Smart Structures and Systems, *Vol. 28, No. 6 (2021) 779-789*
https://doi.org/10.12989/sss.2021.28.6.779

Railway sleeper crack recognition based on edge detection and CNN

Gang Wang and Jiawei Xiang*

College of Mechanical & Electrical Engineering, Wenzhou University, Wenzhou, 325035, P.R. China

(*Received January 24, 2021, Revised June 6, 2021, Accepted August 28, 2021*)

Abstract. Cracks in railway sleeper are an inevitable condition and has a significant influence on the safety of railway system. Although the technology of railway sleeper condition monitoring using machine learning (ML) models has been widely applied, the crack recognition accuracy is still in need of improvement. In this paper, a two-stage method using edge detection and convolutional neural network (CNN) is proposed to reduce the burden of computing for detecting cracks in railway sleepers with high accuracy. In the first stage, the edge detection is carried out by using the 3×3 neighborhood range algorithm to find out the possible crack areas, and a series of mathematical morphology operations are further used to eliminate the influence of noise targets to the edge detection results. In the second stage, a CNN model is employed to classify the results of edge detection. Through the analysis of abundant images of sleepers with cracks, it is proved that the cracks detected by the neighborhood range algorithm are superior to those detected by Sobel and Canny algorithms, which can be classified by proposed CNN model with high accuracy.

Keywords: convolutional neural network; edge detection; mathematical morphology operations; neighborhood range algorithm; railway sleeper cracks

1. Introduction

Railway sleeper is an important cornerstone of the railway system. Damages in sleepers may lead to economic losses even catastrophic casualties, and hence its health status directly affect the safety of the railway system (Siddhartha *et al*. 2018). In recent years, the health management and maintenance of railway sleepers has attracted the attention of many researchers (Zaharah *et al*. 2019, Zeng *et al*. 2020, Sysyn *et al*. 2019).

Damages will affect the strength and life of structures (Ada *et al*. 2018, Celik *et al*. 2018, Isik *et al*. 2020). In order to study the complex environmental effects and structural states in the process of structural operation, Liu *et al*. (2020) proposed a method for structural temperature missing data recovery based on long short time memory (LSTM) network. Do *et al*. (2019) proposed a damage detection method based on time series to detect, locate and quantify the damage of shear structure under variable mass effect. Researchers using finite element method to analyze the dynamic response of cracks for the usage of structural damage detection (Figueira *et al*. 2018, Jiang and Xiang 2020, Xiang *et al*. 2009, 2012, 2014, Haeri *et al*. 2018). Damage detection of beam-like structures has attracted many researchers (Xiang *et al*. 2011, Yang *et al*. 2012, Song *et al*. 2017) using finite element method based measurement techniques. In order to locate the damage location in a structure, Yang *et al*. (2018) proposed a damage localization method based on the singular value decomposition. Zhong and Xiang (2019) proposed an improved linear array to locate the impact on the stiff composite panel. Shemirani *et al*. studied the effect of confining pressure on shear behavior of joint bridge area by using discrete element method.

Timely detection of cracks on railway sleepers is very important to ensure the safety of railway transportation. The method of simulating human inspection process is considered to be a promising way for sleeper health inspection (Yella *et al*. 2009). For a long time, observation and listening to knocking sound are two common methods to check the cracks of railway sleepers, which are time-consuming and inaccurate. Therefore, the railway industry urgently needs a fast and accurate method of sleeper health detection. Janeliukstisa *et al*. (2019a) proposed a bending state monitoring technology for railway prestressed concrete sleepers using acoustic emission. Sengsri *et al*. (2020) used acoustic emission to detect the damage of fiber reinforced foam polyurethane composites railway bearings. Acoustic emission performs well in the concrete structure with uniform and single material. Whereas, it will be disturbed when there are steel bars or stones with different texture in the concrete structure.

Recently, through the analysis of sleeper surface images, it found that surface cracks in railway sleeper can be detected effectively by computer vision technology (Janeliukstis *et al*. 2019b, Xia *et al*. 2020). Classic image processing methods and artificial intelligence models, e.g., convolutional neural network (CNN), are two main methods (Gao *et al*. 2020, 2021) to identify sleeper cracks using sleeper surface images with and without cracks.

Classic image processing is to first extract features and then perform classification. The image features are

*Corresponding author, Professor,
E-mail: wxw8627@163.com

http://www.techno-press.org/?journal=sss&subpage=7
ISSN: 1738-1584 (Print), 1738-1991 (Online)

extracted by the feature extraction methods such as hog (Mizuno *et al*. 2012), sift (Hsu *et al*. 2012) and gray level co-occurrence matrix (Beura *et al*. 2015), and then classified by the support vector machine (Suykens and Vandewalle 1999), decision tree (Friedl and Brodley 1997) or other classification algorithms. The classification performance of the above mentioned models relied on the appropriate extracted features. Although there are many image feature extraction methods, it is still no universal rule to deal with images in different cases. Therefore, in order to select appropriate features for specific problems, a large number of experiments are needed for trial and error.

Among the entire classical image processing methods, edge detection and mathematical morphology operation are two commonly used image pre-processing methods (Chen and Chiu 2019). Edge detection is the process to characterize the brightness changes in the image, which can be used to detect cracks in the surface of structures (Torre and Poggio 1986, Prasath *et al*. 2020). In addition, the importance and influence of mathematical morphology in the fields of automatic vision detection, target recognition, image analysis and pattern recognition have been fully investigated (Shih 2010). The morphological operations of binary image can filter out the noise in the image and connect the adjacent targets (Gonzalez *et al*. 2004). The result of edge detection is a binary image with several noise areas, and a series of reasonable mathematical morphology operations can remove the noise areas in the binary image to improve the results of edge detection. However, image based computer vision technology alone cannot lead a high crack detection precision for the field wide applications, such as railway sleepers.

As a deep learning model, CNN (Lecun and Bengio 1995) ingeniously combines various neural network layers to learn potential crack features for classification. Since Alex team used CNN to finish the task of image classification in 2012 and achieved remarkable results (Krizhevsky *et al*. 2012), the tasks of classification, location, detection and segmentation in the field of computer vision have developed rapidly. In addition, researchers have explored deep learning models with different structures and depths to improve performance. In 2014, visual geometry group net (VGG-net) (Simonyan and Zisserman 2014) and Inception-net (Szegedy *et al*. 2014) were proposed successively and employed to classify images with higher classification accuracy. However, with the wide application of CNN in image classification, its shortcomings are gradually revealed. On the one hand, training a large network usually requires powerful computing equipment and a long time. Therefore, on the premise of ensuring the accuracy, it is very important to find a way to reduce the complexity of CNN model and speed up model training. On the other hand, a large number of raw images are needed to train a neural network with excellent classification ability for railway sleepers.

