ISSN: 2414-1895 DOI: 10.6919/ICJE.202605_12(5).0009

Deep Learning-Based Segmentation and Quantification of Metro Shield Tunnel Lining Cracks

Longxu Chen[1], Yong Wang[1,2], Chuanbin Ren[2]

[1] Beijing University of Technology, Beijing 100124, China

[2] Beijing Urban Construction Exploration & Surveying Design Research Institute Co., Ltd., Beijing 100101, China

Abstract

To address the issues of insufficient segmentation accuracy and high computational complexity in crack detection of metro shield tunnel linings, a lightweight instance segmentation model named YOLO-LCD based on YOLO11 is proposed. The model introduces StarNet into the backbone to reduce computational cost, constructs a C3k2-S module in the neck to enhance crack feature representation, and replaces traditional concatenation with a Modulation Fusion Module (MFM) to improve multi-scale feature fusion efficiency. Meanwhile, a Lightweight Shared Convolution Detection head (LSCD) is adopted to reduce parameter redundancy. Furthermore, a complete crack extraction and quantification framework is developed by integrating image enhancement, Frangi filtering, and skeleton extraction, enabling automatic measurement of crack length, width, and orientation. Experimental results demonstrate that the proposed model significantly reduces parameters and computational cost while maintaining competitive segmentation accuracy, achieving a favorable balance between accuracy and efficiency and meeting the practical requirements of metro tunnel inspection.

Keywords

Metro Shield Tunnel; Crack Detection; Instance Segmentation; Lightweight Model; Crack Quantification.

1. Introduction

With the advancement of urbanization and the continuous development of the economy and society in China, urban population and traffic pressure are increasing day by day. As an important means to alleviate traffic congestion in large cities, urban rail transit has been accelerating the planning and construction of rapid, high-capacity public transportation systems such as subways and light rails in major cities. The shield tunneling method has been widely used due to its fast construction speed, high mechanization degree, and good structural quality. However, metro projects are characterized by large investment, complex construction, and long service life, requiring extremely strict quality control during construction, completion, and operation and maintenance phases. With the end of the construction peak period, China's urban metro is shifting from construction-oriented to construction-and-maintenance-balanced. During operation, tunnel lining cracks and other defects may appear; if not detected in time, they will affect structural safety. Therefore, it is urgent to improve the disease detection and evaluation capabilities of metro shield tunnels to ensure long-term structural safety and operational stability.

With the rapid development of computer hardware, computer vision has been extensively studied in the field of civil engineering structural inspection. Numerous scholars have conducted extensive research and experiments on crack detection. Currently, the main crack detection methods are divided

ISSN: 2414-1895 DOI: 10.6919/ICJE.202605_12(5).0009

into three categories: threshold segmentation methods [1], edge detection methods [2], and deep learning methods [3]. In recent years, with the rapid development of deep learning technology, deep learning has been widely applied in the field of crack detection and segmentation due to its characteristics of not requiring predefined disease features or image preprocessing, but rather learning automatically from large amounts of sample data to extract features. In recent years, many researchers have conducted studies in the field of crack detection and segmentation. Liu et al. [4] proposed DeepCrack, which aggregates multi-scale and multi-level features and directly supervises the features of each convolutional layer, combining guided filtering and conditional random fields for optimization. Liu et al. [5] applied the U-Net network to concrete crack detection tasks, enabling the model to effectively identify crack positions from original images under complex conditions and achieve pixel-level detection. Zhao et al. [6] proposed an approach that extends the path aggregation network model and adds a semantic branch, utilizing PANet's bottom-up path enhancement, adaptive feature pooling, and fully connected fusion modules to address fracture issues in crack segmentation. Xu et al. [7] proposed the YOLO-LD model incorporating LSKA and DSC to enhance crack feature extraction capability. Duan et al. [8] proposed a tunnel lining crack detection model based on improved YOLOv5, which enhances the model's crack feature extraction and recognition capabilities in complex environments by adding semantic context encoding and detail preservation encoding modules. Dang et al. [9] used YOLOv8n as a baseline, integrated the SWC module, MLCA mechanism, and Inner-SIoU loss function, and proposed the SMI-YOLOv8 multi-scale feature attention fusion algorithm, improving the model's crack recognition ability.

Among the above studies, some machine learning-based crack recognition algorithms are object detection algorithms, which cannot effectively extract crack masks for further skeleton extraction to obtain crack length and width information for quantification. Additionally, some crack segmentation algorithms have high computational overhead and cannot meet the lightweight requirements in actual tunnel scenarios.

Therefore, this study constructs an improved YOLO11 shield tunnel lining segmentation model named YOLO-LCD. StarNet [10] is introduced into the backbone network, the C3k2-S module constructed using the StarNet Block replaces the original C3k2 in the neck network, and the traditional concatenation module in the neck network is replaced with the MFM [11] module. The LSCD [12] mechanism is introduced into the segmentation head. Through the above four improvement strategies, the model is further lightweighted while ensuring segmentation accuracy, and a complete set of metro shield tunnel lining crack extraction and quantification methods based on image information is constructed based on the segmentation results.

