Journal of Southeast University (English Edition) Vol. 39, No. 4, pp. 340 - 349 Dec. 2023 ISSN 1003—7985

Intelligent detection of cracks on cement pavements of rural highways

Wang Meng[1] Zhang Xiaoyue[1] Liu Cheng[2] Xu Huitong[1] Yang Yanze[1]

([1] School of Civil Engineering, Beijing Jiaotong University, Beijing 100044, China)

([2] China-Road Transportation Verification and Inspection Hi-Tech Co., Ltd., Beijing 100088, China)

Abstract: Traditional artificial image processing methods suffer from problems in the detection of damage in rural highway pavements, such as low efficiency, nonobjective results, and the inability to process a large amount of data in time. To solve these problems, an intelligent method is proposed for the detection of cracks on rural cement pavements. The proposed method is integrated with ResNet50 for pavement classification and an improved YOLOv5 crack detection algorithm, considering the distribution characteristics of rural highway sections. Different training strategies and different network depth were compared to construct an efficient pavement classification model based on ResNet50 with the aim of automatically identifying cements and asphalt pavements in rural highways. A dataset that contains 18 028 pieces of crack detection data for cement pavements of rural highways was created. A comparative experimental study of single- and two-stage object detection algorithm was performed, and the optimal detection algorithm with both detection accuracy and efficiency was obtained. Furthermore, the adaptive spatial feature fusion strategy and the optimized regression loss function are integrated into the optimization algorithm to effectively solve the problem of multi-scale crack leakage detection in the image, and further improve the overall detection accuracy. The integrated method was applied to the field measurement of real cement pavements of rural highways. The results demonstrate that the accuracy of pavement type classification is 98.4% and that of crack detection is 93.0%, indicating that the proposed method can provide accurate and efficient solutions for the detection of cement pavements of rural highways.

Key words: rural highway; cement pavement; crack; deep learning; image classification; object detection

DOI: 10.3969/j.issn.1003 - 7985.2023.04.003

By the end of 2022, the total mileage of China's highways had reached $5.354\,8\times10^6$ km, and the mileage of highway maintenance was $5.350\,3\times10^6$ km, accounting for 99.9% of the total highway mileage[1]. This shows that China is not only expediting the construction of highways but also paying increasing attention to the challenge of road maintenance[2]. The mileage of rural highways is $4.531\,4\times10^6$ km, accounting for 84.62% of the total length of highways[1]. As an important part of China's road network, the maintenance of rural highways has not received sufficient attention owing to the influence of capital, technology, personnel, and other factors. However, with the guidance of national policies and the continuous enrichment of maintenance methods, the maintenance of rural highways has gradually improved. In 2017, the State Council required in the "Notice on the Issuance of the 13th Five-Year Plan Modern Comprehensive Transportation System Development Plan" to strengthen the maintenance of rural highways, improve protection facilities, and ensure the basic travel conditions in rural areas. The "Development Outline of Road Maintenance Management in the 14th Five-Year Plan," issued by the Ministry of Transport in 2022, was proposed to improve the rural highway management system and further facilitate the reformation of the rural highway management system alongside stating other requirements.

Highways in rural areas are mostly of low grade, and their pavements are mainly constructed using cement. Cement pavement diseases are on the rise owing to compromised construction quality, the rolling of heavy-duty vehicles, and the influence of climate and temperature[3]. Regarding the detection of pavement diseases, the traditional manual operation method suffers from low efficiency and low precision[4], which cannot meet the requirements for large-scale maintenance management at this stage. Therefore, research must be performed on intelligent disease detection methods. In a bid to ensure the accuracy and implementability of such methods, the Highway Performance Assessment Standards (JTG 5210—2018)[5] stipulates that automated testing equipment should be able to distinguish approximately 1.0 mm of pavement cracks and that the recognition accuracy should exceed 90%. The Ministry of Transport proposed several regulations to simplify the classification of diseases in the Guidelines for the Assessment of the Technical Status of

Received 2023-08-13, **Revised** 2023-10-19.

Biography: Wang Meng (1985—), female, doctor, professor, wangmeng@bjtu.edu.cn.

Foundation items: Beijing Nova Program (No. 20220484103), Beijing Municipal Natural Science Foundation (No. 8222027), the Fundamental Research Funds for the Central Universities (No. 2022YJS071).

Citation: Wang Meng, Zhang Xiaoyue, Liu Cheng, et al. Intelligent detection of cracks on cement pavements of rural highways[J]. Journal of Southeast University (English Edition), 2023, 39(4): 340 - 349. DOI: 10.3969/j.issn.1003 - 7985.2023.04.003.

Low-grade Rural Roads[6] for the automated detection of low-grade rural highways.

Cracks on the cement pavement of rural highways often exhibit various forms, such as cross cracks, hairline cracks with less damage, and severe cracks that extend and eventually penetrate the plate[7]. Meanwhile, the pavement conditions of rural highways are complex, and often, many unavoidable disturbances occur, such as horizontal and longitudinal joints, broken pavement markings, shadows, falling branches, pavement stains, and scratches. Hence, cracks on the cement pavements of rural highways have become one of the road diseases difficult to accurately identify.

