Research on Road Defect Detection Method Based on Adaptive Perception Module Improving YOLOv8

Jianze Liu
School of Mathematics and Statistics, Beihua University, Jilin, Jilin, China
**Corresponding Author.*

Abstract: Road defect detection is an important part of road maintenance management. In actual road images, defect targets such as cracks and potholes often have problems such as weak edges, unclear textures, and irregular shapes. At the same time, they are easily interfered with by background factors such as shadows, markings, stains, and repair traces, which makes the detection model prone to missing detections and false detections. To address these issues, this paper proposes a YOLOv8 road defect detection method that integrates an adaptive perception module.

This method is based on the YOLOv8x detection framework and introduces an Adaptive Perception Module (APM) into the network. Through convolutional feature enhancement, residual connections, and attention weighting mechanisms, it improves the model's perception ability of the edge, texture, and local structure information of the diseased area. To analyse the impact of the embedding position of APM on the detection performance, this paper embeds APM at the Backbone P4, Backbone P3, Neck P4, and Neck P3 positions respectively, and conducts comparative experiments with the baseline model that does not incorporate APM.

The experimental results show that the detection gain of the APM module exhibits significant location dependence. Among them, the APM-Neck-P4 model achieves the best overall performance, with precision, recall, F1, mAP@0.5, and mAP@0.5:0.95 being 0.5760, 0.4658, 0.5151, 0.4836, and 0.2416, respectively. Compared with the baseline, the mAP@0.5 of APM-Neck-P4 increases by 0.0316, and the mAP@0.5:0.95 increases by 0.0271, indicating that introducing APM at the Neck P4 position can effectively enhance the multi-scale fusion feature representation and improve the accuracy of road defect detection. The research results can provide a reference for the structural design and embedding position selection of the feature enhancement module in road defect detection.

Keywords: Road defect Detection; YOLOv8; Adaptive Perception Module; Feature Enhancement; Object Detection

1. Introduction

Road defects are significant factors affecting the performance of road usage and traffic safety. If cracks, potholes, rutting and looseness and other defects are not detected and dealt with in time, they may further expand the damage range, increase maintenance costs, and affect the safety of vehicle passage. Therefore, researching efficient and accurate automatic detection methods for road defects is of practical significance for improving the intelligence level of road maintenance.

Early detection of road defects mainly relies on manual inspection and traditional image processing methods. Manual inspection is intuitive, but it is inefficient, and the detection results are easily influenced by the experience of the personnel; traditional image processing methods usually use methods such as grey difference, edge detection, threshold segmentation and texture analysis to identify the defect areas. They have certain effects in simple scenarios, but when facing changes in lighting, shadow occlusion, marking interference and road noise, the detection stability is insufficient.

In recent years, deep learning methods have achieved good results in the target detection task and have also been widely applied to road defect identification. Among them, the YOLO series models have the characteristics of fast detection speed, compact structure and convenient deployment and are suitable for real-time road inspection scenarios. YOLOv8 has good performance in feature extraction and multi-scale detection but still faces certain challenges in road defect detection. Due to the weak edges, unclear textures of defect targets such as cracks, and the visual similarity between potholes, repair areas and shadows, the model is prone to missing detections and false detections.

To address the above problems, this paper introduces an adaptive perception module (APM) into the YOLOv8 model. By enhancing the edges, textures and local structural information of the defect area, the model's ability to express key features in complex road images is improved. At the same time, considering that the embedding position of the module will affect its effect, this paper designs control experiments with different embedding positions to compare the impact of introducing APM in the backbone network and feature fusion stage on detection performance, providing a reference for the structural optimisation of road defect detection models.

2. Research Background and Significance

Road defect detection is an important part of road maintenance management. Yi Mingwei et al. summarised the main types of road defects and detection technologies at home and abroad and pointed out that cracks, potholes, and rutting are common road defects. The detection methods mainly include manual detection and automated detection [1]. With the development of image acquisition equipment, unmanned aircraft, and intelligent maintenance systems, road defect detection is gradually shifting towards automation and intelligence.

In recent years, deep learning methods have been widely applied in road defect detection. Compared with traditional image processing methods, deep learning models can automatically learn multi-level features in images and have stronger adaptability to complex scenes. Sun Lishuang et al. improved the YOLOX network by introducing the CFPN feature fusion structure, ECA attention mechanism, and WIoU loss function and improved the detection effect of defect targets such as cracks and potholes [2]. Jiang Yiwen et al. combined YOLOv8 with a deformable attention transformer structure and achieved good performance in precision rate, recall rate, and mAP indicators, indicating that the YOLO series models combined with attention mechanisms can effectively improve the performance of road defect recognition [3].

