Deep learning based multi-channel road crack detection method

Shengcheng Yang, Pengchao Jing, Zhixiong Guo, Shengxi Wang*, Lei Gong, Dongdong Mu, Xingjun Cai

China Construction Third Engineering Bureau Group Co., LTD, Xian, 710065, China

*E-mail: 690423330@qq.com

Abstract. With the growth of our economy, the construction of road infrastructure in the country is also developing vigorously. Under prolonged use, cracks of varying degrees can appear on road pavements. The detection of cracks on road surfaces is also an important part of road maintenance as well as road safety and security. Typically, traditional manual road crack detection is time-consuming and labour-intensive, and the results can be inaccurate due to differences in individual evaluation criteria. With the increasing maturity of deep learning algorithms, deep learning models are also being used for road crack detection. However, the currently used deep learning-based road crack detection methods are all based on the extraction of colour RGB images, and the information of the images is not fully utilised and the accuracy of the extraction needs to be improved. Therefore, this paper proposes a multi-channel road crack detection method based on deep learning, which improves the road crack detection accuracy by using RGB images and its grey map four channels as model inputs. Comparing the extraction results of deep learning models using only RGB three channels and four channels on the publicly available dataset CRACK500, it is found that the multi-channel road crack detection method proposed in this paper achieves an accuracy of 86.69%, which is ~10% better than the traditional RGB three channel detection method in terms of accuracy. It also provides a new idea for the road crack detection method based on deep learning.

Keywords: Crack detection; Deep learning; Multichannel; Convolutional neural networks

1 Introduction

With the development of economy, China's highway infrastructure is very complete. The detection of road cracks is a very important work for highway maintenance and road safety and security. Therefore, the rapid and automatic detection of road pavement cracks is an urgent engineering problem. At present, the inspection of road cracks mainly relies on manual walking inspection [1], which generally requires professional

C. Chen et al. (eds.), *Proceedings of the 3rd International Conference on Digital Economy and Computer Application (DECA 2023)*, Atlantis Highlights in Computer Sciences 17,
https://doi.org/10.2991/978-94-6463-304-7_59

inspectors to go to the field for judgment and assessment, which is both time-consuming and laborious. At the same time, due to the strong subjectivity of this method, it will lead to the inaccuracy of the judgment results.

In recent years, the rapid development of computer vision and semantic segmentation technology has made the rapid and automatic detection of road cracks possible [2,3,4]. Some image processing based road crack recognition methods have been proposed, such as Salloo et al [5] used suitable thresholds on a gray scale map to extract cracks on roads.Cheng et al [6] also proposed an algorithm for road crack detection based on gray scale map, which is based on the two logics that there is a difference between the gray scale value of the cracks and the background gray scale value as well as the existence of the continuity of the same kind of pixels, which can effectively detect the cracks in road images.Tanaka et al [7] used a morphological approach for road crack detection based on the morphological properties of the cracks. However, these methods are susceptible to external environmental conditions such as lighting, background, texture and shadows. Therefore their detection results lack stability and robustness.

In order to solve the above problems, artificial intelligence technology is applied to the crack detection work of roads. Deep learning-based road crack detection methods can quickly and effectively extract the cracks on the road surface. Currently, all deep learning-based road crack detection methods use color RGB images as input to the model, i.e., a three-channel deep learning model is used. The information of the images may not be fully utilized. Therefore, this study proposes a multi-channel deep learning based road crack detection method. By increasing the input of the model, the image information can be mined more, so that the model can extract the cracks of the road pavement more accurately.

2 Project Overview

This study introduces the multi-channel road crack detection method with the overlap section at the starting position of Maishizha Yellow River Bridge project in Jianzha County, Huangnan Prefecture, Qinghai Province - Shenwu North Road as the engineering background. Huangnan County, Jianzha County, Maishizha Yellow River Bridge and its approach project for the north-south direction, the starting point is located in the north new area of Jianzha County to connect with the built Shenwu North Road, the end of the Tuanjie-2 village in the unity of Hualong County, the unity of the interchanges level crossing; the project start and end piles for the K0+044.625-K1+089.868, the route length of 1045.243m, the main road for the city. Shenwu North Road due to long time use and large vehicles crushed local road surface cracks of different degrees, the detection of road surface cracks is one of the links of construction quality and safety. In order to quickly and accurately detect the cracks on the road surface, this study detects the cracks through image recognition with a multi-channel road crack detection method based on deep learning.