2. Crack Instance Segmentation Model: YOLO-LCD

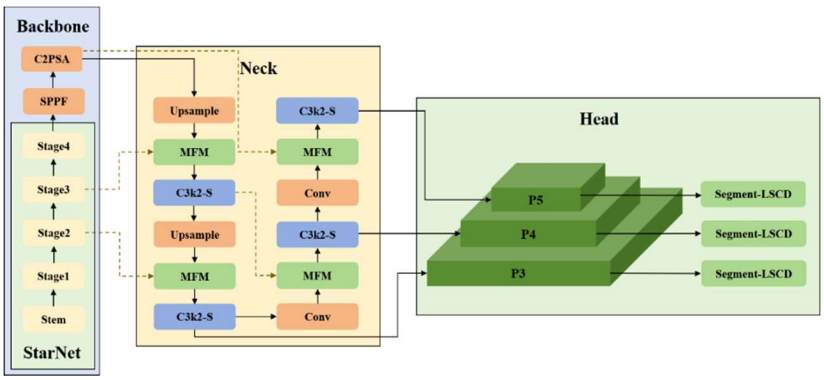


Fig. 1 YOLO-LCD network architecture

Considering the practical requirements of metro shield tunnels, the model should have certain lightweight characteristics. Therefore, this paper selects the YOLO11n algorithm as the baseline model for optimization and proposes the YOLO-LCD (YOLO-Lightweight Crack Detection) model.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202605_12(5).0009

The YOLO-LCD network structure is shown in Fig. 1. YOLO-LCD introduces StarNet into the backbone network, uses the Star Block to construct C3k2-S in the neck network, and replaces the traditional concatenation module with the MFM module. The LSCD mechanism is introduced into the segmentation head. Based on the above four improvements, the complexity of the original network is reduced, further meeting the lightweight requirements in practical scenarios.

2.1 StarNet Module

This paper introduces StarNet into the backbone network of the YOLO11 framework to replace the original structure. The core idea of StarNet is to use element-wise multiplication as the basic operation for feature interaction. As shown in Fig. 2, the basic building block of StarNet is the StarNet Block. First, local spatial information is extracted through depthwise separable convolution, and then the features are fed into two parallel convolutional branches. Element-wise multiplication is performed on the two branch outputs to complete feature fusion. One branch passes through ReLU activation, acting as a dynamic gate control, similar to the attention mechanism but without Softmax or matrix multiplication, with significantly lower computational overhead than self-attention structures. The fused features are output after channel compression and residual connection. The overall structure is simple and efficient. This paper adopts a lightweight variant of StarNet, with a base channel number of 16 and block numbers of 1, 1, 3, 1 at each stage. The network outputs three-scale feature maps P3, P4, and P5 through four downsampling stages, which are sent to the subsequent neck network for multi-scale fusion. Compared with the original YOLO11 backbone network, it has fewer parameters and is suitable for deployment in engineering detection scenarios with limited computing resources.

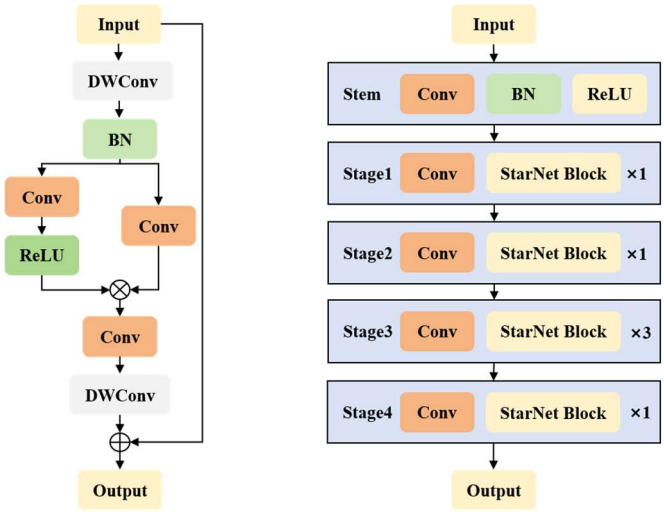


Fig. 2 Architectures of the StarNet Block and StarNet

2.2 C3k2-S Module

To enhance the neck network's extraction capability for tunnel crack features, this paper replaces the C3k2 module in the YOLO11 Neck with the C3k2-S module. As shown in Fig. 3, the basic computational unit of C3k2-S is replaced with the StarNet Block. Compared with the original C3k2, C3k2-S has the following advantages. The multiplicative gate mechanism in the StarNet Block can dynamically suppress background interference and strengthen discriminative local features such as crack edges, providing stronger responses to low-contrast, fine-line crack targets in tunnel environments. In addition, the large-kernel depthwise separable convolution expands the receptive field while maintaining parameter lightweighting, helping to capture long-range contextual information along the crack extension direction, reducing the risk of missed detection caused by local occlusion or uneven illumination, and effectively improving the network's comprehensive detection accuracy for cracks of different widths and orientations without significantly increasing the model's inference overhead.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202605_12(5).0009

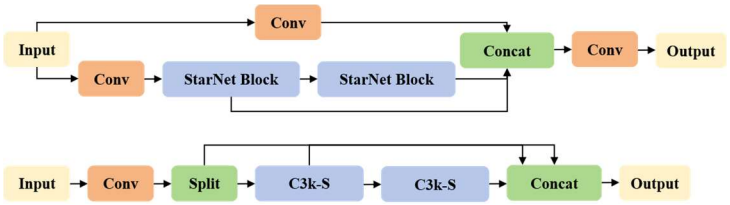


Fig. 3 Architectures of the C3k-S and C3k2-S

2.3 LSCD Segmentation Head

In object detection and instance segmentation tasks, the parameter scale and computational complexity of the detection head directly affect model inference efficiency. Traditional YOLO series detection heads independently deploy complete regression and classification branches on each scale feature layer, with parameters increasing linearly with the number of detection layers, exhibiting obvious structural redundancy, making it difficult to meet the real-time and lightweight requirements for engineering deployment in tunnel crack detection. To this end, this paper introduces the Lightweight Shared Convolution Detection head (LSCD) and extends it to the instance segmentation framework. LSCD replaces independent branches at each scale with shared convolution weights, compressing the detection head parameter scale to approximately constant level. Group normalization is used instead of batch normalization to maintain stable training convergence under the condition of limited samples in the tunnel crack dataset. A learnable scale factor is configured for each detection layer to compensate for expression deviations between shared weights across different scale features, ensuring that the model can maintain effective responses to both fine crack textures and large-area damage regions while achieving effective lightweighting.

