Proceedings of the 4th International Conference on Software Engineering and Machine Learning
DOI: 10.54254/2755-2721/2026.34322

Deep Learning for Mechanical Crack Recognition

Yaozu Pan

Faculty of Civil Engineering and Mechanics, Jiangsu University, Zhenjiang, China
panyaozu0506@163.com

Abstract. Cracks in mechanical structures cause numerous hazards, such as stress concentration, reduced structural strength and fatigue performance; continuous crack propagation easily leads to component fracture and equipment failure; it results in downtime maintenance and increased costs, and even safety accidents in severe cases. Therefore, accurate crack detection is crucial. Traditional detection relies on manual work, which has pain points such as low efficiency, strong subjectivity, obvious missed detection of micro-cracks, and difficulty in identifying internal hidden cracks, failing to meet the needs of real-time large-scale detection. This study takes YOLOv5 as the core and combines deep learning technology to carry out research on mechanical crack detection. A mechanical crack dataset is constructed, and sample annotation and data augmentation are completed to improve sample diversity; parameters are modified based on the YOLOv5 model, and after multiple rounds of training iteration and verification, the model finally achieves a mAP@0.5 of 31.37%, effectively reducing the missed detection rate and false detection rate; after model deployment, it can realize fast, automatic and high-precision identification and location of mechanical surface cracks, effectively alleviating the core pain points of traditional detection, and providing an efficient and feasible technical solution for mechanical structure crack detection.

Keywords: Mechanical crack, Object detection, Deep learning, YOLO, Small object detection, Machine vision

1. Introduction

Mechanical equipment is widely used in manufacturing, transportation, energy, aerospace and other fields. During long-term operation, it is easily affected by alternating loads, vibration shocks and complex working conditions, thus producing crack defects [4-1]. Cracks have weak characteristics in the initial stage, but will continue to expand with operation, causing local stress concentration and structural fatigue, and even leading to equipment failure and safety accidents in severe cases. Therefore, realizing fast and accurate detection of mechanical cracks is of great significance for ensuring industrial production safety. Traditional crack detection methods mainly include manual visual inspection and nondestructive testing technologies such as eddy current, ultrasonic and magnetic particle testing, but they have problems such as low detection efficiency, strong subjectivity and easy missed detection of tiny cracks, which are difficult to meet the real-time and intelligent detection needs of modern industry.

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.34322

With the development of machine learning and deep learning technologies, crack detection has gradually developed towards intelligence. Early methods mainly manually extract features such as texture, grayscale and edge, and combine algorithms such as Support Vector Machine and Random Forest to realize crack recognition [1], but this method has poor adaptability to complex backgrounds and irregular cracks. In contrast, deep learning can automatically extract deep features and shows higher accuracy in crack detection. Song Li et al. proposed an improved deep neural network to realize quantitative crack detection [2]; Yang Zhang et al. combined Bayesian inference to improve detection ability in complex noise environments [3]; Tang et al. proposed a multi-network fusion method to enhance model generalization performance [4]; Saberironaghi et al. proposed DepthCrackNet to improve crack detection effect through attention mechanism [5]. In addition, Song et al. realized real-time tunnel crack detection based on semantic segmentation [6]; Cheng et al. completed pixel-level crack detection based on U-Net [7]; He et al. proposed the BGCrack model to improve the segmentation accuracy of steel structure cracks [8].

In the field of object detection, the YOLO series algorithms are widely used in industrial crack detection due to their fast detection speed and strong real-time performance. Chen et al. used YOLOv8 to realize concrete crack identification and width measurement [9]; Liu et al. combined binocular vision and lightweight network to realize real-time crack detection [10]. However, YOLO still has certain limitations in crack detection tasks: cracks are slender and irregular, and traditional rectangular boxes are difficult to locate accurately; false detection and missed detection are prone to occur under complex backgrounds [11]; features of small target cracks are easy to lose, and attention mechanism and multi-scale feature fusion need to be introduced [12]; meanwhile, industrial crack datasets generally have problems such as insufficient samples and large scene differences, which limit the model generalization ability [13]. Aiming at the problem of insufficient data, Turan Abdullah et al. proposed a sliding window data augmentation method to improve detection effect [14]; Wu et al. proposed an unsupervised crack detection method based on VQ-VAE-2, which still has good performance under few-shot conditions [15].

