



ARTICLE

FDSC-YOLOv8: Advancements in Automated Crack Identification for Enhanced Safety in Underground Engineering

Rui Wang¹, Zhihui Liu^{2,*}, Hongdi Liu³, Baozhong Su⁴ and Chuanyi Ma⁵

¹Key Laboratory of Computing Power Network and Information Security, Ministry of Education, Shandong Computer Science Center (National Supercomputer Center in Jinan), Qilu University of Technology (Shandong Academy of Sciences), Jinan, 250353, China

²School of Qilu Transportation, Shandong University, Jinan, 250100, China

³Geotechnical and Structural Engineering Center, Shandong University, Jinan, 250061, China

⁴Shandong Lairong High-Speed Railway Co., Ltd., Weihai, 264200, China

⁵Shandong High-Speed Group Co., Ltd., Jinan, 250014, China

*Corresponding Author: Zhihui Liu. Email: zhihui.liu@126.com

Received: 19 February 2024 Accepted: 24 April 2024 Published: 08 July 2024

ABSTRACT

In underground engineering, the detection of structural cracks on tunnel surfaces stands as a pivotal task in ensuring the health and reliability of tunnel structures. However, the dim and dusty environment inherent to underground engineering poses considerable challenges to crack segmentation. This paper proposes a crack segmentation algorithm termed as Focused Detection for Subsurface Cracks YOLOv8 (FDSC-YOLOv8) specifically designed for underground engineering structural surfaces. Firstly, to improve the extraction of multi-layer convolutional features, the fixed convolutional module is replaced with a deformable convolutional module. Secondly, the model's receptive field is enhanced by introducing a multi-branch convolutional module, improving the extraction of shallow features for small targets. Next, the Dynamic Snake Convolution module is incorporated to enhance the extraction capability for slender and weak cracks. Finally, the Convolutional Block Attention Module (CBAM) module is employed to achieve better target determination. The FDSC-YOLOv8s algorithm's mAP50 and mAP50-95 reach 96.5% and 66.4%, according to the testing data.

KEYWORDS

Crack segmentation; improved YOLOv8; engineering applications; feature extraction

1 Introduction

The evaluation of underground engineering structural surfaces plays a crucial role in maintaining tunnels' structural health and reliability. During tunnel construction, the disruption of surrounding rock stress equilibrium caused by excavation [1] often leads to the formation of cracks on structural surfaces. Without timely intervention measures for maintenance, the severity of these cracks can escalate, resulting in significant tunnel deformation and collapse accidents, leading to casualties,



economic losses, and project delays [2,3]. As a result, identifying nearby rock fissures and quickly gaining a knowledge of their characteristics are crucial for guaranteeing building site safety.

Since computer technology has advanced, deep learning methods for extracting crack information have gained a lot of interest. Crack identification can be thought of as a fundamental case of line recognition in the field of computer vision since cracks typically display linear or curved structures [4–6]. In general, there are two main groups of algorithms for detecting cracks: the first one is based on conventional picture processing techniques [7] while the other one is based on deep learning techniques. The creation of different filters, like edge detectors, which are totally dependent on the intrinsic properties of optical pictures, is essentially what the classic image processing approaches are all about [8]. In spite of the considerable research conducted on these approaches, their drawbacks are also obvious. Limited by the basic principle of these traditional image processing techniques, it is often difficult to identify irregular targets against intricate backgrounds environments and poor lighting conditions [9]. In other words, the robustness of the traditional image processing techniques is not satisfactory. Deep learning techniques are currently being used extensively in computer vision [10–12], mainly because they have advantages in recognition speed and accuracy compared to traditional target recognition algorithms. Although there are some deep learning-based techniques for detecting cracks in subterranean structures, they are still not very successful and need to be improved.

Within the field of automated crack detection, early studies proposed various methods. Zhao et al. [13] improved the traditional Canny algorithm to enhance the effectiveness of edge detection in road images. Li et al. [14] proposed a deep fusion strategy that combines cyclic residual convolution and context encoder networks to detect cracks, but this method has a relatively high requirement for image brightness. Talab et al. [15] applied the Sobel operator to filter and denoise concrete images, followed by Otsu threshold segmentation for crack edge detection. Although the Sobel operator has certain practicality in the detection of cracks on concrete surfaces, its potential issue lies in the thickening of crack edges, which in turn affects the accuracy of crack recognition. Furthermore, these traditional image processing algorithms are susceptible to disturbances such as lighting variations, stains, and debris, which can hinder the effectiveness and precision of crack detection.

The two main types are one-stage and two-stage methods in deep learning algorithms [16]. The creation of candidate regions containing geographic information about the target is a prerequisite for Two-Stage algorithms, which is the main distinction between the two. In recent years, several studies have proposed CNN-based methods for pavement disease detection. For example, a novel method for measuring and identifying road discomfort based on CNN technology was proposed by Sha et al. [17]. This method utilizes a CNN model to successfully achieve intelligent monitoring of road conditions, identify different types of road distresses, and provide accurate measurements of their dimensions and locations. Similarly, Wang [18] studied the use of the Faster R-CNN algorithm for automatic concrete crack identification. She extracts crack features using the Zeiler&Fergus Net (ZFNet) network and then detects cracks in cement concrete pavements using Faster R-CNN, enabling real-time crack localization from coarse to fine. All things considered, these CNN-based techniques have a lot of potential for effective and precise pavement disease diagnosis.

In terms of real-time performance and detection speed, one-stage algorithms have a considerable advantage over two-stage algorithms. Xie et al. [19] proposed an improved YOLOv5 model that incorporates multi-scale and multi-level information fusion, along with a multi-scale channel attention module. This model achieved a detection accuracy of 60.24% for small targets on UAVs, with a detection time of only 16 ms. To ensure the model's adaptability to multi-size targets and detection

accuracy, Lei et al. [20] suggested an updated YOLO v3 model technique for rural land object identification and categorization. A unique convolutional neural network named MYOLOv3-Tiny was proposed by Xu et al. [21] for the real-time detection of rail fasteners, and it achieved a 99.32% detection accuracy. Chen et al. [22] developed a defect detector based on YOLOv3 for the inspection of surface-mounted LED chips. In their approach, they introduced DenseNet as the backbone network, replacing the original Darknet-53.