Deep learning-based object detection algorithms are widely used in intelligent detection methods. This approach can be used to automatically extract object features by learning a large number of data samples to realize crack detection[8]. Gou et al.[9] proposed an improved, faster region-based convolutional neural network (R-CNN) model for testing complex asphalt pavement cracks. Sun et al.[10] applied this R-CNN model to automate the identification of pavement-sealed cracks. The patent by Xia et al.[11] provided a pavement crack identification method based on the Cascade R-CNN model, which can increase pavement crack identification accuracy. Zhang et al.[12] incorporated an adaptive spatial scale fusion algorithm based on YOLOv3 (You Only Look Once, Version 3) to detect road cracks, which can achieve satisfactory detection results. Hegde's group[13] improved YOLOv5 and proposed three methods, namely integrated prediction, integrated model, and integrated prediction plus integrated model, to detect road damage. Additionally, various other algorithms, such as SSD[14-15], Mask R-CNN[16-17], RetinaNet[18], and YOLOX[19], have been used for the detection of road surface diseases.

However, crack detection of rural highway cement pavements still suffers from the following problems: 1) Research on intelligent crack detection is usually performed based on specific datasets and specific algorithms of the scenarios studied in corresponding papers. Additionally, most optimized models have not been open source, which impedes their usability, and thus, one cannot confirm whether any such algorithm is suitable for other application scenarios. 2) The self-built dataset used in the current research generally contains a small number of images, with the data source being relatively single, resulting in a lack of generalization ability of model detection. 3) A lot of interference occurs during pavement crack detection, which easily results in misidentification by the model. The high complexity of crack morphology complicates the identification of fine cracks.

In this study, a total of 247 812 images were collected from multiple provinces and rural road surfaces to solve the abovementioned problems. A pavement classification model, which is based on ResNet50 (the residual network with a network depth of 50), was constructed to efficiently identify asphalt and cement pavements by adjusting the parameters mixed precision and network depth. A rural highway cement pavement crack dataset with a wide range of data sources and a large amount of data was created based on the classification results. A model comparison test was performed on six typical deep learning-based object detection algorithms, following which the optimal detection model for cement pavement cracks of rural highways was selected. The selected YOLOv5 is enhanced by incorporating adaptive spatial feature fusion (ASFF) and optimizing the regression loss function Wise-IoUv3 (WIOUv3). The integrated, intelligent crack detection method, combined with the ResNet50 classification model and enhanced detection model, was tested continuously on the road, which provides a reference for intelligent crack detection in the cement pavements of rural highways.

1 Deep Learning-Based Object Detection Algorithm

1.1 Description of the deep learning-based object detection algorithm

Deep learning-based object detection algorithms can be generally classified into two categories[20]. One uses candidate regions to identify and locate targets; hence, these are called two-stage object detection algorithms. The other kind is based on direct regression, and the identification and location are performed by a network model; hence, these are called single-stage object detection algorithms.

Typical two-stage detection algorithms include R-CNN[21], Fast R-CNN[22], Faster R-CNN[23], Mask R-CNN[24], and Cascade R-CNN[25]. The commonly used single-stage detection algorithms are YOLO series[26-29] and SSD[30]. In this study, a total of two two-stage detection algorithms and four single-stage detection algorithms are selected (see Table 1), and their performances in crack detection of cement pavements of rural highways are compared to provide the necessary basis for selecting the algorithm most suitable for such applications.

1.2 Evaluation indexes of the object detection algorithms

There are two types of evaluation indexes for object detection algorithms, both of which are used to evaluate the detection speed and detection accuracy of a model, respectively. Frames per second (FPS), which is the number of image frames that can be detected by a model per second, is used to evaluate a model's detection speed.

Table 1 Six selected algorithms and their characteristics

Algorithm type	Algorithm name	Algorithm characteristics	Shortcomings
Two-stage detection algorithm	Faster R-CNN[23]	A region proposal network (RPN), a fully convolutional network, is first introduced, which can apply convolutional features along with detection networks, and the time consumed in the region proposal process is considerably reduced.	The detection speed is too slow to meet the speed requirements of real-time detection.
	Cascade R-CNN[25]	Cascade regression is used as a resampling mechanism to solve the problem that high-threshold classifiers cannot learn effectively, avoid the overfitting problem of the model, and increase the accuracy of detection.	
Single-stage detection algorithm	YOLOv3[28]	A series of changes have been made, such as confidence score prediction for boundary box, object classification for multiple labels, cross-size prediction, and feature extraction using a new backbone network, and the performance has been considerably improved.	Low positioning accuracy; object detection is not sufficiently comprehensive.
	YOLOv5	Mosaic data enhancement, Focus + CSP structure, FPN + PAN structure, and CIOU _Loss are applied, and the detection speed is considerably fast[31].	
	YOLOX[32]	It is an object detection algorithm without an anchor frame that can directly predict object boundary frames. Decoupling heads are used to enhance the classification and regression performances. The label allocation strategy SimOTA was applied to reduce the training time.	
	YOLOv8[33]	It belongs to the class of object detection algorithms without an anchor frame, which uses mosaic data enhancement technology, a C2f module, FPN + PAN structure, soft-NMS post-processing technology, and a decoupling detection head.	

Precision (P), recall (R), average precision (AP), and mean value of average precision (mAP) are commonly used to evaluate the detection accuracy of a model. Precision and recall were calculated from the confusion matrix shown in Table 2. Precision refers to the proportion of TP (samples predicted by the model to be cracks that are indeed cracks) in TP + FP (all samples predicted to be cracks), which represents the detection accuracy of the model. Recall refers to the proportion of TP in TP + FN (all samples that are actually cracked), indicating whether the model is comprehensive or not. Precision and recall are expected to simultaneously attain the optimum value during the model training process; however, the two will contradict each other in some cases. Therefore, the P-R curve is formed with precision as the vertical axis and recall as the horizontal axis. The area surrounded by the P-R curve and the horizontal and vertical axes is defined as AP. The term AP reflects the detection object, and mAP is the average of all detection objects.

