2025 Volume 3, Issue 1
Big Data and Cloud Innovation
ISSN (Online): 2529-7805

Optimization Research of Lightweight YOLOv8 Model in Building Crack Detection

Chenyuan Liang[1], Yuan Gao[1]*, Bingzhang Huang[2,3], Qing Deng[4]

[1]School of Automation, Guangxi University of Science and Technology, Liuzhou 545616, Guangxi, China

[2]School of Civil and Architectural Engineering, Liuzhou Institute of Technology, Liuzhou 545026, Guangxi, China

[3]Guangxi Zhuang Autonomous Region Prefabricated Building Life Cycle Management and Virtual Simulation Engineering Research Center, Liuzhou 545026, Guangxi, China

[4]School of Economics and Management, Liuzhou Institute of Technology, Liuzhou 545026, Guangxi, China

Corresponding author: Yuan Gao, gxgaoy98@21cn.com

Abstract:

In order to improve the detection accuracy and efficiency of building cracks based on the YOLOv8 model, this paper proposes an improved YOLOv8 model. The improved model incorporates the SlimNeck structure, the CSPELAN4 designed based on the GELAN architecture, and the InnerGIoU loss function respectively. Then, an experimental comparative study of this model in building crack detection is carried out. The experimental results show that the precision *P* increases by 2.1%, the recall rate *R* increases by 4.2%, *mAP@0.5* increases by 2.3%, and *mAP@0.5:0.95* increases by 6.0%. At the same time, *Params* and *GFLOPs* are reduced by 21.6% and 23.5%, respectively.

Keywords:

Crack detection
YOLOv8
Lightweight model
SlimNeck
GELAN
InnerGIoU loss function

Online publication: March 4, 2025

1. Introduction

Building cracks, a key sign of building structural health, can expose problems like material aging, foundation settlement, and structural deformation [1,2]. They are often elongated, irregular, and low-contrast small areas in images, demanding high visual feature extraction from detection models [3].

With deep learning progress, YOLO (You Only Look Once) models are mainstream for real-time detection as they can finish target detection in one forward pass. Scholars have optimized early YOLO models for building crack detection. Li *et al.* enhanced YOLOv3 with depthwise separable convolutions, improving the detection of fine and complex cracks [4]. He *et al.* combined YOLOv4 with deformable convolutions for better accuracy and stability [5]. Huang *et al.* proposed a C_CB module to boost the network's feature expression and retain more crack-edge info during downsampling [6].

Existing research indicates YOLOv8 has improved building crack detection accuracy and cut computational complexity [7]. However, it still has issues like low accuracy in complex environments, inability to detect fine cracks precisely, and inefficient deployment in resource-limited settings. Thus, further optimizing YOLOv8 to meet lightweight and accuracy demands in building crack detection is essential [8].

This paper proposes an improved YOLOv8 model to offer an effective idea for its lightweight design and enhance its computational efficiency and detection accuracy in building crack detection. Experimental comparisons verify the model's effectiveness and advancement in this area.

2. YOLOv8 model and its improvement

2.1. Improved YOLOv8 model structure

The YOLOv8 model has the structural feature of a single forward pass and mainly consists of four parts: Input, Backbone, Neck, and Head. The input image size is 640×640 [9].

Figure 1 shows the principle structure of the improved YOLOv8 model. The GSConv layer in SimNeck replaces Conv in layers 16 and 19 of the original YOLOv8. The VoVGSCSP layer replaces layers 18 and 21. To boost detection of fine building cracks in complex backgrounds, the CSPELAN4 layer from GELAN architecture replaces the C2f layer in layers 0–15. To enhance regression accuracy and generalization for complex crack detection, the InnerGIoU loss function replaces the CIoU loss function of the original model.