These aforementioned studies provide effective models for target detection in many situations. However, they are almost designed for overground applications such as pavement and bridges. The environment in the underground space is much more complex than the ground. For example, the images of cracks captured in underground space are usually affected by poor lightness, heavy dust, and high humidity. Therefore, the models for crack detection in underground space should be specially designed to fit these factors.

The challenges for crack segmentation in the underground engineer are manifold. One the one hand, significant obstacles and constraints for related research activities are brought about by the current lack of publicly available datasets specifically designed to address structural surface cracks in subsurface engineering. On the other hand, in the captured images, the cracks are usually slender lines or curves, occupying only a relatively small area in the image, and thus are prone to the problem of information loss during feature extraction. Additionally, the cracks usually have irregular shapes and sizes and are interfered by light, noise, and other factors. Taking into account the aforementioned challenges, this paper presents an enhanced model for crack segmentation in underground engineering, which capitalizes on the strengths of the YOLOv8s model [23]. The architecture of this model has been carefully designed based on the unique features of underground cracks, leading to a notable improvement in identification accuracy. The following are this paper's primary contributions:

1. This study establishes a dataset for crack segmentation in underground engineering, gathering rich and authentic tunnel scene data, thereby providing a benchmark for evaluating algorithm performance.
2. The presented model introduces the fusion method of YOLOv8 and Dynamic Snake Convolution (DSConv) in the field of crack segmentation, so that the model is more suitable to the slender morphology, and the crack segmentation capability is effectively improved.
3. To improve crack segmentation, the paper use the multi-branch convolutional module Receptive Field Module (RFB) to enhance feature extraction and replace fixed convolutional modules with Deformable Convolution for better capturing crack complexities. Additionally, it integrate a Convolutional Block Attention Module (CBAM) module to improve focus and accuracy on crack regions.

2 Crack Segmentation Model

2.1 Baseline

YOLOv8, the latest iteration of the YOLO network family, introduces numerous advancements and innovations that build upon the success of previous YOLO versions, ultimately enhancing its overall performance and adaptability [24]. Illustrated in Fig. 1a, the network architecture of YOLOv8 is comprised of three distinct components: the Backbone, Neck, and Head.

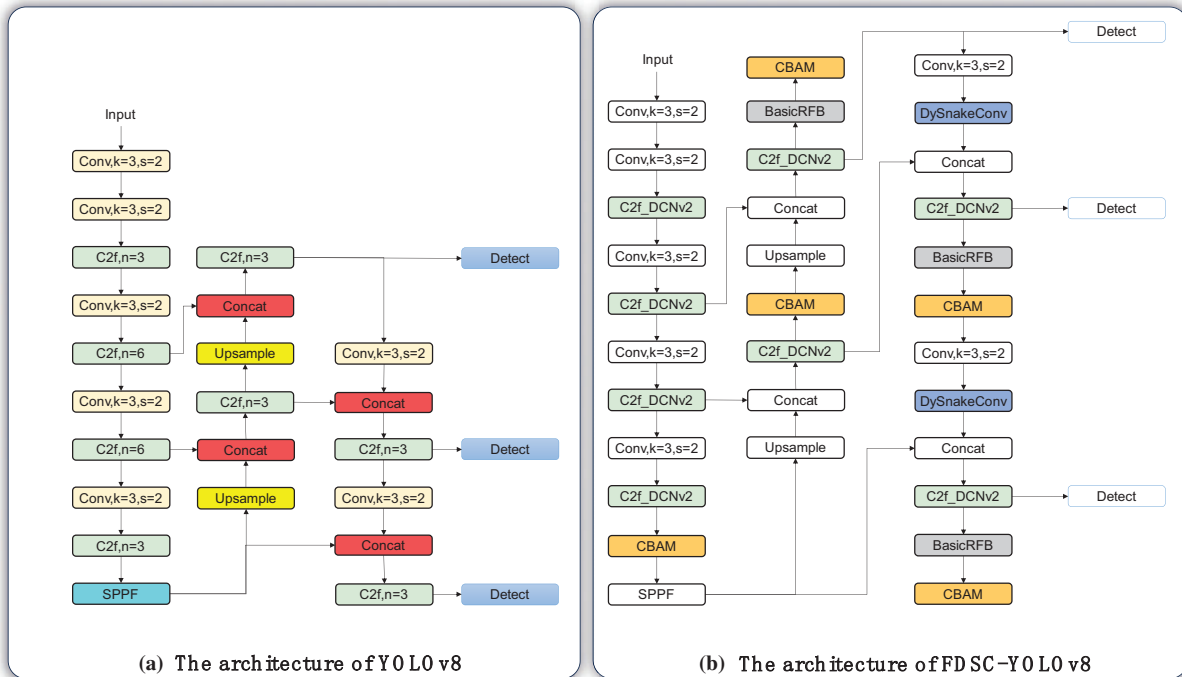


Figure 1: (a) The architecture of YOLOv8; (b) The architecture of Focused Detection for Subsurface Cracks YOLOv8

2.2 Improve with YOLOv8

Existing target segmentation algorithms are effective for crack segmentation but still have drawbacks in the underground engineer. For better segmentation of cracks, this study have improved the YOLO algorithm and termed it Focused Detection for Subsurface Cracks YOLOv8 (FDSC-YOLOv8), which can extract cracks quickly and accurately, and the details of the algorithm are described below.

In the FDSC-YOLOv8 model, the standard convolution in Convolution to FullyConnected (C2f) is replaced by Deformable Convnets v2 (DCNv2) on the original Backbone to form C2f-DCNv2, which can better capture the deformation of the cracks. The RFB is introduced in the Head part to add a BasicRFB layer, which enhances the model sensing field. To bolster the crack feature extraction capabilities, the DySnakeConv layer is constructed using the DSConv. Lastly, the CBAM module is integrated into the YOLOv8s to prioritize crucial features. The enhanced model's structure is depicted in Fig. 1b.