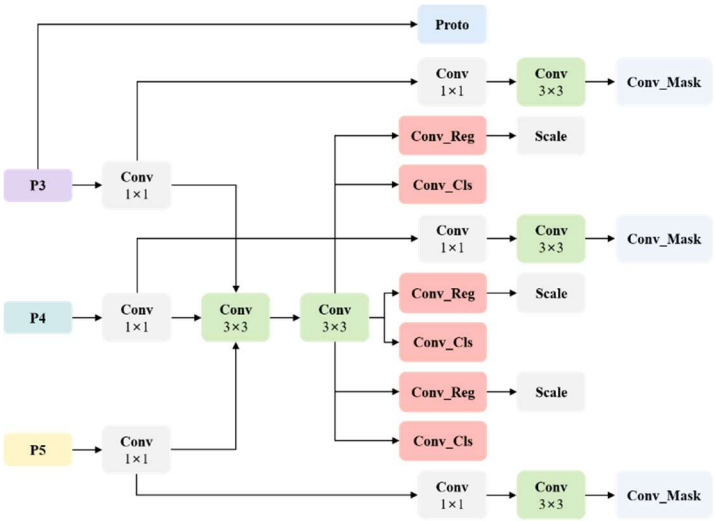


Fig. 4 Architecture of the LSCD Segment Head

2.4 MFM Modulation Fusion Module

In the multi-scale feature fusion stage of the feature pyramid network, traditional concatenation operations directly concatenate feature maps from different levels along the channel dimension. The output channel number increases linearly with the number of input branches, and subsequent convolutional layers need to bear additional channel compression computational burden, with obvious redundancy in the overall structure. To this end, this paper introduces the Modulation Fusion Module (MFM) to replace concatenation operations in the Neck, achieving adaptive weighted fusion through an attention mechanism, improving cross-scale feature integration effectiveness while maintaining the output channel dimension unchanged. Compared with traditional concatenation operations, the attention-driven adaptive fusion mechanism enables the network to dynamically adjust the fusion weights of shallow and deep features based on the current scale and morphology of crack targets, enhancing the comprehensive perception capability for different cracks and reducing

ISSN: 2414-1895 DOI: 10.6919/ICJE.202605_12(5).0009

parameter count and computational overhead, conforming to the lightweight design goal for engineering deployment.

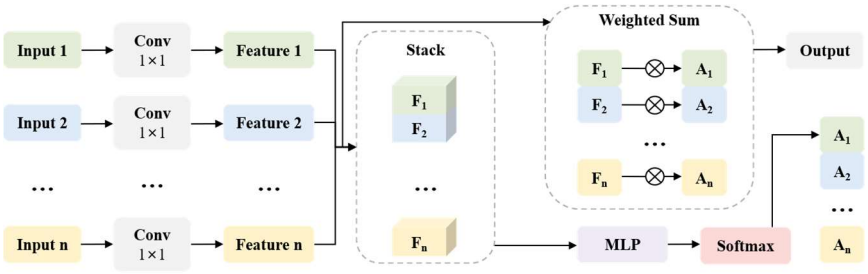


Fig. 5 Architecture of the Modulation Fusion Module

3. Crack Extraction and Quantification Process

Crack extraction and quantification are mainly divided into three stages: target segmentation, crack extraction, and quantitative analysis.

3.1 Target Segmentation Process

The target segmentation stage is based on the YOLO-LCD model, adopting a sliding-window inference strategy to obtain mask predictions for various targets, as shown in Fig. 6.

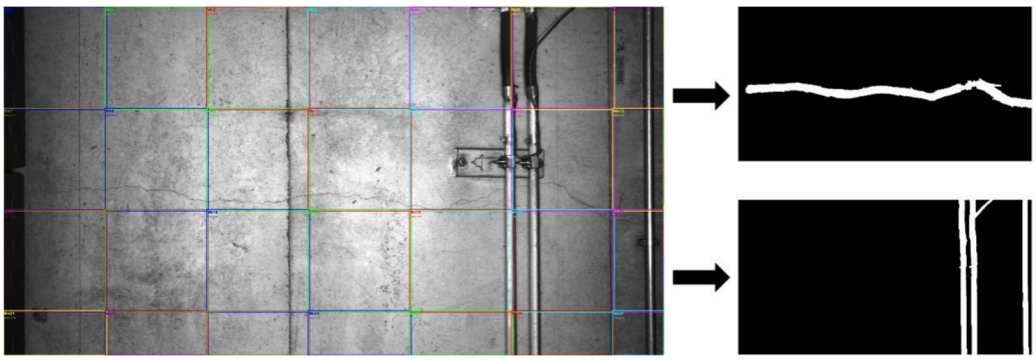


Fig. 6 Sliding-window inference results

3.2 Crack Extraction Process

The crack extraction stage includes image enhancement and illumination correction, Frangi filter-based linear structure enhancement, instance segmentation mask image cropping, connected component gray-level-based background removal, patch-wise adaptive binarization, morphological processing, and skeleton extraction.

(1) For the low illumination and uneven lighting commonly present in original acquired images (Fig. 7(a)), the global enhancement algorithm based on Retinex theory [13] is first applied to the original low-light crack images for global enhancement to improve overall lighting conditions. Subsequently, CLAHE (Contrast Limited Adaptive Histogram Equalization) and adaptive Gamma correction are applied for local contrast optimization to complete image preprocessing (Fig. 7(b)).