Table 2 Confusion matrix

Confusion matrix		Actual	
		Positive(crack)	Negative(noncrack)
Predicted	Positive (crack)	TP	FP
	Negative (noncrack)	FN	TN

2 Construction of the Cement Pavement Crack Dataset of Rural Highways with Typical Large Samples

A total of 247 812 original images of multiple rural highways from many provinces were selected to create a cement pavement crack dataset of rural highways with large samples, with an image size of 4 096 × 2 280 pixels.

2.1 Intelligent recognition of pavement types based on image classification technology

In the process of mass data acquisition of rural highways, it is difficult to ensure that all images belong to the same type of road surface. Hence, one must distinguish the images before building the dataset for detection. In a bid to accurately distinguish cement pavements from asphalt ones, ResNet18 was applied to explore the influence of the single- and mixed-precision training strategies on training duration and classification performance. The indicators P, R, A, and F_1 were used for the performance evaluation. The evaluation index A represents accuracy, that is, the proportion of TP + TN (samples correctly predicted) in TP + TN + FP + FN (all predicted samples) (see Table 2), indicating the probability of overall correct identification of the two samples by the model. The evaluation index F_1 integrates the two indexes of precision and recall, and the higher the value, the better the model's performance. F_1 is calculated as

$$F_1 = \frac{2PR}{P + R}$$

The ratio of single precision time to mixed precision time for training an epoch is 2 : 1. Using mixed precision considerably reduces training time and increases training efficiency. A comparison of the classification performances of various training strategies is presented in Table 3. The results demonstrate that the mixed-precision strategy will not cause the loss of precision, and still maintain a high classification accuracy. Therefore, the mixed-precision strategy can improve the training efficiency on the premise of ensuring high classification accuracy.

Table 3 Comparison of classification performances of various training strategies %

Category	Training strategy	P	R	A	F_1
Asphalt	Single precision	96.70	99.01	97.82	97.84
	Mixed precision	96.97	99.31	98.10	98.13
Cement	Single precision	98.99	97.84	97.82	98.41
	Mixed precision	99.29	96.89	98.10	98.08

To further explore the influence of network depth, based on the mixed precision strategy, the classification performances of ResNet18, ResNet50, and ResNet101 are compared on the two pavement types in Table 4.

As shown in Table 4, ResNet50 classifies two types of images most accurately, and neither increasing nor decreasing the network depth can enhance the model's performance. The change curve of the loss function of the training and the validation set of the ResNet50 classification model is shown in Fig. 1. After training 15 epochs, the loss of both the training set and the validation set remains approximately 0, and there is no overfitting phenomenon of loss increase, which confirms that the model training performance is satisfactory. Therefore, the classification model trained by ResNet50 as the backbone network is selected for image differentiation. The specific classification process is shown in Fig. 2.

Table 4 Comparison of classification performances of various network depths %

Category	Network depth	P	R	A	F_1
Asphalt	18	96.97	99.31	98.10	98.12
	50	97.56	98.89	98.21	98.22
	101	96.52	99.15	97.79	97.81
Cement	18	99.29	96.89	98.10	98.08
	50	98.88	97.53	98.21	98.20
	101	99.12	96.42	97.79	97.75

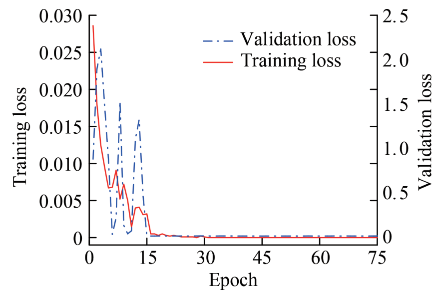


Fig. 1 Change curve of loss function for training and validation sets

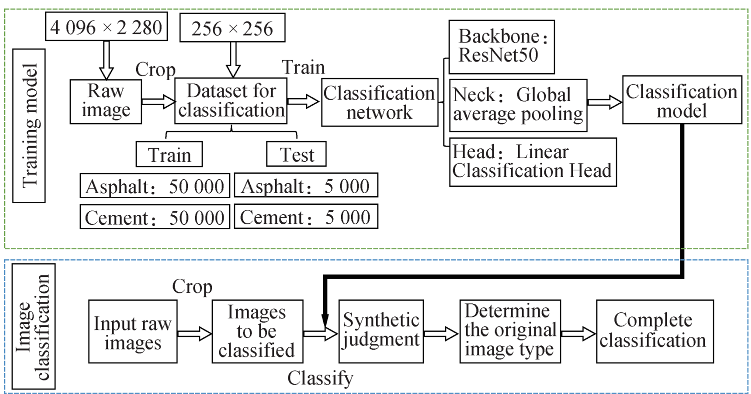


Fig. 2 Image classification process

2.2 Construction of the cement pavement crack dataset of typical rural highways

Using this classification model, 18 028 images of cement pavements were selected as annotated data, and the dataset covered cracks with various shapes, widths, and trends. Meanwhile, taking advantage of the large number of samples, joint and pavement markings (see Fig. 3) are taken as negative samples and marked with cracks at the same time. This is so that the algorithm can learn and identify features and effectively increase the crack detection accuracy.

Images are dividedinto training, validation, and test sets at a ratio of 8 : 1 : 1, and the construction process is shown in Fig. 4. In the process of data set partitioning, the hierarchical sampling method is used to make the proportion of positive and negative samples as similar as possible, so as to ensure the consistency of data distribution. The marked quantities of the three types of objects are presented in Table 5.

