

Research on road damage recognition and classification based on improved VGG-19

Jiaqi Wang¹, Kaihang Wang², Kexin Li³

^{1,2}Landscape Architecture, Beihua University, Jilin, China

³Beihua University, Jilin, China

¹Corresponding author

E-mail: ¹wangjiaqixs@126.com, ²kaihangW@126.com, ³likenefu@126.com

Received 12 June 2023; accepted 22 August 2023; published online 6 October 2023

DOI <https://doi.org/10.21595/mme.2023.23455>



Copyright © 2023 Jiaqi Wang, et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract. In recent years, methods of road damage detection, recognition and classification have achieved remarkable results, but there are still problems of efficient and accurate damage detection, recognition and classification. In order to solve this problem, this paper proposes a road damage VGG-19 model construction method that can be used for road damage detection. The road damage image is processed by digital image processing technology (DIP), and then combined with the improved VGG-19 network model to study the method of improving the recognition speed and accuracy of VGG-19 road damage model. Based on the performance evaluation index of neural network model, the feasibility of the improved VGG-19 method is verified. The results show that compared with the traditional VGG-19 model, the road damage VGG-19 road damage recognition model proposed in this paper shortens the training time by 79 % and the average test time by 68 %. In the performance evaluation of the neural network model, the comprehensive performance index is improved by 2.4 % compared with the traditional VGG-19 network model. The research is helpful to improve the model performance of VGG-19 road damage identification network model and its fit to road damages.

Keywords: road damage recognition, digital image processing, VGG-19, neural network model evaluation index.

1. Introduction

Whether it is concrete road or asphalt road, after a period of driving, various damages, deformations and other defects will occur one after another [1]. The generation and extension of road damages will reduce the performance and life of road [2]. At the same time, road damage can also cause serious traffic jams and accidents, resulting in adverse social impacts. Efficient and accurate recognition, detection and classification of road damages are the focus of researchers [3].

With the continuous development of China's economy, the development trend of highway construction is becoming more and more complicated [4]. Road engineering plays an important role in national economic construction [5]. The number of modern highway construction has risen sharply and road structure system has been continuously improved. At the same time, the identification and detection of road damages are also facing major challenges. Although road damage recognition methods have been widely used, most of these recognition methods rely on artificial vision for recognition and detection, which has the problems of low detection accuracy, long time consumption and high risk. With the progress of the times and the development of science and technology, the method of detecting road damages based on machine vision recognition is gradually rising. It meets the needs of most road damage identification and detection, intelligently and automatically completes the data collection of various road damages, and greatly improves the speed and accuracy of road damage detection. The efficiency of the method based on machine vision recognition to detect road damages is about 40 times that of artificial vision, and the detection rate and accuracy are far more than artificial vision [6-10].

Deep Learning (DL) is a new research direction in the field of Machine Learning (ML). It is introduced into machine learning to make it closer to the original goal – Artificial Intelligence

(AI) [11]. With the rapid development of deep learning theory, AlexNet has made breakthroughs in the field of object detection [12]. Since then, algorithms based on large-scale visual recognition challenges and Convolutional Neural Networks (CNN) have been studied in depth, and great progress has been made in image classification, object detection and semantic segmentation [13]. Target detection technology based on deep learning has been widely used in biomedicine, transportation, safety, civil engineering and other fields [14]. The classification and detection of road damages has always been an active field of civil engineering research [15]. Fan R., Bocus M. J. et al., as the first researchers to use convolutional neural networks in the field of road damage detection and recognition [16]. Taichi Yamada developed an automatic mobile robot that uses a 2D laser scanner to obtain road damage information, which can measure the depth of road damage [17]. With the advancement of imaging technology and the continuous improvement of photo resolution, image-based road damage recognition has gradually become the mainstream [18]. Allen Zhang et al. proposed an efficient architecture CrackNet based on convolutional neural network (CNN), which can realize pixel-level automatic crack detection in 3D asphalt road [19]. Singh, Janpreet et al. proposed Mask-RCNN (Mask-Region Convolutio) [20].

At present, the recognition and detection of road damages based on deep learning has achieved remarkable research results, but in general, the recognition and detection of road damages is mainly aimed at cracks, ignoring the ductility and variability of road damages [21]. After cracks occur on the road surface, failure to deal with them in time will lead to structural damage to the road, decrease the bearing capacity of the road surface, accelerate local or patchy damage to the road surface, and evolve into other types of damage, such as potholes and looseness [22]. Moreover, in the detection of road damage identification, only cracks are considered, and other types of road damages are not considered. There will be over-fitting, false detection, etc., resulting in low accuracy and slow speed in the process of deep learning [23]. However, among many neural networks, VGG-19 (Visual Geometry Group-19, VGG-19) can deal with image classification problems well due to its own structural characteristics. Moreover, in recent years, VGG-19 has achieved remarkable results in the field of structural damage identification, such as medicine, civil engineering, computer and other fields [24-26]. However, there are few studies in the field of pavement disease identification. The VGG-19 network model uses small convolution kernels and small pooling kernels. It has the advantages of simple structure, more channels, deeper layers, and wider feature maps. It is mostly used in image recognition and classification work [27]. However, due to its large amount of computing resources and many parameters used, the VGG-19 network model is time-consuming and low-precision in the process of recognition, detection and classification. In addition, there is still a lack of samples in the current road damage identification, which leads to insufficient model training and inaccurate test results in the process of training and testing of the VGG-19 network model.

Therefore, this paper studies the classification method of road damage recognition based on improved VGG-19. The road damage image samples are obtained by network extraction and camera equipment shooting. Digital Image Processing (DIP) such as Region of Interest (ROI) and Histogram Equalization (HE) were used to remove the noise interference under the influence of environmental factors and obtain the gray image of road damages. In the case of limited sample labels, a VGG-19 network model for identification and classification of various road damages is established based on MATLAB R2022b. Based on statistical data analysis methods, the performance indicators of each network model are sorted and analyzed. On this basis, the VGG-19 network model which can accurately and efficiently train, identify and classify road damages is obtained through experimental comparison.