2.2.1 Modifying the Backbone Network

The intricacies of crack segmentation tasks necessitate heightened sensitivity to target deformations and intricate details. Conventional convolution operations may lack the requisite flexibility to effectively handle such scenarios, resulting in a reduction in target localization accuracy [25]. Hence, we opt to incorporate Deformable Convnets v2 (DCNv2) [26] to augment the backbone network of YOLOv8.

Deformable convolution is a type of convolution operation that dynamically adjusts the shape of the kernel by learning offset for the center position within the receptive field. In comparison to conventional fixed convolution kernels, deformable convolution exhibits enhanced adaptability, proving more adept at accommodating the deformations intricate structures of the target.

In DCNv2, the following is the calculation formula for the output eigenvalue $F(x)$:

$$F(x) = \sum_{n=1}^N \omega_n \cdot y(x + x_n + \Delta x_n) \cdot \Delta m_n \tag{1}$$

where Δx_n and Δm_n stand for the modulation scalar and learnable offset, respectively, for the k -th site. Compared to DCNv1, DCNv2 introduces not only the offset of each sampling point but also a weight coefficient Δm_k to discern whether the introduced region is of interest to us.

The incorporation of deformable convolution into the backbone network of YOLOv8 is pursued to enhance the model’s recognition of crack targets. In particular, the C2f module of YOLOv8 has undergone a modification where the traditional convolution has been substituted with DCNv2, resulting in the creation of the C2f-DCNv2 module. This operation not only introduces stronger nonlinear modeling capabilities to the model but also helps to better capture the changing shape and detailed information of cracks, as shown in the structural diagram in Fig. 2.

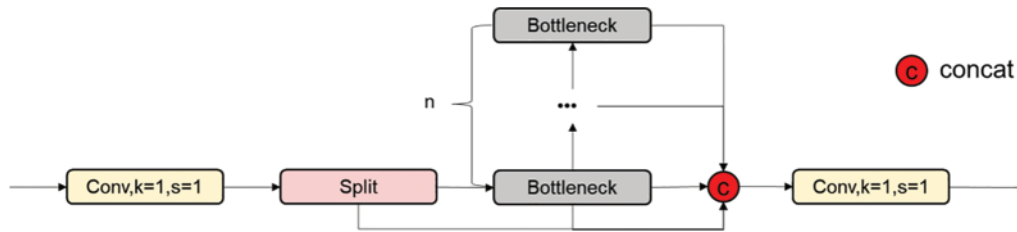


Figure 2: The architecture of C2f-DCNv2

2.2.2 Improved Neck Convolution Block

Following the aforementioned enhancements, there has been a notable improvement in segmentation accuracy. Nevertheless, challenges persist, particularly in dealing with small targets and low-resolution issues within crack segmentation. To address these issues, the adoption of the Receptive Field Block (RFB) module [27] into the Neck network has been selected. This strategic addition aims to amplify the extraction capabilities for small targets, thereby enhancing the model’s discernment of crack targets.

RFB is a dedicated module crafted to broaden the receptive field and enhance feature expression capabilities. Comprising multiple parallel branches [28], each adept at capturing feature information at various scales, this structural design empowers the model with a more exhaustive comprehension of the input image. This capability proves especially vital in tasks like crack segmentation, where dealing with a diverse array of morphological changes is imperative. The detailed structure of the RFB module is in Fig. 3.

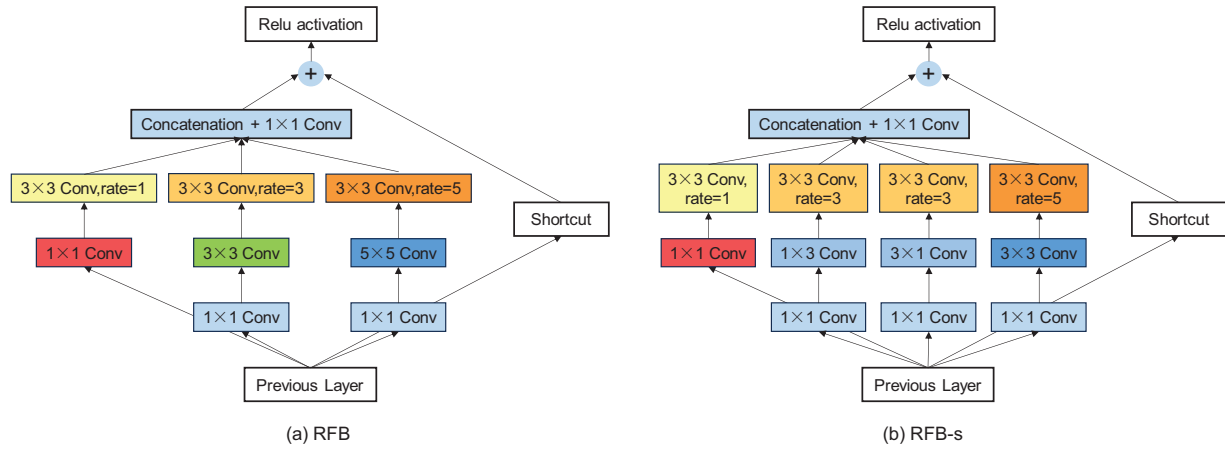


Figure 3: The architecture of RFB. (a) RFB; (b) RFB-s

The integration of the RFB module into the Neck section of YOLOv8 is chosen to extend the receptive field range for feature extraction. The introduction of an RFB module into the Neck enhances the model's capacity to capture contextual information about crack targets, thereby improving overall detection performance. Additionally, this paper have incorporated the receptive field module into the Head section of YOLOv8. This strategic inclusion aims to further enhance the perception of crack targets during the detection phase, ensuring more accurate localization and segmentation of cracks.

2.2.3 Improved Head Convolution Block

To further enhance the perception of complex shaped cracks and improve segmentation accuracy, Dynamic Snake Convolution (DSConv) [29] will be introduced into the Head layer of YOLOv8. During the detection phase, DSConv modifies the convolutional kernel's shape flexibly, improving the model's adaptability to target deformation and complicated structures. This is accomplished during the detection phase by dynamically modifying the convolution kernel's form. This addition is geared towards comprehensive optimization of the crack segmentation model, aiming to improve its adaptability and accuracy across various crack shapes. The DSConv module primarily tackles the difficulty in capturing subtle local structural details and dealing with intricate global morphological variations. The structural diagram of the module is illustrated in Fig. 4.