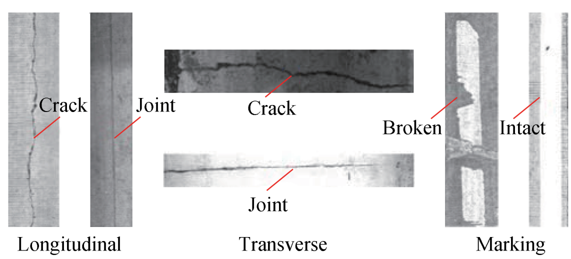


Fig. 3 Joints, cracks, and markings

Table 5 Number of annotations for the three types of objects

Dataset	Positive sample	Negative sample		Positive and negative sample ratio
	Crack	Joint	Marking	
Training set	22 083	8 431	429	2.49 : 1
Validation set	2 686	1 040	54	2.46 : 1
Test set	2 844	1 103	54	2.46 : 1

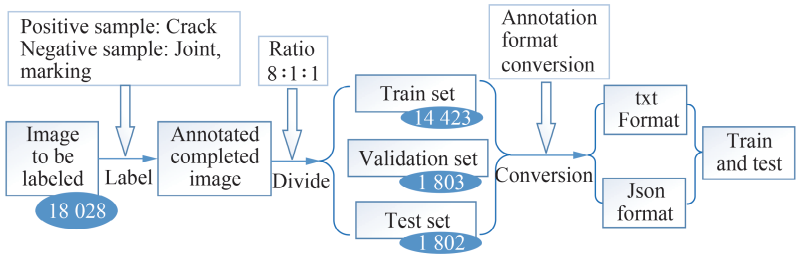


Fig. 4 Building process of the dataset

3 Experimental Study on the Optimized Model of Intelligent Detection for Cracks on Cement Pavements of Rural Highways

3.1 Experimental parameters

The operating system used for algorithm training is Ubuntu 20. 04. 4, the processor is Intel Xeon Processor (Skylake), and the GPU model is Tesla V100 SXM2 16 GB. PyTorch is used as the deep learning framework. Table 6 presents some hyperparameters. The detection model adopts the mixed precision training method, and the optimization algorithm adopts the stochastic gradient descent algorithm. YOLOv5 and YOLOX selected YOLOv5s and YOLOXs as the pretraining models, respectively, to increase the convergence speed of the model and avoid overfitting. Other algorithms applied the corresponding pretraining backbone network. The core components of the eight models are presented in Table 7.

Table 6 Setting of partial hyperparameters

Hyperparameters	Value
Learning rate	0.01
Batch size	16
Epoch	75
Image size/pixel	640 × 640
Intersection over the union threshold	0.50
NMS threshold	0.50
Confidence threshold	0.30

Table 7 Core components of the models

Algorithm	Backbone	Neck	Head	Model number
Faster R-CNN	ResNet50	FPN	RPN + StandardRoI	1
	ResNet101	FPN	RPN + StandardRoI	2
Cascade R-CNN	ResNet50	FPN	RPN + CascadeRoI	3
	ResNet101	FPN	RPN + CascadeRoI	4
YOLOv3	MobileNetV2	YOLOV3Neck	YOLOV3Head	5
YOLOv5	Focus + CSP	FPN + PAN	Coupled-Head + Anchor-base + CIOU_Loss	6
YOLOX	CSPDarknet	YOLOXPAFPN	Decoupled-Head + Anchor-Free	7
YOLOv8	Focus + C2f	FPN + PAN	Decoupled-Head + Anchor-Free	8

3.2 Experimental results and analyses

When the intersection over the union threshold is set to 0.5, the AP values of cracks, joints, and markings are denoted as Crack_AP50, Joint_AP50, and Marking_AP50, respectively, and the corresponding AP is denoted as mAP_0.5. The variation curves of mAP_0.5 of the validation set with the number of epochs are shown in Fig. 5. The overall detection precision of model 6 is higher than those of other models, indicating that model 6 trained by the single-stage detection algorithm YOLOv5 can achieve a more accurate fitting effect at a higher speed.

The results of the evaluation indicators detected by each model on the test set are presented in Table 8. All eight models satisfactorily recognize joints with fixed shapes. However, the crack recognition performance is somewhat reduced compared with the recognition of joints owing to the cracks' complex shapes, wide variation range, random distribution, and irregular number of occurrences.

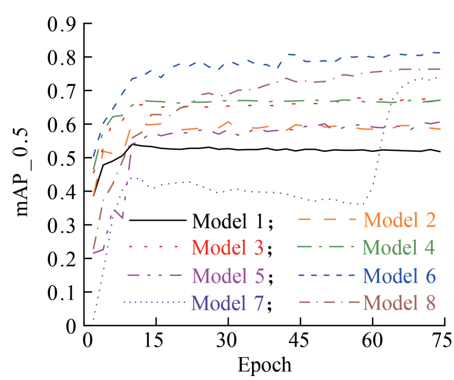


Fig. 5 Variation curve of mAP_0.5 in the validation set

Two possible reasons exist for the model's poor performance vis-à-vis road marking recognition. First, the dataset has fewer images of road markings, following which the model cannot completely learn the features of the markings. Second, the shape of the markings is not complete, and the characteristics are not obvious owing to road wear and other reasons.