Comprehensive use of the above detection methods may bring better results. Franca *et al*. proposed an unsupervised method based on image processing, heuristics and feature fusion to count sleepers, identify types and defects, the accuracy of identifying visible defects of sleepers reaches 93% (Franca and Vassallo 2020).

In order to complete the task of railway sleeper crack recognition with lower calculation cost, the improved version of edge detection method (Canny 1987) might be a useful tool to filter the sleeper image without cracks using 3×3 neighborhood range algorithm and mathematical morphology operation. By the way, a series of data argumentation methods (Krizhevsky *et al*. 2012) are used to simulate and generate more abundant samples, to avoid the adverse effects of small sample number in CNN model training. This method does not care about the complicated feature extraction method and the adjustment of feature classifier, but focuses on how to highlight the crack information in the image. After obtaining the suspicious crack target, the proposed CNN model can accurately judge whether there are cracks in the image.

Based on the above analysis, a two-stage method using a improvement version of edge detection and CNN is proposed to reduce the burden of computing for detecting cracks in railway sleepers with high accuracy.The paper is arranged as follows. In section 2, the theories of 3×3 neighborhood range algorithm, mathematical morphology operation and CNN are introduced in brief. In section 3, the two-stage algorithm for detecting the crack of railway sleeper is described in detail. In section 4, the two-stage method is applied to detect the railway sleeper cracks. Compared with Sobel method and canny method, the superior of the present two-stage method is highlighted with efficient classification accuracy. In section 5, brief conclusions are given.

2. Theoretical background

As mentioned above, researchers have constructed many CNN models with different structures to solve various practical image processing tasks. (Krizhevsky *et al*. 2012, Simonyan and Zisserman 2014, Szegedy *et al*. 2014). The purpose of this paper is to combine the advantages of classical image processing means and the emerging CNN to identify sleeper cracks efficiently and accurately.

2.1 Edge detection

In digital image, the pixel value (or called brightness) corresponding to the edge position of the object will change dramatically. Prewitt operator (Seif *et al*. 2010) and Sobel operator (Gao *et al*. 2010) can detect the changes of pixel value in horizontal or vertical direction. In fact, the edge of an object is not always horizontal or vertical. Because the shape of objects in reality is not regular geometry, there are almost no horizontal or vertical edges of objects in digital images. In order to improve the edge detection in other directions, Canny edge detection algorithm (Canny 1987) are proposed. Canny algorithm combines horizontal and vertical edge detection operators, and calculates the gradient direction of pixel value change, so as to identify more accurate edge. The direction of edge is arbitrary, so when the algorithm can detect the edge in any direction, the detection performance will be better. Inspired by this, we propose a method that can detect edge information in any direction in one step. In the image of railway sleeper, the pixel value of crack location will change dramatically. Therefore, the range of 3×3 neighborhood around each pixel in the image can reflect the change degree of brightness value near the pixel. The 3×3 neighborhood rang algorithm can be used to detect cracks in the image of railway sleepers.

2.1.1 The 3×3 neighborhood range algorithm

The difference between the maximum value and the minimum value of the pixel value in the neighborhood is used to reflect the intensity of the pixel value change on the image.

$$\begin{cases} R(x,y) = \underset{i,j}{Max}[G(x+i,y+j)] \\ \qquad - \underset{i,j}{Min}[G(x+i,y+j)] \\ \qquad i,j = [-1,0,1] \end{cases} \tag{1}$$

where G(x, y) and R(x, y) represent the pixel value of row x and column y in the original image and the corresponding range of the 3×3 neighborhood, respectively.

It is worth noting that the range of i and j is [-1, 0, 1]. In the field of image edge detection, researchers often determine an area around a pixel to study the pixel value change of the image near the pixel. This small area is called neighborhood (Xie *et al*. 2018, Goel *et al*. 2020). The common neighborhood definitions include 4-neighborhood (adjacent pixels in the upper, lower, left and right directions of a pixel) and 8-neighborhood (8 pixels in the range of 3×3 around a pixel). This kind of small area should take a certain pixel in the area as the geometric center, otherwise, the standard of considering the intensity of image brightness change in different directions is not uniform, so it is not allowed to choose the neighborhood where the 2×2, 4×4 symmetrical center cannot fall on a specific pixel in the area. When we study the edge information in the railway sleeper image, we choose to do research in the 8-neighborhood of each pixel. On the one hand, the 8-neighborhood can express a little more direction information than the 4-neighborhood. On the other hand, in order to describe the cracks in sleeper images more accurately, we should use a small 3×3 neighborhood instead of a large 5×5 neighborhood, or larger. Fig. 1 is a schematic diagram of the 3×3 neighborhood range algorithm.

2.1.2 Morphological operations

The mathematical morphology operations mentioned in this paper are all aimed at binary images. The binary dilation (Gonzalez *et al*. 2004) of A by B, denoted A⊕B, is defined as the set operation

$$A \oplus B = \{z|(\hat{B})z \cap A \neq \emptyset\} \tag{2}$$

where $\hat{B}$ is the reflection of the structuring element B. In other words, it is the set of pixel locations z, where the reflected structuring element overlaps with foreground pixels in A when translated to z. The binary erosion (Gonzalez *et al*. 2004) of A by B, denoted A⊖B, is defined as the set operation

$$A \ominus B = \{z|Bz \subseteq A\} \tag{3}$$

Dilation operation and erosion operation are the most basic operations of mathematical morphology operations. All mathematical morphology operations are constructed by reasonable combination of dilation and erosion. For example, the operation of morphological closure consists of two steps: first dilation operation, and then erosion operation, which can remove the small holes on the structure, connect the adjacent structures, and make the structure contour smooth.

2.2 Image classification based on CNN

Among the digital image processing methods, convolution transform is a common operation. It is widely used in image filtering, edge detection and feature extraction (Yin *et al*. 2020, Viswanath *et al*. 2011). CNN is a kind of artificial neural network which uses convolution operation to extract image features. The difference between convolution layer in CNN model and traditional image convolution operation is that the weight of convolution kernel in CNN is learned by the network itself by using gradient descent method. The convolution kernel in other convolution operations is set manually.