In addition to model structure optimisation, there are also studies that enhance detection performance from the perspective of data augmentation. Maeda et al. utilised a generative adversarial network to generate images of road damage and used them as supplementary samples for model training, thereby improving the model's generalisation ability [4]. Feng Xiaoshuo reviewed the development of image data augmentation methods and pointed out that data augmentation can to some extent alleviate the problem of insufficient samples and improve the model's generalisation ability [5]. However, the generated samples may still have differences from the real scene, and when the similarity between the generated image and the original image is high, the improvement in model performance is limited.

In terms of feature enhancement, Li Ao et al. proposed a joint learning framework based on an adaptive perception module. By using large convolution kernels, residual connections, and the CBAM attention mechanism, they enhanced the feature representation of the diseased areas and suppressed irrelevant background interference [6]. This study shows that APM can effectively improve the model's perception ability for diseased areas such as cracks, potholes, and rutting.

Overall, existing studies have promoted the development of road defect detection from aspects such as network structure, feature fusion, attention mechanism, and data augmentation [2-6]. However, the analysis of the differences in the role of the feature enhancement module at different network positions is still insufficient. To address the issues of weak target edge, unclear texture, and strong background interference of road defects, this paper introduces APM into YOLOv8 and conducts experiments at different embedding positions to analyse its impact on detection performance.。

3. Research Methods

This paper addresses the issues of unclear disease area characteristics and strong background interference in road defect images, and designs a YOLOv8 detection method that integrates an adaptive perception module. The research idea is to maintain the basic framework of YOLOv8 detection while introducing APM to enhance the feature extraction and feature fusion process, and analyze the impact of the module on detection performance through experiments at different embedding positions.

The overall research process of this paper includes dataset processing, basic model construction, introduction of the APM module, design of different embedding positions, model training and performance evaluation.

3.1 Data Preparation and Preprocessing

This paper uses a dataset of road defect images as the experimental object. The dataset contains typical road defect targets such as cracks and potholes. To ensure the quality of model training, the original images are first screened to eliminate samples that are blurry, repetitive in content, have missing targets, or have incorrect annotations, thereby reducing the influence of low-quality data on the model training process.

Subsequently, the data annotation files were checked to ensure that the disease category labels were consistent with the positions of the target bounding boxes. After the data cleaning was completed, the dataset was divided into a training set, a validation set and a test set, and the input image size was uniformly standardised to meet the training requirements of the YOLOv8 model.

Considering that the main focus of this study is to compare the impact of different APM embedding positions on model performance, all groups of experiments adopted the same data division and preprocessing methods. This paper did not introduce a class balance strategy to avoid the influence of additional processing methods on the comparison between different models.

The dataset of this paper includes five types of road defect targets: alligator crack, longitudinal crack, other corruption, pothole, and transverse crack. From the distribution of data labels, it can be seen that the number of longitudinal crack samples is the largest, while the numbers of pothole and other corruption samples are relatively smaller. There is a certain imbalance in the category distribution. All the experiments adopt the same data division method to ensure the comparability of the results among different models.

3.2 YOLOv8 Base Model

YOLOv8 is a single-stage object detection model, consisting of three main components: backbone, neck, and head. The backbone is responsible for extracting multi-level features from the image, the neck is used to fuse feature information of different scales, and the head outputs the target category, bounding box position, and confidence. YOLOv8 features a fast detection speed, flexible structure, and convenient deployment, making it suitable for tasks with high real-time requirements, such as road defect detection.

In the road defect detection scenario, YOLOv8 can perform multi-scale target detection quite well. However, when dealing with weak crack edges, complex pothole textures, as well as background interference such as shadows and markings, the model may still have the problem of insufficient feature representation. Especially for crack-related targets, which are usually narrow in width and have little difference from the normal road texture, they are prone to be weakened during feature extraction. Therefore, in YOLOv8, this paper introduces the APM module to enhance the model's ability to extract key features of the defect area.

3.3 Adaptive Perception Module (APM)

The logical structure is shown in Figure 1, APM is a module designed to enhance the expression of local features, and it is mainly composed of three parts: convolutional feature extraction, residual connection, and attention weighting.