Table 8 Model detection results

Model number	mAP _0.5	Crack _AP50	Joint _AP50	Marking _AP50	FPS/ (frame · s^{-1})	Model size/MB
1	0.549	0.696	0.795	0.154	18.5	315
2	0.605	0.705	0.770	0.335	15.5	460
3	0.699	0.743	0.843	0.512	28.6	527
4	0.717	0.742	0.825	0.584	22.6	672
5	0.612	0.670	0.854	0.313	52.7	26.8
6	0.838	0.806	0.959	0.748	133.3	13.7
7	0.761	0.768	0.940	0.576	43.7	102
8	0.761	0.782	0.889	0.612	122.0	21.5

According to the detection results of models 1-4, the deeper the backbone network, the higher the overall detection accuracy of the two-stage detection algorithm; this is because a deeper backbone network can assist the model in learning deeper features of the object more comprehensively. As can be seen from the results of FPS, the detection speed degrades as the network depth increases. This is because greater depth introduces more parameters, increases the complexity of the model, and results in a higher time cost for detection. The detection performance of Cascade R-CNN is better than that of Faster R-CNN, indicating that the use of cascade regression as a resampling mechanism can considerably increase the model's detection accuracy. The comparison results of models 5-8 demonstrate that YOLOv5 outperforms the other three single-stage detection algorithms in terms of detection accuracy, speed, and model lightweightness.

Comprehensive analysis shows that the models trained by the two-stage detection algorithm have high complexity, large consumption of computing resources, and obvious disadvantages in detection speed; thus, the models have difficulty in meeting the requirements of industrial production in terms of lightweightness and real-time performance. The single-stage detection algorithm, particularly YOLOv5, considerably reduces the model size and can obtain high detection accuracy while providing considerable advantages in terms of detection speed.

3.3 Comparative analysis of the image detection results

Fig. 6 directly compares the detection performances of the eight models, of which models 1-4 can correctly locate and classify most cracks. Overall, the models of the two-stage detection algorithms miss less, but the overlap of detection frames is serious, which confirms that this kind of algorithm has a poor ability to suppress the detection frame overlap in the post-processing stage. The redundant detection frame affects the statistics of the number of cracks, which is not conducive to the accurate determination of pavement conditions.

Models 5 and 7 have poor recognition performance, and there is a serious problem of missing recognition.

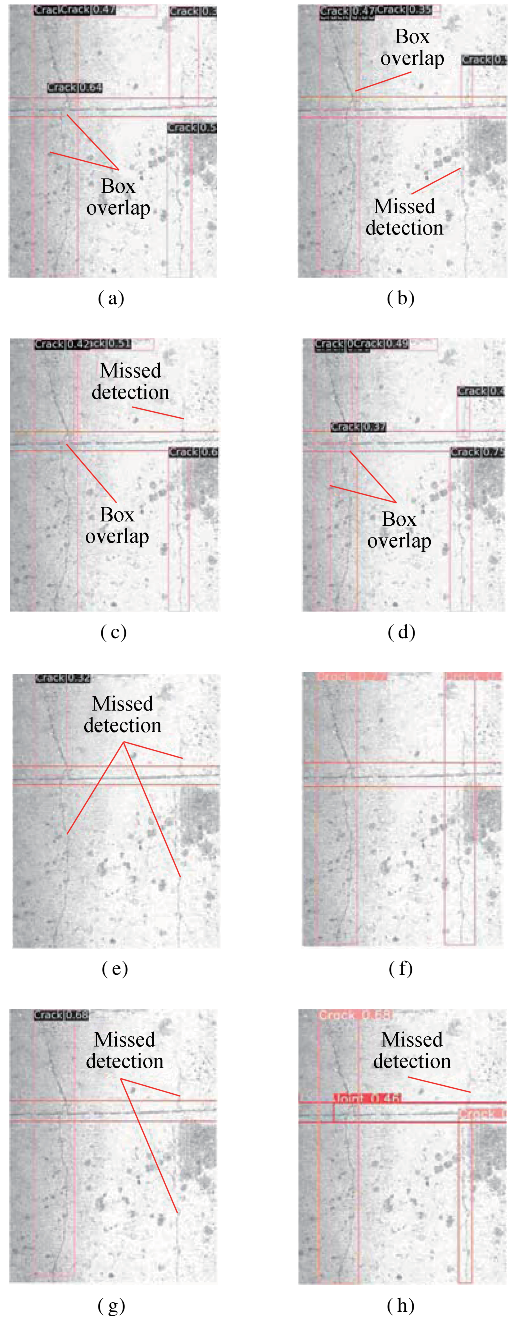


Fig. 6 Comparison of the image detection performances of the eight models. (a) Model 1; (b) Model 2; (c) Model 3; (d) Model 4; (e) Model 5; (f) Model 6; (g) Model 7; (h) Model 8

This shows that the two algorithms have weak learning ability, low detection accuracy, and poor generalization performance. Model 8 does not detect cracks with low resolution, and its anti-interference ability must be further improved. Model 6, trained by YOLOv5, delivers satisfactory results of comprehensive recognition, accurate positioning, and nonoverlapping detection frames, which confirms that the algorithm has strong capabilities in terms of feature extraction and post-processing of detection frames, and it also has the optimum detection performance among the eight models.