3 Deep Learning Models with Multiple Channels

In this study, we use the FANet [8] deep learning network model as a basis on which to build a multi-channel deep learning model.FANet was first proposed by Tomar et al. and used for segmentation of medical images. Its structure can effectively localize the target location, learn its features and add them to the next round of training so that the network can better learn the features of the target. Also, the network is designed with SE-Residual module which effectively solves the gradient vanishing and exploding problems.The whole network of FANet has 7.72 million parameters and 94.75 GMacflops.Compared with other common deep models, FANet has less parameters and the network model is more lightweight. In this study, the original three-channel FANet network was changed to four channels, and in addition to the regular RGB channels, this study also added a grayscale map channel as an input to the model.

As shown in Figure 1., the workflow of this study consists of three aspects: preparing the training set, training the model, and crack extraction. First, the gray map of the RGB images in the public dataset CRACK500 is calculated as shown in Equation (1). The calculated gray map is combined with RGB to form a four-channel dataset. This dataset is used as the training set to train the constructed multi-channel FANet model. Once the model is trained, the crack detection of road images can be performed. The input of the model is still RGB image plus four channels of grayscale map, and after feeding them into the model, the model will output the corresponding detection results. Meanwhile, in order to make the data more representative, this study also uses data enhancement operations to expand the dataset, which include: flipping, rotating, and linear stretching of the image.

$$Gray = 0.299R + 0.587G + 0.114B. \tag{1}$$

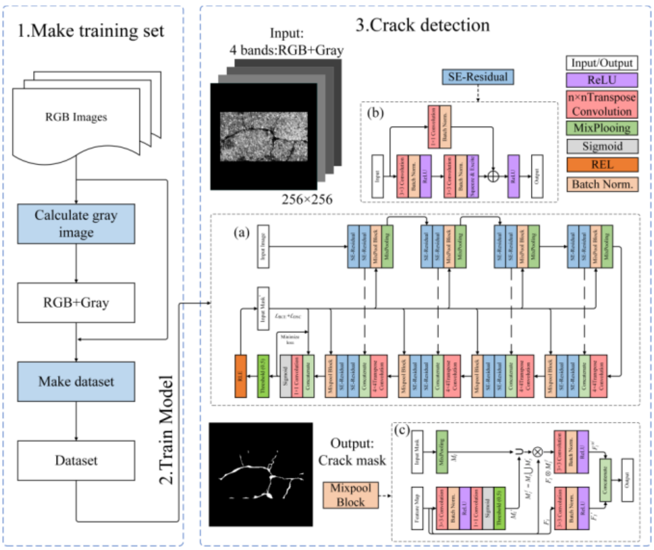

Fig. 1. Workflow

4 Experimental data and environment

The experimental data used in this study is the publicly available dataset CRACK500 for road crack detection.The CRACK500 dataset contains a total of 500 image data of size 2560×1440 pixels. In this study, the training set, test set and validation set data are divided according to the ratio of 6:2:2. The dataset includes images with different lighting and shading conditions to ensure image diversity.

The hardware configuration of the experimental environment in this study is Windows 11 operating system, 64GB of RAM, and NVIDIA GeForce RTX 3060 graphics card.The programming language used is Python 3.7, and the deep learning model is based on the Pytorch deep learning framework, compiled with the VScode platform.The experiments have been performed for 150 rounds of training, and the Adam optimizer module has been used to achieve the initial learning rate. The Adam optimizer module was used with an initial learning rate of 1e-4, and ReduceLROnPlateau was used to monitor the training accuracy and automatically adjust the learning rate. The ReLu activation function is used to suppress the gradient vanishing problem during the training process, which accelerates the model convergence speed and reduces the training time.