(2)After image enhancement and illumination correction, crack contrast may still be insufficient. Therefore, this study employs the Frangi filter [14] to process the input grayscale image (Fig. 7(c)) to enhance regions matching crack characteristics and suppress background non-target structures. This method is based on local Hessian matrix analysis (Eq. 1).

$$H(x,y,\sigma)=\begin{bmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{bmatrix}=\nabla^2 G_\sigma * I \tag{1}$$

ISSN: 2414-1895 DOI: 10.6919/ICJE.202605_12(5).0009

In the formula, $H(x,y,\sigma)$ represents the Hessian matrix at position (x,y) with scale σ; $\nabla^2 G_\sigma$ represents the second-order differential operator formed by Laplacian operation on the Gaussian function; * represents two-dimensional convolution operation; I represents the input grayscale image. Let the eigenvalues satisfy$|\lambda_1| \leq |\lambda_2|$, then the Frangi response function for cracks is defined as Eq. 2:

$$F(x,y)=\begin{cases}0, & ,\lambda_2>0 \\ \left[1-e^{-\frac{R_B^2}{2\beta^2}}\right]\cdot e^{-\frac{S^2}{2c^2}} & ,\lambda_2\leq 0\end{cases} \quad (2)$$

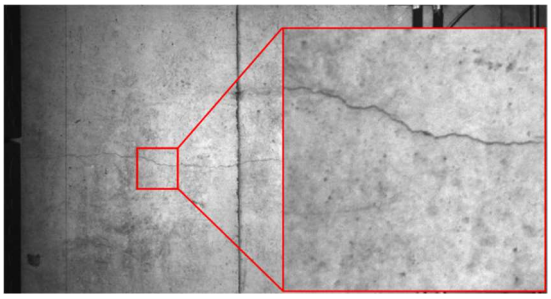
In the formula, $R_B=|\lambda_1|/|\lambda_2|$represents tubularness; $S=\sqrt{\lambda_1^2+\lambda_2^2}$ represents structure strength; β, c are control parameters.

(3)Crack masks are output through the YOLO-LCD model, and crack regions are cropped to remove the background (Fig. 7(d)) to reduce subsequent computational overhead.

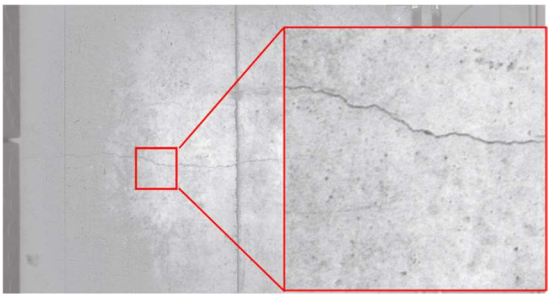
(4) Patch-wise adaptive threshold segmentation combined with connected component analysis is used for preliminary crack region extraction. For each connected component in Fig. 7(d), the average gray level in the original Frangi map is calculated. Background information is filtered out through area constraints and gray-level constraints, retaining only regions meeting the conditions, finally synthesizing Fig. 7(e).

(5) Patch-wise adaptive binarization is then applied (Fig. 7(f)), where the image is divided into local sub-blocks and binarization thresholds are independently calculated for each sub-block to adapt to uneven illumination and local contrast phenomena. Morphological processing operations and secondary connected component analysis are further applied to purify the results, obtaining Fig. 7(g).

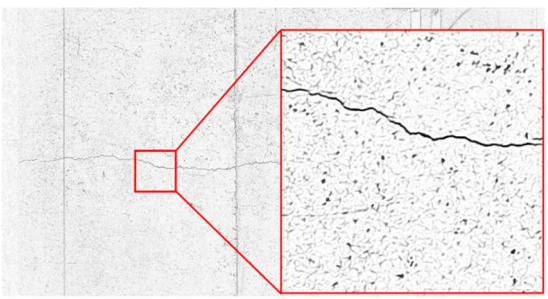
(6) The Zhang-Suen algorithm [15] is employed for skeleton extraction on the binarized image, and graph-theoretic pruning optimization is applied to remove redundant short branches, finally outputting the crack skeleton image Fig. 7(h).



