



OPEN An intelligent prediction method for rock core integrity based on deep learning

Zhaoxia Hu¹, Hua Mei¹ & Lei Yu²✉

To address the issue of serious inefficiency in the traditional manual evaluation methods of rock core integrity, a deep learning-based algorithm named IDA-RCF (Intelligent detection algorithm for Rock Core Fissure) is proposed in this paper, which realizes the automatic evaluation of rock core integrity in accordance with the fissure identification results. In IDA-RCF, a two-branch feature extraction network is firstly proposed, in which branch one is used to fully extract the complex and variable local detail fissure features by Deformable convolution, and branch two is used to capture the global context information of the rock core images by EfficientViT network based on the self-attention. Then a multi-level feature fusion network is proposed for adaptively fusing local and global features from the same level and the fused feature information from the previous level, thereby capturing more valid information and eliminating redundancies. Then the fused feature layer is decoded by the feature decoder to output the detection results of rock core fissure. Finally, the fissure rate is automatically calculated based on the detection results to predict the degree of rock core integrity. The experimental results show that the accuracy indexes F1, mAP@0.5 and mAP@0.5:0.95 of IDA-RCF are 93.09%, 94.44% and 84.61%, respectively. The relative error between the prediction results and the manual statistical results of the fissure rate is only 4.38%, and the prediction accuracy for the degree of rock core integrity is 93.8%, indicating that the proposed method in this paper is able to accomplish the intelligent evaluation task of rock core integrity with high precision.

Keywords Geotechnical engineering, Deep learning, Intelligent prediction, Rock core integrity, Fissure

Rock core integrity (denoted as RCI in this paper) directly reflects the stability and homogeneity of underground rock layers. In underground projects such as tunnels, mines and pits, the advance acquisition of RCI helps engineers to identify potentially unstable areas and to anticipate possible geologic hazards such as landslides, landslides and rockbursts. Thus, disaster prevention and mitigation measures can be deployed in advance during engineering design and construction to ensure the safety of construction personnel and equipment^{1,2}. Most of the traditional RCI evaluation methods rely on manual observation, experimental testing and empirical judgment, which are not only time-consuming and costly, but also susceptible to the influence of human factors, resulting in subjectivity and uncertainty in the evaluation results. Especially when a large number of rock core samples need to be processed, the efficiency is obviously insufficient³.

Image processing technology has been gradually applied to RCI evaluation by virtue of its advantages of automation and long distance. For example, Hall et al.⁴ analyzed the rock core geometry through the Hough transform algorithm, which simplified the rock core photographs into relevant image texture attributes for the judgment of rock core integrity. Saricam et al.⁵ first used the edge tracking algorithm to detect the cross-section of the rock core block, then used the thresholding algorithm to eliminate the shadows produced by the light source and segment the rock core block, and finally calculated the RCI by measuring the length of the rock core's central axis. Image processing techniques can quickly process a large number of rock core images, greatly improving the efficiency of RCI evaluation. However, this type of algorithm is very sensitive to lighting conditions and core features (surface color, texture, etc.). Rock core images under different conditions require different parameter settings, resulting in poor robustness of the algorithms. Secondly, the recognition accuracy of image processing techniques still needs to be improved when dealing with complex backgrounds, low-contrast images, and fine fissures^{6,7}. In recent years, deep learning has achieved a large number of significant results in data detection, recognition, and classification of geotechnical engineering fields, which provides a new and effective solution for evaluating RCI. Alzubaidi et al.⁸ established a rock core classification model based on a Convolutional Neural

¹School of Civil and Architecture Engineering, Hunan Institute of Technology, Hengyang 421002, China. ²College of Urban and Rural Construction, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China. ✉email: yuleizhku@163.com

