Research on intelligent detection of pavement damage based on CNN

Yuanhang Tang[1]**, Kexin Li**[2]**, Kaihang Wang**[3]
College of Civil Engineering Transportation, BeiHua University, Jilin, China
[2]Corresponding author
E-mail: [1]*tyh18844655238@163.com*, [2]*likenefu@126.com*, [3]*kaihangW@126.com*

Received 6 September 2022; received in revised form 2 October 2022; accepted 27 October 2022
DOI https://doi.org/10.21595/mme.2022.22918

Check for updates

Abstract. In order to timely detect road damage under complex constraints, the road damage mechanism was used to reclassify the types of road damage, the VGG-19 model was applied to identify and detect less road damage images intelligently through convolutional neural network and transfer learning and at the same time, the road damage situation was detected using the Softmax classifier and the feasibility and accuracy of the method were verified on the basis of the detection set. The results show that the detection and recognition accuracy of the model proposed in this paper reaches 86.2 %, an increase of 6.2-49.8 percentage points than the detection results of other convolutional neural network models. Therefore, it can be concluded that the transfer learning- and convolutional neural network-based road damage intelligent detection methods proposed in this paper are feasible, and this research is helpful to realize high-precision real-time intelligent detection of road damage.

Keywords: convolutional neural network (CNN) transfer learning pavement damage image processing.

1. Introduction

In recent years, the construction of roads has continued to develop and progress, and the total number of roads has continued to increase. The upward trend is quite evident. The maintenance project that goes along with such a huge road project is also unprecedentedly huge. Especially under the action of cold weather, the number of road damage will increase to a considerable level. Once the damage to the road surface increases, the resulting road safety hazards also increase substantially. Therefore, timely repair and maintenance of the road surface have become the prime focus of road workers. The current road detection methods can be roughly divided into three types: manual detection, instrument detection, and artificial intelligence detection [1-3]. Due to the growing number of roads along with the shortage of professional technical personnel, manual inspection requires a significant amount of professional and technical personnel. The second is the instrument detection method. The instrument detection method cannot guarantee the degree of complete popularisation and it is currently confronted with expensive instruments. Contrarily, artificial intelligence methods are convenient and affordable, and they have received widespread attention from scholars [4-8].

As an intelligent learning method, the convolutional neural network employs the convolution layer, the pooling layer, and the fully connected layer In order to identify and extract image features multiple times, then learn and train them [8-12]. A number of scholars are still exploring the direction of pavement damage images, and using a variety of convolutional neural network models to train pavement damage images [13-16].

For instance, the first team to employ convolutional neural networks in the field of pavement damage detection was Fan R., Bocus M. J. et al. [17]. Rui Fan, Mohammud [18] et al. proposed a novel road crack detection algorithm based on deep learning and adaptive image segmentation. Singh, Janpreet [19] et al. suggested a MaskRCNN-based pavement damage classification and detection method. Mask-RCNN is one of the state-of-the-art algorithms for object detection and semantic segmentation in natural images. Allen Zhang, Kelvin C. P. Wang [20], et al. put forward a CrackNet, an efficient convolutional neural network (CNN)-based architecture for pixel-level automatic crack detection in 3D asphalt pavements. On the other hand, K. C. Wang [21] et al. proposed a pavement automation system on the basis of convolutional neural network research.

Due to the fact that images of pavement damage have various forms, several scholars have not been able to classify and train pavement damage images well. We propose a method to reasonably classify images utilising the pavement damage mechanism, and use a transfer learning method for training with multiple convolutional neural network models.

In the second chapter, the overall experimental framework is introduced, in the third chapter the key models used have been described, and in the fourth and fifth chapters we explained the experiments and discuss the results, finally in the sixth chapter summary is given:

$$m\ddot{x} = (p - p_a)F - c_1 x - c_2 \dot{x} - P. \tag{1}$$

2. Experimental method

2.1. Image acquisition of pavement damage mechanism

Currently, there are various types of road damage, such as cracks, subsidences, channels, pads, pits, disks, etc., sometimes each damage occurs alone, sometimes in several complex forms [17-20]. However, it can be found that there is a certain law: all kinds of damage phenomena are caused by the interaction between driving and natural factors as well as the road surface, and it varies with external influencing factors and road surface working characteristics. The ability of the subgrade and pavement as a whole to deform or fail under the influence of an external force is referred to as pavement strength, and it reflects the mechanical properties of the pavement. The stability of pavement denotes the variability of the internal structure of the pavement and the strength of the pavement under the influence of external factors and the range of the change. The common road damage phenomenon refers to the specific manifestation of insufficient strength and road strength as well as stability.