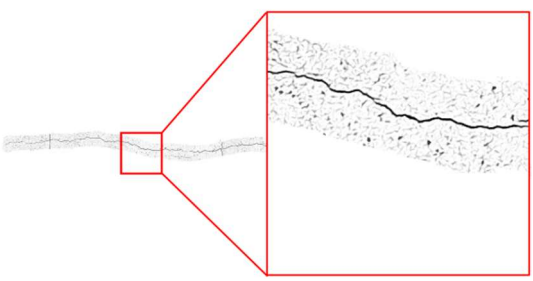
(a) Original image



(b) Image preprocess

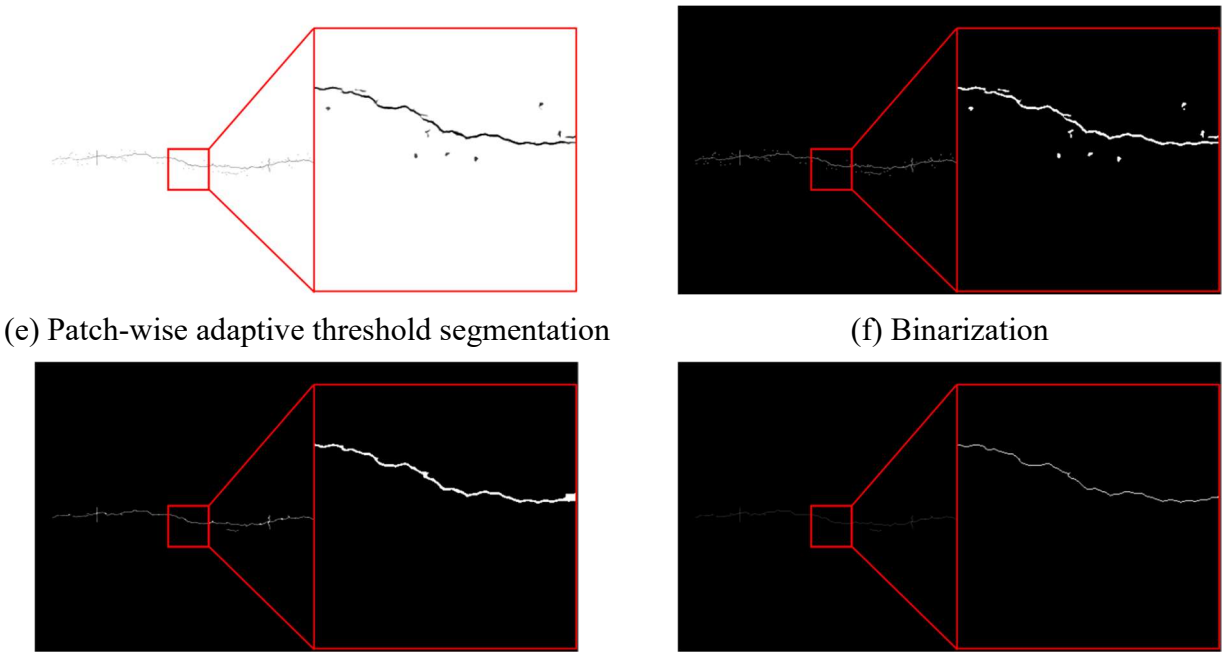


(c) Frangi filter



(d) Crop crack regions

ISSN: 2414-1895 DOI: 10.6919/ICJE.202605_12(5).0009



(e) Patch-wise adaptive threshold segmentation (f) Binarization

(g) Morphological process (h) Skeleton extraction

Fig. 7 Flowchart of Crack Skeleton Extraction

3.3 Crack Quantification Method

The crack length and width quantification process (Fig. 8) is based on skeletonization of the crack binary image, with crack length calculated through 8-connectivity analysis. The crack length expression is Eq. 3.

$$L=\sum_{i=1}^{n}\left(\sum_{j=1}^{m}S(i,j)\right) \tag{3}$$

In the formula, S(i,j) is the pixel value of the skeleton binary image. On this basis, each connected component is further traversed, and diagonal connections and straight connections are distinguished respectively to make the calculation result more accurate.

Crack width measurement adopts a method combining multi-directional search and local window analysis, with the calculation formula as Eq. 4:

$$w=\max\left(\sum_{k=1}^{36}\left(d_k^{+}+d_k^{-}\right)\right) \tag{4}$$

In the formula, d_k^{+} and d_k^{-} represent the positive and negative distances along angle k, respectively. This method searches in 36 directions and calculates the sum of positive and negative direction distances to ensure the accuracy of width measurement. Finally, distance transform is used to verify the measurement results to ensure the reliability of width measurement.

The above length and width calculation results are in pixels. To convert them to actual physical scales, this paper performs scale calibration based on the known geometric dimensions of the segments. This

ISSN: 2414-1895 DOI: 10.6919/ICJE.202605_12(5).0009

calibration method requires no external targets and directly uses the segment's own dimensions as a reference, possessing strong engineering practicality and operational simplicity.