2.2. Design of SlimNeck structure

The SlimNeck structure in this paper's design is based on the GSConv operation module, the GS-bottleneck layer, and VoVGSCSP layer. In GSConv, depthwise separable convolution (DWConv) cuts computation while keeping key feature info. **Figure 2** shows the structural principles of the VoVGSCSP layer and GS-bottleneck. This layer combines the GS-bottleneck layer with multiple Conv layers for feature map information fusion, boosting

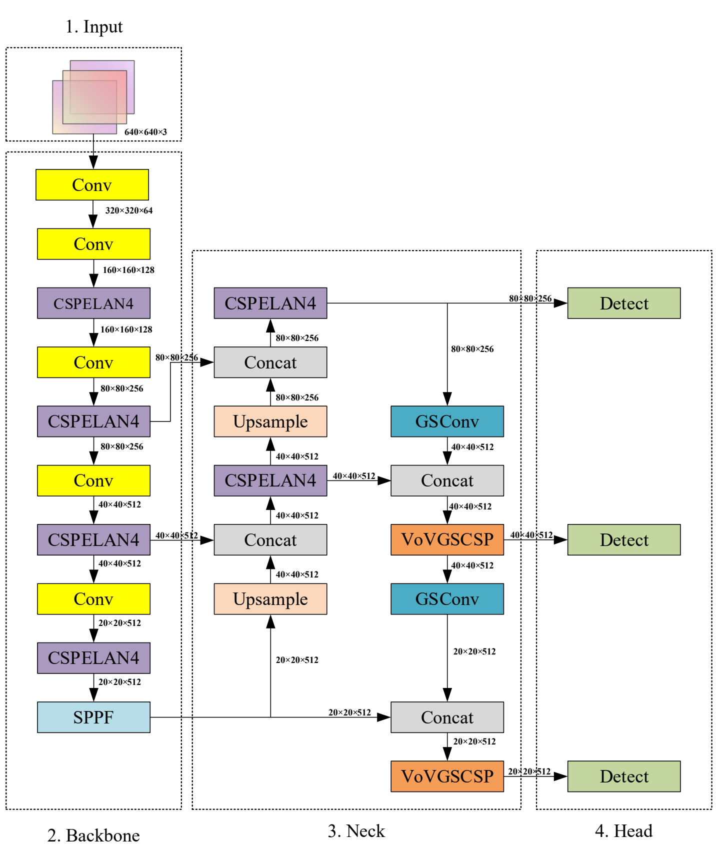

Figure 1. Improved YOLOv8 model structure

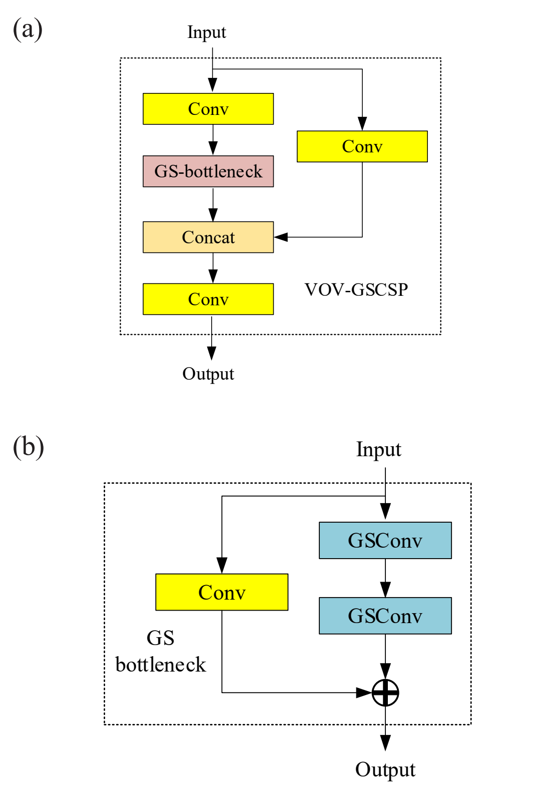

Figure 2. (a) VoVGSCSP layer; (b) GS-bottleneck layer

feature reuse efficiency.

2.3. Design of GELAN architecture

Figure 3 shows the schematic diagrams of the CSPNet architecture, the ELAN architecture, and the GELAN architecture, respectively. By comparing **Figure 3**, it can be seen that the GELAN architecture can be regarded as a combination of CSPNet and ELAN, having both the segmentation and recombination function of CSPNet and the hierarchical stacking function of ELAN [10].