As a result, in the face of complex types of pavement damage, we broadly classify pavement damage into five types: potholes, depressions, one-way cracks, looseness, and cracks, which greatly enhance the identification efficiency during the identification process. Moreover, these damage types can provide maintenance personnel with a clearer maintenance direction. Pavement damage image classification is illustrated in Fig. 1.

a) Loose b) Pit c) Depression d) Rut e) Crack

Fig. 1. Image of pavement damage types

2.2. Convolutional neural networks

A convolutional neural network is a very special kind of deep feedforward neural network. It is a multi-level non-fully connected neural network, which simulates the structure of the human brain. It can directly recognise visual patterns from original images via supervised deep learning [20-24]. The CNN model generally has four layers of stacking, including a convolutional layer, pooling layer, fully connected layer, and SoftMax classification layer. Its structure is shown in Fig. 2.

Each layer of the neural network is constructed on multiple planes composed of independently distributed neurons. The connection between layers is a non-fully connected convolution calculation. Each neuron represents a partial dimension value of the input unit weighting. The current relatively successful CNN models consist of the Alex network, Google network, and VGG-19 network, which are capable of accurately identifying generalised objects.

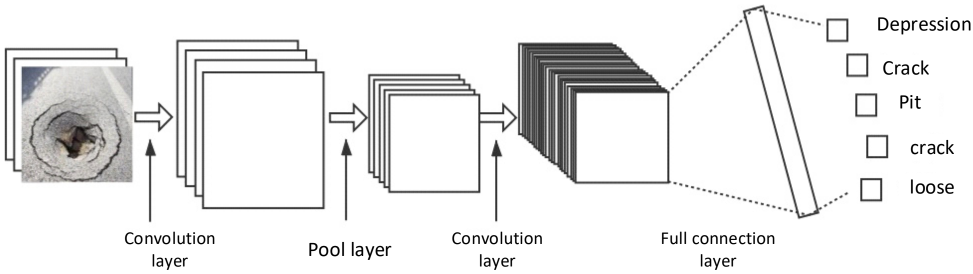


Fig. 2. Structure of convolutional neural network for pavement damage image

2.3. The transfer learning structure of the VGG-16 model

In this section, a transfer learning method is put forward to apply the VGG-19 network to pavement damage detection. Transfer learning is a powerful deep learning technique. We can frequently employ a well-trained convolutional neural network such as the VGG-19 network when the dataset is not large enough to train a complete CNN network, which is shown in Fig. 3.

There are two types of VGG-19 model transfer learning: one is to freeze the weights trained prior to a fully connected layer of the model, only by changing the last fully connected layer, train its own data in the fully connected layer, and finally get the recognition. The other is to modify the weights in the intermediate layers and make the training results more accurate by continuously adjusting the weights. The first form of transfer learning is used in this experiment. First, the VGG-19 network model is introduced by transfer learning. Secondly, freeze the other layers with the exception of the fully connected layer. After importing the data, adjust the weight parameters of the data in the fully connected layer, and finally, train the data to achieve the result.

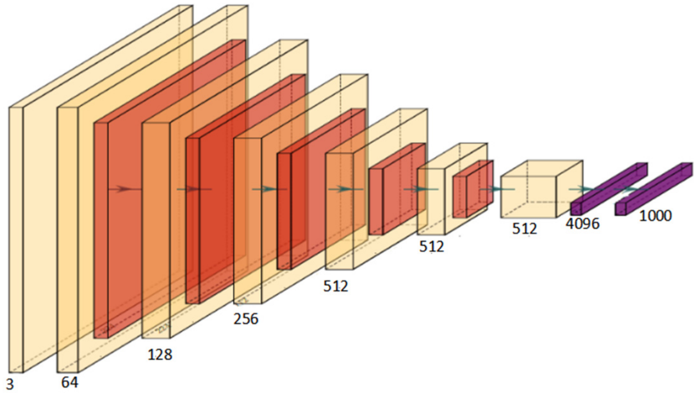


Fig. 3. VGG-19 model structure diagram

3. Experimental process

3.1. Technical route

A pavement damage identification model based on a convolutional neural network is established, and the classification results of pavement damage identification are achieved. Fig. 4 depicts the steps of the research process.

