ISSN 2663-1563 Vol. 8, Issue 1: 128-133, DOI: 10.25236/AJAGE.2026.080116

Research on Road Crack Detection Technology Based on Improved YOLOv8n

Guizhuo Ai, Lan Ma*

School of Civil and Transportation Engineering, Qinghai Minzu University, Xining, China
**Corresponding Author*

Abstract: *The detection of road and cracks is crucial for the safe operation and maintenance of infrastructure, but in practical scenarios, complex backgrounds and multi-scale morphological changes lead to insufficient detection accuracy of YOLOv8n baseline model. In response to this issue, this article introduces the HPA (Hybrid Pooling Attention) module in the YOLOv8n Neck stage, which combines average pooling and max pooling with cross space learning to enhance the transmission and fusion of multi-scale crack features and alleviate information attenuation in feature propagation. The improved model significantly improves the detection rate and positioning accuracy of small cracks in complex backgrounds while maintaining real-time performance, providing an efficient solution for intelligent detection of road and cracks.*

Keywords: *Road and crack detection; YOLOv8n; Hybrid Pooling Attention (HPA); Complex background; Multi scale feature fusion; Lightweight detection*

1. Introduction

Road and cracks are the main early signs of pavement structure degradation[1], and their accurate detection is crucial for ensuring the safe operation of transportation infrastructure. The traditional manual inspection method is inefficient and subjective, making it difficult to meet the needs of large-scale road network maintenance; However, methods such as threshold segmentation and morphological analysis based on image processing have insufficient adaptability to environmental changes and crack morphology diversity [2]. In recent years, deep learning object detection algorithms represented by the YOLO series have shown great potential in the field of automatic crack detection by leveraging end-to-end structures and efficient reasoning capabilities. However, existing methods still face problems such as insufficient feature discrimination and high missed detection rates for small targets when facing complex background textures, uneven lighting, and multi-scale crack targets. In response to the above challenges, this paper takes YOLOv8n as the baseline model with a balanced detection accuracy and inference speed, introduces HPA (Hybrid Pooling Attention) hybrid pooling attention module in the Neck stage, and uses a dual parallel structure of average pooling and maximum pooling, combined with cross space learning to achieve feature interaction fusion, strengthen the transmission and retention ability of multi-scale crack features, effectively alleviate information attenuation in feature propagation, and improve the recognition accuracy and localization robustness of the model for small cracks in complex environments.

2. Theoretical Analysis of Deep Learning

The One Stage object detection algorithm in the field of computer vision is favored by people for its fast detection speed and good detection accuracy. The YOLOv8 algorithm is an excellent version of this technology, which has made multiple adjustments to the network structure. The C2f module is used to reconstruct the backbone network in the network structure, and the previous anchor box mechanism is eliminated. A combination of anchor box free strategy and decoupled detection head is adopted [3]. The direct result of these adjustments is a significant improvement in feature extraction efficiency, coupled with the assistance of multi-scale feature fusion mechanisms, making the model still highly sensitive to diseases with large scale differences (micro cracks, large-scale cracks). In addition, after using dynamic labels and an improved loss function, YOLOv8 has strong robustness to complex and changing engineering backgrounds, and can meet the dual requirements of timeliness and accuracy in detecting cable-stayed diseases.

3. Data Augmentation and Apparent Disease Analysis

3.1 Research on Integrating Manually Labeled Data

This article uses supervised learning to train an improved YOLOv8 model, which automatically learns the feature distribution patterns and semantic information of diseases from a large number of learning samples. The initial dataset consists of 1406 high-quality surface images of road cracks[4] under varying lighting conditions, shooting angles, and complex backgrounds, covering six typical types of images: Alligator cracks, Longitudinal cracks, Oblique cracks, Pothole, Repair, and Transverse cracks. In order to improve the generalization boundary and sample size of the model, data augmentation techniques such as rotation, mirror flipping, and color jitter were used to increase the original sample size to 2812 images. During the annotation phase, researchers employ the LabelImg tool to perform fine-grained manual bounding-box annotation (as is shown in Figure 1), and store the annotated results in XML format for subsequent secondary verification and category revision. The input end of the model uses preprocessing scripts to batch convert XML files into normalized TXT format supported by YOLOv8. From high-quality manual annotation to standardized data input, it ensures the accuracy and diversity of training data, enabling the model to quickly identify the characteristics of various diseases, greatly improving the efficiency and accuracy of engineering detection.