In APM, the convolution feature extraction part is used to enhance the edge, texture and local structural information of the damaged areas. Considering that the shapes of road damage targets are irregular, and the differences between cracks, potholes and other targets and the background are sometimes weak, the convolution structure can improve the model's ability to model local details to a certain extent.

Residual connections are used to retain the input features and reduce the loss of effective information during the feature transformation process. When the dimensions of the input features and the output features are inconsistent, a 1×1 convolution can be used for channel adjustment to enable the two sets of features to complete fusion. The attention-weighting part is used to redistribute the feature weights from both the channel and spatial dimensions, enabling the model to pay more attention to the regions related to the disease target and reducing the influence of irrelevant background information on the detection results.

Through the above structure, APM can, while maintaining the original feature information, further enhance the local response of the damaged area, providing more effective feature representation for target detection in complex road scenarios.

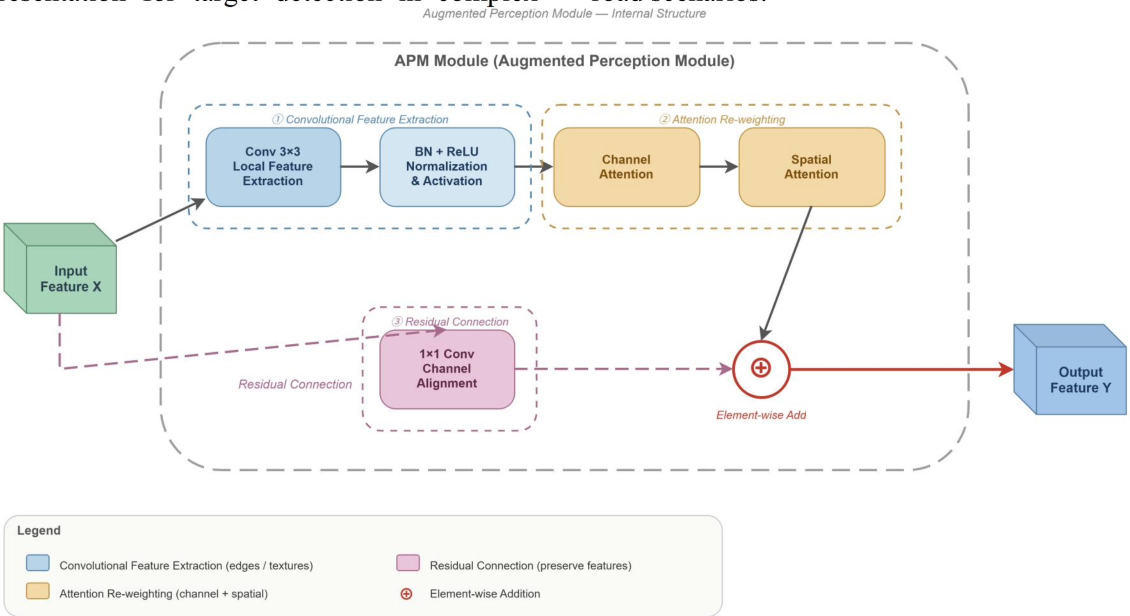

Figure 1. Overall Architecture of APM

3.4 APM-YOLOv8 Fusion Model

This paper introduces APM into the YOLOv8 model to construct the APM-YOLOv8 road defect detection model. The overall structure of the model still retains the basic framework of YOLOv8's backbone, neck and head, but the APM module is embedded at different network positions to enhance the feature representation at corresponding levels.

The Backbone stage is mainly responsible for extracting the basic and semantic features of the image. Among them, the shallow features usually contain more edge, texture and spatial position information, while the mid-high level features have stronger semantic expression capabilities. The neck stage is mainly responsible for the fusion of features of different scales and plays an important role in the detection of small targets and targets with large-scale variations.

Due to the fact that different network layers contain different information, the embedding position of APM may directly affect its feature enhancement effect. Therefore, this paper does not merely increase the number of modules but conducts a comparative experiment around the adaptability of APM at different positions to analyse the differences in its role in the road defect detection task.

3.5 Experimental Design for Different Embedding Positions

To investigate the impact of APM embedding position on the performance of YOLOv8 in road defect detection, this paper designs a set of baseline models and four sets of APM-YOLOv8 improved models. All five sets of models use the same dataset division, input size, training epochs, optimiser, and data augmentation strategy, only changing the embedding position of APM in the network. By keeping other variables constant, it is possible to objectively analyse the adaptability of different network layers to the APM module. The five experimental schemes set by this paper are shown in Table 1.