2. Road damage image processing based on improved VGG-19

In order to realize the accurate identification of road damage in the case of a small number of samples, the VGG-19 network model is constructed based on MATLAB. It is necessary to process the obtained road damage images to make the illumination on the images uniform, reduce noise

interference, eliminate irrelevant information in the images, restore useful real information, enhance the detectability of relevant information and simplify the data to the greatest extent, obtain samples that can effectively identify road damage parameter information, and construct a VGG-19 network model [28].

2.1. Road damage image

The road images used in this study include cracks, potholes, rutting, loose and normal road images, a total of 1051 road damage images were obtained. The road crack image is from the public dataset of road damage image of Crack Forest; potholes and normal images are from public source datasets in CSDN; road rut and loose images are from the road damage data set in ExtremeMart. In addition, in order to expand the number of training samples, the camera equipment is used to shoot the road damage image, a total of 150 road damage images, and a limited road damage image sample library is constructed. Randomly selected images of non-damage and different damage types from the sample library, as shown in Fig. 1.

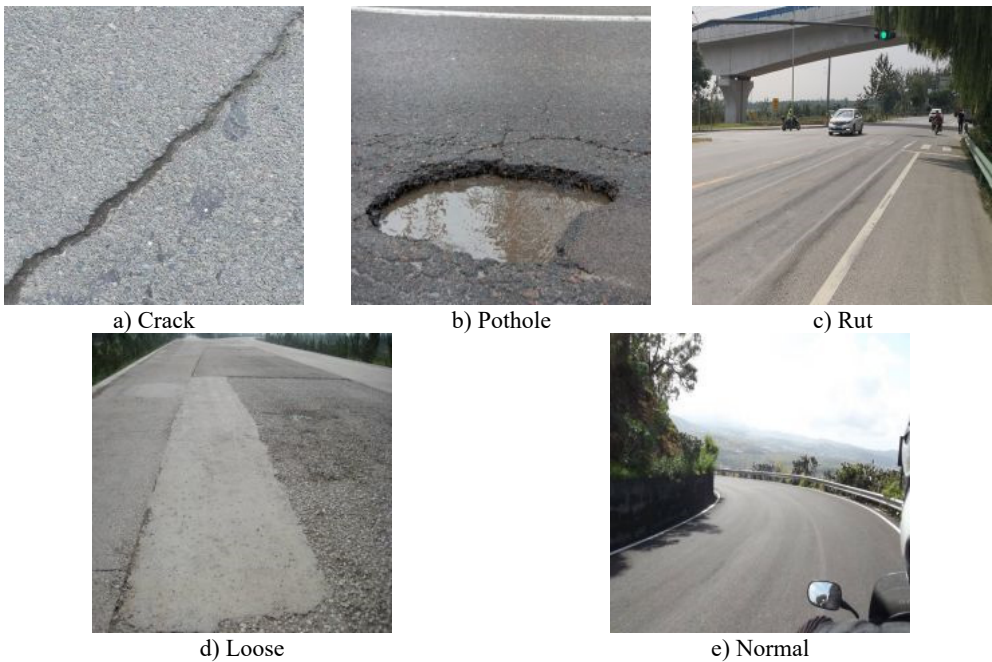


Fig. 1. Road damage image

The road surface damage contained in Fig. 1 are analyzed. The contrast is blurred, and the unique texture features of various damage types are not obvious. They are seriously disturbed by environmental noise such as automobiles and illumination. Therefore, in order to better fit the VGG-19 network model, it is necessary to perform image denoising and image enhancement processing on road damage images.

2.2. Road damage image processing

The quality of the directly obtained pavement disease image is poor, and the network model is difficult to capture the feature information of the disease image [29]. In order to make the pavement disease image better fit the VGG-19 network model, eliminate the redundant noise interference, and enhance the feature information beneficial to the network model training in the pavement disease image, the pavement disease image is processed based on image processing

methods such as gray level changes (GLC) normalization (Normalize), region of interest extraction (ROI), normalization (Normalize), histogram equalization (HE), and median filtering (MF) to obtain pavement disease images with strong recognition, high contrast, and obvious texture features. Improve the efficiency of VGG-19 classification and recognition images. The image processing flow of pavement disease in this study is shown in Fig. 2.