A typical CNN model is usually composed of convolution layer, subsampling layer (also known as pooling layer), fully connected layer, etc. (Wang and Xiang 2019). The convolution layer uses convolution kernel to extract image features, and the subsampling layer can reduce the computation and improve the generalization

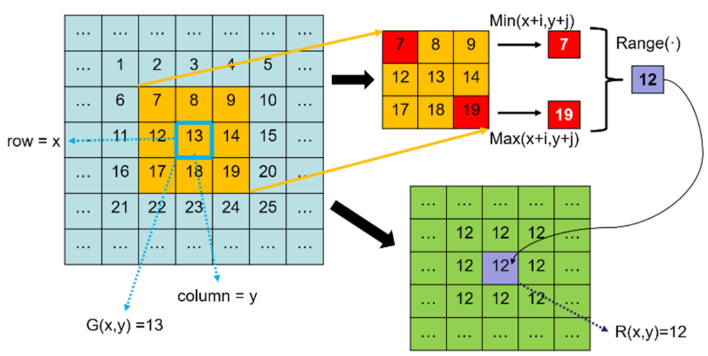

Fig. 1 Schematic diagram of the 3×3 neighborhood range algorithm

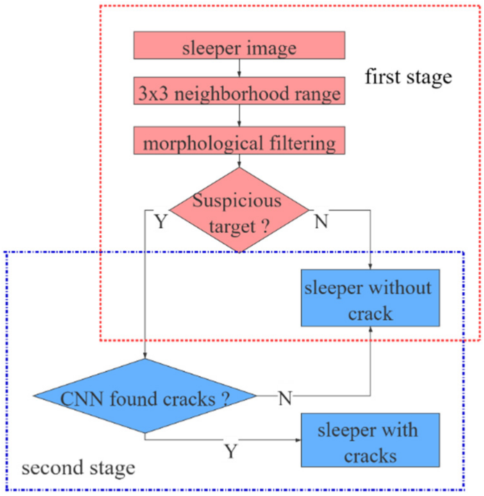

Fig. 2 The algorithm flow chart

ability of the model, the fully connected layer uses the feature classification image learned from the previous layers.

3. The two-stage method to detect cracks in railway sleepers

A two-stage algorithm based on edge detection and CNN is proposed to detect railway sleeper crack. The algorithm flow chart is shown in Fig. 2.

In the first stage, the 3×3 neighborhood range algorithm is employed to find out the possible edge of sleeper crack, and mathematical morphology operations are used to eliminate the noise target in the edge detection results. In the second stage, CNN is used to classify sleeper with or without crack.

3.1 Edge detection

Image pre-processing means to highlight the content of interest in the image through some simple methods. In this paper, the 3×3 neighborhood range algorithm is used to extract the crack edge in the sleeper image. The specific steps are as follows: Firstly, the range value of the 3×3 neighborhood range. The range of 9 pixels in the 3×3 neighborhood of each pixel are calculated, and the result is preserved as a new feature map. Secondly, fixed threshold binarization. Fixed threshold (with the threshold increases, the foreground content in the binary results decreases, and vice versa.) is used to binarize the image.

It must be note that the adaptive threshold cannot used here. Because the magnitude of each pixel represents the intensity of the change of the brightness value at the location in the original image, rather than the brightness of the pixel in the real image. If adaptive thresholds are used, the adaptive thresholds in different pictures are likely to be different, which lead to the failure to unify the criteria for judging cracks in different pictures.

3.2 Image morphological filtering

The operation of edge detection can detect the position where the brightness changes greatly. In the sleeper image, in addition to the edge of the crack, there are also many noise locations which will cause dramatic changes in the surrounding brightness. A series of mathematical morphology operations are used to eliminate the noise in the feature image obtained by edge detection. Elimination of noise target includes such operations:

Step 1: Eliminate noise points. All the 8-connected regions less than 4 pixels in the binary image are defined as noise, and these regions are filtered out.

Step 2: Morphological closed operation. As shown in Figs. 3(b) and (c), the actual crack is often a continuous and irregular curve, but the detected targets may have small holes or discontinuity. Morphological closed operation links the close targets into a continuous curve, not including the distant target. In addition, closed operation eliminates the small holes on the target.

Step 3: Eliminate noise area. The real crack will not be too small or too large. Small targets are likely to be caused by random noise. They cannot be connected into a curve after closing operation, so they are independent small object. If some noises are close to each other, the closed operation will connect these noises to form a big object. It is easy to observe that the width of the crack in the image is about 10 pixels (the height and width of the image are 1200 and 1400, respectively), so a small crack contains at least 10×50 pixels and a large crack contains at most 20×1400 pixels. Therefore, it is necessary to filter targets with too small and too large area caused by noise.

Step 4: Fill. Morphological closure can eliminate

(a) (b) (c)

Fig. 3 Sleeper images. (a) without crack; (b) with crack; (c) with crack and waterpipe

Table 1 The proposed CNN structure

Type	Filters number	Kernel size	Strides	Activation	Output shape
Input					(512,640,1)
Convolution	16	(5,5)	(2,2)	Relu	(512,640,16)
Max Pooling		(2,2)	(2,2)		(256,320,16)
Convolution	32	(5,5)	(2,2)	Relu	(128,160,16)
Max Pooling		(2,2)	(2,2)		(64,80,32)
Convolution	64	(3,3)	(1,1)	Relu	(32,40,32)
Max Pooling		(2,2)	(2,2)		(32,40,64)
Convolution	128	(3,3)	(1,1)	Relu	(16,20,64)
Max Pooling		(2,2)	(2,2)		(16,20,128)
Convolution	256	(3,3)	(1,1)	Relu	(8,10,128)
Max Pooling		(2,2)	(2,2)		(8,10,256)
Flatten					(4,5,256)
Dropout					(5120,)
Dense				Sigmoid	(1,)

micropores, but not macropores. As shown in Fig. 3(b), there are some railway sleepers with water-pipe with edge detection result is a closed circle, which can be effectively distinguished from ordinary crack targets according to the area after filling operation.

The aim of this paper is to determine whether there are cracks in railway sleeper images. Through the edge detection and noise suppression described above, we get a binary image, in which the foreground represents that the suspicious target is detected in the original image, and the background represents that there is no target in the original image, shown in Fig. 4. If the edge detection result of a sleeper image does not contain any suspicious target, then the sleeper image cannot have cracks. However, if there are suspicious targets in a sleeper crack image, these targets are not necessarily crack. Therefore, the proposed edge detection method can filter out some sleeper images without cracks, but it is unable to determine whether the detected object is really a crack.

3.3 CNN model

A simple and reliable edge detection method is proposed to detect suspicious targets in sleeper cracks. Although it cannot identify whether the suspicious target is a crack or not, the suspicious target in the original image is clearly represented by the edge detection result, which is benefit to further distinguish the classes of suspicious targets. Firstly, many background pixels unrelated to cracks become zero, which highlights the target pixels that may be cracks. Secondly, the value space of pixel value in the original image is 0-255, however, edge detection algorithm makes the value space of pixel become zero or, which is a binary space. On this basis, a CNN model is built to judge whether the detected suspicious target is a real crack.

A CNN model with five convolution layers stacked in sequence is proposed to identify the cracks. The structure of proposed CNN model is shown in Table 1. In this CNN model, the number of filters, kernel size and strides of each convolution layer are (16-32-64-128-256), (5,5,3,3,3) and (2,2,1,1,1) respectively, and each convolution layer is followed by a subsampling layer.

At the end of the model, the three dimensional convolution output data is changed into one dimensional data. Then a fully connected layer with 2 neurons is stacked, whose activation function is sigmoid. The sigmoid function is defined as follows

$$f(x) = \frac{1}{1 + e^{-x}} \tag{4}$$

where x represent input, and f(x) represent the output when input is x. This activation can adjust the range of the CNN model output to (0~1). It will benefit to model training and predicting.