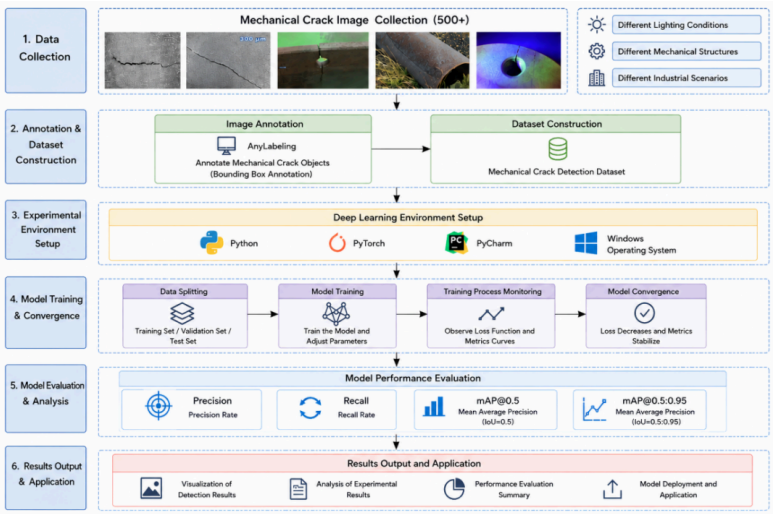

Figure 1. Technical route map

Based on the above research background, this paper carries out research on mechanical crack detection based on YOLOv5 as the basic model. A mechanical crack detection dataset is constructed by collecting crack images in different scenarios and completing data annotation; an experimental platform based on Python and PyTorch is built in the Windows environment, and YOLOv5 model

DOI: 10.54254/2755-2721/2026.34322

training and testing are completed by adopting the transfer learning strategy. Finally, the model performance is comprehensively evaluated through indicators such as Precision, Recall, mAP@0.5 and mAP@0.5:0.95, and the model convergence and detection stability are analyzed combined with the change of loss function, providing a reference for intelligent detection and safety operation and maintenance of industrial equipment.

2. Theoretical basis of mechanical crack detection

2.1. Mechanical crack characteristics and imaging features

Mechanical cracks seriously affect the reliability and service life of mechanical systems, and the detection of mechanical cracks faces many challenges. Cracks are mostly linear, slender, intermittent and diverse in branches on the structural surface, often accompanied by small through-cracks, with narrow width and small area, belonging to small objects. They are low-contrast and weak-texture defects with blurred edges and gentle grayscale changes, easily confused with the background. In actual industrial imaging, many factors interfere with crack detection. Uneven illumination causes large differences in image brightness, cracks are covered in over-bright areas, and the contrast is lower in over-dark areas; light reflection forms spots to blur crack edges, and shadows cover part of the features; image noise superposition reduces image quality, making crack features more difficult to extract; oil stains and rust form coverings to change crack appearance, scratches are similar to cracks and easy to be misdetected, and complex surface textures constitute a complex background. These interference factors lead to covered crack features, low signal-to-noise ratio and unclear boundaries, which greatly increase the difficulty of automatic crack detection. It is urgent and significant to develop automatic detection technology that can overcome interferences and accurately detect mechanical cracks.

2.2. Fundamentals of deep learning-based object detection

In the research direction of object detection, convolutional neural network has become the most basic and core architecture. The extraction of features in images and the accurate recognition of objects largely depend on it. The structure of the entire network can be divided into convolutional layer, pooling layer, activation function and batch normalization, which cooperate with each other. The convolutional layer is the most critical module in the entire network. Its working mode is not complicated: it slides the convolution kernel on the image repeatedly to accurately capture important features such as edges, textures and cracks; moreover, due to the local feature perception mechanism, it focuses on a local area of the image to mine detailed information therein, and weight sharing reduces the number of model parameters, making calculation much faster. After the convolutional layer processing, the pooling layer takes over, which is responsible for downsampling the feature map, that is, reducing the data dimension and the number of parameters, thus naturally reducing the consumption of computing resources, while retaining the key features in the image, and making the model more resistant to slight changes in the image. The activation function mainly adds nonlinear transformation to the network, which enables the convolutional neural network to fit more complex features and functional relationships. Finally, batch normalization standardizes the feature distribution of each batch of data, accelerating convergence and stabilizing the training process, alleviating the overfitting problem to a certain extent, and often improving the generalization performance of the model when encountering new unseen data.