Network (CNN), which realizes automated classification of rock cores. Su et al.⁹ realized the segmentation of rock core blocks with Mask RCNN and Unet algorithms, and the segmentation results was used for the intelligent evaluation of RCI. Fu et al.¹⁰ improved the YOLOv5 target detection algorithm by using channel and spatial attention to realize the accurate identification of rock core blocks, and automatically predicted the RCI by using the developed post-processing program. Compared with image processing techniques, deep learning algorithms automatically extract features from core images through multilayer neural networks, which avoids the tedious process of manually setting thresholds and improves the accuracy and comprehensiveness of feature extraction¹¹. In addition, the robustness of the deep learning algorithm is significantly enhanced, which is mainly reflected in the ability to adapt to different lighting conditions, changes in color and texture features, and complex backgrounds¹².

However, most studies directly apply the existing deep learning algorithms to RCI evaluation, which lacks the analysis for the characteristics of the rock core images and the targeted design of the network, as reflected in the following three aspects:

(1) Existing studies generally recognize intact rock core blocks by deep learning algorithms, and then calculate the proportion of intact rock core blocks occupying the total rock cores. However, the color and striations of rock in different properties vary greatly (Fig. 1a), which makes it difficult to construct a model that can accurately identify all types of rock, resulting in a narrow application scenario. (2) The shape of the convolution kernel of the traditional CNN is a fixed square (generally 1×1 , 3×3), which makes the network not well adapted to the target deformation or targets with different angles when performing feature extraction. The rock core fissures are characterized by features with variable geometry, extension direction and edge features (Fig. 1b). It is difficult to accurately recognize such targets with square convolution kernels. (3) There are often handwritten numbers, lighting, shadows and other external interference in the rock core image (Fig. 1c). Therefore, it is critical to combine the global information and local information to distinguish the target and the background accurately. Currently, the commonly used intelligent detection algorithms are mainly divided into CNN-based and Self Attention (SA)-based target detection algorithms^{13,14}. CNN extract local features layer by layer by convolution operation, which is suitable for processing image data with strong local correlation, but it is hard to model long-distance dependencies. SA can directly capture the global features by calculating the relationship between all positions in the input sequence, but the ability of local feature extraction is insufficient. The construction of network structures that can fully extract both global and local information from images still needs to be further improved in the existing research. In summary, the rock cores obtained from field sampling have complex characteristics including obvious color and striation differences, multiple external interferences, and variable variable geometry features of fissure, etc. The existing intelligent algorithms for rock core integrity often suffer from poor resistance to environmental interference, bad generalization performance, and difficulty in extracting global and local features of the target simultaneously, which leads to poor recognition accuracy in practical engineering applications.

Aiming at solving the above issues, a deep learning-based algorithm named IDA-RCF (Intelligent detection algorithm for Rock Core Fissure) is proposed in this paper, which realizes the automatic evaluation of rock core integrity in accordance with the fissure identification results. The main contributions of this paper are as follows:

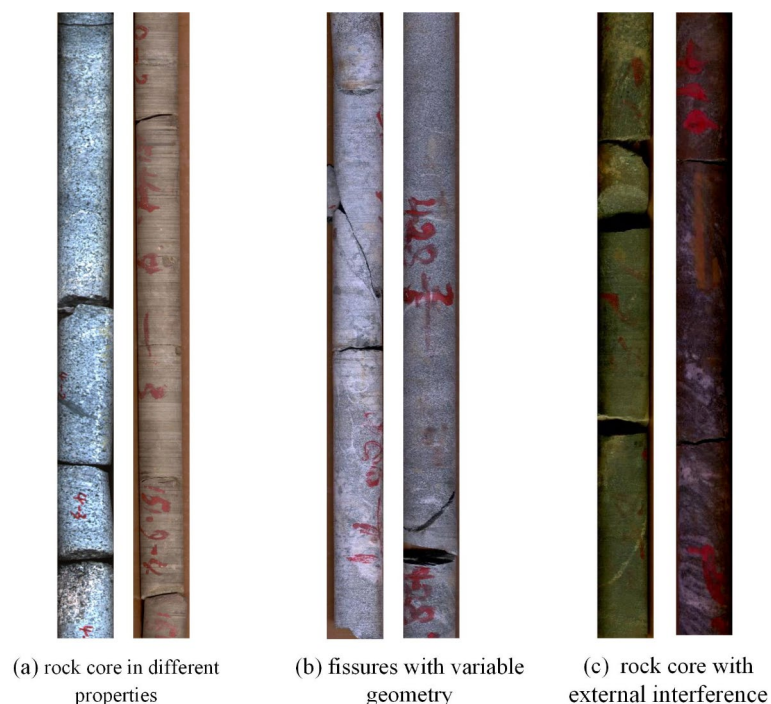


Fig. 1. Images of drilled rock cores.

- (1) A two-branch feature extraction network is constructed, in which branch one fully extracts the complex and varied local details of the rock core fissures through Deformable Convolution, and branch two captures the global context information through EfficientViT based on SA, so as to effectively improve the network's ability to locate the fissures and resist the external interference.
- (2) A multi-level feature fusion network is proposed for adaptively fusing local and global features from the same level and the fused feature information from the previous level, thereby capturing more valid information and eliminating redundancies.
- (3) A high-quality rock core fissure dataset is built by collecting, organizing and labeling the drill hole images collected at the construction site, which contains rich rock core samples with different geological, integrity, and background conditions.
- (4) The experimental results show that the prediction accuracy of rock core integrity automatically calculated according to the fissure identification results of IDA-RCF is 93.8%, which is suitable to accomplish the task of rock core integrity evaluation with high efficiency and precision.

Intelligent detection algorithm for rock core fissures

Methodology

In this paper, a deep learning-based algorithm for intelligent detection algorithm of core fissures, named IDA-RCF (Fig. 2), is proposed. In this paper, a deep learning-based algorithm for intelligent detection algorithm of rock core fissures is constructed, named IDA-RCF (Fig. 2). IDA-RCF consists of three modules: Two-branch feature extraction network, multi-level feature fusion network, and feature decoder. Among them, two-branch feature extraction network is used to extract local and global features of the rock core image, the multi-level feature fusion network is used to adaptively fuse feature information from different levels, the feature decoder module is used to decode the fused feature information and output the identification results of rock core fissures. In the following, the principles of two-branch feature extraction network and multi-level feature fusion network will be elaborated. The feature decoding module is consistent with YOLOv8¹⁵, which will not be described in this paper.

Two-branch feature extraction network

Two-branch feature extraction network consists of an EfficientViT-based global feature extraction branch (denoted as branch one) and a deformable convolution-based local feature extraction branch (denoted as branch two). The rock core images are input into the two branches in parallel. In branch one, the image is extracted by multiple EfficientViT modules to obtain three global feature layers C_1 , C_2 , and C_3 with sizes of 13×13 , 26×26 , and 52×52 . In branch two, the image is extracted by multiple DBS (Deformable convolution-Batch Normalization-SiLu) modules consisting of Deformable convolution, Batch Normalization and SiLu activation function to obtain three local feature layers S_1 , S_2 , and S_3 with sizes of 13×13 , 26×26 , and 52×52 . The principles