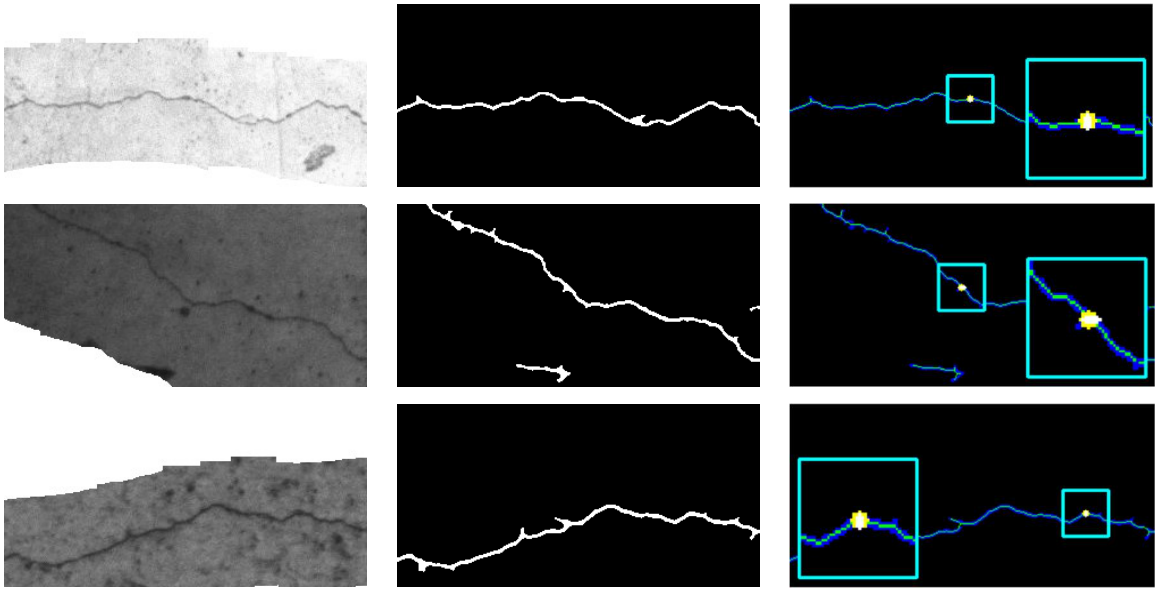


Fig. 8 Crack length and width measurement

Crack orientation quantification (Fig. 9) employs Principal Component Analysis (PCA). First, all points in the binarized crack mask are extracted as the set $X = \{(x_i,y_i)\}_{i=1}^{N}$, where N is the total number of crack mask pixels. Subsequently, the crack data is centralized by calculating the centroid $\bar{x}=(\bar{x},\bar{y})^{\top}=\frac{1}{N}\sum_{i=1}^{N}(x_i,y_i)^{\top}$, so that the centralized coordinates are $X_c=\{(x_i-\bar{x},y_i-\bar{y})\}_{i=1}^{N}$. On this basis, the covariance matrix $C = \frac{1}{N-1}X_c^{\top}X_c$ is constructed, and its eigenvalue decomposition is solved. The first principal component direction is the unit eigenvector corresponding to the maximum eigenvalue, which characterizes the axis along which the crack pixels extend with the maximum variance in space, i.e., the overall orientation of the crack. As shown in Fig. 9, this crack orientation belongs to a circumferential crack. This method identifies the geometric principal axis of cracks in a data-driven manner, avoiding subjectivity in manual interpretation and achieving objective quantification of orientation parameters.

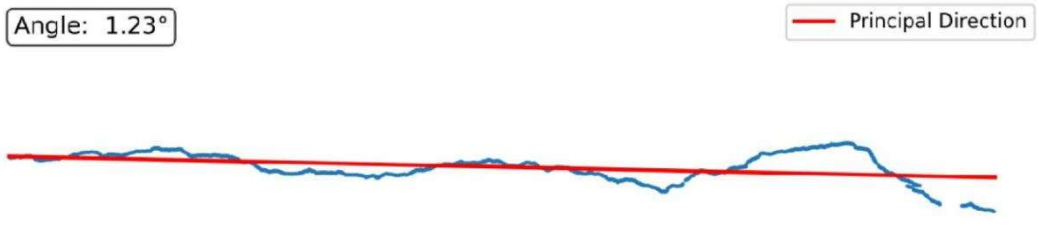


Fig. 9 Crack orientation measurement

4. Experimental Design

4.1 Dataset Processing

The crack dataset was constructed based on real metro tunnel images captured by CCD industrial cameras installed on a mobile monitoring vehicle (as shown in Fig. 10). The camera system consists of 8 independent cameras and auxiliary lighting systems, with a total coverage angle of 290 degrees.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202605_12(5).0009

The camera resolution is 4096 x 2168 pixels, the lens focal length is 16 mm, the effective camera viewing angle is 36.25 degrees, the camera overlap angle is 5.7 degrees, and the overlap length is 0.23 m, capable of covering the field of view range above the track slab for all shield segment angles.

Fig. 10 Mobile inspection vehicle

The dataset was annotated by domain experts for cracks. Due to the long and thin characteristics of cracks, appropriate surrounding regions were selected during annotation to ensure complete crack segmentation. Meanwhile, targets with characteristics similar to cracks, such as pipelines and equipment, were also annotated. This study randomly divided the dataset into training, validation, and test sets in an 8:1:1 ratio. The crack dataset includes cracks in various environments such as blur and occlusion, enabling the model to comprehensively adapt to the metro tunnel environment and enhancing the model's segmentation performance in tunnels.

Due to the extremely high image resolution, to save computational resources during crack detection, the dataset images were segmented into 640×640 patches, with sliding-window processing applied at the bottom and right edges. As shown in Fig. 11, the blue grid represents standard segmentation regions, and the yellow grid represents sliding-window cropping regions. Images containing cracks but with crack area less than 5% of the image area were deleted after cropping. After cropping, the dataset contains a total of 10584 images and their corresponding annotation files.