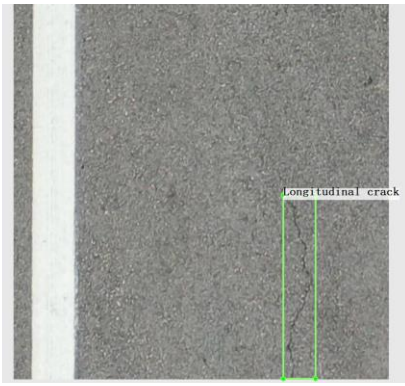

Figure 1. Example Demonstration of Disease Feature Annotation Files

3.2 Analysis of Typical Crack Diseases in Roads

 *Figure 2. Alligator crack*	 *Figure 3. Longitudinal crack*	 *Figure 4. Oblique crack*
 *Figure 5. Pothole*	 *Figure 6. Repair*	 *Figure 7. Transverse crack*

The accurate identification of road and cracks first requires a systematic analysis of their causes and morphological characteristics. Different types of cracks reflect different stress states and degrees of damage, and their morphological differences directly affect the feature extraction strategy of detection algorithms. Based on relevant literature, this article selects six typical apparent diseases for the study of common crack diseases on roads and s, namely: Alligator crack, Longitudinal crack, Oblique crack,

ISSN 2663-1563 Vol. 8, Issue 1: 128-133, DOI: 10.25236/AJAGE.2026.080116

Pothole, Repair, and Transverse crack. The specific morphological characteristics are shown in the Figures 2-7.

4. YOLOv8 Realizes Target Detection of Road Crack Diseases

4.1 Improving the YOLOv8 Algorithm

To address the issues of severe interference from complex backgrounds and low detection rates of multi-scale crack targets (especially small cracks) in road and crack detection[5], this paper proposes a crack detection method based on improved YOLOv8n. The innovation of this method lies in the introduction of the Hybrid Pooling Attention Module (HPA), which draws on its cross space learning mechanism to capture global statistical features and local saliency responses through a dual parallel structure of average pooling and maximum pooling, and integrates them into the Neck stage of YOLOv8n to achieve deep interaction and enhancement of local texture details and global semantic information. The experimental results show that the improved algorithm significantly improves the detection accuracy and localization robustness of the model for multi-scale cracks in complex environments while maintaining a high detection speed. Especially, the recognition ability for small cracks is significantly enhanced, and the detection is more comprehensive and accurate.

4.2 Algorithm Implementation Based on the Improved YOLOv8

The hybrid pooling attention module (HPA) structure proposed in this article is shown in Figure 8. This module uses a parallel dual stream structure, which first standardizes the input features, then uses mixed pooling to capture local fine-grained texture and global semantic information, followed by adaptive weighted fusion of multi-scale features using channel attention mechanism, and finally adjusts the channel dimension using convolutional layers, and establishes residual connections with the original input. Without changing the spatial resolution of the feature map, the fast interaction between local details and global dependencies is achieved, improving the model's ability to perceive subtle cracks on the surface of road surface and complex environmental disturbances.