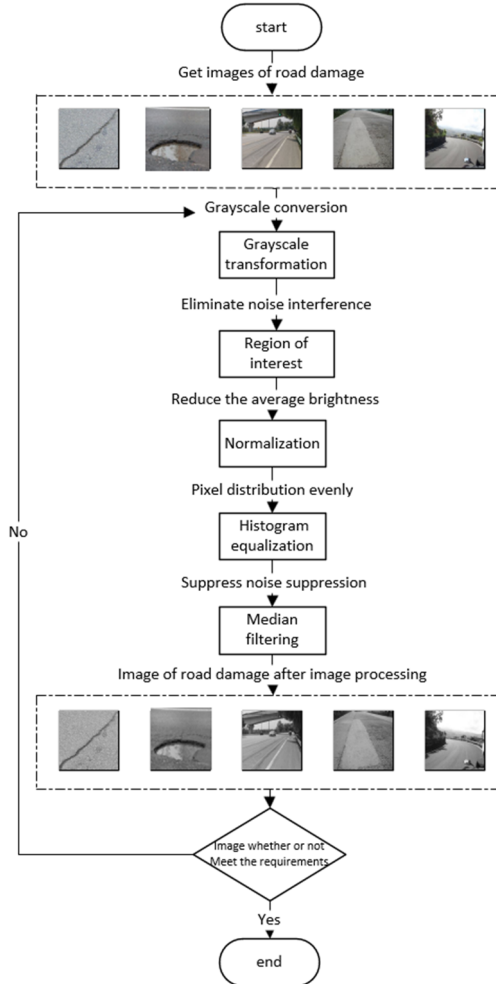


Fig. 2. Road damage image processing

2.2.1. Gray level change

The gray level transformation of the image refers to the method of changing the gray value of each pixel in the source image point by point according to a certain transformation relationship according to a certain target condition. The purpose is to improve the image quality and make the image display effect clearer. The gray-scale transformation method used in the research is a gray-scale linear transformation, which belongs to a kind of gray-scale transformation. By establishing a gray-scale mapping to adjust the gray level of the source image, the purpose of image enhancement is achieved. The principle is to enhance or weaken the gray level of the image by transforming the pixel value of the image through a specified linear function. The formula of linear transformation of gray scale is a common one-dimensional linear function, as shown:

$$g(x, y) = k \cdot f(x, y) + b. \quad (1)$$

Let x be the original gray value, then the calculation formula of the gray value of the transformed y is as shown:

$$y = k \cdot x + b, \quad (0 \leq y \leq 255), \quad (2)$$

where k represents the slope of the line, that is, the degree of tilt, b represents the intercept of the linear function in the y -axis. In the formula, the values of different k and b represent different meanings: when $k > 1$, it means the contrast of the enhanced image, the pixel value of the image increases after linear transformation, and the overall effect is enhanced. When $k = 1$, it means that the image brightness can be adjusted by adjusting the size of b . When $0 < k < 1$, the overall contrast of the image is weakened; when $k < 0$, it means that the bright area in the original image darkens and the dark area in the original image brightens.

2.2.2. Region of interest extraction

After converting road damage image into a grayscale image, the image part containing the texture features of the road damage is set as the “region of interest” (ROI) of the image. ROI interception is a very critical step in image processing. The interception of ROI region can reduce the amount of data in subsequent image processing while eliminating some noise. The study selected a window of 2440×2600 to intercept the ROI region of five road damage images.

2.2.3. Road damage image normalization

In the process of image processing, image normalization is a necessary preprocessing process. There are three purposes to study image normalization. First, if image normalization is not performed, the size of the image cannot perfectly fit the input layer size of the VGG-19 network model, resulting in the network model cannot be trained. Second, if the input layer size is set to be large, the gradient transmitted to the input layer will become large during back propagation. If the gradient is large, a smaller learning rate must be set, otherwise it will cross the optimal. In this case, the selection of the learning rate needs to refer to the size of the input layer value, and the data normalization operation can easily select the learning rate, so as to facilitate the network model training; third, image normalization can remove the average brightness value of the image. Subtracting the statistical average of the data on each image can remove the common part and highlight individual differences.

Based on a common image normalization principle, the road damage image is normalized (Normalize) batch processing. The principle is shown in Eq. (3):

$$y = \frac{x - MinValue}{MaxValue - MinValue}. \quad (3)$$

Among them, x and y are the values before and after normalization, respectively. MaxValue and MinValue are the maximum and minimum gray values of the original image, respectively. On this basis, the image size is modified to 244×244 to perfectly match the input layer of the network model.

2.2.4. Histogram equalization

Histogram plays an important role in image processing. The gray histogram of the image describes the gray distribution in the image, which can intuitively show the amount of each gray level in the image. Histogram equalization (HE) is a method to enhance the contrast of images. The main idea is to transform the histogram distribution of an image into an approximate uniform

distribution through the cumulative distribution function, thereby enhancing the contrast of the image. In order to extend the brightness range of the original image, a mapping function is needed to evenly map the pixel values of the original image to the new histogram, but the mapping function has two necessary conditions : First, the order of the original pixel values cannot be disrupted, and the size relationship between bright and dark after mapping cannot be changed; second, the mapping must be in the original range, that is, the range of the pixel mapping function should be between 0 and 255.

Since the image is composed of pixels, the image histogram equalization is solved by the discrete cumulative distribution function. In the process of histogram equalization, the mapping method is shown in Eq. (4):

$$S_k = \sum_{j=0}^k \frac{n_j}{n}, \quad (k = 0, 1, 2, \dots, L - 1). \quad (4)$$

Among them, S_k refers to the value of the current gray level after the cumulative distribution function mapping, n is the sum of pixels in the image, n_j is the number of pixels in the current gray level, and L is the total number of gray levels in the image.

2.2.5. Median filtering

The median filtering (MF) is a nonlinear smoothing image processing technology, which is a nonlinear image processing technology based on the order statistics theory and can effectively suppress noise. The basic principle of median filtering is to replace the value of a point in a digital image or digital sequence with the median value of each point in a neighborhood of the point, so that the surrounding pixel values are close to the true value, thereby eliminating isolated noise points.

Based on the two-dimensional median filtering function (medfilt2) of MATLAB, the median filtering (MF) processing of road damage images is carried out by filling the image edges, adding salt and pepper noise, and data conversion. Set the one-dimensional sequence f_n , take the window length m (m is odd), and perform median filtering on it. The number of m is successively taken out from the input sequence, and the number of m is sorted by size, and the number whose serial number is the center point is taken as the filter output. The principle is as shown in Eq. (5):

$$y(i) = \text{Med}\{f(i - v), \dots, f(i), \dots, f(i + v)\}, \quad i \in N, \quad v = \frac{m - 1}{2}. \quad (5)$$

Taking the road crack image as an example, the road crack image after image processing is shown in Fig. 3.