4. Experimental investigations

In order to test the actual performance of the proposed algorithm in crack detection, we have carried out the experiment with a railway sleeper crack data set, which is from Engineering failure prediction and health management technology Specialized Committee. In order to balance the number of different classes of samples, 500 cracked sleeper images and 500 healthy sleeper images were selected. Each image is a 1200×1400 gray image. By the way, the images of railway sleepers contain downpipes, impurities, and stains.

The algorithm proposed in this paper is consists of two stages, so the experiment is also divided into two steps. The first step of railway sleeper crack detection was completed in MATLAB 2016, and the second step of CNN model construction and training was completed in tensorflow2.0. Our code and data are as follows: https://github.com/famer3riots/crack-detection.

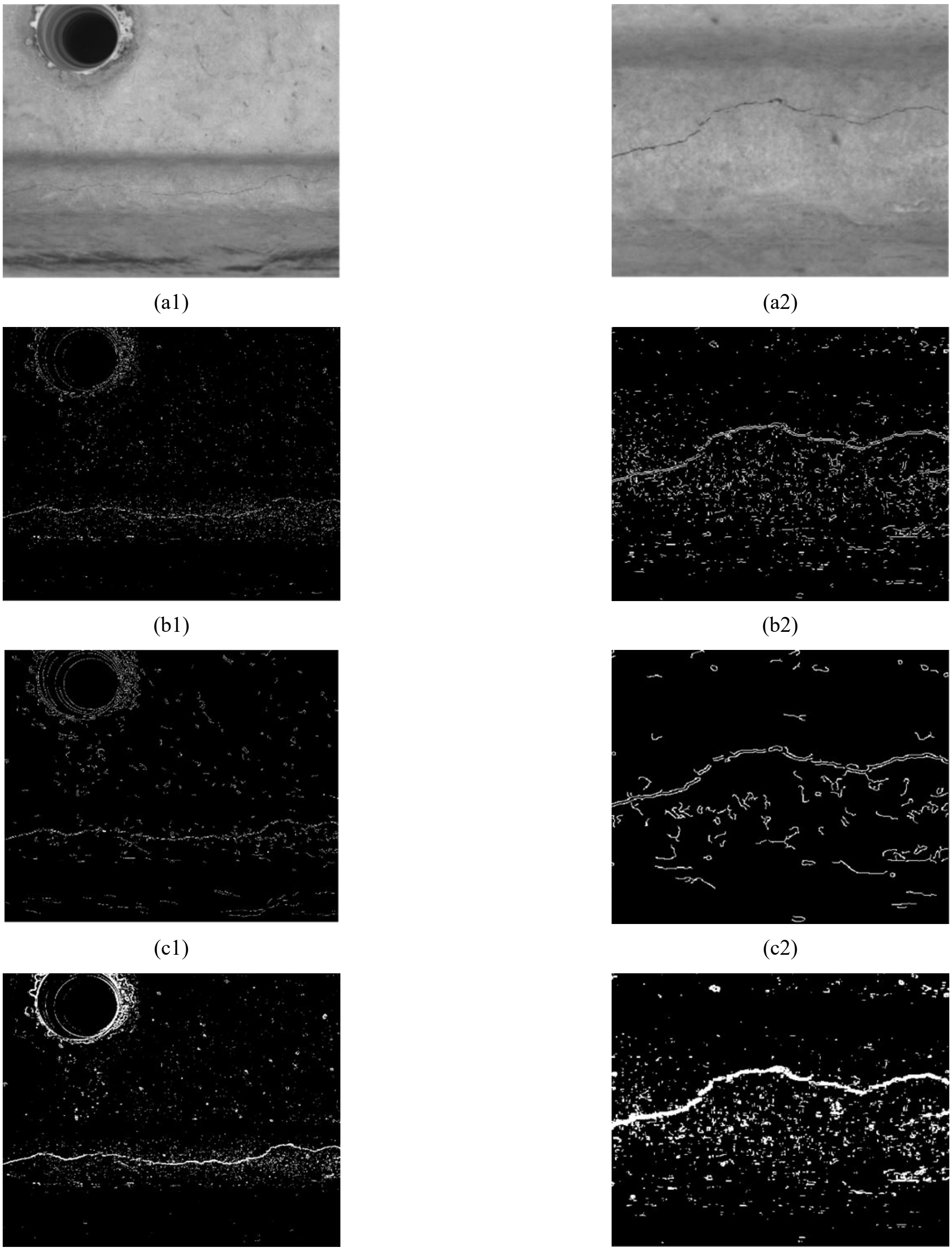

Fig. 4 Edge detection. (a) original image; (b), (c), (d) edge detection results of by sobel/canny/the 3×3 neighborhood range

4.1 The 3×3 neighborhood range algorithm detects cracks edge

As shown in Fig. 4, a1 is a sleeper image with cracks, b1/c1/d1 is the result of edge detection of a1 through sobel/canny/the 3×3 neighborhood range, and a2/b2/c2/d2 is the local magnification of the same position on the left side of a1/b1/c1/d1, respectively. The 3×3 neighborhood range algorithm combined with the fixed threshold binarization method can detect stronger, clearer, and more continuous crack edges than Sobel algorithm and Canny algorithm. Due to the extremum of neighborhood range algorithm appears randomly in any position of 8-neighborhood, this method does not tend to detect the edge information in a specific direction. Therefore, the cracks detected are robust and continuous.

4.2 Image morphological filtering

The 3×3 neighborhood range algorithm is used to detect the position where the brightness of sleeper image changes dramatically, and the result is filtered by the proposed