The research is roughly divided into two categories: research on the collection and selection of pavement damage images by means of photography and the selection of network databases. The data set on the experimental subjects were processed by the study following the acquisition. For acquiring a more reasonable classification of pavement damage, the study makes use of the basic theoretical knowledge of pavement damage mechanism so as to reclassify the research object, that is, the pavement damage image. The results obtained from the classification types are displayed in Fig. 1. There are five types: pits, depressions, one-way cracks, looseness, and cracks. If these five forms of damage can be found and resolved immediately, more serious structural damage can be avoided on the road surface and the basic hazards of road surface damage can be mitigated. To ensure better outcomes when the images are fed into the model, the classified dataset is augmented by image processing. For the expansion and processing of the data set, a variety of image processing methods have been adopted in the research, so that the images can be extracted better.

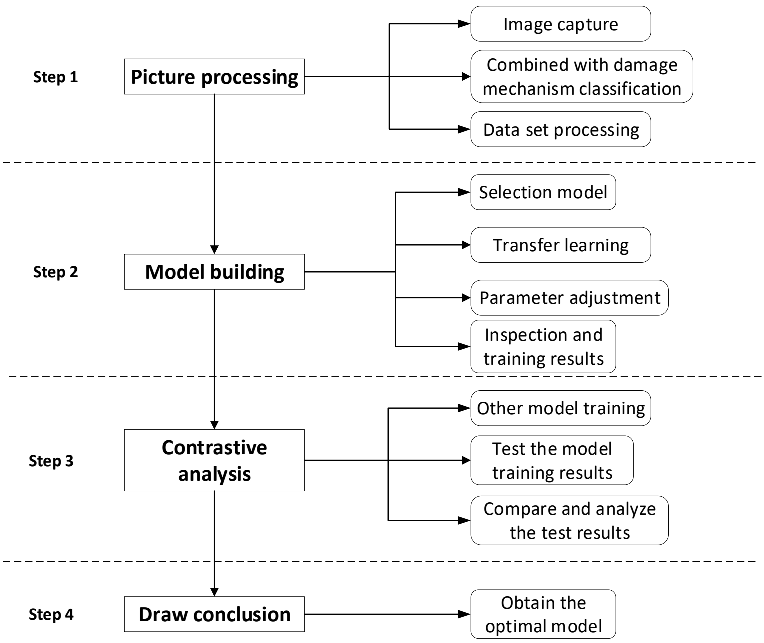


Fig. 4. Schematic diagram of identification process

For the well-processed images, the method of transfer learning is utilised to substitute the pavement damage image data set into the research selection model. After adjusting the details of the selected model, we were able to obtain an experimental model that could be adapted to this dataset, so that this model can be fully applied to the study of pavement damage images.

Since a number of existing network models of convolutional neural networks can be used in the experiments on pavement damage, for the purpose of verifying the rationality of the research model in the selection of pavement damage models, this study carried out models of the same data set for various models, training, and validation testing. The results produced can demonstrate that for the data set, the model selected in this study is more reasonable and can be applied in future research on artificial intelligence detection.

4. Experiment and result analysis result analysis

4.1. Data set processing

Prior to a model being trained, preparatory work or the establishment of the data set is performed. Due to the complexity of pavement conditions and multiple safety factors, pavement damage images cannot get a large number of samples, hence the transfer learning method resolves the difficulty of having a smaller sample library.

For the types of pavement damage we have categorised in this sample database, 20 samples of each type of sample data are gathered, and a total of 100 samples of five types of pavement damage images are collected such as potholes, cracks, and depressions. For these 100 images, a data set was established. We expanded the smaller data set using techniques such as rotation extension and edge extension in order to get better data results. Its rotation formula is as follows:

$$\begin{bmatrix} x_{i'} \\ y_{i'} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} x_i \\ y_i \end{bmatrix}, \quad i = 1,2,3,4. \tag{2}$$

Once the database is expanded, a total of 3000 datasets are obtained, with 600 samples of each type after classification. Secondly, the data set is split into a test set and validation set with a ratio of 4:1, and finally, 2400 test set data as well as 600 validation set data are acquired. Next, we perform region-of-interest interception (ROI) on the classified dataset.