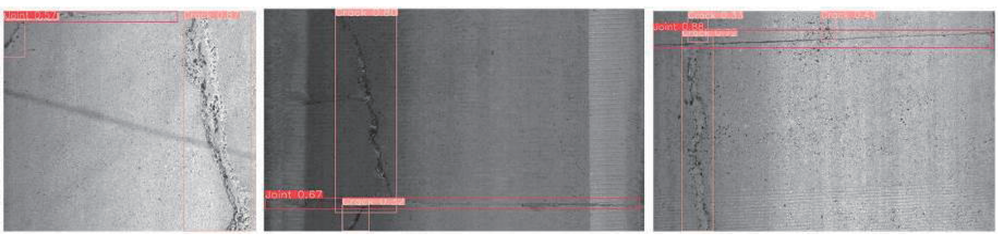
4 Improvement and Application of the Optimization Model

4.1 Improved YOLOv5 and analysis of its performance

In a bid to further increase the detection accuracy of the YOLOv5 algorithm for images featuring multiscale cracks, the feature fusion module in the algorithm framework was improved by adding ASFF[34]. Additionally, the regression loss function WIOUv3[35] was applied to replace the original loss function CIOU in the post-processing stage. The values of the hyperparameters remained unchanged before and after improvement.

Fig. 7 compares the detection performances of the model before and after adding ASFF for both long cracks (large target) and short cracks (small target) in an image. The comparison shows that in terms of ensuring the comprehensiveness of the original model in the detection of large targets, ASFF can more effectively fuse features of different levels. Following this, the feature map contains more effective information and can more efficaciously solve the problem of incomplete detection by YOLOv5 when an image contains multiscale cracks.

(a)



(b)

Fig. 7 Comparison of the detection performances. (a) Before adding ASFF; (b) After adding ASFF

WIOUv3[35] uses the gradient gain distribution strategy in the dynamic nonmonotonic focusing mechanism to reduce the interference caused by inaccurately labeled samples in the training samples. It focuses more on ordinary quality boxes to enhance the overall performance of the detector by balancing the learning of low- and high-quality samples.

After WIOUv3 replacement, regression loss decreased faster than CIOU, but the detection performance is not as good as that of CIOU, as shown in Fig. 8. From Table 9, one can see that WIOUv3 slightly increases the detection accuracy of the test set and increases the speed of model detection to a certain extent. The number of model parameters remains unchanged because no additional parameters are introduced. WIOUv3 delivers a considerable improvement effect only when the quality of the annotated samples is poor, and the improvement effect is not obvious here, which confirms that the large sample dataset constructed in this study has a high annotation quality.

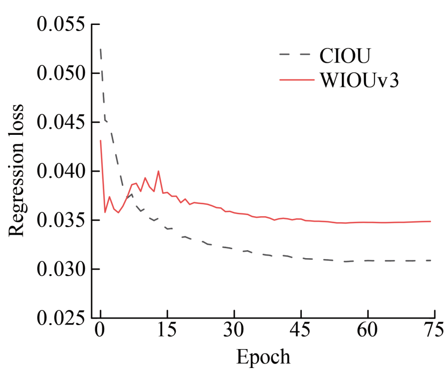


Fig. 8 Loss curve of bounding box regression

Table 9 Comparison of the detection results of various improvement measures

Improvement	mAP_0.5	Crack_AP50	Joint_AP50	Marking_AP50	FPS/(frame $\cdot s^{-1}$)	Parameter M
Original edition	0.838	0.806	0.959	0.748	133.3	7.018 216
WIOUv3	0.827	0.810	0.957	0.714	135.1	7.018 216
ASFF	0.845	0.805	0.957	0.771	103.1	12.457 665
ASFF + WIOUv3	0.846	0.802	0.954	0.783	102.0	12.457 665

After combining YOLOv5 with ASFF and WIOUv3, according to the detection performance of the test set in Table 9, the overall detection accuracy of the model is increased by 0.8%, and the target detection accuracy of marking is increased by 3.5%. Although the number of model parameters is increased, the detection speed is reduced; however, real-time detection requirements can still be ensured.

4.2 Application of intelligent models for classification and detection

Next, we intended to test the classification and detection performances of the ResNet50 image classification model trained on a large dataset and the improved YOLOv5 model, combined with ASFF and WIOUv3, on an actual complete rural highway. For this, a total of 500 continuous 1-km cement pavement images were selected from a highway in Nongxing Village, Harbin City, for application testing. The results of manual labeling were used as the reference standard for testing. The confidence threshold was set to 0.3, and thus, detection results with values less than 0.3 were considered unrecognized.

A total of 492 images with correct classification were obtained, and the accuracy was 98.4%. The accuracy rate of the improved detection model was 93.0%, and the image identification errors included 4 categories (see Fig. 9); the corresponding number is presented in Table 10. The correctly identified images included: 1) there are diseases in the image, and the model has comprehensively and accurately identified them; 2) there is no disease in the image, and the model does not misidentify it. It can be seen from the error types that the negative samples that cause interference with the detection model also include irregular objects such as stains, scratches, and branches. The model has a strong anti-interference ability to negative samples such as joints, road markings, and grooves. In summary, the detection performance of the model meets the requirement stated in the Highway Performance Assessment Standards (JTG 5210—2018)[5] that the identification accuracy rate of automatic detection should exceed 90%, indicating that the improved, optimal model can be used for intelligent crack detection in the cement pavements of rural highways.

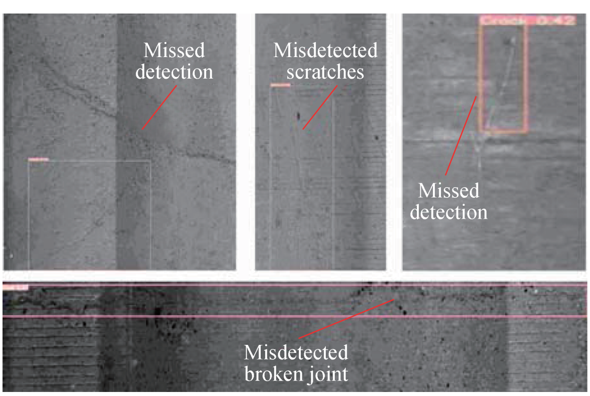


Fig. 9 Examples of misrecognition

Table 10 Type and number of instances of misrecognition

Type of error	Number
Identified road stains or scratches as cracks	14
Missed detection crack	11
Identified branches as cracks	5
Identified broken joints as cracks	5

5 Conclusions

1) In asphalt and cement pavement image classification, the incorporation of the mixed precision strategy can effectively increase the classification accuracy. ResNet50 delivers a better classification performance than ResNet18 and ResNet101.