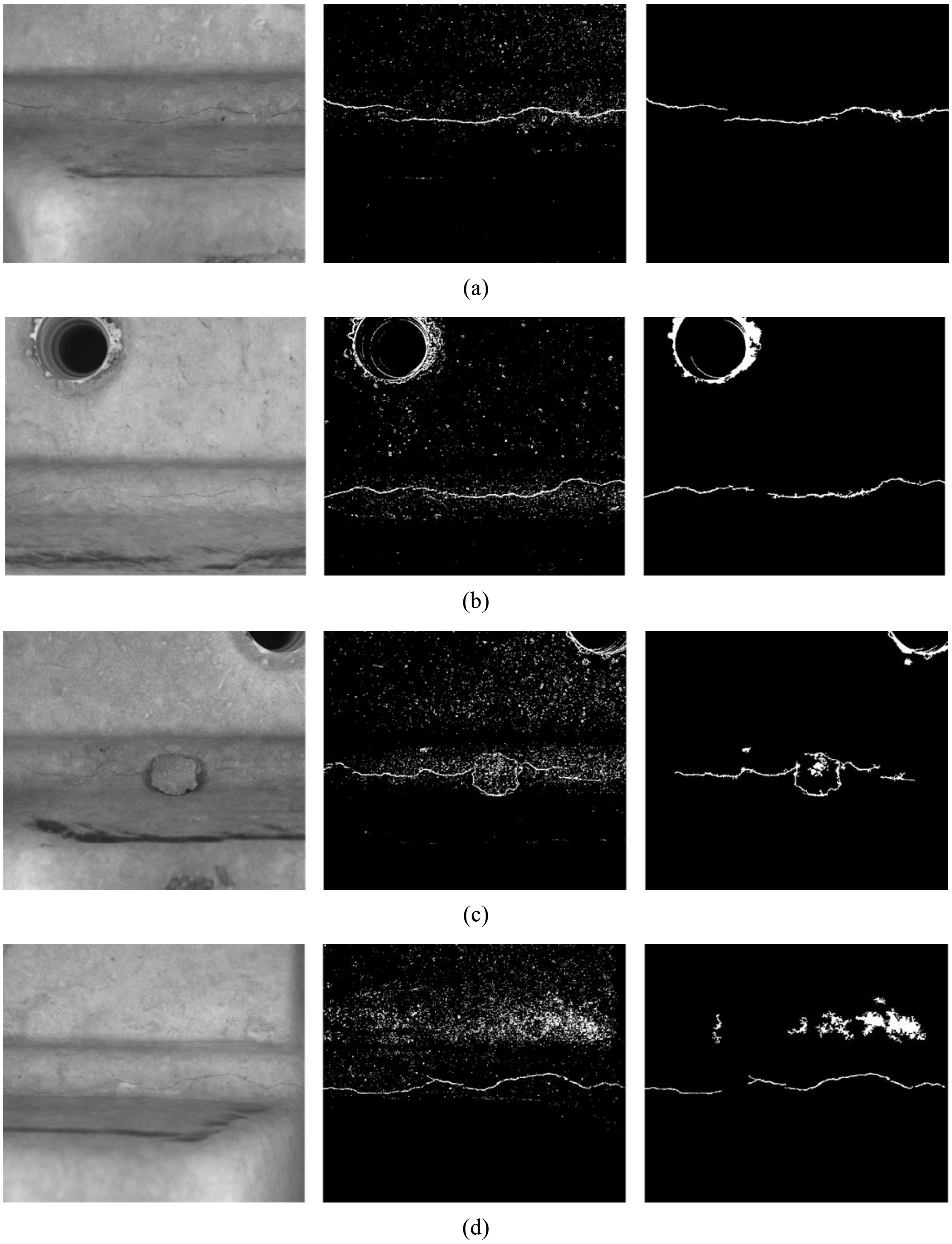

Fig. 5 The proposed method detects cracks in sleeper images. Left is the original image; Middle is the edge detection result; Right is the morphological filtering result

morphological filtering algorithm. These experiments are used to demonstrate the advantages of our algorithm in edge detection and filtering. As shown in the Fig. 5, the left side is the original image, the middle part is the 3×3 neighborhood range edge detection result, and the right side is the morphological filtering result.

The experiment shows that the proposed 3×3 neighborhood range algorithm can detect the cracks accurately, but also detect the pipe and many random noises. Mathematical morphology operation is used to filter out the noise in edge detection results. On the one hand, morphological filtering can eliminate the isolated random noise in the edge detection result, but cannot eliminate the real cracks and pipes. On the other hand, for those random noises with concentrated distribution, they become connected noise regions after morphological filtering and cannot be eliminated.

After testing the 500 sleeper images with cracks and 500 images without cracks in the railway sleeper crack data set described above, the number of sleeper images with or without cracks predicted in the detection results are recorded in Table 2 respectively.

Table 2 Edge detection result

True \ Prediction	**Cracks**	**No crack**
Cracks	479	21
No crack	225	275

According to the different detection results and real images, there are four cases:

(1) There are cracks in the image and cracks in the detection results.
(2) There are cracks in the image and no cracks in the detection results.
(3) There is no crack in the image and there is a crack in the detection result.
(4) There is no crack in the image and no crack in the detection result.

In these four cases, (1) and (4) are correct, while (2) and (3) are false detection results. In the detection process, the number of foreground pixels in the detection results can be artificially controlled by adjusting the binarization threshold in edge detection operation and the definition of noise size in morphological filtering, so as to adjust the proportion of the above four situations. In the detection of railway sleeper cracks, our purpose is to detect as many cracks as possible, do not omit cracks. This increases the proportion of cases (1) and (3) in the test results. The probability of detecting crack from the image with crack is 0.9580, and that from image without crack is 0.45.

4.3 CNN identifies cracks

In the previous experiments, some images are detected as suspicious targets, which need to be further used to determine whether there are cracks in the images. Under the condition that the proportion of the images with cracks in the training samples and the test samples is the same, these images are divided into training samples and test samples according to the ratio of 4:1.

4.3.1 Data argumentation

Data argumentation is a common method to expand datasets and improve the generalization ability of models, in machine learning filed. (Han *et al*. 2018). Image data argumentation is to simulate multiple different images of an object in different directions, positions and illumination conditions by means of flipping and adding noise to an image of an object (Shorten and Khoshgoftaar 2019, Perez and Wang 2017). Before inputting the image of steel into CNN model, we use the method of data argumentation to the image.

In this paper, we choose random clipping, compression, and random flipping to generate more samples. For training sample data enhancement, the steps are as follows: input image (1200,1400), random clipping (1024,1280), compression (512,640), random left and right flip (512,640), random up and down flip (512,640); for the test sample, there is no need for data enhancement, but to ensure that the image size is consistent when inputting CNN, the test sample needs the following steps: input image (1200,1400), center position clipping (1024,1280), compression (512,640). As shown in Fig. 6, image (a) is the result image of edge detection, and the remain three images are the images with different postures obtained by the above method.

In data argumentation procedure, the size of the original image will be compressed, which reduces the burden of computing devices. As shown in (b), (c) and (d) in Fig. 6, more positions and directions are simulated by random