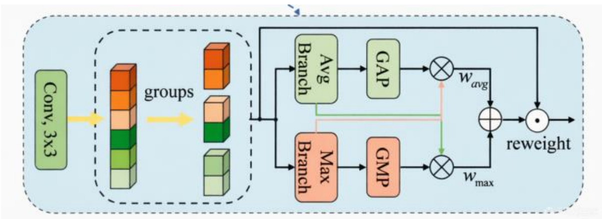

Figure 8. Schematic diagram of the internal structure of HPA module

4.3 Model Training Results

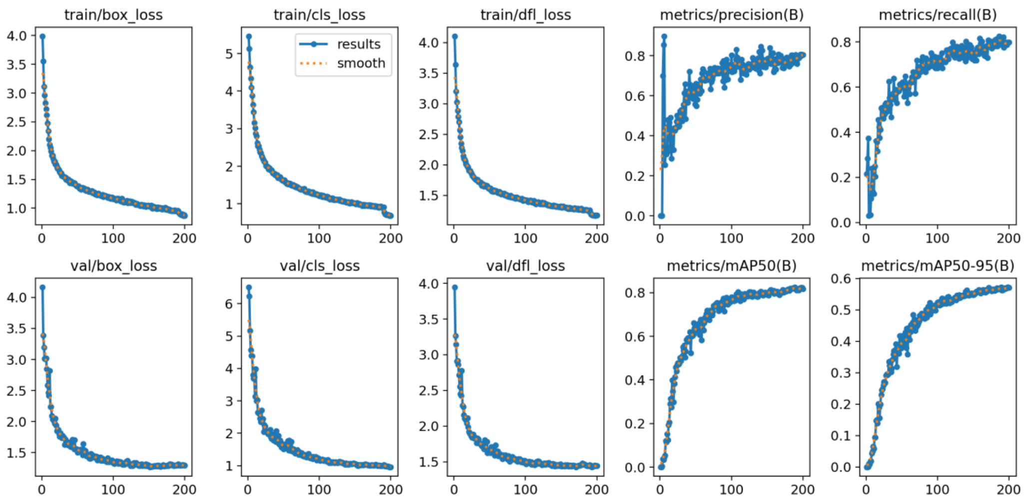

Figure 9. Changes in the loss function and accuracy during each cycle of model training

ISSN 2663-1563 Vol. 8, Issue 1: 128-133, DOI: 10.25236/AJAGE.2026.080116

The loss function and accuracy changes of each cycle of model training are shown in Figure 9. As the number of training iterations increases, the loss function value of the model continuously decreases, and the accuracy of the model prediction also continuously improves.

The precision recall curve (P-R curve) of the training results is shown in Figure 10, and there are certain differences in the detection performance of various types of crack diseases. Among them, Pothole has the highest AP value, reaching 0.993, followed by Repair at 0.955, Oblique crack at 0.836, Transverse crack at 0.825, Alligator crack at 0.726, and Longitudinal crack at 0.612. The mean average precision at IoU 0.5 (mAP@0.5) for all six categories reaches 0.824. The results show that the model has excellent detection performance for targets with distinct morphological features and clear edges (such as potholes and repair areas), while its performance in detecting longitudinal cracks with strong texture continuity and low discrimination from the background is relatively weak. This is mainly attributed to the slender and easily confused longitudinal cracks with road texture, which provides direction for subsequent optimization for this category.

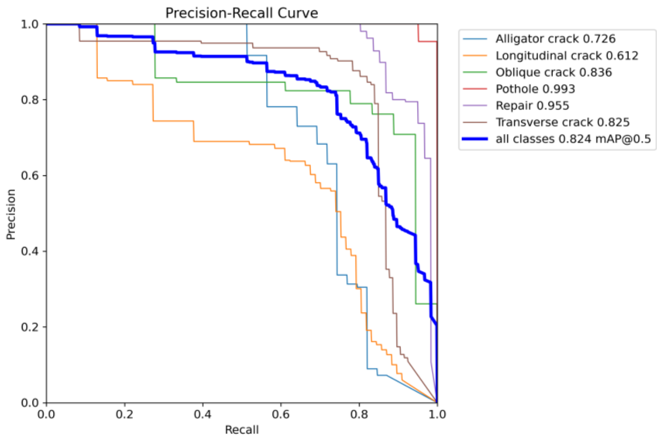

Figure 10. Disease Identification Accuracy and Recall Curve (P-R Curve)

4.4 Model Comparison

To verify the superiority of the method in this paper, mAP is used as the main evaluation metric. Since the task is a binary classification, the results are defined through the confusion matrix, as is shown in Table 1:

Table 1. Algorithm Evaluation Binary Classification Confusion Matrix

Binary label	Algorithm true value	
	Negative	Positive
Negative	TN	FN
Positive	TP	FP