DOI: 10.54254/2755-2721/2026.34322

2.3. YOLOv5 network structure and principle

YOLOv5 is a representative single-stage object detection algorithm. Compared with traditional two-stage detection algorithms, a major difference of YOLOv5 is that it does not need to generate region proposals first, and can directly complete object location and classification in one step, so it can meet the high real-time requirements of many industrial scenarios. The entire network structure can be roughly divided into four parts: input end, backbone network, neck network and detection head. The input end mainly includes image preprocessing and data augmentation, such as Mosaic augmentation, adaptive anchor box calculation and image scaling, to improve the generalization ability of the model; the backbone network uses modules such as Focus, C3 and SPPF to extract crack features at different levels layer by layer; the neck network adopts the FPN+PAN structure for multi-scale feature fusion, which significantly enhances the model's ability to detect tiny or irregular cracks; the detection head is responsible for outputting information such as object position, category and confidence, and finally obtains the final detection result after non-maximum suppression processing. In general, YOLOv5 adopts an end-to-end detection approach, which can quickly complete the identification and location of crack targets. Based on these considerations, this paper selects YOLOv5 as the basic model for mechanical crack detection.

3. Construction and preprocessing of mechanical crack dataset

3.1. Dataset source and data annotation

To build a high-quality and widely applicable special detection dataset for mechanical cracks, a large number of image selections were carried out. The selection criteria were clear images, obvious crack directivity, and no excessive fine cross through-cracks. Qualified crack images were selected one by one from the Google Images search engine, and 514 images were finally obtained. During the selection process, the images were made to cover more scenarios, different types of mechanical components and various crack morphologies as much as possible. In terms of component types, common components in industrial production such as metal components, welded joints, frames and shafts were involved, and cracks on their surfaces have a great impact on the safe operation of equipment. The images not only contain real cracks, but also select complex backgrounds with interference factors such as uneven illumination, oil stains and rust, to restore various complex environments encountered in actual detection as much as possible. In terms of crack types, common mechanical cracks such as fatigue cracks, welded cracks, cast cracks and heat treatment cracks are basically covered, which can reflect the crack characteristics under different causes and formation mechanisms. The dataset selected in this way has guaranteed authenticity, diversity and generalization ability, and can provide a solid and reliable data support for the subsequent crack detection algorithm research and model training.

3.2. Data annotation

The annotation of crack images was completed with the help of the AnyLabeling annotation tool, and each image was manually and accurately annotated one by one to mark the position and boundary of the target as accurately as possible. During annotation, all crack targets were uniformly assigned a single category label of *crack*, and the label files were saved in JSON format after annotation, which completely recorded information such as image name, crack target coordinates

DOI: 10.54254/2755-2721/2026.34322

and category information. In this way, the annotation format of the entire dataset is standardized and unified, meeting the input requirements of subsequent deep learning model training.

3.3. Data preprocessing

To achieve good model training effect, this study randomly divided the constructed mechanical crack dataset into three parts: training set, validation set and test set in a ratio of 7:2:1. The training set accounts for 70%, which is mainly used for model parameter learning and feature fitting to drive the network to update weights round by round; the validation set accounts for 20%, which is used for model performance monitoring, hyperparameter tuning and judging whether overfitting occurs during training; the test set accounts for 10%, which does not participate in training and validation, and is finally used to objectively evaluate the generalization ability and detection accuracy of the model. This not only makes full use of sample data, but also ensures the training stability, evaluation rationality and test fairness of the model on different data subsets.

3.4. Dataset analysis

To fully grasp the feature distribution law of the mechanical crack dataset and provide data support for model improvement and training, this study conducted statistical analysis from multiple dimensions, and focused on completing the small object proportion statistics. Crack targets were divided into large, medium and small sizes according to pixel area, and the number and proportion of large, medium and small cracks were counted. Mechanical cracks mostly exist in fine and slender forms, with small overall size and high proportion of small objects, putting high requirements on the model's small object detection ability. Statistics of the horizontal and vertical coordinates of cracks in the image show that cracks are evenly distributed in the image without obvious regional aggregation, which can ensure that the model learns crack features in the global range and improves generalization ability. Statistics of the width-to-height ratio of crack targets show that mechanical cracks are slender linear, with generally large aspect ratios, concentrated in high-proportion intervals. This feature provides an important data basis for anchor box clustering optimization and loss function improvement. Small objects are judged according to the proportion of target pixels in the total image area, and the statistical results show that the dataset has a high proportion of tiny cracks, belonging to a typical small object detection scenario, prone to missed detection, which is the key direction of model improvement in this study.

4. Experimental design and result analysis

4.1. Experimental environment

To ensure the stability of model training and the reliability of detection performance, this study built a unified experimental environment for training and testing. Each training input 8 images, with a total of 100 training rounds. At the hardware level, the experiment adopted a high-performance independent GPU (NVIDIA Geforce RTX 4060 Laptop) as the core computing unit, relying on its powerful parallel computing ability to accelerate the model training iteration and inference process, effectively shortening the experimental cycle and meeting the computing power requirements of deep learning models. At the software framework level, a deep learning environment was built based on Python 3.12 and PyTorch 1.13.1+cu116. This framework has the advantages of simplicity, ease of

DOI: 10.54254/2755-2721/2026.34322

use, high flexibility and complete ecology, and can efficiently realize model construction, parameter tuning and function expansion, providing stable software support for the experiment.