4.2. Image processing

The image is intercepted by the region of interest (ROI), and the image is initially cropped into a 256×256 two-dimensional image. The road damage image is then further preprocessed, the road damage image is flipped horizontally and diagonally, and the random parts are enlarged. Enhance image contrast is based on the histogram equalization method. And fill up the blank parts of the preprocessed image. As a result, the image can be extracted with more accuracy in the subsequent feature extraction process. A specific example picture is indicated in Fig. 5.

Fig. 5 shows a featured image of road damage and surface damage that will be put to the test. It is obtained by performing feature extraction on the region of interest, adjusting the contrast, normalising the image, and performing edge-filling processing on the periphery of the image. This series of operations is carried out on the data set, and a more reasonable sample library is ultimately obtained.

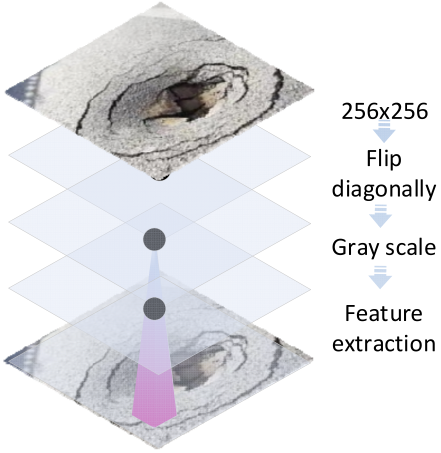


Fig. 5. Image processing

4.3. Establishment of VGG-19 model structure

4.3.1. Model establishment

The processed data set is classified by label, and converted into an electronic data set, followed by inputting into the VGG-19 network that has been migrated. Freeze the convolutional layers and pooling layers with the exception of the fully connected layer, and re-establish the fully connected layer. VGG-19 consists of two parts, VGG19.features, and VGG19.classifier. VGG19.features has convolutional layers and pooling layers as well as three linear layers and finally a Softmax classifier. After loading VGG19 with Tensorflow 1.4 and setting the pretraining weights to True, mark the convolutional and pooling layers as frozen layers, which makes these layers untrainable. Subsequent pairs use CrossEntropyLoss for a multi-class loss function. The basic idea can be found in Fig. 6.

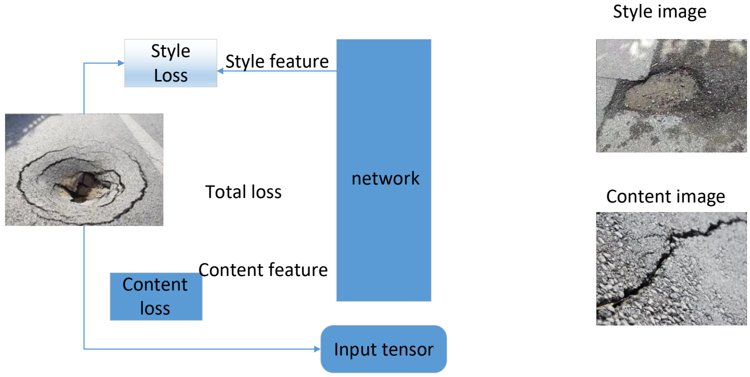


Fig. 6. Basic idea of VGG-19 model based on pavement damage images

4.3.2. Image classification based on softmax classifier

Since the original VGG-19 network model trains 1000 types of images, but then there are only 5 types of training in this paper. Following the modification of the two linear layers of the fully connected layer, the Softmax classifier is likewise utilised for this experiment. The schematic of the Softmax classifier is depicted in Fig. 7.