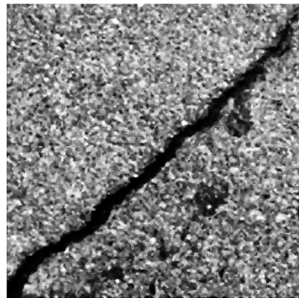


Fig. 3. Final image of road cracks after image processing

By analyzing the road crack image information in Fig. 3, it can be seen that the crack image

processed by image processing methods such as gray level changes (GLC), region of interest extraction (ROI), normalization (Normalize), histogram equalization (HE) and median filtering (MF) has uniform illumination, the contrast of interference noise is significantly reduced, and the contrast of cracks is significantly increased. This image can perfectly fit the training and recognition of the neural network model. This road damage image can form a road damage image data set for use by the neural network model.

3. Establishment of road damage recognition model based on VGG-19

3.1. Improved methods of VGG-19 model

In convolutional neural networks, VGG-19 is often used in the field of object recognition. Each layer of the VGG-19 neural network will use the output of the previous layer to further extract more complex features until it is complex enough to be used to identify objects, so each layer can be seen as many local feature extractors. However, there are still some shortcomings and defects in the identification of road damages. The ordinary VGG-19 network model often appears in the training and testing of neural networks, such as slow speed, low training accuracy and over-fitting. In order to solve the above problems, combined with the characteristics of road damage images, an improved model based on VGG-19 network model structure is proposed: Road damage VGG-19.

The construction of Road damage VGG-19 network model adopts the method of adjusting batch size and adding batch normalization layer. Because the most important factor affecting the training speed and test accuracy of the VGG-19 network model is the size of the batch size, the batch size can be defined as the number of samples taken in one training, which belongs to a hyperparameter, usually set when the model is trained, generally $2n$. The size of the Batch size affects the optimization degree and speed of the model. At the same time, it directly affects the use of GPU memory. Too large or too small Batch size has a great impact on the model. Therefore, it is necessary to set an appropriate batch size value to find the best balance between training speed and memory capacity. In the training process of the neural network, as the depth deepens, the input value distribution will shift to the upper and lower ends of the value interval, which leads to the disappearance of the gradient of the low-level neural network during back propagation, which is an important reason for the slower convergence of the deep network. The batch Normalization can force the distribution of the input values of each layer of the neural network back to the standard normal distribution with a mean value of 0 and a variance of 1 through a certain standardization method. Its role can accelerate the convergence speed of the model training, so that the model can detect and identify the road damage image faster and more accurately with a small number of sample labels. The Road damage VGG-19 network model is obtained by adding the batch normalization layer, and its network structure is shown in Fig. 4.

By analyzing the network structure shown in Fig. 4, it can be seen that the road damage image enters the VGG-19 network input layer as input data. After five convolution layers, five pooling layers, data normalization, and full connection layers, the network model recognizes the image as a crack image. Then, the final output layer of the VGG-19 network model outputs the image as a road crack image.

Based on the original VGG-19 convolutional neural network model and the same equipment, the batch size is adjusted and the batch normalization layer is added to adjust the model structure to improve the training speed and verification accuracy of the model. Adjusting the size of the batch size can significantly improve the speed and verification accuracy of the model in identifying road damage images. In addition, with the continuous increase of the training depth of the model, the road damage image is used as the input of the model, and the input value distribution will be offset and close to the upper and lower ends of the value range. The addition of the batch normalization layer can force the distribution of the input value of each layer of the road damage image back to the standard normal distribution with a mean of 0 and a variance of 1, thereby

accelerating the convergence speed of the model training and making the model training process more stable. Five improvement schemes were designed: (1) traditional VGG-19 model (2) adding a batch normalization layer at the end of the traditional VGG-19 model structure (3) adding a batch normalization layer before each activation function in the traditional VGG-19 model (4) adding only two batch normalization layers on the basis of the traditional VGG-19 model. Add and modify the batch size to 64 before the last two fully connected layers of the model structure. (5) Based on the traditional VGG-19 model, the model structure is modified to 19 layers, which is the same as Model (4). Two batch normalization layers are added at the same position, and the batch size is 64. The training information of the improved VGG-19 network model is shown in the Table 1.

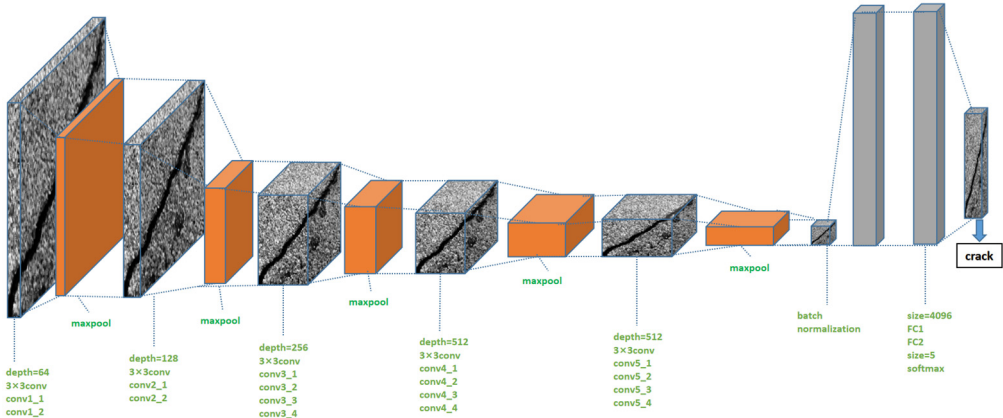


Fig. 4. Road damage VGG-19 network structure diagram

Table 1. Improved VGG-19 network model training information

Model serial number	Training time	Accuracy of validation
Model (1)	210 min 28 s	94.79 %
Model (2)	185 min 25 s	94.79 %
Model (3)	202 min 05 s	92.89 %
Model (4)	186 min 20 s	94.31 %
Model (5)	44 min 41 s	97.20 %

By analyzing the data in Table 1, it can be seen that Model (5) has much higher efficiency than other models in pavement disease identification, and the verification accuracy of Model (5) for road damage images is higher than that of other models. It can be seen that the ability of Model (5) to recognize road damage images is higher than other models. Therefore, Model (5) is used as the VGG-19 network model of this study, and it is named Road damage VGG-19 road damage identification model. In addition, according to the type of road damage recognition during model training, the final output size of the model is defined as 5.