4.2. Training process and convergence analysis

To verify the training effect and convergence performance of the YOLOv5 mechanical crack detection model, this paper carried out model training according to the set hyperparameter configuration based on the Windows and PyTorch experimental environment built above. YOLOv5s was adopted as the basic model during training, transfer learning was carried out with yolov5s.pt pre-trained weights, the input image size was set to 640×640, the Batch Size was set to 8, the training epochs were 100, and the SGD optimizer was used to complete network parameter update. Meanwhile, the Mosaic data augmentation method was combined to improve the model's detection ability for complex backgrounds and small-sized crack targets.

During the model training process, the changes of key indicators of the training set and validation set were monitored in real time. Through comparative analysis of the results of the training set and validation set, the generalization ability and crack detection performance of the model were verified, providing a reliable basis for subsequent mechanical crack detection result analysis and model optimization.

The specific parameter values are shown in Table 1 below.

Table 1. Summary of parameter values

Parameter Name	Parameter value	Parameter Explanation
Model	YOLOv5s	YOLOv5s is adopted as the basic detection model
Input image size	640×640	Network input image resolution
Epochs	100	Number of model training rounds
Batch Size	8	The number of training samples in each batch
Initial learning rate	0.01	Initial learning rate of the model
Training platform	PyTorch	Deep learning framework
Operating System	Windows	Model training runtime environment
Training equipment	GPU	Used to accelerate model training

4.2.1. Loss curve analysis

The loss value during model training is the core indicator to measure the model convergence state. The total loss of the YOLOv5 model consists of three parts: bounding box loss (box_loss), object loss (obj_loss) and classification loss (cls_loss). Combined with the loss changes of the training set and validation set, the model learning effect and generalization ability can be fully reflected, and the specific analysis is as follows:

Training set loss curve analysis: According to the training set loss curve (Figure 2), the initial value of train/box_loss (localization loss) is 0.1047, which continues to decrease during training and finally stabilizes at about 0.03, indicating that the model's ability to locate crack target boxes continues to improve, the fitting effect is normal, it can gradually master the position characteristics of cracks, and the positioning accuracy is continuously optimized. The initial value of train/obj_loss (object confidence loss) is 0.0275, which continues to decrease with training rounds and finally stabilizes at about 0.01, indicating that the model's ability to judge "whether there is a crack target in

DOI: 10.54254/2755-2721/2026.34322

the image" is continuously optimized, and it can effectively distinguish cracks from background interference. The initial value of train/cls_loss (classification loss) is 0.0467, with the fastest decrease rate, stabilizing at about 0.001 after the 3rd round of training, which is the fastest converging item among the three types of losses, indicating that the model's classification ability for crack categories is quickly fitted, and the classification task learning effect is excellent.

Validation set loss curve analysis: According to the validation set loss curve (Figure 3), the initial value of val/box_loss (validation set localization loss) is 0.0986, showing a continuous downward trend as a whole, and finally stabilizes at about 0.058, which is completely consistent with the downward trend of the training set loss without obvious deviation, indicating that the generalization of the model's positioning ability is good and no overfitting occurs. Val/obj_loss (validation set object confidence loss) fluctuates stably in the range of 0.013~0.017 throughout the process without obvious upward trend, indicating that the generalization of the model's object judgment ability is stable, and it can maintain a good recognition effect in unseen validation samples. Val/cls_loss (validation set classification loss) has an initial value of 0.0085, which decreases rapidly and finally stabilizes at about 0.0005, which is highly consistent with the convergence trend of the training set, verifying the excellent generalization effect of the model's classification ability.