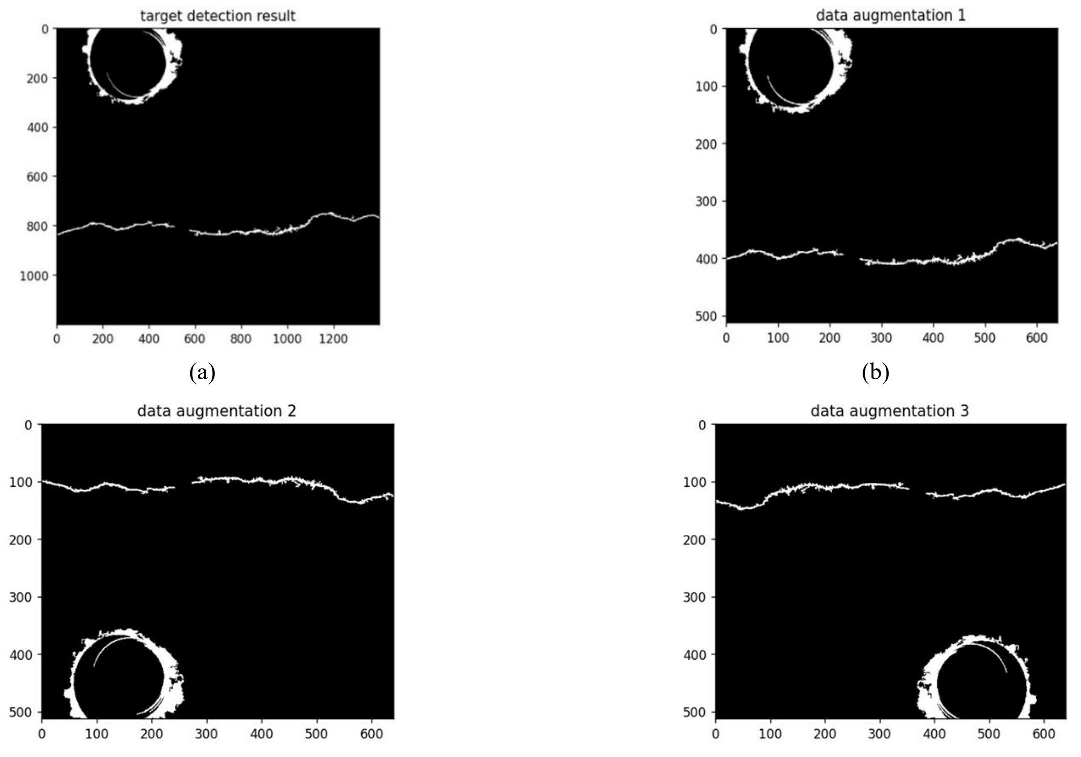

Fig. 6 Data argumentation: (a) edge detection result (1200,1400); (b), (c), (d) data argumentation 1/2/3 (512,640)

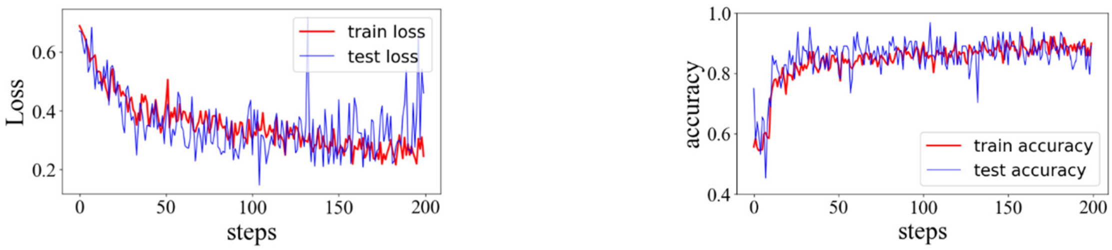

Fig. 7 Loss and accuracy curve of CNN model training

clipping and random flipping. Finally, the image size after data argumentation is (512,640) and its number of channels is one.

4.3.2 CNN model training procedure

In CNN model optimization, we choose the mini batch optimization strategy and Adam optimizer (Kingma and Ba 2014). Moreover, the batch size is 40, and the learning rate of the optimizer is 1e-4. In order to optimize the proposed CNN model, the cross entropy function is taken as the loss function of the optimization task. It is defined as follows

$$CE(y,p) = -(ylog(p) + (1-y)\log(1-p)) \qquad (5)$$

where y represents the true value and p represents the predicted value. CE represents the cross-entropy loss when the true value is y and the predicted value is p.

The output of the last layer of our model is a vector with only one element, and this value is adjusted nonlinearly by sigmoid function so that the range of output value is (0 ~ 1). This value is exactly the prediction value of CNN model for image class. Therefore, when the predicted value is greater than 0.5, it is considered that there are cracks in the image, otherwise, no cracks are considered.

As shown in the Fig. 7, the x-axis represents the training steps, and the blue curve and red curve represent the performance of the model on the training data and test data respectively. With the increase of training steps, the loss of training data and test data decreases, while the accuracy increases. The accuracy of classification training data can reach 0.97 after training in 200 steps. And, the accuracy of CNN model in the test data is about 0.95. As a contrast, we use CNN model with the same structure to directly classify the original railway sleeper image, and the final accuracy is about 0.95. But this method needs to use CNN model to judge every railway sleeper image, which requires more computation.

5. Conclusions

An algorithm of sleeper crack detection based on edge detection and morphological filtering is proposed. On the one hand, many sleeper images without cracks are pre filtered. In the actual image of railway sleeper, about 10% of the images contain cracks. Our proposed 3×3 neighborhood range algorithm can filter out 95.8% of the crack images. Therefore, in the actual detection of railway sleeper cracks, about 86% of the sleeper images can be directly determined as healthy sleepers by this method, without complex CNN model detection.

On the other hand, the cracks in the sleeper images are highlighted, which is helpful for CNN to identify the crack information. For cracked sleepers, the probability of detecting suspicious targets is more than 0.95, so the probability of missing detection of cracks is still very low. In addition, it is found that the leakage cracks are very small cracks. Small cracks mean that the health of the sleeper is better than that of the large crack. When the small crack sleeper continues to be damaged to a certain extent, it can be detected. This shows that the crack detection algorithm proposed in this paper can detect the cracks in railway standardization in time and effectively prevent the occurrence of railway accidents.

Finally, a simple CNN model is built to accurately classify the detected suspicious targets. The classification accuracy of the proposed model to image with suspicious targets is 0.95. The experimental results show that, compared with using CNN model directly, this method can identify many images without cracks in advance with simple image processing technology, thus reducing the computational cost.

In addition, if there are cracks on the surface of other objects, the detection method proposed in this paper can be used. For example, if a color image of a surface crack is obtained, it is only necessary to convert the color image into a gray image before the first stage of edge detection. However, our method can only detect the length and width of the crack, but not the depth. Acoustic emission techniques may be required to detect the depth of the crack (Leaman *et al*. 2020). In our study, the definition of crack is that the naked eye can see cracks in the image, while ignoring those very small cracks.

Acknowledgments

This work is supported by the NSFC (No. U1909217), the ZJNSF (No. LD21E050001), the Zhejiang Zhejiang Special Support Program for High-level Personnel Recruitment of China (No. 2018R52034) and the Wenzhou Major Science and Technology Innovation Project of China (No. ZG2020051).