2) The crack detection accuracy of the two-stage detection algorithms "Faster R-CNN" and "Cascade R-CNN" is less affected by the depth of the network. Although the crack detection is comprehensive, the redundant detection frames affect the crack number statistics. YOLOv3, YOLOX, and YOLOv8 have higher detection speeds than the two-stage algorithms, but they have poor feature learning ability, lack generalization ability, and have problems with crack detection. The selected YOLOv5 algorithm offers the combined advantages of the other five algorithms.

3) Upon the incorporation of ASFF and WIOUv3, the improved YOLOv5 detection model can more effectively address the problem of incomplete recognition of multi-scale cracks in images and enjoys an increase in overall detection accuracy.

4) The integrated ResNet50 classification model and the improved YOLOv5 intelligent inspection model achieved accuracies of 98.4% and 93.0%, respectively, in the field measurement, which can meet the relevant requirements stated in the Highway Performance Assessment Standards (JTG 5210—2018)[5], which necessitate that the accuracy of intelligent detection should be more than 90%.

References

[1] Ministry of Transport of the People's Republic of China. 2022 Transport Industry Development Statistical Bulletin [R]. Beijing: Ministry of Transport of the People's Republic of China, 2023. (in Chinese)

[2] Ma J, Zhao X M, He S H, et al. Review of pavement detection technology[J]. *Journal of Traffic and Transportation Engineering*, 2017, **17**(5): 121 – 137. DOI: 10.3969/j.issn.1671-1637.2017.05.012. (in Chinese)

[3] Zhang D. Common rural cement concrete pavement diseases and treatment[J]. *Shanxi Architecture*, 2014, **40**(25): 175 – 176. DOI: 10.13719/j.cnki.cn14-1279/tu.2014.25.092. (in Chinese)

[4] Zhang W G, Zhong J T, Yu J X, et al. Research on pavement crack detection technology based on convolution neural network[J]. *Journal of Central South University (Science and Technology)*, 2021, **52**(7): 2402 –

2415. DOI: 10.11817/j.issn.1672-7207.2021.07.026. (in Chinese)

[5] Ministry of Transport of the People's Republic of China. Highway performance assessment standards: JTG 5210—2018[S]. Beijing: People's Communications Publishing House, 2018. (in Chinese)

[6] Ministry of Transport of the People's Republic of China. Low-class rural highway performance assessment guide [R]. Beijing: Ministry of Transport of the People's Republic of China, 2023. (in Chinese)

[7] Pei M. On concrete pavement cracks and prevention measures[J]. *Shanxi Architecture*, 2010, **36**(7): 258, 286. DOI: 10.13719/j.cnki.cn14-1279/tu.2010.07.082. (in Chinese)

[8] Liu Y F, Fan J S, Nie J G, et al. Review and prospect of digital-image-based crack detection of structure surface [J]. *China Civil Engineering Journal*, 2021, **54**(6): 79−98. DOI: 10.15951/j.tmgcxb.2021.06.008. (in Chinese)

[9] Gou C, Peng B, Li T R, et al. Pavement crack detection based on the improved faster-RCNN[C]//2019 *IEEE 14th International Conference on Intelligent Systems and Knowledge Engineering (ISKE)*. Dalian, China, 2020: 962-967. DOI: 10.1109/ISKE47853.2019.9170456.

[10] Sun Z Y, Pei L L, Li W, et al. Pavement sealed crack detection method based on improved faster R-CNN[J]. *Journal of South China University of Technology (Natural Science Edition)*, 2020, **48**(2): 84−93. (in Chinese)

[11] Xia Y, Zhao M, Chen Y F, et al. Pavement crack disease identification method based on cascade RCNN: CN114170511A [P]. 2022-03-11. (in Chinese)

[12] Zhang R, Shi Y X, Yu X Z. Pavement crack detection based on deep learning[C]//2021 *33rd Chinese Control and Decision Conference (CCDC)*. Kunming, China, 2021: 7367−7372. DOI: 10.1109/CCDC52312.2021.9602216.

[13] Hegde V, Trivedi D, Alfarrarjeh A, et al. Yet another deep learning approach for road damage detection using ensemble learning[C]//2020 *IEEE International Conference on Big Data*. Atlanta, GA, USA, 2021: 5553−5558. DOI: 10.1109/BigData50022.2020.9377833.

[14] Sun Z Y, Ma Z D, Li W, et al. Pavement crack identification method based on deep convolutional neural network fusion model[J]. *Journal of Chang'an University (Natural Science Edition)*, 2020, **40**(4): 1−13. DOI: 10.19721/j.cnki.1671-8879.2020.04.001. (in Chinese)

[15] Guo W T. Intelligent detection device of pavement disease based on image recognition technology[J]. *Journal of Physics: Conference Series*, 2021, **1884**(1): 012032. DOI: 10.1088/1742-6596/1884/1/012032.