3.2. Improved VGG-19 network model training

The configuration of the running device is one of the main factors affecting the training time of the model, that is, the CPU and GPU of the device. Experimental environment: CPU: AMD Ryzen 7 6800H with Radeon Graphics, 3.20 GHz; gPU: NVIDIA GeForce RTX3060 Lap Top GPU, combined with MATLAB deep learning framework, completed training in MATLAB R2022b environment.

In order to improve the training speed and verification accuracy of the model, the matching ability and recognition ability of the model training for the road damage image are strengthened. The training parameters of the model (1-5) are uniformly set to 64 mini batch size, 0.0001 Initial

Learn Rate and 30 Validation Frequency. The setting of the mini batch size can divide the road damage image sample set into equal subsets, and train each subset to improve the training speed of the road damage recognition model. Setting the initial learning rate reasonably can make the road damage recognition model converge to the local minimum in a suitable time when training and recognizing the road damage image, so as to improve the training speed of the road damage recognition model. In addition, reasonably setting the number of iterations (Validation Frequency) can prevent the over-matching problem of the road damage recognition model when training and recognizing the road damage image.

In addition to the above unified training parameters, the fundamental difference between model (1-5) training is whether the batch normalization layer and the position of the batch normalization layer are added, which determines the speed of training time and the accuracy of training. The specific settings can be seen in Table 1 in Section 2.1.

After the VGG-19 network model training is completed, the trained model parameters are saved for subsequent model testing.

3.3. Improved VGG-19 network model testing

In this study, in the MATLAB R2022b test environment, after training the model (1-5), the trained model parameters are loaded, and the accuracy of classification and recognition is obtained by model test classification. In the test set, the cracks, potholes, ruts, loose and normal roads are tested and identified in turn, and the average time and accuracy of each road damage identification test are counted to form the model (1-5) test data comparison table, see Table 2.

Table 2. Improved VGG-19 network model test information

Model serial number	Average test time	Accuracy
Model (1)	0.59 s	97 %
Model (2)	0.45 s	98 %
Model (3)	0.69 s	98 %
Model (4)	0.47 s	99 %
Model (5)	0.19 s	98 %

By analyzing the data in Table 2, it can be seen that average test time of model (1-4) is more than 0.45 s. However, the average test time of model (5), namely Road damage VGG-19 road damage identification model, is only 0.19 s, and the average test time of Road damage VGG-19 road damage identification model is the shortest. Secondly, the test accuracy of Road damage VGG-19 road damage recognition model is 98 %, which is the second in all test models, which is in line with the significance of model modification, and has played a role in improving the accuracy of the model and reducing the test time.

It can be concluded that the efficiency, verify accuracy and average test time of Model (5), that is, Road damage VGG-19 pavement disease recognition model, are better than other models.

4. Feasibility evaluation of road damage VGG-19 network model

4.1. Construction of network model performance evaluation index

In order to verify the effectiveness of the network model, the feasibility of model (1-5) is evaluated after the training and testing stages are completed. Performance evaluation indicators such as Accuracy, Sensitivity, Specificity, Precision, and F1 Score were used to verify the performance of the model. These values depend on the true positive (TP), false negative (FN), true negative (TN) and false positive (FP) values, which can be calculated from the confusion matrix of the five models. The calculation formula and corresponding description of the performance index are as follows:

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \quad (6)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (7)$$

$$Specificity = \frac{TN}{TN + FP} \quad (8)$$

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

$$F1Score = \frac{2TP}{2TP + FN + FP} \quad (10)$$

Accuracy represents the proportion of all samples with correct prediction in all test samples, reflecting the recognition accuracy of road damage recognition model for road damage images. Sensitivity, also known as recall rate, represents the proportion of correct prediction results in all positive events, reflecting the recognition ability of road damage recognition model for road damage images. Specificity expresses the recognition ability of negative samples, which is similar to sensitivity, and also reflects the recognition ability of road damage recognition model for road damage images. Precision represents the proportion of true positive samples in all predicted positive samples, reflecting the recognition accuracy of road damage recognition model for road damage images. F1 score. In the competition of some classification problems, F1 score is used as the final evaluation method, which represents the harmonic average of precision and recall. The maximum is 1 and the minimum is 0. For the road damage recognition model, F1 score is often used to reflect the performance of the model.

The essence of the classification confusion matrix is a table. Through the collection of samples, we can directly know the positive and negative situations of the predicted road damage images in the real situation [30]. At the same time, the results of the classification model can be obtained from the sample data, and the Positive and Negative in the data can also be grasped. Therefore, four basic indicators are obtained, which are collectively referred to as the first-level indicators. The four indicators are presented in the table to obtain the following matrix, namely the classification confusion matrix, which defines TP, FN, TN and FP based on True, False, Positive and Negative. For the road damage recognition model, the classification confusion matrix reflects the number of observations of the road damage model for the road damage image misclassification and classification. The classification confusion matrix of (1-5) is shown in Fig. 5.

By analyzing the confusion matrix in Fig. 5, it can be seen that the output type of the network model is five types. Therefore, the classification confusion matrix of models (1-5) belongs to the multi-classification confusion matrix. In each classification confusion matrix, the real class and the prediction class are included. The real class represents the real value, and the prediction class represents the predicted value. The serial numbers 1-5 represent the cracks, potholes, rutting, loose and normal road images respectively. Taking the classification confusion matrix of Model (5) as an example, the blue number in the 5×5 matrix represents the correct number of road damages identified for each type. Taking the first column of the matrix as an example, it represents that 292 images are identified as road crack damages, 6 images are identified as road pothole damages, 2 images are identified as road rutting damages, 3 images are identified as road loose damages and 5 images are identified as normal road. The blue percentage below the first column represents the proportion of the damage with the real label as the road crack being correctly predicted, and the yellow percentage represents the proportion of the image with the real label as the road crack being incorrectly predicted as other types of road damages. According to the data obtained in the matrix, the values of TP, FN, TN and FP can be calculated, and then the evaluation indexes can be calculated.