The convolution kernel in DSConv is linearized in both the x and y directions. The following is the change along the x-axis:

$$P_{i\pm k} = \begin{cases} (x_{i+k}, y_{i+k}) = (x_i + k, y_i + \sum_i^{i+k} \Delta y) \\ (x_{i-k}, y_{i-k}) = (x_i - k, y_i + \sum_{i-k}^i \Delta y) \end{cases} \quad (2)$$

The following is the change along the y-axis:

$$P_{j\pm k} = \begin{cases} (x_{j+k}, y_{j+k}) = (x_j + \sum_j^{j+k} \Delta x, y_j + k) \\ (x_{j-k}, y_{j-k}) = (x_j + \sum_{j-k}^j \Delta x, y_j - k) \end{cases} \quad (3)$$

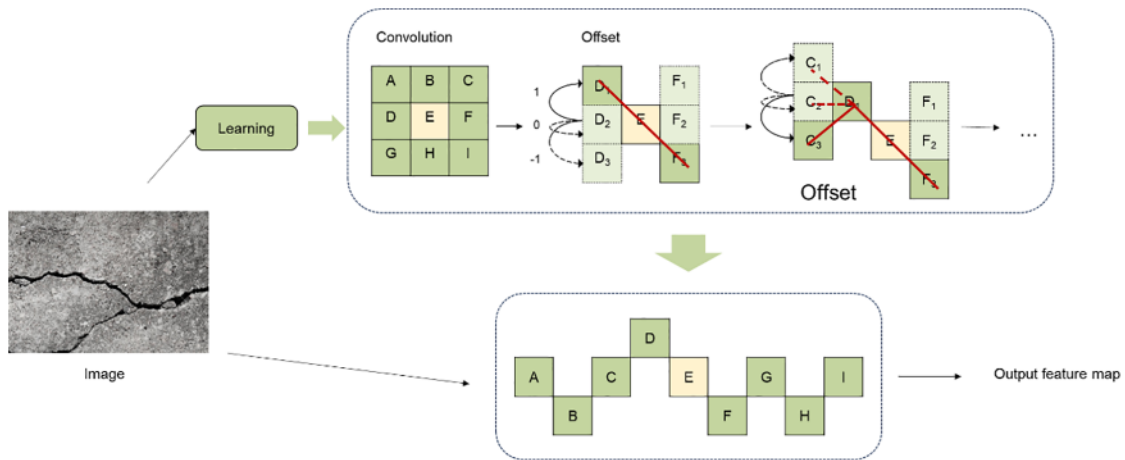


Figure 4: The architecture of DSConv

Among them, $P_{i\pm k} = (x_{i\pm k}, y_{i\pm k})$ represents the specific position of each grid, $k = \{0, 1, 2, 3, 4\}$ represents the horizontal separation from the grid center, $\Delta = \{\lambda \mid \lambda \in [-1, 1]\}$ is the accumulated offset amount.

Given that cracks typically exhibit characteristics similar to tubular structures, such as being elongated, irregular, and twisted. This paper introduces the novel application of DSConv to the field of crack segmentation. A pioneering step, the primary enhancement involves integrating DySnakeConv layers after each convolutional layer in the Head network of YOLOv8. This strategic addition aims to align better with the direction of cracks, thereby augmenting the model’s perception of cracks. This adaptation contributes to a more profound understanding and segmentation of crack regions, ultimately improving the performance of crack segmentation, particularly for slender and twisted crack structures. The structure is delineated in Fig. 5. This innovative design is poised to bring heightened adaptability and accuracy to the crack segmentation model.

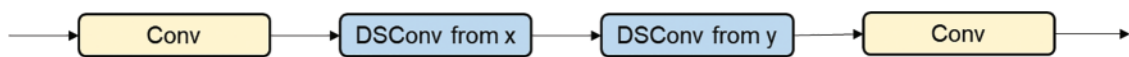


Figure 5: The architecture of DySnakeConv

2.2.4 Add CBAM Attention Mechanism

This article integrates the CBAM [30] module to improve the perception of fracture regions, building on the previously reported improvements and improving feature expression capabilities. As a result, the model can focus more on target attributes, improving the targets’ segmentation accuracy.

CBAM represents a prominent attention mechanism devised to dynamically modulate the feature responses of each layer within the network. The Spatial Attention Module (SAM) and the Channel Attention Module (CAM) are its two attention modules. The structural diagram of CBAM is elucidated in Fig. 6. The SAM module in CBAM extracts the maximum and average feature values for each spatial position in the feature map using two different strategies. To be more precise, this module performs these two pooling techniques separately to every channel in the feature map. This produces two distinct matrices that enhance the model’s capacity for detection and feature representation.

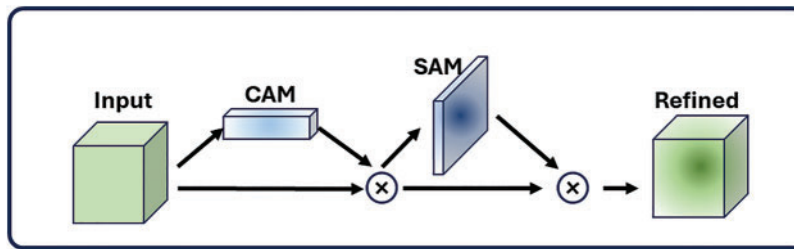


Figure 6: The architecture of CBAM

Feature map F first goes through global average pooling and global maximum pooling following CAM processing. The weights of CAM will then be computed from these pooled findings using multi-layer perceptrons. Subsequently, in order to standardize and normalize these weights, the Sigmoid function was used to normalize them. Finally, these normalized attention weights will be applied to the original input feature map one by one according to the channel. In particular, each channel's weight is multiplied by the appropriate feature map channel throughout the computation process, which improves attention and efficiently filters features.