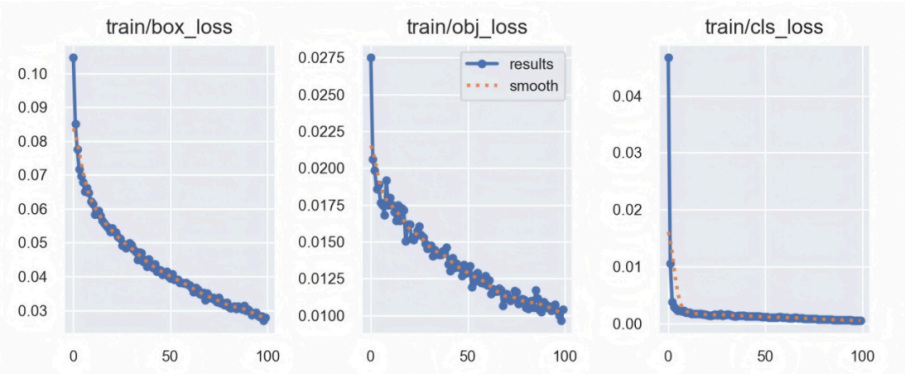

Figure 2. Training set loss curve

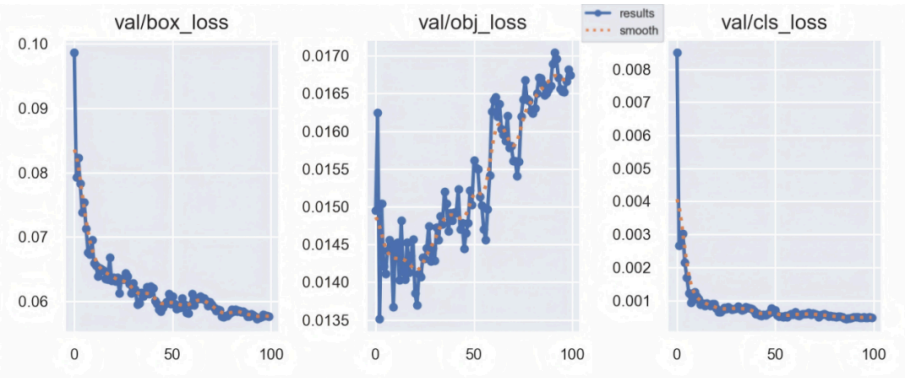

Figure 3. Validation set loss curve

Overall, the three types of loss values of the training set and validation set all show a trend of "rapid decrease—slow stability" with the increase of training rounds, without obvious fluctuation or upward phenomenon, and the change trends of the two are consistent with small deviation, indicating that the model training process converges well, can effectively learn the feature information of mechanical cracks, and has no signs of underfitting or overfitting.

DOI: 10.54254/2755-2721/2026.34322

4.2.2. Core evaluation metric analysis

It can be seen from the core evaluation metric curve (Figure 4) that the core evaluation indicators such as precision, recall and mAP all show a continuous upward trend as a whole, and tend to be stable in the later stage of training. The change law of each indicator is highly consistent with the model learning process. Precision mainly measures the accuracy of model prediction, which rises slowly from the initial 0.037 and finally stabilizes in the range of 0.4 to 0.5, with a maximum of 0.52. This upward trend indicates that the model's ability to distinguish crack targets from background interference is continuously enhanced, and false detection is gradually reduced; recall measures the model's coverage ability for targets, with an initial value of only 0.04, then steadily rising to between 0.3 and 0.4, with a maximum of 0.43, which means that the model can capture more and more real mechanical cracks, the probability of missed detection is decreasing, and the recognition ability for tiny and hidden cracks is gradually optimized; as the core indicator for comprehensively evaluating the overall detection performance of the model, mAP is subdivided into two dimensions: mAP@0.5 and mAP@0.5:0.95. Among them, mAP@0.5 rises from the initial 0.007 to the range of 0.25 to 0.3, with a maximum of 0.3137. This indicator takes IOU equal to 0.5 as the threshold, reflecting the overall performance of the model under conventional detection accuracy requirements; mAP@0.5:0.95 covers multiple IOU thresholds from 0.5 to 0.95, and finally stabilizes between 0.1 and 0.15, reflecting the model's generalization ability under different detection accuracy requirements. Comprehensively, it shows that the overall detection accuracy of the model is steadily improving.