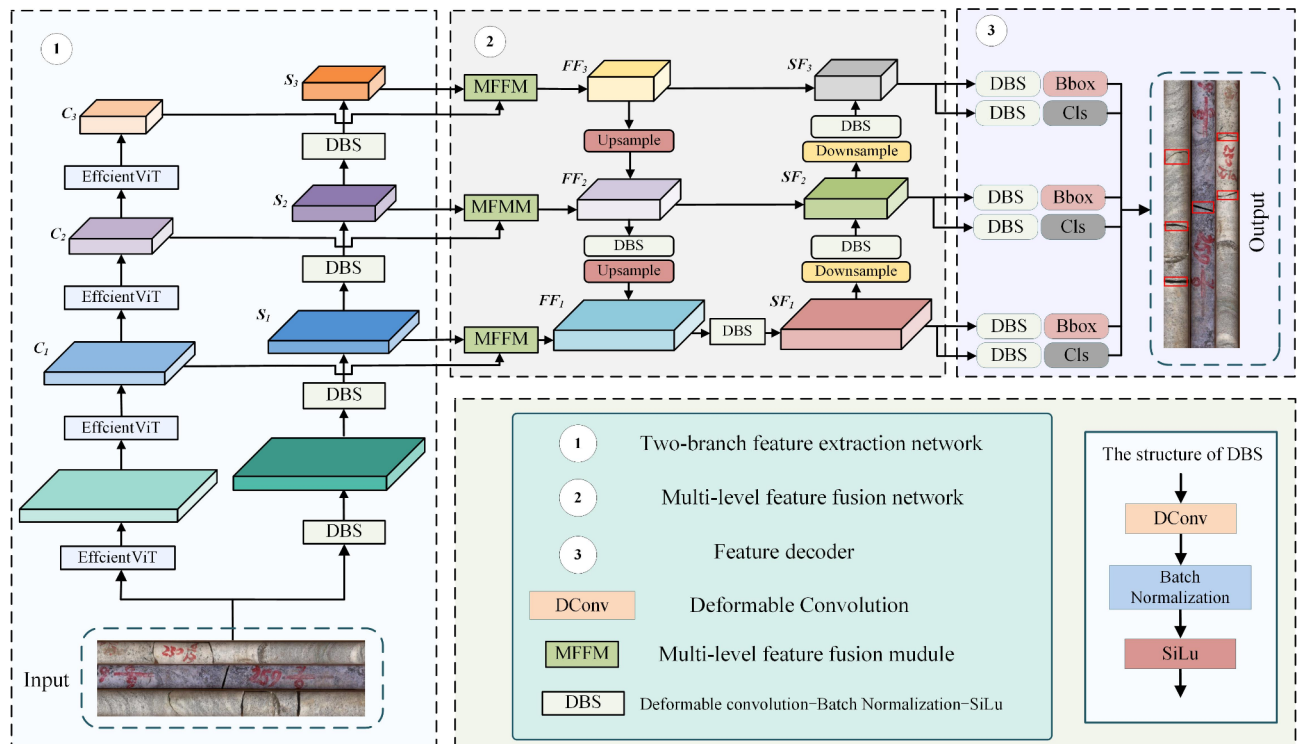


Fig. 2. IDA-RCF network structure.

of EfficientViT which is the core module in branch one, and the deformable convolution which is the core module in branch two are introduced in the following.

EfficientViT

The network structure of EfficientViT is shown in Fig. 3, which consists of Token Interaction, Feed forward network and Cascaded Group Attention. Among them, Token Interaction is used to enhance the model's ability to process local information, Feed forward network is used to enhance the model's nonlinear representation ability, and Cascaded Group Attention is used to extract the global context information of the image. The core structure of EfficientViT is the Cascaded Group Attention (CGA), which reduces the computational cost of the SA and enhances the feature representation capability of the model by means of feature segmentation and cross-head cascading.

In CGA, the input feature layer (assuming that the dimension is $[N, C, H, W]$, where N is the batch size, C is the number of channels, H and W are the height and width, respectively) is first segmented into G sub-feature layers, each of which has a dimension of $[N, C/G, H, W]$. Each sub-feature layer is then input independently into an attention head. In the attention head, sub-feature layer S is first multiplied with three different weight matrices (W^Q, W^K, W^V) to generate Query matrix Q , Key matrix K , and Value matrix V . Then the weighted context vectors are obtained by performing SA computation via Eq. (2), which represent the degree of attention that each position in the input sequence pays to the other positions:

$$Q = S \cdot W_Q, K = S \cdot W_K, V = S \cdot W_V \quad (1)$$

$$Attention(Q, K, V) = Softmax\left(\frac{Q \cdot K^T}{\sqrt{d_k}}\right) \cdot V \quad (2)$$

where d_k is the dimension of K .