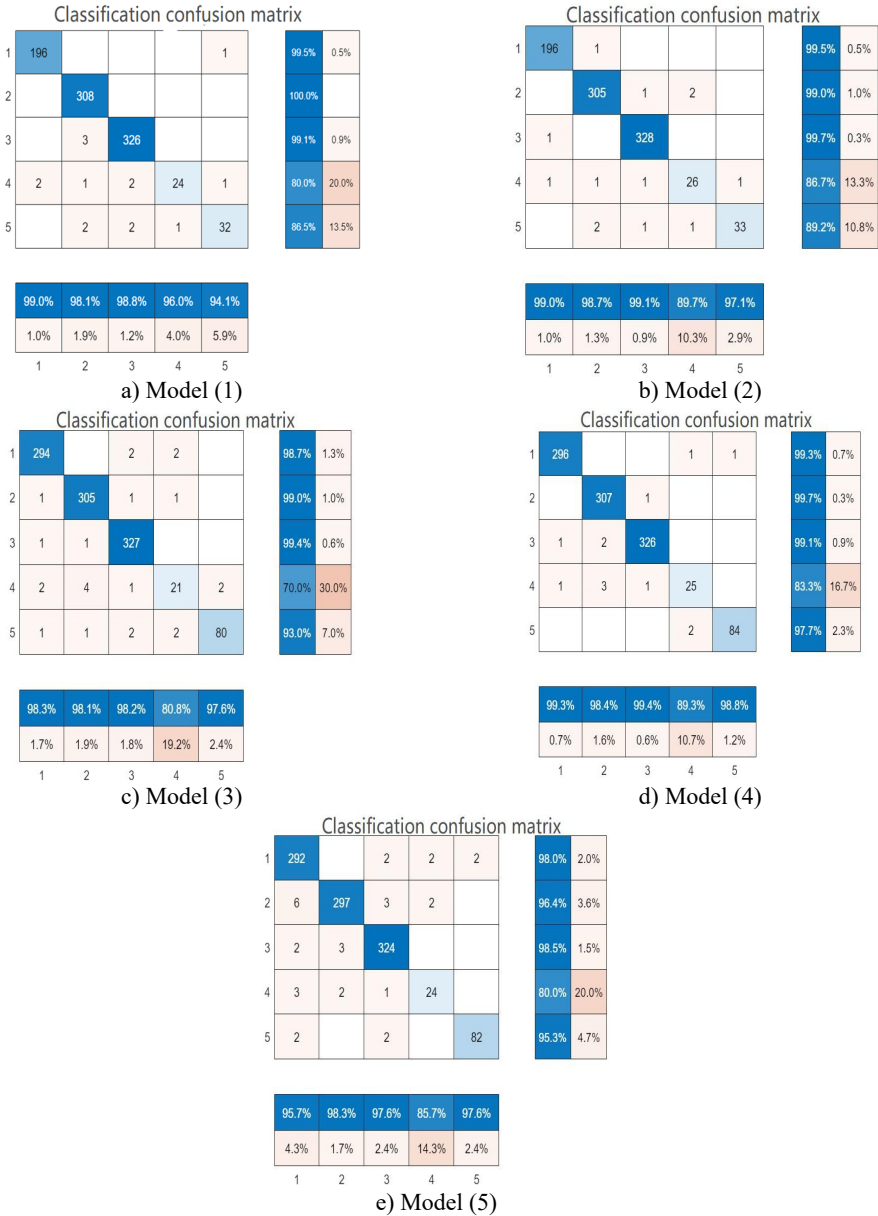


Fig. 5. Classification confusion matrix

4.2. Performance comparison of road damage recognition network models

According to the classification confusion matrix in Fig. 6, the values of TP, FN, TN and FP are calculated. TP, FN, TN and FP are all related to the recognition ability of the model to road damage images. The evaluation indexes were calculated by combining the Eqs. (6-10) in Section 3.1 with the values of TP, FN, TN and FP. The value of accuracy represents the recognition accuracy of the model for all types. However, other performance evaluation indicators represent the performance values of the model for a certain type of recognition. Therefore, this study corresponds other performance evaluation indicators to road crack identification [30].

In the performance evaluation of the model, the values of TP, FN, TN and FP are the first-

level indicators, and the values of Accuracy, Sensitivity, Specificity and Precision calculated by them are the second-level indicators. On the basis of these four second-level indicators, a third-level indicator, F1 score, will be generated. The comparison of each evaluation index is shown in Fig. 6.

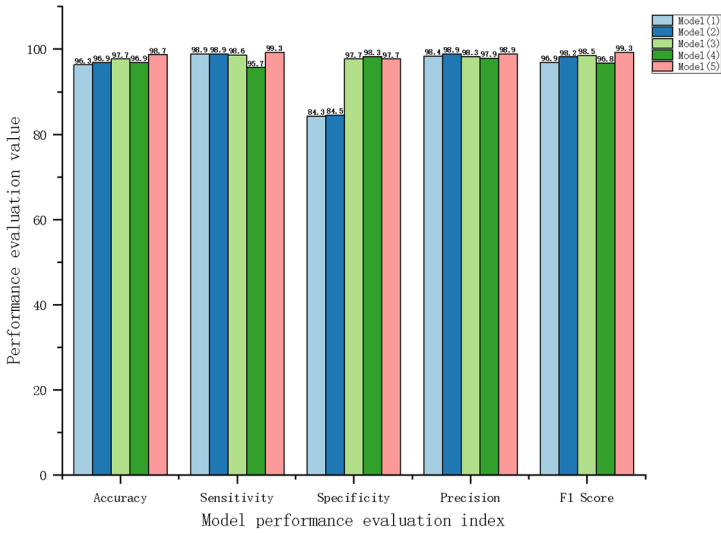


Fig. 6. Comprehensive comparison chart of neural network model performance

By analyzing the image information in Fig. 6, it can be seen that the accuracy, sensitivity, precision and F1 score of Model (5), namely Road damage VGG-19 road damage recognition model, are higher than those of the other four models. However, the specificity is lower than the model (4). Through comprehensive analysis, the research improves the performance of the model by modifying the parameters and structure of the VGG-19 network model.