References

Ada, M., Sevim, B., Yüzer, N. and Ayvaz, Y. (2018), “Assessment of damages on a RC building after a big fire", *Adv. Concr. Constr., Int. J.*, **6**(2), 177-197. https://doi.org/10.12989/acc.2018.6.2.177

Beura, S., Majhi, B. and Dash, R. (2015), "Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer", *Neurocomputing*, **154**, 1-14. https://doi.org/10.1016/j.neucom.2014.12.032

Canny, J. (1987), "A computational approach to edge detection", *IEEE. T Pattern. Anal.*, **8**, 679-698. 10.1109/TPAMI.1986.4767851

Celik, O., Terrell, T., Gul, M. and Catbas, F.N. (2018), "Sensor clustering technique for practical structural monitoring and maintenance", *Struct. Monit. Maint., Int. J.*, **5**(2), 273-295. https://doi.org/10.12989/smm.2018.5.2.273

Chen, S.C. and Chiu, C.C. (2019), "Texture construction edge detection algorithm", *Appl. Sci-Basel*, **9**(5), 897. https://doi.org/10.3390/app9050897

Do, N.T., Mei, Q. and Gul, M. (2019), "Damage assessment of shear-type structures under varying mass effects", *Struct. Monit. Maint., Int. J.*, **6**(3), 237-254. https://doi.org/10.12989/smm.2019.6.3.237

Figueira, D., Sousa, C. and Serra Neves, A. (2018), "Winkler spring behavior in FE analyses of dowel action in statically loaded RC cracks", *Comput. Concrete, Int. J.*, **21**(5), 593-605. https://doi.org/10.12989/cac.2018.21.5.593

Franca, A.S. and Vassallo, R.F. (2020), "A method of classifying railway sleepers and surface defects in real environment", *IEEE Sens. J.*, **21**(10), 11301-11309. https://doi.org/10.1109/JSEN.2020.3026173

Friedl, M.A. and Brodley, C.E. (1997), "Decision tree classification of land cover from remotely sensed data", *Remote Sens. Environ.*, **61**, 399-409. https://doi.org/10.1016/S0034-4257(97)00049-7

Gao, W., Zhang, X., Yang, L. and Liu, H. (2010), "An improved sobel edge detection", *Proceedings of the 3rd International Conference on Computer Science and Information, Technology*, **5**, 67-71. https://doi.org/10.1109/ICCSIT.2010.5563693

Gao, Y., Liu, X.Y. and Xiang, J.W. (2020), "FEM simulation-based generative adversarial networks to detect bearing faults", *IEEE T. Ind. Inform.*, **16**(7), 4961-4971. https://doi.org/10.1109/TII.2020.2968370

Gao, Y., Liu, X.Y., Huang, H.Z. and Xiang, J.W. (2021), "A hybrid of finite element simulation and generative adversarial networks to classify faults in rotor-bearing systems", *ISA Transact.*, **108**, 356-366. https://doi.org/10.1016/j.isatra.2020.08.012

Goel, N., Kaur, H. and Saxena, J. (2020), "Modified decision based unsymmetric adaptive neighborhood trimmed mean filter for removal of very high density salt and pepper noise", *Multimed.Tools Appl.*, **79**(27), 19739-19768. https://doi.org/10.1007/s11042-020-08687-y

Gonzalez, R.C., Woods, R.E. and Eddins, S.L. (2004), *Digital Image Processing Using Matlab*, Gatesmark Publishing, New York, USA.

Haeri, H., Sarfarazi, V., Zhu, Z. and Moradizadeh, M. (2018), "The effect of ball size on the hollow center cracked disc (HCCD) in Brazilian test", *Comput. Concrete, Int. J.*, **22**(4), 373-381. https://doi.org/10.12989/cac.2018.22.4.373

Han, D., Liu, Q. and Fan, W. (2018), "A new image classification method using CNN transfer learning and web data augmentation", *Expert. Syst. Appl.*, **95**, 43-56. https://doi.org/10.1016/j.eswa.2017.11.028

Hsu, C., Lu, C. and Pei, S. (2012), "Image feature extraction in encrypted domain with privacy-preserving SIFT", *IEEE T. Image Process.*, **21**, 4593-4607. https://doi.org/10.1109/TIP.2012.2204272

Isik, E., Aydin, M.C. and Buyuksarac, A. (2020), "24 January 2020 Sivrice (Elazığ) earthquake damages and determination of earthquake parameters in the region", *Earthq. Struct., Int. J.*, **19**(2), 145-156. https://doi.org/10.12989/eas.2020.19.2.145

Janeliukstis, R., Ručevskis, S. and Kaewunruen, S. (2019a), "Mode shape curvature squares method for crack detection in railway prestressed concrete sleepers", *Eng. Fail. Anal.*, **105**, 386-401. https://doi.org/10.1016/j.engfailanal.2019.07.020

Janeliukstisa, R., Clarkb, A., Papaeliasc, M. and Kaewunruen, S. (2019b), "Flexural cracking-induced acoustic emission peak frequency shift in railway prestressed concrete sleepers", *Eng. Struct*, **178**, 493-505. https://doi.org/10.1016/j.engstruct.2018.10.058

Jiang, Z. and Xiang, J. (2020), "Method using XFEM and SVR to predict the fatigue life of plate-like structures", *Struct. Eng. Mech., Int. J.*, **73**(4), 455-462. https://doi.org/10.12989/sem.2020.73.4.455

Kingma, D.P. and Ba, J. (2014), "Adam: A method for stochastic optimization", *arXiv preprint* arXiv:1412.6980.

Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012), "ImageNet classification with deep convolutional neural networks", *Proceedings of the 26th Annual Conference on Neural Information Processing Systems*, Lake Tahoe, NV, USA.

Leaman, F., Herz, A., Brinnel, V., Baltes, R. and Clausen, E. (2020), "Analysis of acoustic emission signals during fatigue testing of a M36 bolt using the Hilbert-Huang spectrum", *Struct. Monit. Maint., Int. J.*, **7**(1), 13-25. https://doi.org/10.12989/smm.2020.7.1.013

Lecun, Y. and Bengio, Y. (1995), "Convolutional networks for images, speech, and time series", (Arbib, M.A. Eds.), In: *The Handbook of Brain Theory and Neural Networks*, MIT Press, Cambridge, USA.

Liu, H., Ding, Y.L., Zhao, H.W., Wang, M.Y. and Geng, F.F. (2020), "Deep learning-based recovery method for missing structural temperature data using LSTM network", *Struct. Monit. Maint., Int. J.*, **7**(2), 109-124. https://doi.org/10.12989/smm.2020.7.2.109

Mizuno, K., Terachi, Y., Takagi, K., Izumi, S., Kawaguchi, H. and Yoshimoto, M. (2012), "Architectural study of HOG feature extraction processor for real-time object detection", *2012 IEEE Workshop on Signal Processing Systems*, Quebec City, QC, Canada, October, pp. 197-202. https://doi.org/10.1109/SiPS.2012.57

Perez, L. and Wang, J. (2017), "The effectiveness of data augmentation in image classification using deep learning", arXiv preprint arXiv:1712.04621.