Then a cascade operation is performed on the outputs of all the attention heads, as shown in Fig. 9b, which adds the outputs of each head to the subsequent heads, gradually optimizing the feature representation to form a feature layer containing richer information. The cascaded features are then projected through a linear layer to make the cascaded feature representation consistent with the original input dimensions with the following formula:

$$F = Linear(Concatenate[Attention1, Attention2, Attention3, \dots, AttentionG]) \quad (3)$$

Deformable convolution

In the traditional convolution operation, the value of each position in the output feature layer is computed by applying a convolution kernel to a localized region on the input feature layer. Setting the input feature layer as X , the convolution kernel as W , and the output feature layer as Y , the convolution operation can be denoted as:

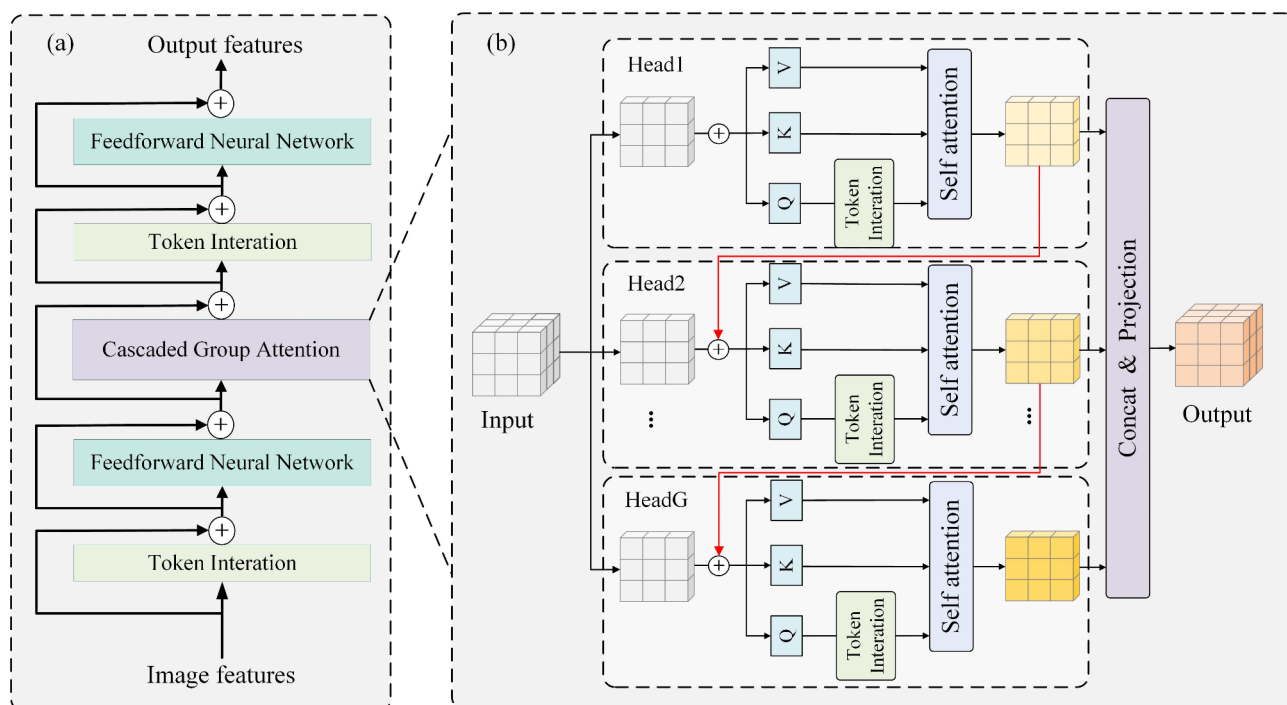


Fig. 3. EfficientViT network structure.

$$y(p) = \sum_{p_n \in R} W(p_n)X(p + p_n) \quad (4)$$

where p is the position of the output feature layer, R is the receptive field of the convolution kernel, and p_n is the position of each element in the convolution kernel with respect to the center point.