$$R_{ReLu}(x) = \begin{cases} 0, X < 0 \\ x, x \geq 0. \end{cases} \tag{2}$$

5 Results

5.1 Visualization and analysis of experimental results

In this study, the original deep learning model as well as the multi-channel deep learning model proposed in this study were used to extract road cracks from the images in CRACK500, respectively. Figure 2. shows a comparison of some results of the two models for road crack extraction. The comparison reveals that there are a large number of misclassifications in the results extracted by the original convolutional neural network based only on the RGB three-channel, and the misclassifications mostly occur at the image edges. On the one hand, this is due to the fact that the neighborhood information at the edge of the image is not as complete as that in the middle, and on the other hand, it also shows that the three-channel deep learning model needs to be improved. In contrast, similar problems no longer exist in the results extracted by the multi-channel based deep learning model in this study. It indicates that the increased channel input provides more effective information for the deep learning model, which improves the problem of being mistakenly mentioned as road cracks at the edge of the image, and effectively improves the extraction of cracks.

The experiment verifies the feasibility of deep learning to detect road cracks on the one hand, and on the other hand, it illustrates qualitatively that the multi-channel based deep learning method proposed in this paper can improve the effect of the depth model to detect road cracks. The method in this paper can be directly applied to this project to reduce the workload of manual inspection of road cracks and realize the automation and intelligence of engineering maintenance.

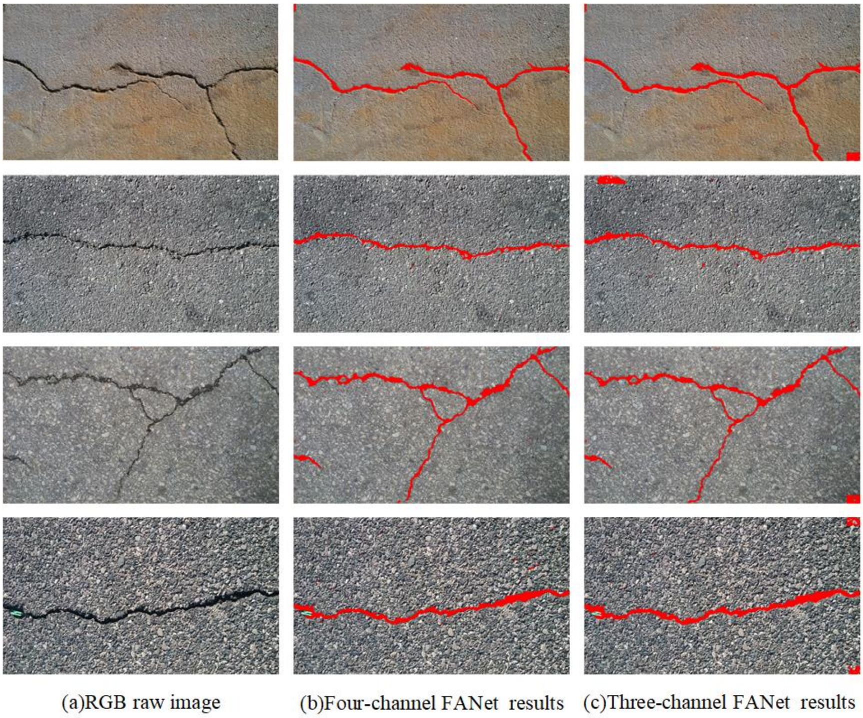

Fig. 2. Comparison of model extraction results

5.2 Quantitative assessment of experimental results

In order to quantitatively evaluate the accuracy of the multi-channel FANet and three-channel FANet models for extracting road cracks, the accuracy metrics were calculated in this study using the results of both on the validation set of CRACK500, which includes all the typical images in the dataset to ensure the validity of the validation data. Table 1 lists the accuracy of the two models for road crack extraction. The evaluation metrics of accuracy include: precision, recall, F1 score, and overall accuracy. These evaluation metrics are calculated as follows:

$$Precision = \frac{TP}{TP+FP} \times 100\%. \tag{3}$$

$$Recall = \frac{TP}{TP+FN} \times 100\%. \tag{4}$$

$$F1\ score = \frac{2 \times Precision \times Recall}{Precision+Recall} \times 100\%. \tag{5}$$

$$Overall\ Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%. \tag{6}$$

Where TP (true-positive) represents the number of pixels in which road cracks were correctly detected; TN (true-negative) represents the number of pixels in which "clean" backgrounds were correctly detected; FP (false-positive) represents the number of pixels in which "clean" backgrounds were incorrectly detected as cracks; and FN (false-negative) represents the number of pixels in which road cracks were incorrectly recognized as "clean" backgrounds.