Table 1 presents the five APM embedding schemes set up in this paper. Different schemes only change the position of APM, while the rest of the network structure and training conditions remain the same. This design helps to observe the differences in the roles of APM in shallow texture features, mid-to-high-level semantic features, and multi-scale fusion features.

The schematic diagram of the four embedding positions is shown in Figure 2.

The APM at the shallow layer focuses more on enhancing the crack edges, texture details and local structural information; the APM at the medium and high layers focuses more on enhancing the semantic expression of the defect targets; the APM at the neck stage mainly acts on multi-scale fusion features, which helps to improve the detection effect of defect targets at different scales. By comparing the detection indicators of the five groups of models, the differences in the role of APM at different network positions can be analysed.

Table 1. Experimental Design Scheme for APM Embedding Position

Serial Number	Experimental group	Model Name	The embedding position of APM	Experimental objective
1	exp1_baseline	Baseline	No APM	As the base comparison model
2	exp2_apm_backbone_p4	APM-Backbone-P4	Backbone P4	Analysis of the enhancement effect of APM on the features of Backbone P4
3	exp3_apm_backbone_p3	APM-Backbone-P3	Backbone P3	Analysis of the enhancement effect of APM on the features of Backbone P3
4	exp4_apm_neck_p4	APM-Neck-P4	Neck P4	Analysis of the enhancement effect of APM on the fused features of Neck P4
5	exp5_apm_neck_p3	APM-Neck-P3	Neck P3	Analysis of the impact of APM on the fusion features of Neck P3

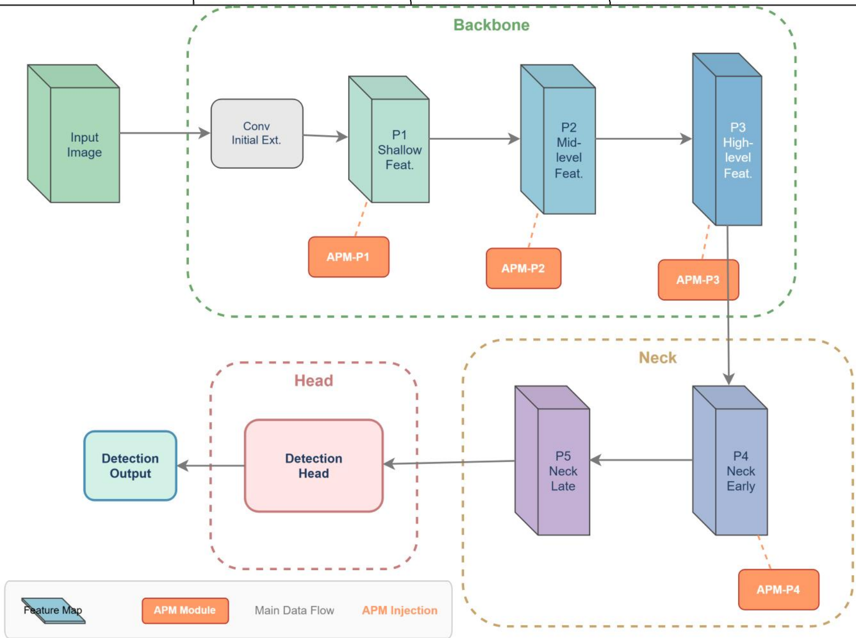

Figure 2. Design of APM Embedding Position

4. Experimental Setup

To ensure the comparability of the experimental results, all the models in this paper adopt the same data set division, input image size, number of training rounds, optimiser settings and data augmentation strategies. Each group of experiments only changes the embedding position of APM, while keeping all other conditions consistent.

4.1 Experimental Environment and Parameter Settings

All models were trained for 80 rounds. The input image size was uniformly set to 640×640, and the batch size was set to 8. During the training process, an automatic optimiser configuration was adopted, and AMP mixed-precision training was enabled to improve training efficiency and reduce memory consumption. In each group of experiments, a random seed of 0 was set, and the deterministic training mode was enabled to enhance the reproducibility of the experimental results. In terms of data augmentation, strategies such as Mosaic, horizontal flipping, HSV colour perturbation, and RandAugment were used during the training stage. Among them, Mosaic was turned off in the last 10 rounds to help the model better adapt to the real image distribution in the later training stage.