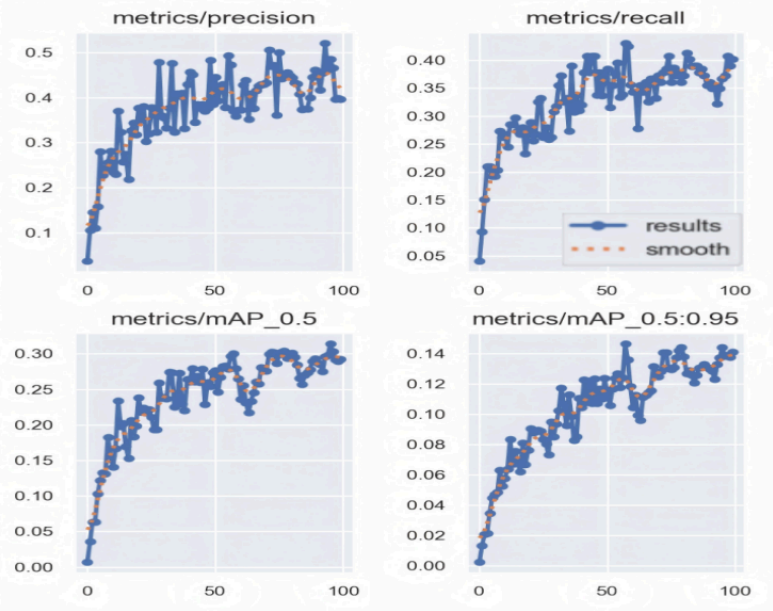

Figure 4. Core evaluation metric curve

Although all indicators of the model show a steady growth trend, the overall precision, recall and mAP values are still relatively low, indicating that the current model still has certain limitations in the mechanical crack detection task. The main reasons for this phenomenon include the following aspects: first, the dataset scale constructed in this paper is relatively limited, containing only more than 500 crack images, and insufficient data samples will limit the model's ability to learn complex crack features, easily leading to insufficient feature expression; second, mechanical cracks themselves have the characteristics of slender, irregular and small target area, accounting for a low

DOI: 10.54254/2755-2721/2026.34322

proportion in the image, belonging to a typical small object detection problem. YOLOv5 is prone to loss of detailed information in the shallow feature extraction process, thus affecting the recognition effect of tiny cracks; in addition, there are complex background factors such as oil stains, rust, texture interference and illumination changes in industrial scenarios, which easily make the model misidentify background textures as cracks, thereby affecting detection accuracy; meanwhile, some cracks have blurred edges and low contrast, which will also increase the difficulty of model feature extraction.

In view of the above problems, the model performance can be further optimized in the future: first, expand the scale of the mechanical crack dataset, add crack samples of different working conditions and types, and improve the model generalization ability; second, introduce multi-scale feature fusion or attention mechanism for small object cracks to enhance the fine crack feature extraction ability; third, optimize anchor box parameters and loss function to improve the model's detection effect on slender cracks; fourth, improve data preprocessing and image enhancement methods to reduce the interference of complex backgrounds on detection results, thereby improving the overall detection accuracy and robustness of the model.

4.2.3. Training stability analysis

During the training process, it is observed that all loss values of the training set and validation set decrease without obvious oscillation or rebound, and the deviation between the two is controlled within a small range, indicating that the model parameter update process is relatively stable, problems such as gradient explosion or gradient disappearance do not occur, and the model continuously learns the feature information of mechanical cracks. With the gradual attenuation of the learning rate, core indicators such as Precision, Recall and mAP continue to rise and tend to be stable in the later stage of training; in addition, the accuracy gap between the training set and validation set has been kept within a small range, no obvious overfitting phenomenon occurs, and the model's generalization ability is relatively good. In addition, no abnormal situations such as data loading errors, video memory overflow or program interruption were encountered during the entire training process, and the training time of each batch was relatively stable. The configuration of the experimental environment can fully meet the model training requirements. Comprehensively considering the changes of loss curves and evaluation indicators, the model training has good convergence effect and strong stability, and these training results can provide reliable support for subsequent crack detection and result analysis.

4.3. Detection results and visualization

To intuitively present the actual detection effect, practicability and reliability of the YOLOv5 mechanical crack detection model, this section comprehensively analyzes the model's detection ability through sample detection image visualization and core performance curve visualization combined with detection indicators, providing a reference for the practical engineering application of the model.

4.3.1. Sample detection image visualization

The sample detection visualization results (Figure 5) can intuitively reflect the actual detection effect of the model. The figure contains mechanical crack samples of different scenarios, different

DOI: 10.54254/2755-2721/2026.34322

illumination conditions and different sizes, which are used to evaluate the model's positioning accuracy, background interference suppression ability and tiny crack detection ability.

It can be seen from the detection results that the model can accurately locate most crack targets, the predicted bounding box is basically consistent with the actual crack position, and can effectively distinguish cracks from background interference such as stains and textures on the mechanical surface without a large number of false detections. The model has a good detection effect on cracks with obvious features and large sizes; for some tiny cracks, due to the small target size and inconspicuous features, there are still a few missed detections or bounding box deviations. In addition, the model can still maintain a relatively stable detection effect under different illumination and complex background conditions, indicating that it has a certain environmental adaptability. Overall, the YOLOv5 mechanical crack detection model can meet basic crack detection needs, and also provides a reference for subsequent model optimization.