2.4. Design of InnerGIoU loss function

The calculation formula of the InnerGIoU loss function is as follows:

$$GIoU = 1 - IoU - \frac{\left|C - B^{p} \cap B^{gt}\right|}{\left|C\right|} \tag{1}$$

$$InnerGIoU = GIoU + IoU - InnerIoU \tag{2}$$

In the formula, C is the smallest enclosing area covering B^{p} and B^{gt}. The InnerGIoU loss function uses a scaling factor ratio in [0.5, 1.5]. When ratio > 1, the predicted bounding box is smaller than the ground truth one, aiding high-IoU sample convergence; when ratio < 1, it is larger, better for low-IoU sample convergence [12].

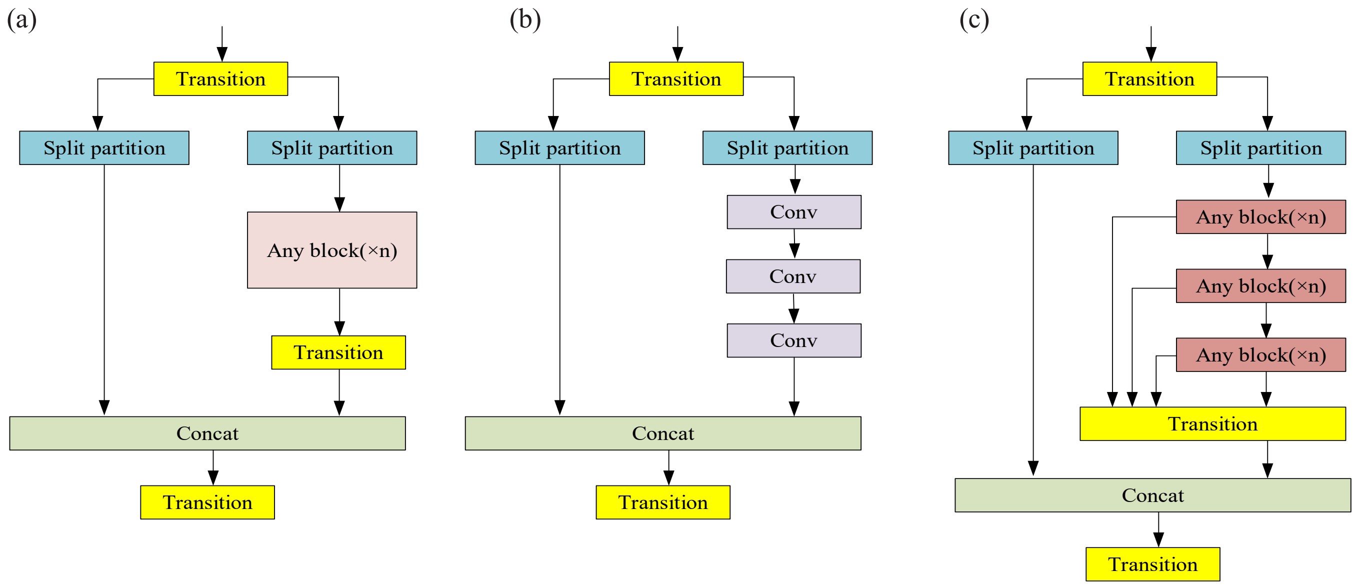

Figure 3. (a) CSPNet architecture; (b) ELAN architecture; (c) GELAN architecture

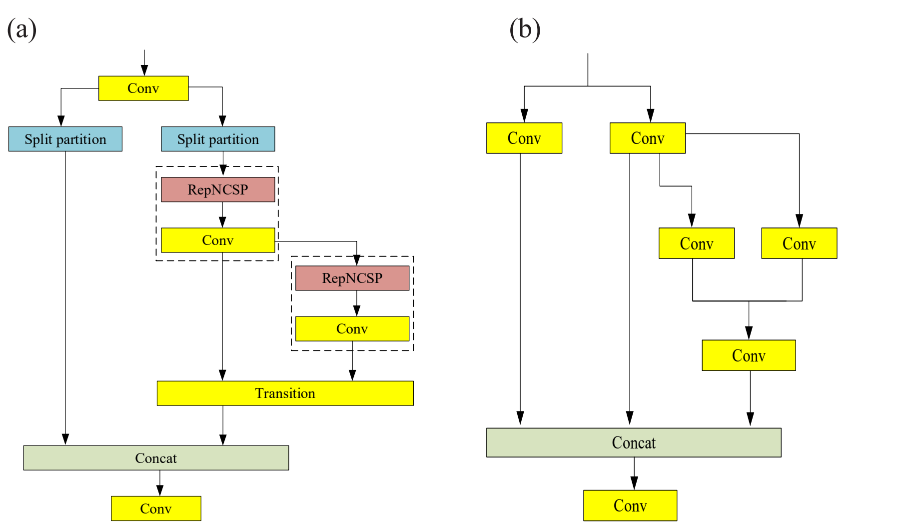

Figure 4(a) shows the structural principle diagram of the CSPELAN4 layer designed based on the GELAN architecture in this paper [11]. The CSPELAN4 layer uses two RepNCSP layers and a Conv layer for stacking to extract deeper and more abstract features. The structural principle of the RepNCSP layer is shown in **Figure 4(b)**.

3. Experimental environment and analysis of experimental results

3.1. Dataset

To verify the improved YOLOv8n model's effectiveness and correctness, an experimental dataset of 3,888 finely-annotated building crack images was used, covering horizontal and vertical cracks. The dataset annotation was done on the Roboflow platform and went through data augmentation like rotation, scaling, and brightness adjustment.

3.2. Experimental result evaluation indicators

The evaluation indicators of this experiment mainly include crack detection precision *P*, recall rate *R*, average precision *mAP@0.5*, average precision *mAP@0.5:0.95*, *Params*, *GFLOPs*, and *FPS*. The specific formulas are as follows:

$$P = \frac{TP}{TP + FP} \tag{3}$$

$$R = \frac{TP}{TP + FN} \tag{4}$$

$$AP = \int_0^1 P(R)dR \tag{5}$$