In deformable convolution, the sampling positions in the convolution operation are learned by introducing learnable offsets Δp_n . The mathematical expression for deformable convolution is:

$$y(p) = \sum_{p_n \in R} W(p_n)X(p + p_n + \Delta p_n) \quad (5)$$

where Δp_n is the offset corresponding to position p_n in the convolutional kernel, which are learned from the input feature layer X through a separate convolutional layer. Assuming that the convolutional kernel of the offset generation layer is W_{offset} , ΔP is computed as follows.

$$\Delta p_n = W_{\text{offset}} * X \quad (6)$$

The number of output channels of W_{offset} is $2 \times k \times k$, where $k \times k$ is the size of the convolution kernel, and 2 indicates that there are two offsets for each position (corresponding to the x and y directions, respectively).

Since the offset Δp_n may be a floating point number, $p + p_n + \Delta p_n$ may not be an integer position. In this regard, deformable convolution uses bilinear interpolation to compute the feature values at non-integer positions. Setting $p' = p + p_n + \Delta p_n$, the bilinear interpolation is computed as follows: four neighboring integer points p_0 , p_1 , p_2 , and p_3 of p' are firstly localized, then a weighted average is carried out according to the distances from p' to the four integer points. During the training process, the offset Δp_n and the convolution kernel W_{offset} are updated simultaneously to minimize the loss of the network.

By introducing learnable offsets, the deformable convolution is able to sample at irregular locations, which enhances the network's ability to handle geometric deformations and different scale objects. Figure 4 shows the difference in the feature extraction process with deformable convolution and standard convolution. As the number of network layers increases, the deformable convolution is able to perform better shape feature extraction for fissures.

Multi-level feature fusion network

To effectively obtain different levels of local features and global features, enhancing the network's anti-interference ability for complex backgrounds and capturing the detailed fissure features. A multi-level feature fusion network is constructed in this paper, which consists of two parts: upward fusion module and downward fusion module.

In the upward fusion module, the global feature layer $C_i (i = 1, 2, 3)$ and the local feature layer S_i outputted from the two-branch feature extraction network are fused through the MFFM to obtain the one-stage fusion feature layer FF_i , then the smaller-size FF_i is sequentially fused with the larger-size FF_{i-1} through the bottom-