From the results, it can be seen that all the precision metrics are improved with the addition of the grayscale map channel, and the precision of detection is improved from ~76% to ~87% compared to the three-channel model. The recall of both models is greater than 85%, indicating that both can detect most of the road cracks. However, the multi-channel model has a recall and overall precision of 90%, which indicates that almost all road crack pixels are detected. Also a 4% improvement in F1 score exists.

It can be illustrated through experiments that the original RGB three-channel deep learning model does not acquire enough image information, and there is still room for improving the performance of the model. The multi-channel deep learning model can effectively improve the accuracy of the convolutional neural network in the road crack detection work, and has achieved considerable results. The road crack extraction model designed in this study adds the grayscale map channel as the model input. The grayscale map is also a common way in road crack detection, in which the location of the cracks can be clearly seen, which makes the detection of road cracks easier. Therefore, the input of multiple channels can provide the deep learning model with additional information of the RGB image, which can fully utilize the information of the original image, making the deep learning model learn the spatial, spectral, and geometric features of the cracks more fully, and effectively improving the detection accuracy.

Table 1. Detection accuracy of the two models

	Precision	**Recall**	**F1 score**	**Overall accuracy**
Three-channel model	75.98%	87.91%	83.06%	**86.98%**
multichannel model	86.69%	91.60%	87.30%	**90.61%**

6 Conclusions

In this study, a multi-channel deep learning based road crack detection method is proposed to increase the input information by increasing the number of channels of the deep learning model so that the model can accept more valid information and mine the road image more fully, which ultimately improves the accuracy of road crack extraction. Compared to the traditional RGB three-channel deep learning model, the RGB as well as the four-channel model for gray-scale maps used in this study improved the accuracy on the validation set by ~10% and achieved 86.69%. It shows that the multi-channel method proposed in this study can provide more effective information and fully explore the information contained in the target image, which can effectively improve the detection accuracy of road cracks and provide a new idea for deep learning-based road crack detection. It also demonstrates the feasibility of deep learning for road crack detection work, which provides a basis for automating road crack detection and road maintenance in construction. For the overhaul and maintenance of the project, it reduces the manual workload, realizes the "intelligent construction site", saves time and economic costs, and ensures the effective, smooth and rapid progress of the project.

References

1. Oliveira H, Correia P L. Supervised crack detection and classification in images of road pavement flexible surfaces[J]. Recent advances in signal processing, 2009: 159-184.
2. Hu G X, Hu B L, Yang Z, et al. Pavement crack detection method based on deep learning models[J]. Wireless Communications and Mobile Computing, 2021, 2021: 1-13.
3. Yang F, Zhang L, Yu S, et al. Feature pyramid and hierarchical boosting network for pavement crack detection[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 21(4): 1525-1535.
4. Gao Z, Peng B, Li T, et al. Generative adversarial networks for road crack image segmentation[C]//2019 International Joint Conference on Neural Networks (IJCNN). IEEE, 2019: 1-8.
5. Shi Y, Cui L, Qi Z, et al. Automatic road crack detection using random structured forests[J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(12): 3434-3445.
6. Cheng H D, Chen J R, Glazier C, et al. Novel approach to pavement cracking detection based on fuzzy set theory[J]. Journal of Computing in Civil Engineering, 1999, 13(4): 270-280.
7. Tanaka N, Uematsu K. A Crack Detection Method in Road Surface Images Using Morphology[J]. MVA, 1998, 98: 17-19.
8. Tomar N K, Jha D, Riegler M A, et al. Fanet: A feedback attention network for improved biomedical image segmentation[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022.