The output and input image classification types are divided into 5 types. Due to the fact that the Softmax classifier can realise multi-classification of data, it can ensure that the accuracy of the classified data is more accurate. The formula is as follows:

$$y_1 = w_{11}x_1 + w_{12}x_2 + w_{13}x_3 + w_{14}x_4 + w_{15}x_5 + b_1. \tag{3}$$

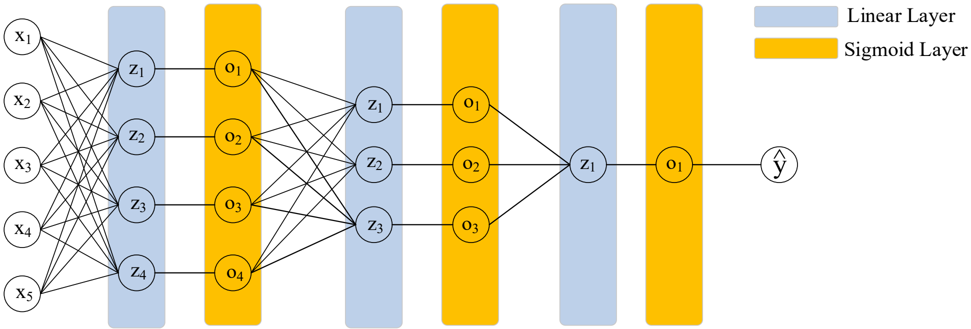


Fig. 7. Softmax classification structure diagram

The SoftMax classifier is an extension of the logistic regression model for multi-classification problems. In the study, the class label y can choose 5 different values. For the training set $\{(x(1), y(1)), \dots, (x(n), y(n))\}$, the class label is $y(n) \in \{1,2, \dots, r\}$. In this paper, the pavement damage images are categorised into five types: potholes, depressions, ruts, cracks, and looseness.

4.3.3. Setting of training parameters

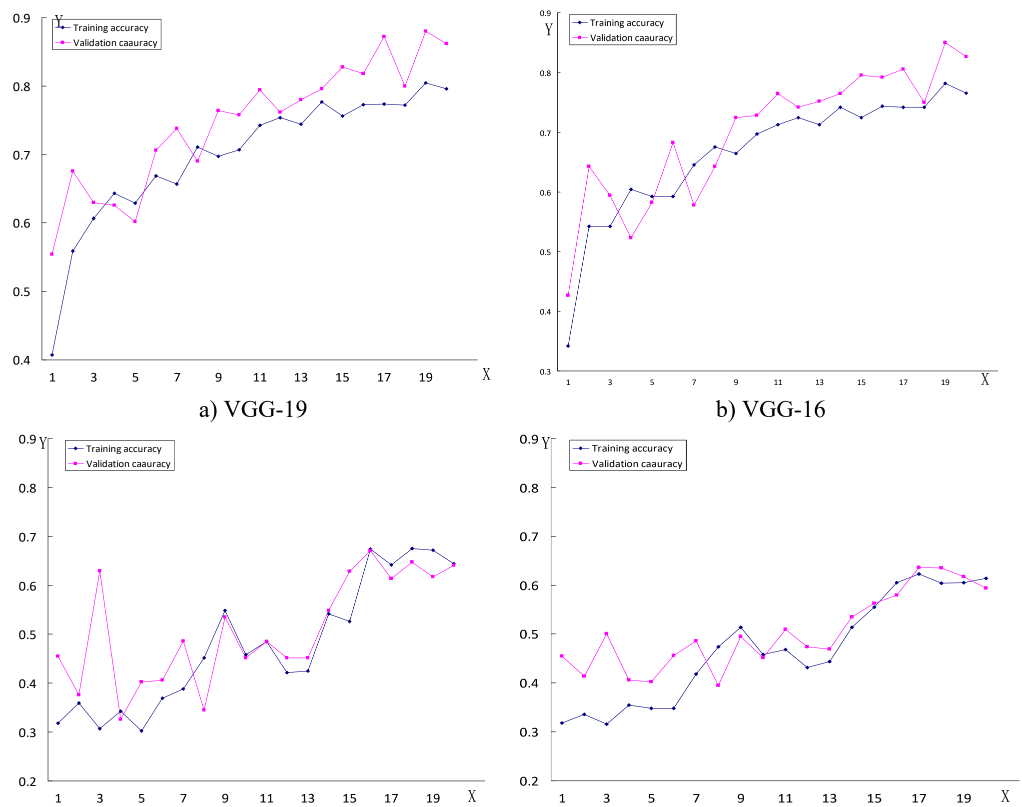
For the training parameters, the image parameters taken in this experiment are 256×256. The tensors for each training of the test set are 100, the tensors for each batch are 20, the total number of training times is 20, and the output verbose = 2; the tensors for each training of the validation set are 50, and the tensors for each batch are 20, the total number of training is 20, and the output verbose = 2. For the optimiser, set the number of neurons in the feature map to 1024, use SGD with a learning rate of 0.0005 and a momentum of 0.9, and use the cross-entropy loss function as the loss function. The formula for the cross-entropy loss function is given below:

$$C = -\frac{1}{n}\sum\nolimits_{x}[y\ln a + (1-y)\ln(1-a)]. \tag{4}$$

4.4. Analysis of results

4.4.1. Model training results

In order to verify the accuracy and reliability of the VGG-19 model, the study employed the same transfer learning method for the purpose of training the VGG-16 model, the ResNet121 model, and the ResNet50 model on the same dataset. The training results are illustrated in Fig. 8. The accuracy of each model can be found in Table 1.



a) VGG-19 b) VGG-16

c) Resnet 121 d) Resnet 50

Fig. 8. The line chart of the accuracy of road damage images trained by each model

Table 1. Comparison of the accuracy rates of various models

Network model	Validation accuracy	Validate the loss rate
Custom model	80.0 %	NUA
VGG-19	86.2 %	0.4043
VGG-16	82.6 %	0.4562
Resnet121	65.7 %	1.3346
Resnet50	60.4 %	1.5826

The diamond dotted line in the figure represents the accuracy of the test set, and the square dotted line reflects the accuracy of the validation set.

The X-axis is the number of detections, while the Y-axis is the accuracy of each detection.

4.4.2. Model comparison analysis

As can be seen from the figure, the training findings of the VGG-19 model are obviously compared with the training process of other models. The phenomenon of overfitting happens because the custom model's loss rate is too high and there are problems of less training set data and fewer convolutional layers of the model during training. Even though the theResNet121, ResNet50, and VGG-16 models do not have overfitting, their accuracy is lower than that of the VGG-19 model, and the loss rate of the ResNet121 model and the ResNet50 model is too high. This demonstrates that these types of models cannot guarantee the training results of graphics for this dataset.

Table 1 presents that the accuracy of the VGG-19 model can reach 86.2 %, while on the other hand, the accuracy of other models is 6 % to 30 % higher in comparison to this model. Compared with other models, the accuracy of this data is more in line with the requirements, the detection of road damage images is more accurate, and it is more suitable for future applications in practical engineering.

5. Conclusions

Aiming at the problem of road damage identification, this paper put forward a transfer learning convolutional neural network utilising VGG-19. After comparing and analysing the test results of three convolutional neural network models of VGG-16, Densnet121, and Resnet50, the following conclusions are drawn.

The road damage detection and recognition methods proposed in this paper add the types of road damage recognition and crank up the accuracy of intelligent road damage identification, which compare with simple road damage crack identification method.

Using less experimental data, this study realizes road damage identification, addresses samples difficulty amid road damage identification and detection, and proves the feasibility of the methods proposed.

The research findings reveal that for the preliminary pavement damage, the research has comparative advantages as well as feasibility for the engineering. In actual road detection, this intelligent detection method can be applied, which is of great help in reducing the work of technicians.

Acknowledgements

The authors have not disclosed any funding.

Data availability

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

Author contributions

Yuanhang Tang: writing – original draft preparation; project administration; software. Kexin Li: writing – review and editing; supervision. Kaihang Wang: visualization; investigation.

Conflict of interest

The authors declare that they have no conflict of interest.

References

[1] J. Lin and Y. Liu, "Potholes detection based on SVM in the pavement distress image," in *2010 Ninth International Symposium on Distributed Computing and Applications to Business, Engineering and Science (DCABES)*, Aug. 2010, https://doi.org/10.1109/dcabes.2010.115

[2] C. Koch and I. Brilakis, "Pothole detection in asphalt pavement images," *Advanced Engineering Informatics*, Vol. 25, No. 3, pp. 507–515, Aug. 2011, https://doi.org/10.1016/j.aei.2011.01.002

[3] L. Zhang, F. Yang, Y. Daniel Zhang, and Y. J. Zhu, "Road crack detection using deep convolutional neural network," in *2016 IEEE International Conference on Image Processing (ICIP)*, Sep. 2016, https://doi.org/10.1109/icip.2016.7533052

[4] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *The Journal of Physiology*, Vol. 160, No. 1, pp. 106–154, Jan. 1962, https://doi.org/10.1113/jphysiol.1962.sp006837

[5] K. Fukushima, "Neocognitron: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics*, Vol. 36, No. 4, pp. 193–202, Apr. 1980, https://doi.org/10.1007/bf00344251

[6] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, Vol. 86, No. 11, pp. 2278–2324, 1998, https://doi.org/10.1109/5.726791

[7] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshops)*, pp. 248–255, Jun. 2009, https://doi.org/10.1109/cvpr.2009.5206848

[8] T. Li and X. Fan, "Road crack identification based on faster R-CNN," (in Chinese), *Application of electronic technology*, Vol. 46, No. 7, pp. 53–56, 2020.