3 Experiments

3.1 Dataset

The data used in this study comes from the DeepCrack [31]. There are 527 photos in the original dataset sample, with a resolution of 544 by 384 pixels. Data augmentation methods are used on the obtained images to enhance the possibility of network overfitting and increase the model's ability to generalize. These techniques effectively expand the dataset to 3162 images by introducing random noise, panning, rotation, mirroring, and other transformations. Considering the complex environment, many noise points, and harsh environment in the tunnel, Gaussian noise and average blur are added to better adapt to the actual tunnel environment. To further improve the stability of the data, Gaussian noise, average blur, pan, rotate, and flip are applied to the original photos. Fig. 7 displays the improved data images.

An 8:2 ratio is used to divide the enlarged dataset into two categories. In particular, there are 633 images were utilized for validation and 2529 photographs were used for training. The crack images are annotated using the Labelme software, which concurrently generates corresponding label files. The outcomes of this image labeling process are depicted in Fig. 8.

In this study, a specialized tunnel crack image test set was constructed to evaluate algorithm performance, implementing a series of measures to ensure its effectiveness. Firstly, we collected a range of images from different regions and under various conditions to ensure the representativeness and diversity of the test set, thus simulating various scenarios encountered in actual engineering projects. Secondly, the collected images underwent meticulous screening to guarantee the quality and accuracy of the test set. Moreover, in order to enrich the dataset and improve its adaptability, data augmentation techniques were employed on the images, including but not limited to rotation, translation, and scaling, thereby expanding the coverage of the test set to encompass a wide range of crack morphologies and environmental variations.

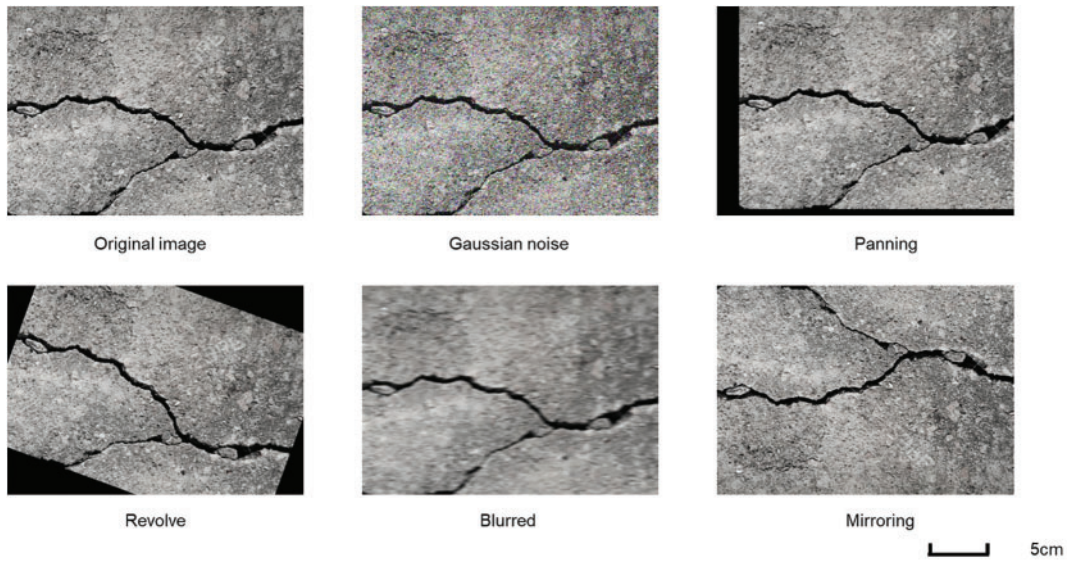


Figure 7: Image of the dataset after data enhancement

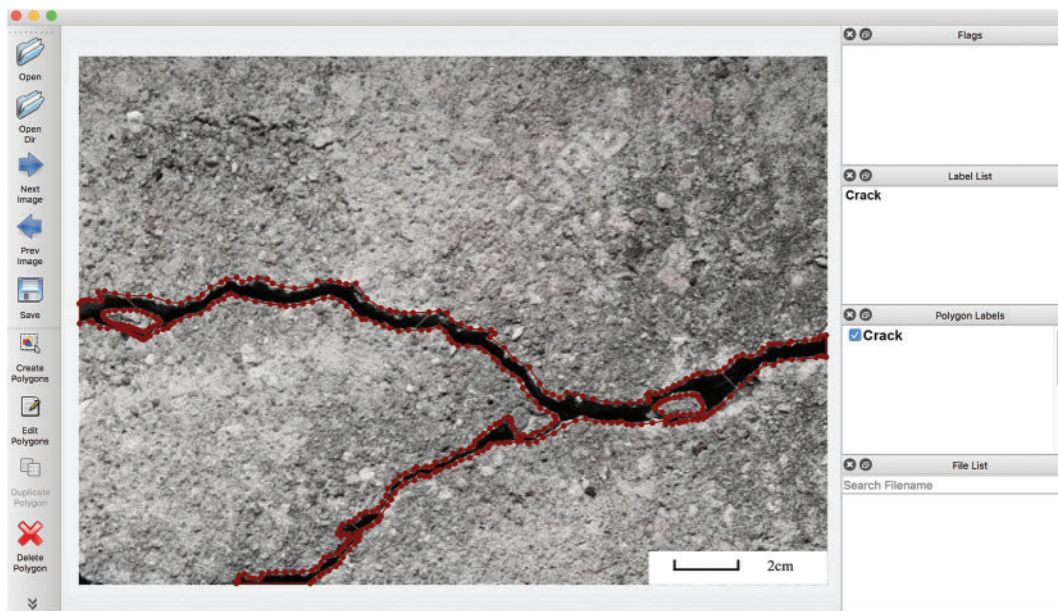


Figure 8: Results of labelme labeling

Through these measures, the study ensured the representativeness, diversity, and generalization capability of the test set, enabling a more accurate evaluation of crack segmentation algorithm performance. The main advantage of this approach lies in considering not only algorithm performance under specific conditions but also its evaluation under broader circumstances, providing a more comprehensive understanding of algorithm robustness and generalization capability. With this evaluation method, the paper can reliably determine the applicability and reliability of the algorithm in

practical engineering applications, thereby providing a reliable benchmark and reference for research on underground engineering crack segmentation algorithms.