[16] Zhang S X, Zhang H C, Li X Z, et al. Study on multi-objective identification of pavement cracks based on machine vision[J]. *Journal of Highway and Transportation Research and Development*, 2021, **38**(3): 30−39. (in Chinese)

[17] Tran T S, Tran V P, Lee H J, et al. A two-step sequential automated crack detection and severity classification process for asphalt pavements[J]. *International Journal of Pavement Engineering*, 2022, **23**(6): 2019-2033. DOI: 10.1080/10298436.2020.1836561.

[18] Shi L, Pei L L, Chen H, et al. Detection of exposed distress of cement pavement based on improved RetinaNet [J]. *Computer Systems and Applications*, 2022, **31**(4): 352−359. DOI: 10.15888/j.cnki.csa.008436. (in Chinese)

[19] Mohan Prakash B, Sriharipriya K C. Enhanced pothole detection system using YOLOX algorithm[J]. *Autonomous Intelligent Systems*, 2022, **2**(1): 1−16. DOI: 10.1007/s43684-022-00037-z.

[20] Huang J, Zhang G. A survey of current research on image target detection algorithms based on deep convolutional neural networks [J]. *Digital Technology and Application*, 2020, **56**(17): 12−23. DOI: 10.3778/j.issn.1002-8331.2005-0021. (in Chinese)

[21] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 *IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, OH, USA, 2014: 580−587. DOI: 10.1109/CVPR.2014.81.

[22] Girshick R. Fast R-CNN[C]//2015 *IEEE International Conference on Computer Vision (ICCV)*. Santiago, Chile, 2016: 1440−1448. DOI: 10.1109/ICCV.2015.169.

[23] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **39**(6): 1137−1149. DOI: 10.1109/TPAMI.2016.2577031.

[24] He K M, Gkioxari G, Dollár P, et al. Mask R-CNN [C]//2017 *IEEE International Conference on Computer Vision (ICCV)*. Venice, Italy, 2017: 2980−2988. DOI: 10.1109/ICCV.2017.322.

[25] Cai Z W, Vasconcelos N. Cascade R-CNN: Delving into high quality object detection[C]//2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT, USA, 2018: 6154−6162. DOI: 10.1109/CVPR.2018.00644.

[26] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection [C]//2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV, USA, 2016: 779−788. DOI: 10.1109/CVPR.2016.91.

[27] Redmon J, Farhadi A. YOLO9000: Better, faster, stronger[C]//2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI, USA, 2017: 6517−6525. DOI: 10.1109/CVPR.2017.690.

[28] Redmon J, Farhadi A. Yolov3: An incremental improvement[EB/OL]. (2018-04-08) [2023-07-20]. https://arxiv.org/abs/1804.02767.pdf.

[29] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: Optimal speed and accuracy of object detection [EB/OL]. (2020-04-23) [2023-05-20]. https://arxiv.org/abs/2004.10934.pdf.
[30] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot MultiBox detector[C]//*European Conference on Computer Vision*. Cham: Springer, 2016: 21 − 37. 10.1007/978-3-319-46448-0_2.
[31] Xu D G, Wang L, Li F. Review of typical object detection algorithms for deep learning[J]. *Computer Engineering and Applications*, 2021, **57**(8): 10 − 25. (in Chinese)
[32] Ge Z, Liu S T, Wang F, et al. YOLOX: Exceeding YOLO series in 2021[EB/OL]. (2021-07-18)[2023-06-10]. https://arxiv.org/abs/2107.08430.pdf.
[33] Reis D, Kupec J, Hong J, et al. Real-time flying object detection with YOLOv8[EB/OL]. (2023-05-17)[2023-07-10]. https://arxiv.org/abs/2305.09972.pdf.
[34] Liu S T, Huang D, Wang Y H. Learning spatial fusion for single-shot object detection [EB/OL]. (2019-11-21) [2023-06-10]. https://arxiv.org/abs/1911.09516.pdf.
[35] Tong Z J, Chen Y H, Xu Z W, et al. Wise-IoU: Bounding box regression loss with dynamic focusing mechanism[EB/OL]. (2023-01-24) [2023-07-20]. https://arxiv.org/abs/2301.10051.pdf.

农村公路水泥路面裂缝智能检测

王　萌[1]　张小月[1]　刘　诚[2]　徐慧通[1]　杨燕泽[1]

([1]北京交通大学土木建筑工程学院, 北京 100044)
([2]中路高科交通检测检验认证有限公司, 北京 100088)

摘要:为解决传统人工图像处理方法在农村公路路面病害检测中存在的效率低、结果不客观、大量数据无法及时处理等问题,考虑农村公路路段分布特征,集成 ResNet50 路面分类和改进的 YOLOv5 裂缝检测算法,提出了一种农村水泥路面裂缝智能检测方法. 利用不同训练策略、不同网络深度进行对比,构建了基于 ResNet50 的路面高效分类模型,实现农村公路水泥和沥青路面的自动判别. 创建了包含 18 028 张农村公路水泥路面裂缝图片的检测数据集,开展单阶段和两阶段目标检测算法对比试验研究,获得兼顾检测精度和效率的优选检测算法. 在优选算法中融入自适应空间特征融合策略和优化回归损失函数,有效解决了图像中多尺度裂缝漏检问题,并进一步提高了整体检测精度. 应用所提集成方法对农村公路水泥路面进行现场实测,结果表明路面类型分类准确率为 98.4%,裂缝检测准确率为 93.0%,表明所提方法能够准确高效地运用于农村公路水泥路面裂缝检测.

关键词:农村公路; 水泥路面; 裂缝; 深度学习; 图像分类; 目标检测
中图分类号:U416.216