The key indicators are defined as follows:

$$\mathrm{Precision(P)} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \tag{1}$$

$$\mathrm{Recall\,(R)} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \tag{2}$$

$$AP = \frac{\sum_{i=1}^{N} P_i}{N} \tag{3}$$

ISSN 2663-1563 Vol. 8, Issue 1: 128-133, DOI: 10.25236/AJAGE.2026.080116

$$\mathrm{mAP} = \frac{\sum_{i=1}^{N} P_i}{M} \tag{4}$$

Further distinction:

$$\mathrm{mAP@0.5} = \frac{\sum_{i=1}^{M} \mathrm{mAP@0.5}_i}{M} \tag{5}$$

$$\mathrm{mAP@}\,0.5:0.95 = \frac{\sum_{i=0}^{9} \mathrm{mAP@(0.5+0.05i)}}{10} \tag{6}$$

In the experiment, YOLOv5, YOLOv8, and the improved YOLOv8 (embedded with HPA module) in this paper were compared. The results showed that the improved model significantly outperformed the baseline in recognition rate and accuracy, especially in detecting weak cracks and low contrast diseases, which verified the effectiveness of the frequency-domain spatial dual domain attention mechanism, as is shown in Table 2.

Table 2.Compare and verify the recognition effects of different detection algorithms

Comparison algorithm	AP/%						mAP@ 0.5:0.95 /%	FPS /Hz
	Alligator crack	Longitudinal crack	Oblique crack	Pothole	Repair	Transverse crack		
YOLOv5	69.2	60.7	74.9	98.0	91.2	83.5	79.5	135.1
YOLOv8	70.4	61.8	76.1	99.3	92.5	84.6	80.8	108.7
YOLOv8—HPA	72.6	61.2	83.6	99.3	95.5	82.5	82.4	104.3

5. Conclusion

In response to the low efficiency of manual inspection, low recognition accuracy caused by complex background interference, and inaccurate multi-scale crack target detection in road crack detection, this paper conducts in-depth research on the application of deep learning object detection algorithms in road crack recognition. This article constructs a road crack dataset containing six typical types of crack diseases and proposes a crack detection and recognition method based on improved YOLOv8n. The main conclusions drawn from the study are as follows:

(1) There has been a significant improvement in performance. In response to issues such as slender road crack targets, large scale differences, and complex lighting background interference, this paper uses a hybrid pooling attention module (HPA) to improve the YOLOv8n benchmark model. The improved model has significantly improved the feature extraction ability for multi-scale crack targets (especially small cracks) while ensuring normal inference speed. The average accuracy mean (mAP) and recall rate (Recall) are better than the original benchmark detection algorithm.

(2) This article establishes an identification system for six typical types of road crack diseases, covering different morphological features such as crocodile pattern cracks, longitudinal cracks, oblique cracks, potholes, repair areas, and transverse cracks. After training, validation, and learning with large-scale annotated data, the model can accurately distinguish different types of crack diseases, indicating that this method has high accuracy and good generalization ability for complex and diverse crack classification tasks.

(3) Effective data augmentation methods can increase the sample size of the training set, simulate differences in imaging conditions such as lighting conditions, shooting angles, and road textures. Multiple methods have been adopted to improve the adaptability of the model to crack detection in complex environments, and stable detection results can be obtained in practical engineering applications.

References

[1] Deng Shengbin, Ye Lijia, Xu Na, et al. Research on Road Crack Identification Method Based on

ISSN 2663-1563 Vol. 8, Issue 1: 128-133, DOI: 10.25236/AJAGE.2026.080116

Lightweight Algorithm [J]. Highway Engineering, 2025, 50 (05): 150-159.
[2] Xu Zhenkuan, Gao Hang, Li Xuelian, et al. Study on crack propagation behavior of composite base asphalt pavement under dynamic load [J]. Highway Engineering, 2024, 49 (05): 86-92.
[3] Chen Hui, Fu Tao, Yu Yongjie, Zhang Shunxiang. Steel surface defect detection based on HMF-YOLO model [J/OL]. Optoelectronics and Laser, 2023,(05),1-13.
[4] Fu Yong. Research on Road Crack Detection Based on YOLOv5s Surveying Data Analysis [J]. China New Technology and New Products, 2026, (9): 110-112.
[5] Yang Qing, Liang Yi, Tao Jueqiang. Lightweight pavement disease detection method based on wavelet transform and deformable attention [J]. Journal of Zhejiang Normal University (Natural Science Edition), 2026, 49 (2): 141-151.