5. Conclusions

In the research, the problems existing in the current road damage detection and recognition are discussed, which may be caused by the model itself and the image itself. Therefore, the research first uses digital image processing techniques such as grayscale conversion, ROI interception, histogram equalization, and median filtering to remove environmental noise interference in road damage images, and obtain road damage images with high contrast and clear targets. Then, the research starts with optimizing the parameters and structural mechanism of the network model, and proposes a total of five road damage identification VGG-19 network model improvement schemes. The five models are trained and tested in turn. Finally, the Road damage VGG-19 road damage identification model with the shortest average test time (0.19 s) and the highest test accuracy (98 %) is selected as the final model. Finally, based on the first-level evaluation index of neural network model, the second-level and third-level indexes of performance evaluation are calculated, and the comprehensive performance evaluation of five road damage identification models is carried out.

The verification analysis shows that the accuracy, sensitivity, precision and F1 score of the road damage VGG-19 road damage recognition model proposed in this paper are higher than other improved models in the comprehensive performance evaluation link, and the comprehensive performance is improved by 2.4 % compared with the traditional VGG-19 network model. It is proved that the model performance of road damage recognition model can be improved by modifying the parameters and structure of VGG-19 neural network.

Acknowledgements

2023 Beihua University Graduate Program “Research on road damage recognition and classification based on improved VGG-19”, [2023] 058. 2022 Beihua College Students Innovation and Entrepreneurship Training Program Project “Development and Application of Pavement Damage Identification System Based on CNN”, Project No. 202210201115.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Author contributions

Jiaqi Wang performed the conceptualization, data Curation, formal Analysis, project Administration, software and writing; Kaihang Wang performed the validation and visualization; Kexin Li performed supervision.

Conflict of interest

The authors declare that they have no conflict of interest.

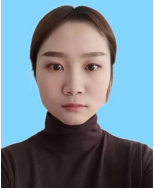
References

- [1] Zhang Jing, Tong Zheng, and Huang Baoshan., “A deeper generative adversarial network for grooved cement concrete pavement crack detection,” *Engineering Applications of Artificial Intelligence*, Vol. 2, No. 2, p. 112, May 2023, <https://doi.org/10.1016/app.22105808>
- [2] L. Wang, J. Li, and F. Kang, “Crack location and degree detection method based on YOLOX model,” *Applied Sciences*, Vol. 12, No. 24, p. 12572, Dec. 2022, <https://doi.org/10.3390/app122412572>
- [3] J. Yang, K. Ruan, J. Gao, S. Yang, and L. Zhang, “Pavement distress detection using three-dimension ground penetrating radar and deep learning,” *Applied Sciences*, Vol. 12, No. 11, p. 5738, Jun. 2022, <https://doi.org/10.3390/app12115738>
- [4] A. Ragnoli, M. de Blasiis, and A. Di Benedetto, “Pavement distress detection methods: a review,” *Infrastructures*, Vol. 3, No. 4, p. 58, Dec. 2018, <https://doi.org/10.3390/infrastructures3040058>
- [5] F.-J. Du and S.-J. Jiao, “Improvement of lightweight convolutional neural network model based on YOLO algorithm and its research in pavement defect detection,” *Sensors*, Vol. 22, No. 9, p. 3537, May 2022, <https://doi.org/10.3390/s22093537>
- [6] P. Manjula and S. B. Priya, “An effective network intrusion detection and classification system for securing WSN using VGG-19 and hybrid deep neural network techniques,” *Journal of Intelligent and Fuzzy Systems*, Vol. 43, No. 5, pp. 6419–6432, Sep. 2022, <https://doi.org/10.3233/jifs-220444>
- [7] M. Ren, X. Zhang, X. Chen, B. Zhou, and Z. Feng, “YOLOv5s-M: A deep learning network model for road pavement damage detection from urban street-view imagery,” *International Journal of Applied Earth Observation and Geoinformation*, Vol. 120, p. 103335, Jun. 2023, <https://doi.org/10.1016/j.jag.2023.103335>
- [8] J. Hu, M.-C. Huang, and X. B. Yu, “Deep learning based on connected vehicles for icing pavement detection,” *AI in Civil Engineering*, Vol. 2, No. 1, pp. 1–14, Apr. 2023, <https://doi.org/10.1007/s43503-023-00010-6>
- [9] L. Ali, H. A. Jassmi, W. Khan, and F. Alnajjar, “Crack45K: integration of vision transformer with tubularity flow field (TuFF) and sliding-window approach for crack-segmentation in pavement structures,” *Buildings*, Vol. 13, No. 1, p. 55, Dec. 2022, <https://doi.org/10.3390/buildings13010055>
- [10] W.-W. Jin et al., “Road pavement damage detection based on local minimum of grayscale and feature fusion,” *Applied Sciences*, Vol. 12, No. 24, p. 13006, Dec. 2022, <https://doi.org/10.3390/app122413006>
- [11] J. Ren, G. Zhao, Y. Ma, Zhao, T. Liu, and J. Yan, “Automatic pavement crack detection fusing attention mechanism,” *Electronics*, Vol. 11, No. 21, p. 3622, Nov. 2022, <https://doi.org/10.3390/electronics11213622>