$$mAP = \frac{\sum_{c=1}^{C} AP_c}{C} \tag{6}$$

P is precision, the ratio of correctly predicted crack samples in all predicted ones. *R* is recall rate, the ratio of correctly predicted crack samples in all actual crack ones. *TP* are correctly predicted crack samples, *FP* are wrongly predicted crack samples, *FN* are wrongly predicted non-crack samples, *C* is the number of crack categories, *AP* is the average precision of a certain crack category, and *mAP* is that of all crack categories.

3.3. Ablation experiment

To show the advantages of the improved model in building crack detection, YOLOv8n, the smallest original YOLOv8 model fitting this paper's lightweight research, was chosen. Ablation experiments were done on the model and loss function improvement schemes. The SlimNeck structure, CSPELAN4 layer from GELAN, and InnerGIoU loss function were introduced one by one for ablation. The results are presented in **Table 1**.

Comparing performance and resource data in **Table 1**, compared to the original YOLOv8n, precision *P* rose by 2.1%, recall rate *R* by 4.2%, *mAP@0.5* by 2.3%, and *mAP@0.5:0.95* by 6.0%, improving overall performance. It has better generalization than a single module. Plus, *Params* and *GFLOPs* decreased by 21.6% and 23.5%, respectively, with an *FPS* of 172.2.

3.4. Comparison experiment of mainstream models

To further verify the performance of the improved model in this paper compared with other models, comparative experimental evaluations were carried out using current mainstream models such as Faster-RCNN, YOLOv5n, and YOLOv6n. The relevant experimental results are shown in **Table 2**.

By comparing the performance indicators and data in **Table 2**, the improved YOLOv8n model not only has good lightweight characteristics and maintains a fast detection speed but can also obtain high detection accuracy.