3.2 Environment and Parameters

A 50 epoch early terminating patience parameter was used to track training progress during the 300 epoch training phase. The operating system and deep learning framework used in the experimental setup are Ubuntu 20.04.5 Long Term Support (LTS) and Pytorch, respectively. Table 1 lists all of the hardware and software combinations.

Table 1: Experimental environment equipment and software configuration

Name	Version
CPU	Intel(R) Core(TM) i5-10400 CPU @ 2.90 GHz
GPU	NVIDIA GeForce RTX 3090
Python	3.8.0
YOLO	v8.0.203
Pytoch	1.12.1

3.3 Evaluation Metrics

The segmentation performance of the suggested method is assessed using the following metrics. The following are the formulas:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

$$\text{F1} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$\text{AP} = \frac{\sum \text{Precision}}{N(\text{TotalImages})}$$

$$\text{mAP} = \frac{\sum \text{AP}}{N(\text{classes})}$$

where TP signifies a true positive, indicating a positive sample that has been accurately detected. FP, on the other hand, stands for a false positive, referring to a negative sample mistakenly identified as positive. TN represents a true negative, which means a negative sample that has been correctly identified as such. Lastly, FN denotes a false negative, where a positive sample is erroneously labeled as negative. The F1-score serves as a statistical measure utilized to assess the accuracy of a test. It encapsulates both precision and recall by computing their harmonic mean, thereby offering a unified metric that accounts. AP stands for the average precision of a category, i.e., it speaks to the proportion of all photographs that have this kind of target present, calculated as the average accuracy over all images in the test dataset for this category. mAP stands for the ratio of the average precision of all the categories.

3.4 Experimental Results and Analysis

3.4.1 Ablation Experiments

This paper used ablation experiments on RFB, DCNv2, DSCConv, and CBAM modules to confirm the efficacy of the enhancements. To evaluate how the changes have affected performance, this paper perform ablation using YOLOv8s as a benchmark. Table 2 presents the findings.

Table 2: The table of FDSC-YOLOv8s model for ablation experiments

Baseline	RFB	DCNv2	DSCConv	CBAM	Precision	Recall	mAP50	mAP50-95
YOLOv8s	–	–	–	–	0.952	0.900	0.942	0.623
YOLOv8s	✓	–	–	–	0.979	0.925	0.958	0.655
YOLOv8s	–	✓	–	–	0.973	0.914	0.954	0.642
YOLOv8s	–	–	✓	–	0.978	0.923	0.961	0.651
YOLOv8s	–	–	–	✓	0.966	0.896	0.944	0.634
Presented model	✓	✓	✓	✓	0.977	0.934	0.965	0.664

The YOLOv8 algorithm, with the addition of the RFB, DCNv2, DySnake, and CBAM modules, has demonstrated improvements in Precision, Recall, mAP50, mAP50-95, and other three indicators. Compared to the YOLOv8 algorithm, FDSC-YOLOv8 method, which incorporates RFB, DCNv2, DySnake, and CBAM, has an increase of 2.3% in mAP50 and 3.8% in mAP50-95. The data have demonstrated that the modules of the proposed method all have an improvement effect on the model, and the effect of using them together is better than the effect of using them alone, which shows that new model outperforms the original YOLOv8s model on the crack segmentation task.

3.4.2 Comparison Experiment

To conduct an objective evaluation of the FDSC-YOLOv8's superiority, the performance is compared and examined with the one-stage detector YOLOv5 and YOLOv7 models. To ensure fairness in the experimental outcomes, the same dataset was utilized under consistent hardware device. Table 3 presents the related experimental findings.

Table 3: The table of the FDSC-YOLOv8s model for comparison experiments

Name	Precision	Recall	mAP50	mAP50-95	F1-score	FPS
YOLOv5s-seg	0.935	0.876	0.936	0.552	0.904	297
YOLOv7-seg	0.971	0.913	0.952	0.609	0.941	235
YOLOv8s-seg	0.952	0.900	0.942	0.623	0.925	285
Presented model	0.977	0.934	0.965	0.664	0.955	272

Compared with the YOLOv5s seg algorithm, Precision has increased by 4.2%, Recall has increased by 5.8%, mAP50 has increased by 2.9%, mAP50-95 has increased by 11.2%, and F1 value has increased by 5.1%; Compared to the YOLOv7 seg algorithm, the presented method has increased the Recall value by 2.1%, mAP50 by 1.3%, mAP50-95 by 5.5%, and F1-score by 1.4%. The enhanced YOLOv8 algorithm has improved Precision by 2.5%, Recall value by 3.4%, mAP50 by 2.3%, mAP50-95 by 4.1%,

and F1 value by 3% when compared to the YOLOv8s-seg method. In addition, although FPS is lower than YOLOv5, its accuracy value is much higher than it. This ensures that FDSC-YOLOv8 remains effective while maintaining consistent performance. The data demonstrates that the FDSC-YOLOv8 outperforms earlier segmentation models in crack segmentation tasks, suggesting that the suggested strategy has a considerable improvement effect in crack segmentation.

3.4.3 The Analysis of the Experiment

The loss curve and mAP curve of the FDSC-YOLOv8 crack segmentation model are shown in Fig. 9. From the data results, the loss value gradually decreases and gradually stabilizes, particularly in the latter stages of the training process. The loss value is almost unchanged.