To explore the role of the APM module in different network layers, this study embedded it into the backbone and neck parts of YOLOv8x and selected two typical feature layers, P3 and P4, for comparative experiments. Through the above grouping, the effectiveness of the APM module can be analysed from two perspectives:

Backbone hierarchical comparison: Analyze the enhancement effect of APM on the feature extraction stage;

Neck level comparison: Analyze the impact of APM on the multi-scale feature fusion stage;

Comparison between P3 and P4: Analyze the sensitivity of shallow detail features and mid-to-high-level semantic features to APM.

4.2 Evaluation Indicators

This paper uses precision, recall, F1, mAP@0.5 and mAP@0.5:0.95 as the evaluation indicators for the model performance in the target detection task and combines PR curves, F1-confidence curves and confusion matrices to conduct a comprehensive analysis of the model's detection effect. Precision represents the proportion of the targets predicted as positive by the model that are actually positive and is used to measure the model's ability to control false detections. Its calculation formula is:

$$Precision=\frac{TP}{TP+FP} \quad (1)$$

"Recall" represents the proportion of all actual targets that are correctly detected by the model, and is used to measure the model's ability to control false negatives. Its calculation formula is:

$$Recall=\frac{TP}{TP+FN} \quad (2)$$

Among them, TP represents the number of correctly detected targets, FP represents the number of samples that were wrongly identified as targets, and FN represents the number of targets that truly existed but were not detected. The $F1$ value takes into account precision and recall and can reflect the balance relationship between false detection and missed detection of the model. Its calculation formula is:

$$F1=\frac{2\times Precision\times Recall}{Precision+Recall} \quad (3)$$

mAP@0.5 represents the average value of AP for all categories when the IoU threshold is 0.5 and is used to measure the model's comprehensive detection capability under relatively relaxed positioning conditions. mAP@0.5:0.95 indicates the average value of mAP calculated when the IoU threshold ranges from 0.5 to 0.95 with a step size of 0.05. Compared to mAP@0.5, it has higher requirements for the accuracy of target positioning and can more comprehensively reflect the positioning quality of the model's detection boxes.

5. Experimental Results and Analysis

To verify the effectiveness of the APM module in the road defect detection task, this paper uses YOLOv8x as the baseline and conducts comparative experiments by embedding the APM module at four positions: Backbone P4, Backbone P3, Neck P4, and Neck P3. The experimental results show that the APM module can improve the model's detection performance at the appropriate positions, but its gain is not consistently present at all levels; instead, it exhibits a significant position dependence.

5.1 Overall Performance Comparison

The overall detection results of the five groups of models are shown in Table 2 and Figure 3. It can be seen that APM-Neck-P4 achieves the best results in four indicators: precision, F1, mAP@0.5, and mAP@0.5:0.95. Among them, mAP@0.5 reaches 0.4836, which is an increase of 0.0316 compared to 0.4520 of the baseline; mAP@0.5:0.95 reaches 0.2416, which is an increase of 0.0271 compared to the baseline. This result indicates that introducing APM at the Neck P4 position can effectively enhance the comprehensive detection performance of the model.

From the overall performance, it can be seen that APM-Neck-P4 has the best comprehensive performance, achieving the highest values in precision, F1, mAP@0.5, and mAP@0.5:0.95; APM-Backbone-P4 has the highest recall, indicating that it is more inclined to detect more targets, but its precision has significantly decreased; APM-Backbone-P3 performs relatively evenly, maintaining a high precision while improving recall and mAP; APM-Neck-P3 fails to surpass the baseline, and its recall, F1, and mAP@0.5 are all slightly lower than those of the base model.

This indicates that the performance gain of the APM module is closely related to the embedding position. Simply adding the module does not necessarily improve the detection performance; only when the module is embedded at the appropriate feature level can it effectively enhance the model's ability to express the target.

Table 2. Comparison of Overall Detection Performance of Different Models

Serial Number	Model Name	Embedding position	Precision	Recall	F1	mAP@0.5	mAP@0.5:0.95
1	Baseline	NO·APM	0.5547	0.4502	0.4970	0.4520	0.2145
2	APM-Backbon e-P4	Backbone·P4	0.5177	0.4813	0.4988	0.4655	0.2290
3	APM-Backbon e-P3	Backbone·P3	0.5507	0.4752	0.5102	0.4801	0.2380
4	APM-Neck-P4	Neck·P4	0.5760	0.4658	0.5151	0.4836	0.2416
5	APM-Neck-P3	Neck P3	0.5547	0.4442	0.4932	0.4490	0.2201

Figure 3. Bar Chart Showing the Overall Performance of Different Models

5.2 Abandonment Analysis of Different Embedding Positions

To further analyse the changes of each improved model compared to the base model, the Table 3 presents the difference in indicators between four sets of APM models and the baseline.