Table 1. Performance and resource evaluation of ablation experiment models

Model	*P*	*R*	*mAP@0.5*	*mAP@0.5:0.95*	*Params*(M)	*GFLOPs*	*FPS*(s)
YOLOv8n	0.793	0.766	0.799	0.567	3.01	8.1	367.6
+Slim-Neck	0.855	0.729	0.822	0.594	2.80	7.3	238.6
+GELAN	0.840	0.758	0.814	0.581	2.64	7.2	221.3
+InnerGIoU	0.825	0.743	0.808	0.588	3.01	8.1	364.6
Improved YOLOv8n	0.810	0.798	0.817	0.601	2.36	6.2	172.2

Table 2. Performance and resource evaluation of mainstream models in comparative experiments

Model	*P*	*R*	*mAP@0.5*	*mAP@0.5:0.95*	*Params*(M)	*GFLOPs*	*FPS*(s)
Faster-RCNN	0.677	0.876	0.918	0.613	136.7	369.8	29.4
YOLOv5n	0.792	0.748	0.801	0.493	1.76	4.1	372.8
YOLOv6n	0.828	0.713	0.817	0.571	4.73	11.4	273.5
YOLOv8n	0.793	0.766	0.799	0.567	3.01	8.1	367.6
Improved YOLOv8n	0.810	0.798	0.828	0.601	2.36	6.2	172.2

4. Conclusion

This study proposed an improved YOLOv8 model for high-accuracy and efficient building crack detection. It uses SlimNeck structure, CSPELAN4 layer, and InnerGIoU loss function for lightweight improvement. Results showed that the optimized model outperforms the original in *mAP@0.5:0.95*, computational complexity, and parameter number. Detection accuracy rises by 2.3% and recall by 4.2%, while model parameters and computational complexity drop by 21.6% and 23.5%, respectively. These boost the model's performance in fine-crack and complex-background detection. The model suits resource-constrained scenarios and offers references for further optimization.

Funding

Guike AB23026065; Guike AB22080083; 2022CCC0101

Disclosure statement

The authors declare no conflict of interest.

References

[1] Xu YC, Li HK, 2024, Research on Innovation and Application of Civil Engineering Structure Health Monitoring Technology. New Urban Construction Science and Technology, 33(10): 57–59.

[2] Zeng W, Gu LS, Chen SJ, et al., 2023, Research on the Influence of Fibers on the Mechanical Properties and Microstructure of Ultra-High Performance Concrete. Journal of Guangxi University of Science and Technology, 34(01): 20–27.

[3] Wang ZX, 2024, Causes Control and Treatment Methods of Concrete Cracks in Construction Engineering. Engineering Research and Practice, 5(7): 181–183.

[4] Li Y, Han Z, Xu H, et al., 2019, YOLOv3-Lite: A Lightweight Crack Detection Network for Aircraft Structure Based on Depthwise Separable Convolutions. Applied Sciences, 9(18): 3781.

[5] He Z, Su C, Deng Y, 2024, A Novel Hybrid Approach for Concrete Crack Segmentation Based on Deformable Oriented-YOLOv4 and Image Processing Techniques. Applied Sciences, 14(5): 1892.

[6] Huang KY, Zhao Y, Hu N, et al., 2024, Research on the Identification of Concrete Bridge Cracks by the YOLOv5 Network Based on the Attention Mechanism. Bulletin of Science and Technology, 40(09): 71–76.

[7] Wang D, Zhang Y, Zhang R, et al., 2024, Detection and Assessment of Post-Earthquake Functional Building Ceiling Damage Based on Improved YOLOv8. Journal of Building Engineering, 98: 111315.

[8] Dong X, Liu Y, Dai J, 2024, Concrete Surface Crack Detection Algorithm Based on Improved YOLOv8. Sensors, 24(16):

5252.

[9] Li QF, Zhang YH, Xie PP, et al., 2025, Rainy-Day Pedestrian Detection Method Based on Improved YOLOv8. Journal of Guangxi University of Science and Technology, 2025: 1–11.

[10] Zhang C, Chen X, Liu P, et al., 2024, Automated Detection and Segmentation of Tunnel Defects and Objects Using YOLOv8-CM. Tunnelling and Underground Space Technology, 150: 105857.

[11] Xu W, Zhu D, Deng R, et al., 2024, Violence-YOLO: Enhanced GELAN Algorithm for Violence Detection. Applied Sciences, 14(15): 6712.

[12] Wang CY, Yeh IH, Liao HYM, 2024, Yolov9: Learning What You Want to Learn Using Programmable Gradient Information. arXiv preprint, 2402.13616.