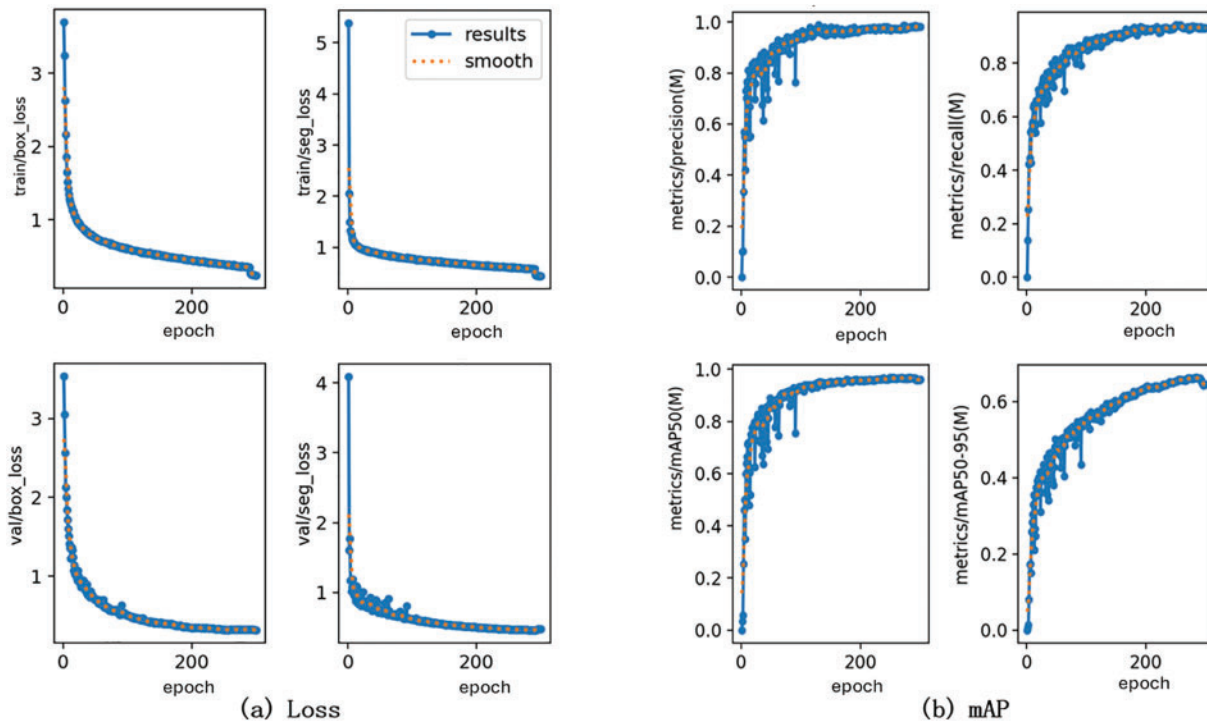


Figure 9: The results of FDSC-YOLOv8. (a) The loss curve of FDSC-YOLOv8. (b) The mAP curve of FDSC-YOLOv8

To further confirm the real-world effectiveness of the upgraded mode, verification was conducted using actual engineering images collected before and after the improvements. In this study, it was found that the improved model can identify cracks with widths as low as a few pixels, as shown in Fig. 10. This further demonstrates the superiority of the model in crack identification.

Numerous tests have decisively shown that the improved YOLOv8 performs better in crack segmentation tasks, confirming its increased ability to capture the structure and morphology of cracks. This achievement holds significant importance in addressing the challenges posed by the diversity and complexity of cracks.



Figure 10: The results before and after improvement. (a) The results before improvement; (b) The results after improvement

4 Conclusion

The FDSC-YOLOv8 algorithm developed in this study demonstrates improved performance in the identification and segmentation of cracks in underground engineering. By integrating RFB, DCNv2, and CBAM attention mechanisms, along with the incorporation of the DSCConv network tailored for slender and twisted structures, FDSC-YOLOv8 significantly improves the precision and detail detection capability for crack features. This improved recognition capacity delivers superior performance with mAP50 and mAP50-95 scores of 96.5% and 66.4%, and lowers misidentification in challenging construction sites, such as those with high levels of noise and cable interference. These achievements highlight the crucial value of the FDSC-YOLOv8 algorithm in enhancing the safety of underground engineering, particularly by early identification and localization of tiny cracks to effectively prevent potential structural issues, thereby reducing the risk of accidents and maintenance costs. Future research will focus on further optimization of algorithm parameters and testing in a broader range of underground engineering scenarios to ensure widespread practical application of the presented technology, offering enhanced safety guarantees in the field of underground engineering. Through such efforts, the FDSC-YOLOv8s will not only serve as a powerful tool for crack identification but also play a key role in advancing the safety of underground engineering.

Acknowledgement: Sincere thanks to all who supported and assisted in this research.

Funding Statement: This research was funded by the National Key R&D Program of China (Project: Key Technologies and Equipment for Multi-View Stereoscopic Disaster Detection and Emergency Response to Derived Disasters in Underground Spaces, 2022YFC3005600), the National Natural Science Foundation of China (52378402), Shandong Provincial Natural Science Foundation Youth Project (ZR2022QE021 and ZR202211100077), Shandong Province Higher Education Young Innovative Team Project (2022KJ037), State Key Laboratory of Precision Blasting and Hubei Key Laboratory of Blasting Engineering, Jiangnan University (PBSKL2022C03), funding from Shandong Railway Investment Holding Group Co., Ltd. (“Key Technologies for Rapid and Intelligent Construction of Large Section High-Speed Railway Tunnels in Low Mountain and Hilly Areas” and “Intelligent Construction Trolley Equipment and Key Technologies for the Lining of Ultra-Long Open Tunnel Sections”).

Author Contributions: The authors confirm contribution to the paper as follows: study conception and design: Z.L. and R.W.; data collection: R.W. and H.L.; analysis and interpretation of results: B.S. and C.M.; draft manuscript preparation: R.W.; writing—review and editing, Z.L. After reviewing the findings, all authors gave their approval to the manuscript’s final draft.

Availability of Data and Materials: On request, the data from this study can be obtained from the corresponding author. Part of the data comes from ourself and another from <https://github.com/yhllleo/DeepCrack>. Since the study team individually created the dataset and has not yet made it public, our data are not available to the general public.

Conflicts of Interest: Given that the two authors are from the same company, potential conflicts of interest may exist. However, we hereby declare that, regarding this study, the authors have not been involved in any relationships or situations that may lead to conflicts of interest.