It can be seen from Figure 4 that APM-Neck-P4 has the most comprehensive improvement compared to the baseline: precision has increased by 0.0213, recall by 0.0157, F1 by 0.0181, mAP@0.5 by 0.0316, and mAP@0.5:0.95 by 0.0271. The APM at the Neck P4 position not only enhances the model's ability to detect road defect targets but also improves the model's false detection control ability and positioning accuracy. In contrast, APM-Backbone-P4, although it has the most significant recall improvement, has a considerable decrease in precision, indicating that it is more prone to false detections. APM-Neck-P3 fails to bring stable gains, indicating that the Neck P3 position is not suitable for the module to function. Analysing the influence of different embedding positions, the performance of the APM module varies significantly at different positions, which indicates that feature enhancement in road defect detection requires a reasonable design that combines network-level features.

Table 3. The Performance Changes of Each Improved Model Compared to the Baseline.

Model Name	Precision change	Recall Changes	F1 Changes	mAP@0.5 change	mAP@0.5:0.95 change
APM-Backbone-P4	-0.0370	+0.0312	+0.0018	+0.0135	+0.0145
APM-Backbone-P3	-0.0040	+0.0251	+0.0132	+0.0281	+0.0235
APM-Neck-P4	+0.0213	+0.0157	+0.0181	+0.0316	+0.0271
APM-Neck-P3	About 0.0000	-0.0060	-0.0038	-0.0030	+0.0056

Figure 4. Performance Improvement of Each Model Compared to the Baseline

Table 4. Classification-Wise AP of Each Improved Model

Model Name	alligator crack	longitudinal crack	other corruption	pothole	transverse crack	all classes
Baseline	0.507	0.436	0.628	0.282	0.402	0.451
APM-Backbone-P4	0.513	0.481	0.633	0.278	0.431	0.467
APM-Backbone-P3	0.538	0.479	0.678	0.285	0.438	0.483
APM-Neck-P4	0.544	0.484	0.666	0.303	0.433	0.486
APM-Neck-P3	0.514	0.426	0.643	0.275	0.396	0.451

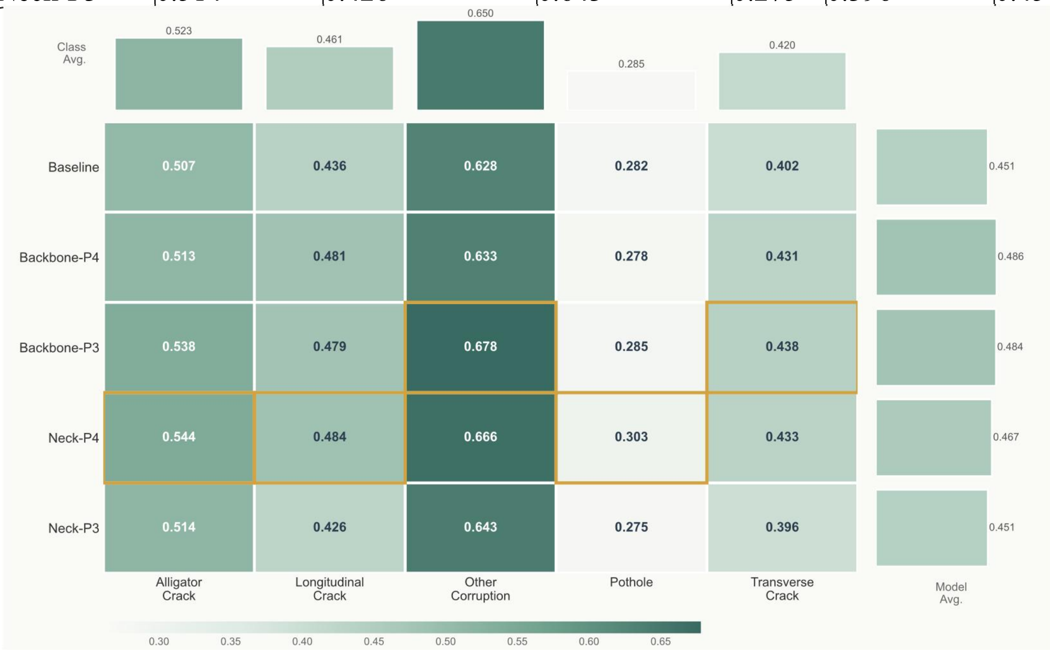

Figure 5. Heatmap of AP@0.5 by Classification for Different Models

5.3 Performance Analysis by Categories

The AP@0.5 results of different models for each road defect category are shown in Table 4. Overall, APM-Neck-P4 achieved the best results for alligator crack, longitudinal crack, and pothole categories, while APM-Backbone-P3 performed better in other corruption and transverse crack categories.