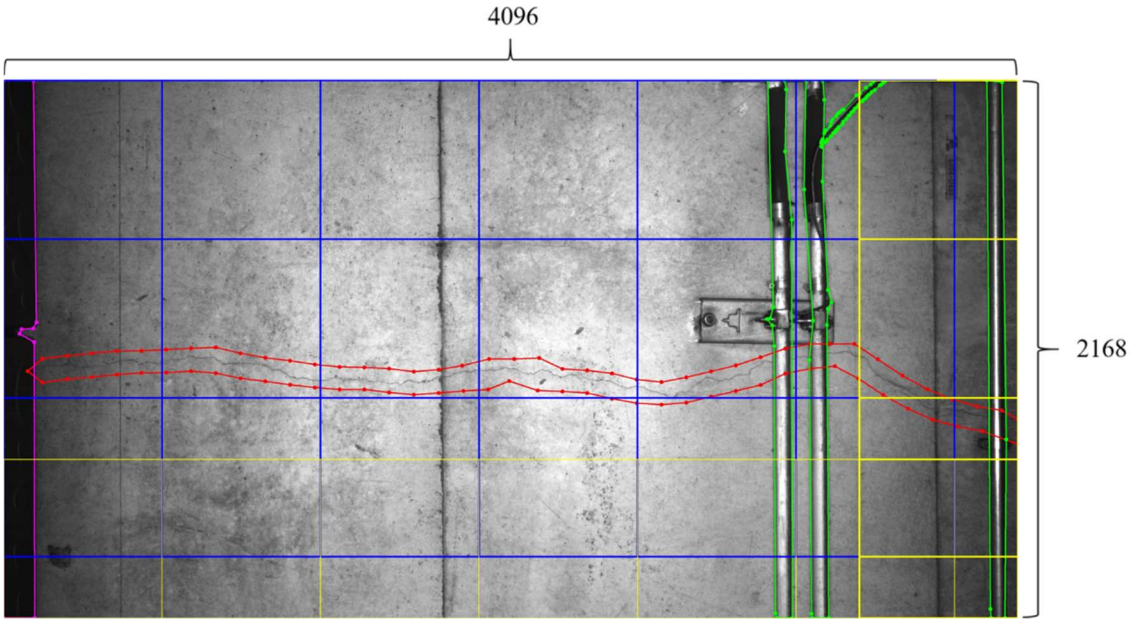


Fig. 11 Dataset cropping example

ISSN: 2414-1895 DOI: 10.6919/ICJE.202605_12(5).0009

4.2 Experimental Hardware and Software Environment

All experiments in this paper were conducted on a computer equipped with an i9-13900K processor, 128GB memory, and Nvidia GeForce RTX 4090. The operating system is Ubuntu 24.04. In terms of software environment, Python version 3.10.12 and PyTorch version 2.2.1 were used.

4.3 Model Evaluation Metrics

This paper adopts Precision (P), Recall (R), F1-Score, and Intersection over Union (IoU) as model performance evaluation metrics. The calculation formulas for each metric are shown in Eqs. (5) to (8):

$$P = \frac{TP}{TP + FP} \tag{5}$$

$$R = \frac{TP}{TP + FN} \tag{6}$$

$$F1 = \frac{2 \times P \times R}{P + R} = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{7}$$

$$IoU == \frac{TP}{TP + FP + FN} \tag{8}$$

In the formulas, TP represents the number of positive samples correctly predicted as positive by the model, FP represents the number of negative samples incorrectly predicted as positive by the model, and FN represents the number of positive samples incorrectly predicted as negative by the model. Precision P represents the proportion of samples predicted as cracks that are truly cracks. Recall R represents the proportion of all real crack samples that are correctly predicted as cracks by the model. F1-Score is the harmonic mean of precision and recall, used to comprehensively balance the performance of both, providing a single metric to evaluate the overall performance of the model, avoiding bias caused by solely pursuing high precision or high recall. The closer its value is to 1, the more ideal the model. Intersection over Union (IoU) measures the overlap between the predicted region and the true region, commonly used in segmentation tasks. The higher its value, the better the coincidence between the predicted region and the true region.

4.4 Ablation Study

Ablation study results show that each improved module forms an effective balance between performance and lightweighting (Table 1). The LSCD detection head improves recall and IoU while reducing parameters and computational cost, demonstrating good feature sharing and generalization capabilities. StarNet significantly compresses model complexity, but when used alone, it somewhat weakens feature expression. After introducing C3k2-S, precision is significantly improved, indicating that its multiplicative gate mechanism effectively enhances crack detail representation. Further combining the MFM module achieves overall improvement in F1 and IoU while maintaining the lowest parameters and computational cost, enabling the model to reach a better balance between accuracy and efficiency.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202605_12(5).0009

Table 1. Performance evaluation of the ablation study

Strategy	LSCD	StarNet	C3k2-S	MFM	P	R	F1	IoU	GFLOPs	Params
Baseline	×	×	×	×	0.7633	0.5690	0.6520	0.4837	10.2	2.8M
Imp. 1	√	×	×	×	0.7374	**0.6027**	0.6633	0.4962	9.1	2.6M
Imp. 2	√	√	×	×	0.7321	0.5601	0.6347	0.4648	7.8	1.9M
Imp. 3	√	√	√	×	**0.7928**	0.5564	0.6539	0.4858	7.8	1.9M
Imp. 4	√	√	√	√	0.7852	0.5622	**0.6552**	**0.4873**	**7.5**	**1.7M**