References

1. Ruan S, Liu D, Gu Q, Jing Y. An intelligent detection method for open-pit slope fracture based on the improved Mask R-CNN. *J Min Sci.* 2022;58(3):503–18. doi:10.1134/S1062739122030176.
2. Gong JF, Zhu Y, Zhang GZ. The grade discrimination of large deformation of layered surrounding rock in tunnel and its treatment counter measures. *J Railway Eng Soc.* 2018;35(12):51–5 (In Chinese).
3. Xiong ZM, Lu H, Wang MY, Qian QH, Rong XL. Research progress on safety risk management for large scale geotechnical engineering construction in China. *Rock Soil Mech.* 2018;39(10):3703–16 (In Chinese).
4. Geusebroek JM, Smeulders AWM, Geerts H. A minimum cost approach for segmenting networks of lines. *Int J Comput Vis.* 2001;43:99–111. doi:10.1023/A:1011118718821.
5. Sironi A, Türetken E, Lepetit V, Fua P. Multiscale centerline detection. *IEEE T Pattern Anal.* 2015;38(7):1327–41.
6. Zhang Z, Xing F, Shi X, Yang L. Semicontour: a semi-supervised learning approach for contour detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016; Las Vegas, NV, USA.* p. 251–9.
7. Tang F, Han C, Ma T, Chen T, Jia Y. Quantitative analysis and visual presentation of segregation in asphalt mixture based on image processing and BIM. *Automat Constr.* 2021;121:103461. doi:10.1016/j.autcon.2020.103461.
8. Li S, Cao Y, Cai H. Automatic pavement-crack detection and segmentation based on steerable matched filtering and an active contour model. *J Comput Civil Eng.* 2017;31(5):04017045. doi:10.1061/(ASCE)CP.1943-5487.0000695.
9. Han C, Ma T, Huyan J, Huang X, Zhang Y. CrackW-Net: a novel pavement crack image segmentation convolutional neural network. *IEEE T Intell Transp.* 2021;23(11):22135–44.
10. Felzenszwalb PF, Girshick RB, McAllester D, Ramanan D. Object detection with discriminatively trained part-based models. *IEEE T Pattern Anal.* 2009;32(9):1627–45.
11. LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *P IEEE.* 1998;86(11):2278–2324. doi:10.1109/5.726791.
12. Zhiqiang W, Jun L. A review of object detection based on convolutional neural network. In: *2017 36th Chinese Control Conference (CCC); 2017; Dalian, China.* p. 11104–9.
13. Zhao H, Qin G, Wang X. Improvement of canny algorithm based on pavement edge detection. In: *2010 3rd International Congress on Image and Signal Processing; 2010; Yantai, China.* p. 964–7.
14. Li G, Li X, Zhou J, Liu D, Ren W. Pixel-level bridge crack detection using a deep fusion about recurrent residual convolution and context encoder network. *Measurement.* 2021;176:109171. doi:10.1016/j.measurement.2021.109171.

15. Talab AMA, Huang Z, Xi F, Liu HM. Detection crack in image using Otsu method and multiple filtering in image processing techniques. *Optik*. 2016;127(3):1030–3. doi:10.1016/j.ijleo.2015.09.147.
16. Yu Y, Bao G, Can L, Lin L. Highway bridge crack detection method based on improved YOLO V5. *J Shandong Univ Technol (Nat Sci Ed)*. 2023;37:1–7 (In Chinese).
17. Sha AM, Tong Z, Gao J. Recognition and measurement of pavement disasters based on convolutional neural networks. *China J Highw Transp*. 2018;31(1):1–10 (In Chinese).
18. Wang LP. Deep learning-based crack detection of concrete pavement (Master's Thesis). Hebei University of Engineering; China; 2018 (In Chinese).
19. Xie CH, Wu JM, Xu HY. Small object detection algorithm based on improved YOLOv5 in UAV image. *Comput Eng Appl*. 2023;59(9):198–206 (In Chinese).
20. Lei RZ, Yang WF, Su XN. Rural feature classification of UAV remote sensing image based on improved YOLO-v3. *Electron Design Eng*. 2023;31(3):178–84 (In Chinese).
21. Qi H, Xu T, Wang G, Cheng Y, Chen C. MYOLOv3-Tiny: a new convolutional neural network architecture for real-time detection of track fasteners. *Comput Ind*. 2020;123:103303. doi:10.1016/j.compind.2020.103303.
22. Chen SH, Tsai CC. SMD LED chips defect detection using a YOLOv3-dense model. *Adv Eng Inform*. 2021;47:101255. doi:10.1016/j.aei.2021.101255.
23. Dumitriu A, Tatui F, Miron F, Ionescu RT, Timofte R. Rip current segmentation: a novel benchmark and yolov8 baseline results. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*; 2023; Vancouver, Canada. p. 1261–71.
24. Zhang ZJ, Li JZ, Wu WL, Zhong JF, Wu YR, Lin QY. Red fire ant nest recognition based on component segmentation and YOLOv8s. *Modern Inf Technol*. 2023;7(20):66–74 (In Chinese).
25. Li Z, Ma Y, Chen Y, Zhang X, Sun J. Joint COCO and mapillary workshop at ICCV 2019: COCO instance segmentation challenge track. *arXiv preprint arXiv: 2010.02475*. 2020.
26. Zhu X, Hu H, Lin S, Dai J. Deformable ConvNets V2: more deformable, better results. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*; 2019; Los Angeles, California, USA. p. 9308–16.
27. Liu S, Huang D. Receptive field block net for accurate and fast object detection. In: *Proceedings of the European Conference on Computer Vision (ECCV)*; 2018; Munich, Germany. p. 385–400.
28. Deng L, Yang M, Li T, He Y, Wang C. RFBNet: deep multimodal networks with residual fusion blocks for RGB-D semantic segmentation. *arXiv preprint arXiv:1907.00135*. 2019.
29. Qi Y, He Y, Qi X, Zhang Y, Yang G. Dynamic snake convolution based on topological geometric constraints for tubular structure segmentation. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*; 2023; Vancouver, Canada. p. 6070–9.
30. Woo S, Park J, Lee JY, Kweon IS. Convolutional block attention module. In: *Proceedings of the European Conference on Computer Vision (ECCV)*; 2018; Munich, Germany. p. 3–19.
31. Liu Y, Yao J, Lu X, Xie R, Li L. DeepCrack: a deep hierarchical feature learning architecture for crack segmentation. *Neurocomputing*. 2019;338:139–53. doi:10.1016/j.neucom.2019.01.036.