From the perspective of category performance in Figure 5, a pothole is the most difficult category to detect. The AP@0.5 of the baseline is only 0.282. This might be related to the small size of the pothole target, the large shape variation, and the low contrast between the edge and the road background. APM-Neck-P4 has improved the AP@0.5 of potholes to 0.303, which is the most significant improvement for this category among all the models.

For longitudinal cracks, APM-Neck-P4 increased its AP@0.5 from 0.436 to 0.484, indicating that the APM at the neck P4 position can enhance the continuous structural expression of slender cracks. For alligator cracks, APM-Neck-P4 also achieved the highest AP, suggesting that this model has better recognition ability for complex networked cracks. It is worth noting that APM-Neck-P3 was lower than the baseline in longitudinal cracks, potholes, and transverse cracks, indicating that introducing APM too early or too shallowly during the fusion stage may not be beneficial for road defect detection. This further validates the importance of the selection of module embedding position.

5.4 Inspection Failure Analysis

The main problem with the baseline model is that it has a significant rate of missed detections, especially for categories such as pothole, longitudinal crack, and transverse crack, which are prone to being misclassified as background. After introducing APM, the improvement in the missed detection issue varies depending on the embedding position.

Take "pothole" as an example. In the Baseline, the proportion of "pothole" being classified as "background" was 0.72. However, APM-Neck-P4 reduced this proportion to 0.66, while simultaneously increasing the AP@0.5 of "pothole" from 0.282 to 0.303. This indicates that the APM at the Neck P4 position can, to a certain extent, enhance the model's perception ability for pothole-like targets.

Comprehensive analysis of Figure 6. Although both APM-Backbone-P4 and APM-Neck-P4 have reduced the rate of pothole missed detections, APM-Neck-P4 has higher precision and mAP, indicating that it better controls false detections while alleviating missed detections. Therefore, from the perspective of overall performance, Neck P4 is the superior embedding position.

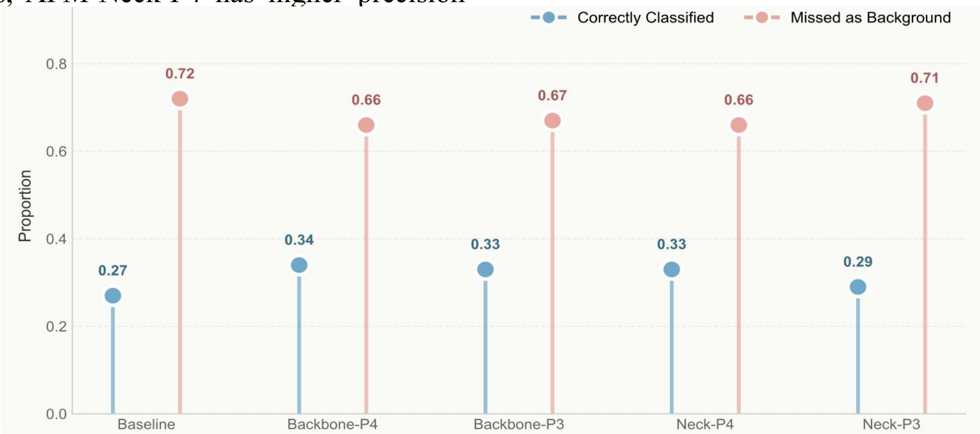

Figure 6. Comparison of Missed Detections for Pothole Categories

Table 5. Comparison of PR Curves and F1-Confidence Indicators for Each Model

Model	PR curve mAP@0.5	Optimal F1	Corresponding confidence level threshold
Baseline	0.451	0.49	0.297
APM-Backbone-P4	0.467	0.50	0.297
APM-Backbone-P3	0.483	0.51	0.302
APM-Neck-P4	0.486	0.51	0.303
APM-Neck-P3	0.451	0.49	0.258

5.5 PR Curve and F1-Confidence Analysis

Combined with the analysis of Table 5, the PR curve results show that the curve of APM-Neck-P4 is overall above the baseline model, indicating that this model can maintain higher detection accuracy at different recall levels. Further analysis of the F1-Confidence curve reveals that APM-Neck-P4 achieves an optimal F1 value of 0.51 at a confidence threshold of approximately 0.303, which is higher than the 0.49 of the baseline model.

