150 pixels, change the size of the filter windows, so the moving step is changed at the same time. Results examples of actual images shown in Fig.7.

(a)The length of the filter window is 48

(b)The length of the filter window is 48*2

(c)The length of the filter window is 48*3

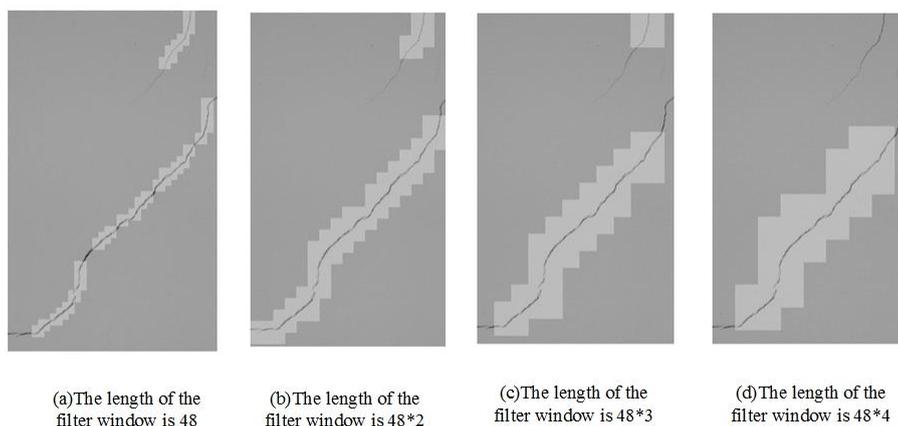(d)The length of the filter window is 48*4

Fig. 7 Examples of experimental results

The data of experimental results is shown in Table 2.

Table 2 The test results of 100 actual images on the filter window length

The length of the filter window(pixels)	The number of images with cracks	The number of images without cracks	The total test time of 100 images(s)	The test time of each image(s)
48*1	58	42	1513.14	15.1341
48*2	61	39	487.51	4.8751
48*3	58	42	214.06	2.1406
48*4	57	43	96.46	0.9646

It can be seen from Fig.7 and table 2 that when the length of filter window is 48, there may be fracture in the connected domain, causing the miss of the crack, and its test time is significantly longer than the other three groups. Meanwhile, when the filter window length chooses 48*4, though the speed of recognition is very fast, only 0.96s per image, but the accuracy is the lowest. So, considering both the accuracy and the efficiency, finally we choose 48*2 filter window, the group (b).

5.2 Experiment of threshold setting

Similarly as 5.1, use the actual images with size of 1000×1666 or 1666×1000 to test, of which 60 have cracks, 40 no cracks. The filter window length is set to 48*2 pixels, so the move step is 48 pixels. Change the crack determination threshold, the experimental results shown in Table 3.

Table 3 The test results of 100 actual images on the crack determination threshold

Determination threshold (pixels)	The number of images with cracks	The number of images without cracks	The total test time of 100 images(s)	The test time of each image(s)
50	64	36	378.13	3.7813
100	63	37	418.54	4.1854
150	61	39	394.87	3.9487
200	57	43	451.20	4.5120

As can be seen from table 3, we get the highest accuracy when the determination threshold is 150 pixels. So 150 pixels can be considered as a more appropriate choice in this image size. And there is no significant correlation between the threshold and the efficiency of image detection.

6.   Conclusion

This paper proposes a bridge crack identification algorithm based on CNN. Besides, proposed an algorithm about the controllable filter window, and carried out the experiments to set the threshold of the connected domain, realized the fast identification of the crack images. This method avoid the complex pre-processing of the input image, improves the recognition efficiency, solves the problem of low identification efficiency and high risk at present. It's beneficial to the development of automation and intelligence in the bridge crack detection. Future research is focused on how to improve the recognition speed under the premise of ensuring the recognition accuracy. We expect to realize the real-time identification in the future, which can greatly reduce the storage capacity of the database and make the bridge crack detection system more practical.

References

[1]Zhang Weifeng, Liu Meng, Yang Minghui. Bridge Crack Detection Technology Based on Digital Image Processing[J]. Modern Transpotation Technology 2008(5):34-36

[2]R Amhaz,S Chambon,J Idier,V Baltazart. A New Minimal Path Selection Algorithm for Automatic Crack Detection on Pavement Images[J]. Proc.IEEE ICIP, 2014:788-792

[3]M Avila,S Begot,F Duclty,T S Nguyen,2D Image Based Road Pavement Crack Detection by Calculating Minimal Paths and Dynamic Programming[J]. Proc.IEEE ICIP,2014:783-787

[4] R Amhaz,S Chambon,J Idier,V Baltazart. Automatic Crack Detection on 2D Pavement Images: An Algorithm Based on Minimal Path Selection[J]. IEEE Trans. Intell. Transp. Syst. 2015, 24p, DOI:10.1109/TITS.2015.2477675.hal-01206038.

[5] Q Zou,Y Cao,Q Li,Q Mao,S Wang. Crack Tree: Automatic Crack Detection from Pavement Images [J]. Pattern Recognit. Lett,Feb. 2012,vol.33,no.3:227-238

[6] K Fernandes,L Ciobanu. Pavement Pathologies Classification Using Grapg-Based Features[J]. Proc.IEEE ICIP,2014:793-797

[7] Y C J,C Jiang,Y Huang. Multiscale Crack Fundamental Element Model for Real-World Pavement Crack Classification[J]. J.Comput.CivilEng,2012,6(4):388-396,

[8] Y J Tsai,C Jiang,Z Wang. Implementation of Automatic Crack Evaluation Using Crack Fundamental Element[J]. Proc.IEEE ICIP,2014:773-777

[9] Arel I,Rose D C,Karnowski T P. Deep machine learning-A new frontier in artificial intelligence research[Research Frontier[J]. Computational Intelligence Magazine,IEEE,2010,5(4):13-18.

[10]Chen Xianchang. Research on Algorithm and Application of Deep Learning Based on Convolution Neural Network[D]. Zhejiang: Zhejiang Gongshang University,2013.

[11] Zhao Zhihong,Yang Shaopu,Ma Zengqiang. License Plate Character Recognition Based on Convolution Neural Network LeNet-5[J]. Journal of System Simulation,2010,22(3):638-641.

[12]Xu Shanshan. Research and Application of the Convolution Neural Network[D]. Jiangsu: Nanjing Forestry University,2013.

[13]LeCun Y,Bottou L, Bengio Y,Haffner P. Gradient-based learning applied to document recognition[C]//Proc.IEEE,1998. USA:IEEE,1998:2278-2324.