Prasath, V.S., Thanh, D.N.H. and Hung, N.Q. (2020), "Multiscale gradient maps augmented Fisher information-based image edge detection", *IEEE Access*, **8**, 141104-141110. https://doi.org/10.1109/ACCESS.2020.3013888

Seif, A., Salut, M.M. and Marsono, M.N. (2010), "A hardware architecture of prewitt edge detection", *Proceedings of 2010 IEEE Conference on Sustainable Utilization and Development in Engineering and Technology*, Kuala Lumpur, Malaysia, November, pp. 99-101.

Sengsri, P., Ngamkhanong, C., Melo, A.L.O.D., Papaelias, M. and Kaewunruen, S. (2020), "Damage detection in fiber-reinforced foamed urethane composite railway bearers using acoustic emissions", *Infrastruct.*, **5**(6), 50. https://doi.org/10.3390/infrastructures5060050

Shemirani, A.B., Sarfarazi, V., Haeri, H. and Marji, M.F. (2018), "A discrete element simulation of a punch-through shear test to investigate the confining pressure effects on the shear behaviour of concrete cracks", *Comput. Concrete, Int. J.*, **21**(2), 189-197. https://doi.org/10.12989/cac.2018.21.2.189

Shih, F.Y. (2010), *Image processing and mathematical morphology*, Archives of Dermatology.

Shorten, C. and Khoshgoftaar, T.M. (2019), "A survey on image data augmentation for deep learning", *J. Big Data*, **6**(1), 1-48. https://doi.org/10.1186/s40537-019-0197-0

Siddhartha, S., Yu, C., Qing, H., Reza, M. and Zhiguo, L. (2018), "Data-driven optimization of railway maintenance for track geometry", *Transp. Res. Part C: Emerg. Technol*, **90**, 34-58. https://doi.org/10.1016/j.trc.2018.02.019

Simonyan, K. and Zisserman, A. (2014), "Very deep convolutional networks for large-scale image recognition", arXiv preprint arXiv:1409.1556.

Song, W., Xiang, J. and Zhong, Y. (2017), "Mechanical parameters detection in stepped shafts using the FEM based IET", *Smart Struct. Syst., Int. J.*, **20**(4), 473-481. https://doi.org/10.12989/sss.2017.20.4.473

Suykens, J.A.K. and Vandewalle, J. (1999), "Least squares support vector machine classifiers", *Neural Process. Lett.*, **9**, 293-300. https://doi.org/10.1023/A:1018628609742

Sysyn, M., Nabochenko, O., Kovalchuk, V., Gruen, D. and Pentsak, A. (2019), "Improvement of inspection system for common crossings by track side monitoring and prognostics", *Struct. Monit. Maint., Int. J.*, **6**(3), 219-235. https://doi.org/10.12989/smm.2019.6.3.219

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A. (2014), "Going deeper with convolutions", *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Boston, MA, USA, pp. 1-9.

Torre, V. and Poggio, T.A. (1986), "On edge detection", *IEEE T. Pattern. Anal.*, **8**, 147-163.

Viswanath, K., Mukherjee, J. and Biswas, P.K. (2011), "Image filtering in the block DCT domain using symmetric convolution", *J. Vis. Commun. Image R.*, **22**, 141-152. https://doi.org/10.1016/j.jvcir.2010.11.005

Wang, S. and Xiang, J. (2019), "A minimum entropy deconvolution-enhanced convolutional neural networks for fault diagnosis of axial piston pumps", *Soft Comput.*, **24**, 1-15. https://doi.org/10.1007/s00500-019-04076-2

Yella, S., Dougherty, M. and Gupta, N.K. (2009), "Condition monitoring of wooden railway sleepers", *Transp. Res. Part C: Emerg. Technol.*, **17**, 38-55. https://doi.org/10.1016/j.trc.2008.06.002

Xia, B., Cao, J., Zhang, X. and Peng, Y. (2020), "Automatic concrete sleeper crack detection using a one-stage detector", *Int. J. Intell. Robot. Appl.*, **4**(3), 319-327. https://doi.org/10.1007/s41315-020-00141-4

Xiang, J., Chen, X. and Yang, L. (2009), "Crack identification in short shafts using wavelet-based element and neural networks", *Struct. Eng. Mech., Int. J.*, **33**(5), 543-560. https://doi.org/10.12989/sem.2009.33.5.543

Xiang, J., Jiang, Z., Wang, Y. and Chen, X. (2011), "Study on damage detection software of beam-like structures", *Struct. Eng. Mech., Int. J.*, **39**(1), 77-91. https://doi.org/10.12989/sem.2011.39.1.077

Xiang, J., Matsumoto, T., Long, J., Wang, Y. and Jiang, Z. (2012), "A simple method to detect cracks in beam-like structures", *Smart Struct. Syst., Int. J.*, **9**(4), 335-353. https://doi.org/10.12989/sss.2012.9.4.335

Xiang, J., Nackenhorst, U., Wang, Y., Jiang, Y., Gao, H. and He, Y. (2014), "A new method to detect cracks in plate-like structures with though-thickness cracks", *Smart Struct. Syst.*, **14**(3), 397-418. https://doi.org/10.12989/sss.2014.14.3.397

Xie, X., Xie, G. and Xu, X. (2018), "High precision image segmentation algorithm using SLIC and neighborhood rough set", *Multimed.Tools Appl.*, **77**(24), 31525-31543. https://doi.org/10.1007/s11042-018-6150-y

Yang, Z., Chen, X., Tian, S. and He, Z. (2012), "Multiple damages detection in beam based approximate waveform capacity dimension", *Struct. Eng. Mech., Int. J.*, **41**(5), 663. https://doi.org/10.12989/sem.2012.41.5.663

Yang, Z.B., Yu, J.T., Tian, S.H., Chen, X.F. and Xu, G.J. (2018), "A damage localization method based on the singular value decomposition (SVD) for plates", *Smart Struct. Syst., Int. J.*, **22**(5), 621-630. https://doi.org/10.12989/sss.2018.22.5.621

Yin, H., Gong, Y. and Qiu, G. (2020), "Fast and efficient implementation of image filtering using a side window convolutional neural network", *Signal Process.*, **176**, 107717. https://doi.org/10.1016/j.sigpro.2020.107717

Zaharah, A.B., Aaqib, S., Irina, S. and Andre, G.D. (2019), "Predictive maintenance using tree-based classification techniques: A case of railway switches", *Transp. Res. Part C: Emerg. Technol.*, **101**, 35-54. https://doi.org/10.1016/j.trc.2019.02.001

Zeng, Z., Shuaibu, A.A., Liu, F., Ye, M. and Wang, W. (2020), "Experimental study on the vibration reduction characteristics of the ballasted track with rubber composite sleepers", *Constr. Building Mater.*, **262**, 120766. https://doi.org/10.1016/j.conbuildmat.2020.120766

Zhong, Y. and Xiang, J. (2019), "Impact location on a stiffened composite panel using improved linear array", *Smart Struct. Syst., Int. J.*, **24**(2), 173-182. https://doi.org/10.12989/sss.2019.24.2.173

CC