- [12] Liu Fangyu, Liu Jian, and Wang Linbing, "Asphalt pavement fatigue crack severity classification by infrared thermography and deep learning," *Automation in Construction*, Vol. 21, No. 12, p. 143, May 2022, <https://doi.org/10.1016/j.autcon.22104575>
- [13] L. Tian, J. Xing, H. Zhao, and J. Chang, "The research on intelligent extraction of furnace mouth flame characteristics based on DNN," *Mathematical Models in Engineering*, Vol. 4, No. 1, pp. 42–48, Mar. 2018, <https://doi.org/10.21595/mme.2018.19765>
- [14] H. Xu, J. Liu, and Z. Lu, "Structural damage identification based on cuckoo search algorithm," *Advances in Structural Engineering*, Vol. 19, No. 5, pp. 849–859, May 2016, <https://doi.org/10.1177/1369433216630128>
- [15] Z. He, H. Zhao, J. Wang, and W. Feng, "Pose matters: Pose guided graph attention network for person re-identification," *Chinese Journal of Aeronautics*, Vol. 36, No. 5, pp. 447–464, May 2023, <https://doi.org/10.1016/j.cja.2022.11.017>
- [16] E. A. Martínez-Ríos, R. Bustamante-Bello, and S. A. Navarro-Tuch, "Generalized Morse wavelets parameter selection and transfer learning for pavement transverse cracking detection," *Engineering Applications of Artificial Intelligence*, Vol. 123, No. 3, p. 106355, Aug. 2023, <https://doi.org/10.1016/j.engappai.2023.106355>
- [17] H. Dong, K. Song, Q. Wang, Y. Yan, and P. Jiang, "Deep metric learning-based for multi-target few-shot pavement distress classification," *IEEE Transactions on Industrial Informatics*, Vol. 18, No. 3, pp. 1801–1810, Mar. 2022, <https://doi.org/10.1109/tii.2021.3090036>
- [18] R. R. and D. Park, "A multiclass deep convolutional neural network classifier for detection of common rice plant anomalies," *International Journal of Advanced Computer Science and Applications*, Vol. 9, No. 1, 2018, <https://doi.org/10.14569/ijacsa.2018.090109>
- [19] V. Gampala, M. Sunil Kumar, C. Sushama, and E. Fantin Irudaya Raj, "WITHDRAWN: deep learning based image processing approaches for image deblurring," *Materials Today: Proceedings*, Dec. 2020, <https://doi.org/10.1016/j.matpr.2020.11.076>
- [20] G. de León, J. Cesbron, P. Klein, P. Leandri, and M. Losa, "Novel methodology to recover road surface height maps from illuminated scene through convolutional neural networks," *Sensors*, Vol. 22, No. 17, p. 6603, Sep. 2022, <https://doi.org/10.3390/s22176603>
- [21] N. Hnoohom, S. Mekruksavanich, and A. Jitpattanukul, "A comprehensive evaluation of state-of-the-art deep learning models for road surface type classification," *Intelligent Automation and Soft Computing*, Vol. 37, No. 2, pp. 1275–1291, 2023, <https://doi.org/10.32604/iasc.2023.038584>
- [22] D. Qiao et al., "Pavement crack detection based on point cloud data and data fusion," *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, Vol. 27, No. 3, p. 381, Jul. 2023, <https://doi.org/10.1098/rsta.20220165>
- [23] D. Yuchuan et al., "Modeling automatic pavement crack object detection and pixel-level segmentation," *Automation in Construction*, Vol. 7, No. 1, p. 150, 2023, <https://doi.org/10.1016/j.autcon.23104840>
- [24] M. Mateen, J. Wen, Nasrullah, S. Song, and Z. Huang, "Fundus image classification using VGG-19 architecture with PCA and SVD," *Symmetry*, Vol. 11, No. 1, Dec. 2018, <https://doi.org/10.3390/sym11010001>
- [25] J. Xiao, J. Wang, S. Cao, and B. Li, "Application of a novel and improved VGG-19 network in the detection of workers wearing masks," in *Journal of Physics: Conference Series*, Vol. 1518, No. 1, p. 012041, Apr. 2020, <https://doi.org/10.1088/1742-6596/1518/1/012041>
- [26] H. Yuan, T. Jin, and X. Ye, "Modification and evaluation of attention-based deep neural network for structural crack detection," *Sensors*, Vol. 23, No. 14, p. 6295, Jul. 2023, <https://doi.org/10.3390/s23146295>
- [27] Z. Q. Yue, W. Bekking, and I. Morin, "Application of digital image processing to quantitative study of asphalt concrete microstructure," *Transportation Research Record*, No. 1492, 1995.
- [28] A. Hijazi and V. Madhavan, "A novel ultra-high speed camera for digital image processing applications," *Measurement Science and Technology*, Vol. 19, No. 8, p. 085503, Aug. 2008, <https://doi.org/10.1088/0957-0233/19/8/085503>
- [29] P. Machart and L. Ralaivola, "Confusion matrix stability bounds for multiclass classification," *arXiv:1202.6221*, 2023.



Jiaqi Wang received B.E. degree in Project engineering Institute from Beihua University, Jilin, China, in 2022. Now he studies at Beihua University as a postgraduate. His current research interests include deep learning, intelligent detection of pavement disease.



Kaihang Wang received B.E. degree in Project engineering Institute from Beihua University, Jilin, China, in 2021. Now she studies at Beihua University as a postgraduate. Her current research interests include directional structural damage, deep learning, and road crack identification



Kexin Li received Ph.D. degree in Road and railway engineering Institute from Northeast Forestry University, Heilongjiang, China, in 2020. Now she works at Beihua University. Her current research interests include structural health monitoring, non-destructing testing technology, new method of pattern recognition.