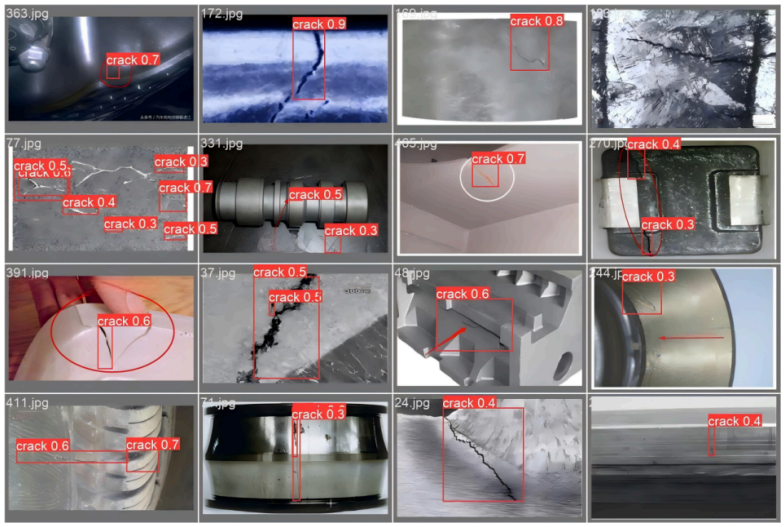

Figure 5. Sample detection images

4.3.2. Performance curve visualization analysis

To further analyze the detection performance of the model, a comprehensive analysis was carried out combined with the PR curve, F1 curve, precision-confidence curve and recall-confidence curve generated during training (Figure 6), mainly to see the model's performance under different confidence thresholds. The PR curve can reflect the relationship between precision and recall. During the process of recall from 0 to 1, precision gradually decreases; when the recall is less than 0.2, the precision can still be maintained above 0.7, but once the recall exceeds 0.2, the precision drops rapidly. The mAP@0.5 calculated by the area under the curve is 0.300, which is basically consistent with the training results, indicating that the model has guaranteed detection accuracy under low recall. Then the F1 curve: during the process of confidence threshold from 0 to 1, the F1 value first rises, reaches a peak of 0.40 at 0.173, and then begins to decrease, indicating that the balance between precision and recall is the best at this point; after the confidence exceeds 0.8, the F1 value basically approaches 0. The precision-confidence curve shows that the higher the confidence threshold, the higher the precision, and the precision directly reaches 1.00 at about 0.904, indicating that the model has few false detections under high confidence and the results are more reliable; the recall-confidence curve is just the opposite, as the confidence threshold increases, the recall keeps decreasing, the maximum recall can reach 0.70 when the confidence is 0, and the recall is basically

DOI: 10.54254/2755-2721/2026.34322

close to 0 after 0.8, which means that setting a lower confidence threshold can find more real cracks, but at the cost of more false detections.

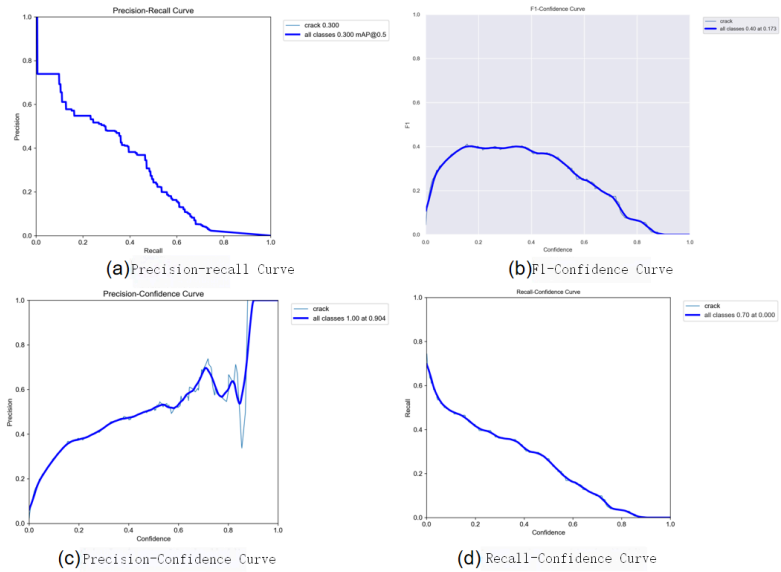

Figure 6. Precision-recall curve, F1 curve, precision-confidence curve and recall-confidence curve

In summary, through the visualization analysis of sample detection images and various performance curves, it is verified that the YOLOv5 mechanical crack detection model has good detection accuracy and generalization ability, and can adjust the confidence threshold according to the needs of different practical application scenarios to achieve the balance between precision and recall, meeting the engineering needs of mechanical crack detection.