[9] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *European Conference on Computer Vision*, 2013.

[10] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proceedings of ICLR*, 2015.

[11] C. Szegedy et al., "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9, Jun. 2015, https://doi.org/10.1109/cvpr.2015.7298594

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Jun. 2016, https://doi.org/10.1109/cvpr.2016.90

[13] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4700–4708, Jul. 2017, https://doi.org/10.1109/cvpr.2017.243

[14] S. Loffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *Proceedings of the 32nd International Conference on Machine Learning*, pp. 448–456, 2015.

[15] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2818–2826, Jun. 2016, https://doi.org/10.1109/cvpr.2016.308

[16] C. Szegedy, S. Loffe, V. Vanhoucke, and A. Alemi, "Inception-v4, Inception-ResNet and the impact of residual connections on learning," *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 31, No. 1, pp. 1–12, Feb. 2017, https://doi.org/10.1609/aaai.v31i1.11231

[17] J. Dai et al., "Deformable convolutional networks," in *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 764–773, Oct. 2017, https://doi.org/10.1109/iccv.2017.89

[18] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, No. 4, pp. 640–651, Apr. 2017, https://doi.org/10.1109/tpami.2016.2572683

[19] F. Chollet, "Xception: deep learning with depthwise separable convolutions," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1800–1807, Jul. 2017, https://doi.org/10.1109/cvpr.2017.195

[20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, Vol. 60, No. 6, pp. 84–90, May 2017, https://doi.org/10.1145/3065386

[21] X. Zhang, X. Zhou, M. Lin, and J. Sun, "ShuffleNet: an extremely efficient convolutional neural network for mobile devices," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9, Jun. 2018, https://doi.org/10.1109/cvpr.2018.00716

[22] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-excitation networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 42, No. 8, pp. 2011–2023, Aug. 2020, https://doi.org/10.1109/tpami.2019.2913372

[23] O. Russakovsky et al., "ImageNet large scale visual recognition challenge," *International Journal of Computer Vision*, Vol. 115, No. 3, pp. 211–252, Dec. 2015, https://doi.org/10.1007/s11263-015-0816-y

[24] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv:1409*, 2014, https://doi.org/10.48550/arxiv.1409.1556

[25] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: surpassing human-level performance on ImageNet classification," *2015 IEEE International Conference on Computer Vision (ICCV)*, Dec. 2015, https://doi.org/10.1109/iccv.2015.123

[26] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "feature pyramid networks for object detection," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, https://doi.org/10.1109/cvpr.2017.106

[27] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, No. 6, pp. 1137–1149, Jun. 2017, https://doi.org/10.1109/tpami.2016.2577031

[28] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," in *Lecture Notes in Computer Science*, pp. 234–241, 2015, https://doi.org/10.1007/978-3-319-24574-4_28

[29] Ali Morovatdar, Reza Ashtiani, Carlos Licon, and Cesar Tirado, "Development of a mechanistic approach to quantify pavement damage using axle load spectra from south texas overload corridors," in *Geo-Structural Aspects of Pavements, Railways, and Airfields (GAP 2019)*, 2019.

Yuanhang Tang received bachelor's degree in civil engineering from Beihua University, Jilin, in 2019. Now she studies at Beihua University. His current research interests include Road damage and image processing.

Kexin Li received Ph.D. in Road and Railway Engineering from Northeast Forestry University in June 2020. She is a full-time teacher and lecturer in the School of Civil Engineering and Transportation, North Hua University. Her main research areas are structural health monitoring, nondestructive testing, and pattern recognition.

Kaihang Wang degree in civil engineering from Beihua University, Jilin, in 2019. Now she studies at Beihua University. His current research interests include Road damage and image processing.