4.5 Comparative Experiments

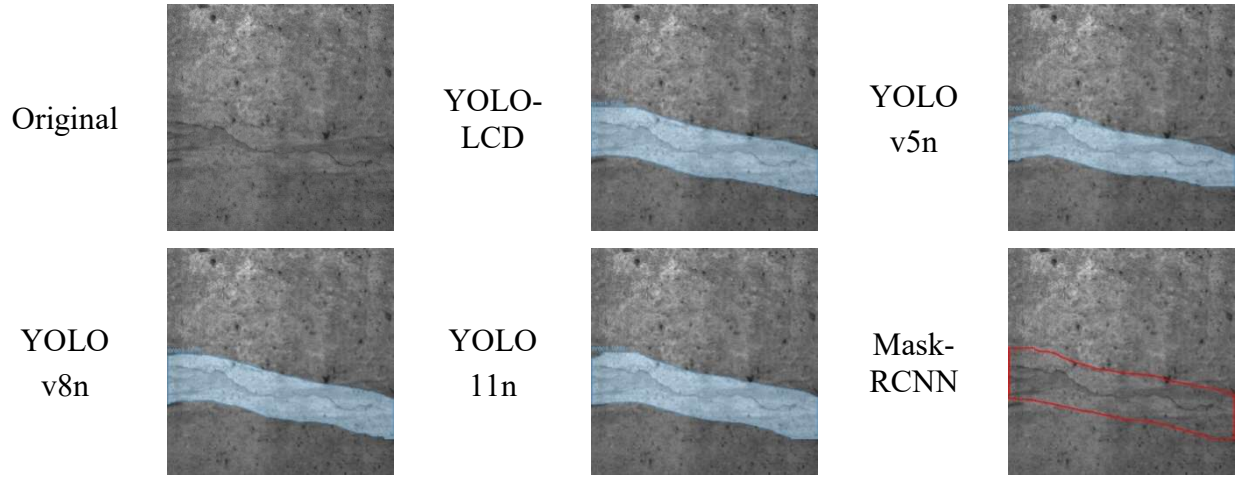


Fig. 12 Comparison with Traditional Models

Comparative experiment results (Table 2) show that YOLO-LCD has obvious advantages in terms of precision and efficiency trade-off. Its precision is superior to YOLOv5n and YOLO11n, and F1 and IoU show no significant gap compared to YOLOv5n and YOLOv8n. Although recall is slightly lower, the overall performance is at the same level. Meanwhile, its computational cost and parameter count are significantly lower than various YOLO series models, with outstanding lightweight advantages. In comparison, although Mask R-CNN has higher recall and F1, its computational complexity is too high for practical engineering deployment. As can be seen from Fig. 12, YOLO-LCD shows no significant gap compared with other traditional models in crack segmentation accuracy, and can significantly reduce computational overhead while ensuring accuracy, better meeting the practical needs for real-time performance and resource constraints in engineering applications.

Table 2. Comparison with Traditional Models

Models	P	R	F1	IoU	GFLOPS	Parameters
YOLO-LCD	0.7852	0.5622	0.6552	0.4873	**7.5**	**1.7M**
YOLOv5n	0.7728	0.5701	0.6562	0.4883	11	2.8M
YOLOv8n	**0.7902**	0.5704	0.6625	0.4954	12.1	3.3M
YOLO11n	0.7633	0.5690	0.6520	0.4837	10.2	2.8M
Mask-RCNN	0.5990	**0.7570**	**0.6688**	**0.5024**	186	44.0M

ISSN: 2414-1895 DOI: 10.6919/ICJE.202605_12(5).0009

5. Conclusion

This paper addresses the requirements for accuracy and efficiency in crack detection of metro shield tunnel linings. Through the construction of a metro shield tunnel crack disease dataset, model comparative experiments, and quantitative analysis, the following core conclusions are obtained:

(1) A lightweight instance segmentation model named YOLO-LCD based on improved YOLO11 is proposed, and an integrated method for crack extraction and geometric quantification is constructed. By introducing StarNet, C3k2-S, MFM, and LSCD modules, the model's parameter count and computational complexity are effectively reduced while enhancing the representation capability for fine cracks. Experimental results show that the proposed model achieves significant lightweighting while maintaining segmentation performance comparable to mainstream models, demonstrating good potential for engineering applications.

(2) A set of methods for crack extraction and quantification based on crack masks is constructed. Through image enhancement and illumination correction, Frangi filter-based linear structure enhancement, instance segmentation mask image cropping, connected component gray-level-based background removal, patch-wise adaptive binarization, morphological processing, skeleton extraction, and multi-directional search quantification, crack extraction and calculation of crack length, width, and orientation are achieved.

References

[1] Kamaliardakani M, Sun L, Ardakani M, et al. Sealed-Crack Detection Algorithm Using Heuristic Thresholding Approach[J]. Journal of Computing in Civil Engineering, 2016, 30(1), Article 04014110.

[2] Wang Q, Zhang N, Jiang K, et al. Tunnel Lining Crack Recognition Based on Improved Multiscale Retinex and Sobel Edge Detection[J]. Mathematical Problems in Engineering, 2021, 1-9, Article 9969464.

[3] Duan S, Zhang M, Xiong J, et al. DSYOLO: A dynamic snake convolution-based segmentation model for tunnel lining cracks in complex environments[J]. Engineering Structures, 2025, 344: 121325.

[4] Liu Y, Yao J, Lu X, et al. DeepCrack: A deep hierarchical feature learning architecture for crack segmentation[J]. Neurocomputing, 2019, 338: 139-153.

[5] Liu Z, Cao Y, Wang Y, et al. Computer vision-based concrete crack detection using U-net fully convolutional networks[J]. Automation in Construction, 2019, 104: 129-139.

[6] Zhao S, Zhang D, Xue Y, et al. A Deep Learning-Based Approach for Refined Crack Evaluation from Shield Tunnel Lining Images[J]. Automation in Construction, 2021, 132: 103934.

[7] Xu H, Wang M, Liu C, et al. Automatic detection of tunnel lining crack based on mobile image acquisition system and deep learning ensemble model[J]. Tunnelling and Underground Space Technology incorporating Trenchless Technology Research, 2024, 154: 106124.

[8] Duan S, Zhang M, Qiu S, et al. Tunnel lining crack detection model based on improved YOLOv5[J]. Tunnelling and Underground Space Technology incorporating Trenchless Technology Research, 2024, 147: 105713.

[9] Dang D, Gong L, Jin C, et al. SMI-YOLOv8: Intelligent detection of tunnel lining cracks via multiscale feature attention fusion[J]. Measurement, 2026, 257: 118833.

[10] Ma X, Dai X, Bai Y, et al. Rewrite the Stars[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.

[11] Zhang Y, Zhou S, Li H. Depth Information Assisted Collaborative Mutual Promotion Network for Single Image Dehazing[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.

[12] Zheng C, Liu L, Fu Q, et al. YOLO-DD: a lightweight framework for UAV detection in complex environments via boundary-aware fusion[J]. EURASIP Journal on Advances in Signal Processing, 2025, 44: 1-15.

[13] Land E H, McCann J J. Lightness and Retinex Theory[J]. Journal of the Optical Society of America, 1971, 61(1): 1-11.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202605_12(5).0009

[14] Frangi A F, Niessen W J, Vincken K L, et al. Multiscale vessel enhancement filtering[C]//Proceedings of the First International Conference on Medical Image Computing and Computer-Assisted Intervention. Cambridge: Springer, 1998: 130-137.

[15] Zhang T Y, Suen C Y. A fast parallel algorithm for thinning digital patterns[J]. Communications of the ACM, 1984, 27(3): 236-239.