5. Conclusion and prospect

5.1. Conclusion

This paper focuses on the automatic detection needs of mechanical cracks, carries out research on mechanical crack detection based on YOLOv5, and realizes automatic identification and location of mechanical cracks through dataset construction, model training and experimental verification. Firstly, a special dataset for mechanical cracks is constructed, covering various types such as fatigue cracks and welded cracks, and image annotation and dataset division are completed combined with common interference factors in industrial sites, providing a data basis for model training. Secondly, model training and verification are completed based on YOLOv5. The loss function shows a steady downward trend during training, the model converges well without obvious overfitting, and has good generalization ability. The experimental results show that the model's maximum precision reaches 0.52, maximum recall reaches 0.43, and maximum mAP@0.5 reaches 0.3137, which can better realize automatic detection and location of mechanical cracks, and has a certain recognition ability for crack targets under complex backgrounds. Finally, the experimental environment built in this paper can meet the needs of model training and inference, and effectively improve the model training efficiency combined with the transfer learning strategy, providing a certain reference for the subsequent engineering application of mechanical crack detection models.

DOI: 10.54254/2755-2721/2026.34322

5.2. Prospect

Although this paper realizes automatic detection of mechanical cracks based on YOLOv5 and achieves certain effects, there are still problems such as limited dataset scale and insufficient tiny crack detection ability. In the future, the generalization ability and anti-interference ability of the model can be improved by expanding crack samples of different working conditions and types and optimizing data augmentation methods; meanwhile, attention mechanism and better feature fusion structure can be introduced to enhance the model's feature extraction ability for fine cracks and further reduce the missed detection rate and false detection rate. In addition, the model can be lightweight compressed and deployed in engineering to meet the real-time detection needs of industrial sites, and combined with nondestructive testing technologies such as ultrasonic and eddy current, the joint detection of surface and internal cracks of mechanical components can be realized, further expanding the application of the model in industrial scenarios such as bridges and wind power equipment.

References

[1] Jie, Z., Zheng, H., & Zhang, L. (2025). Machine learning-based fatigue life prediction of double-sided U-rib welded joints considering multiple factors. International Journal of Fatigue, 109187.
[2] Sun, H., Song, L., & Yu, Z. (2025). Efficient quantifying track structure cracks using deep learning. Computer-Aided Civil and Infrastructure Engineering, 40(19), 2969-2986.
[3] Zhang, Y., Ni, Y. Q., Jia, X., et al. (2023). Identification of concrete surface damage based on probabilistic deep learning of images. Automation in Construction, 156, 105141.
[4] Tang, X., Liu, X., Liu, Y., et al. (2025). A multi-deep learning intelligent surface rock crack identification method for transmission tower siting. Electronics, 14(11), 2255.
[5] Saberironaghi, A., & Ren, J. (2024). DepthCrackNet: A deep learning model for automatic pavement crack detection. Journal of Imaging, 10(5), 100.
[6] Song, Q., Wu, Y., Xin, X., et al. (2019). Real-time tunnel crack analysis system via deep learning. IEEE Access, 7, 64186-64197.
[7] Cheng, J., Xiong, W., Chen, W., et al. (2018). Pixel-level crack detection using U-Net. In TENCON 2018 - 2018 IEEE Region 10 Conference (pp. 0462-0466). IEEE.
[8] He, Z., Chen, W., Zhang, J., et al. (2024). Crack segmentation on steel structures using boundary guidance model. Automation in Construction, 162, 105354.
[9] Chen, G., Bian, Z., Jing, H., et al. (2025). Crack identification of concrete structures based on high-precision multi-level deep learning model. In Structures (Vol. 75, p. 108720). Elsevier.
[10] Liu, L., Shen, B., Huang, S., et al. (2025). Binocular video-based automatic pixel-level crack detection and quantification using deep convolutional neural networks for concrete structures. Buildings, 15(2), 258.
[11] Luo, S., Xie, X., Zhou, B., et al. (2025). Multiscenario generalization crack detection network based on the visual foundation model. Structural Control and Health Monitoring, 2025, 6269747.
[12] Jing, H., Yuan, S., Chen, J., et al. (2025). A multi-scale deep residual network-based guided wave imaging evaluation method for fatigue crack quantification. Measurement Science and Technology, 36(1), 016115.
[13] Zhu, Y., Xu, W., Cai, C. S., et al. (2023). Relative-breakpoint-based crack annotation method for lightweight crack identification using deep learning methods. Applied Sciences, 13(15), 8950.
[14] Turan, A., Gazioglu, E., Aggumus, H., et al. (2026). Using machine learning for crack detection in beams under limited experimental data: The role of sliding window augmentation. Journal of Vibration and Control, 10775463261432650.
[15] Wu, J., & Liu, C. (2025). VQ-VAE-2 based unsupervised algorithm for detecting concrete structural apparent cracks. Materials Today Communications, 44, 112075.