The overall experimental results of the comprehensive indicators show that the APM module can enhance the feature expression ability of YOLOv8 for road defect targets, but its performance gain has a significant positional dependence. Among them, APM-Neck-P4 achieves the best results in precision, F1, mAP@0.5, and mAP@0.5:0.95, indicating that introducing APM in the multi-scale feature fusion stage can more effectively enhance the effective defect features and suppress background interference.

6. Conclusion and Outlook

This paper addresses the issues of weak edges, unclear textures, irregular shapes, and susceptibility to complex backgrounds in road defect images, such as cracks and potholes. To address these problems, a YOLOv8 road defect detection method integrating an adaptive perception module (APM) is proposed. This method enhances the model's ability to represent the edges, textures, and local structural information of the defect areas through convolution feature enhancement, residual connections, and attention weighting mechanisms.

To verify the effectiveness of APM in different network layers, this paper embedded APM at the Backbone P4, Backbone P3, Neck P4, and Neck P3 positions and conducted a comparative experiment with the baseline model. The results show that the performance gain of APM has a significant location dependence. Among them, APM-Neck-P4 achieved the best comprehensive performance, with precision, recall, F1, mAP@0.5, and mAP@0.5:0.95 reaching 0.5760, 0.4658, 0.5151, 0.4836, and 0.2416, respectively. Compared with the baseline, its mAP@0.5 and mAP@0.5:0.95 increased by 0.0316 and 0.0271, respectively, indicating that introducing APM at the neck P4 position can more effectively enhance multi-scale fusion features and improve the accuracy of road defect detection.

The results by category show that APM-Neck-P4 has a good improvement effect on alligator cracks, longitudinal cracks, potholes, etc. Among them, the AP@0.5 of pothole has increased from 0.282 to 0.303. The confusion matrix and PR curve further indicate that this model can improve the overall detection accuracy while alleviating some missed detection problems. Therefore, this paper finally selects APM-Neck-P4 as the optimal solution for improving the YOLOv8 road defect detection model.

This paper still has some shortcomings. To ensure the fairness of the comparative experiments, all the models in each group adopt the same number of training rounds and the same hyperparameter configuration. The training strategies have not yet been systematically optimised, and there is still room for further improvement in the model performance. Future research can combine hyperparameter search, expansion of training rounds, and optimisation of data augmentation strategies to further explore the potential of the model. At the same time, the experiments in this paper are mainly based on the existing road defect data set. The generalisation ability in complex weather conditions, nighttime lighting, and real vehicle inspection scenarios still needs to be further verified. In the future, methods such as lightweight networks, model pruning, and knowledge distillation can be combined to reduce the model's computational cost and enhance its deployment value in the actual road maintenance system.

References

[1] Yi, M. (2025). Research on detection and maintenance methods for major highway diseases. MS, 15(01), 115–123. https://doi.org/10.12677/ms.2025.151014

[2] Sun, L., Liu, Z., & Xie, Z. (2024). Road crack detection in orthophotos based on EC-YOLOX algorithm. Preprints. https://doi.org/10.20944/preprints202403.0680.v1

[3] Jiang, Y. (2024). Road damage detection and classification using deep neural networks. Discover Applied Sciences, 6(8), 421. https://doi.org/10.1007/s42452-024-06129-0

[4] Maeda, H., Kashiyama, T., Sekimoto, Y., Seto, T., & Omata, H. (2021). Generative adversarial networks for road damage detection. Computer-Aided Civil and Infrastructure Engineering, 36(1), 47–60. https://doi.org/10.1111/mice.12561

[5] Feng, X. (2021). A survey on the development of image data augmentation. Computer Science and Application, 11(02), 370–382. https://doi.org/10.12677/CSA.2021.112037

[6] Li, A., Ge, Y., Liu, H., Yang, C., & Zhou, X. (2024). Content-aware interpretable pavement damage detection model. Journal of Computer Research and Development, 61(3), 701–715.