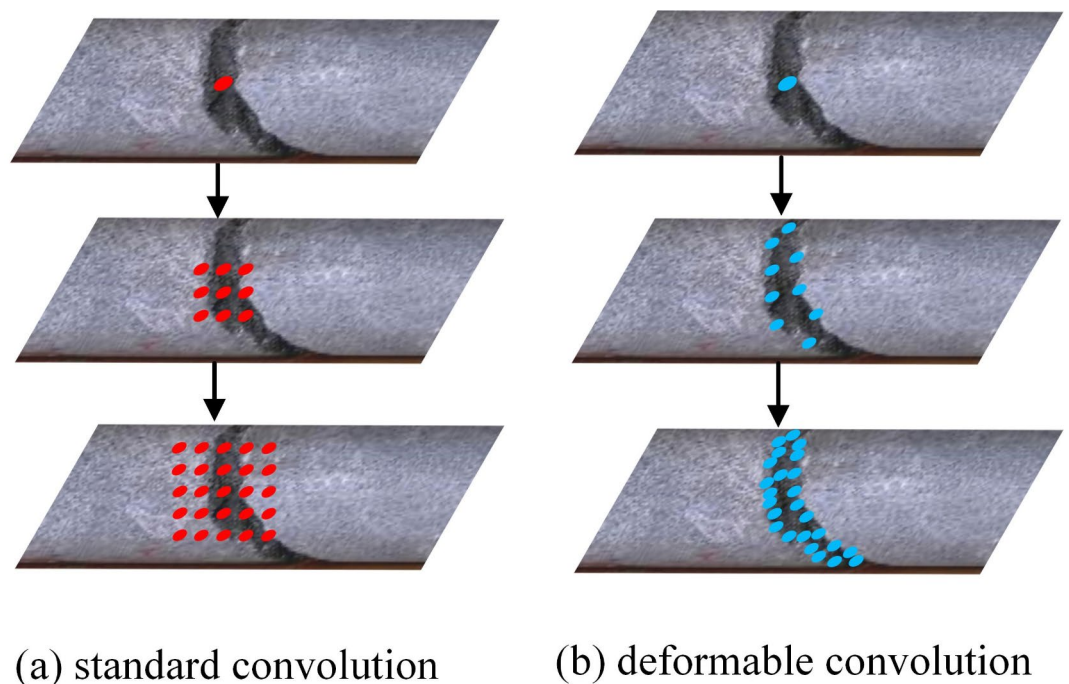


Fig. 4. Standard convolution and deformable convolution.

up path to obtain the two-stage fusion feature layer SF_{i-1} . The core structure of the upper fusion module is the MFFM, as shown in Fig. 5. In MFFM, the global feature layer will be input into the channel attention, which utilizes the interdependence between channel mappings to improve the feature representation of specific semantics, which helps to reduce noise interference and enhance the robustness of image detection. The local feature layer is input into the spatial attention, which enhances the local details and suppresses irrelevant regions by adjusting the weights, thus improving the accuracy of image detection. Finally, the global and local feature layers after applying attention are fused. In the downward fusion module, the larger SF_i is sequentially fused with the smaller SF_{i+1} through a top-down path, which outputs three three-stage fused feature layers with sizes of 13×13 , 26×26 , and 52×52 .

Data set

More than 7000 rock core images with good diversity were collected at the project site, containing rock core types of different integrity, color, texture, and background complexity, which is beneficial for constructing a RCI prediction model with good generalization performance. Some of the rock core images are shown in Fig. 6.

The length and width of the images are adjusted to be the same when inputted into the network. However, the rock core image is generally presented as a long strip, which will lead to distortion in length compression if it is inputted into the network alone, as shown in Fig. 7a. To avoid this phenomenon, the horizontal splicing method is adopted in this paper to preprocess the images, as shown in Fig. 7b. A total of 973 spliced images of rock cores with similar length and width are obtained with a size of 640×640 , which are divided into 60% training set, 20% validation set and 20% test set. Then, the Labelimg annotation tool was used to label the fissure regions in the rock core images (drawing bounding boxes and marking category information) (Fig. 7c) to obtain the final dataset.

Rock core fissure identification experiments

The experiment mainly consists of the following two stages: (1) Comparison experiment of rock core fissure identification: with the data set constructed in this paper as the basis, seven target detection algorithms are selected to conduct a comparison experiment with the algorithms constructed in this paper, to verify the effectiveness of IDA-RCF in identifying the fissures of different geological, integrity, and background conditions. (2) Ablation experiment: the effect of Deformable Convolution, Two-branch feature extraction network, and MFFM adopted in this paper are tested through ablation experiment.

Comparison experiment of rock core fissure identification

Comparison network

To compare the effectiveness of IDA-RCF for rock core fissure identification task, seven well-performing target detection networks were selected for comparison. A brief description of each network is given below:

DETR¹⁶: This network is constructed with the Transformer architecture, which utilizes the global attention mechanism to perform target detection for the first time.

Faster-RCNN¹⁷: a two-stage network which first generates candidate regions by Region Proposal Network and then performs fine-grained classification and regression on these candidate regions.

EfficientDet¹⁸: it combines EfficientNet as the feature extractor and BiFPN as the feature fusion module, which drastically reduces the computational complexity while maintaining high accuracy.

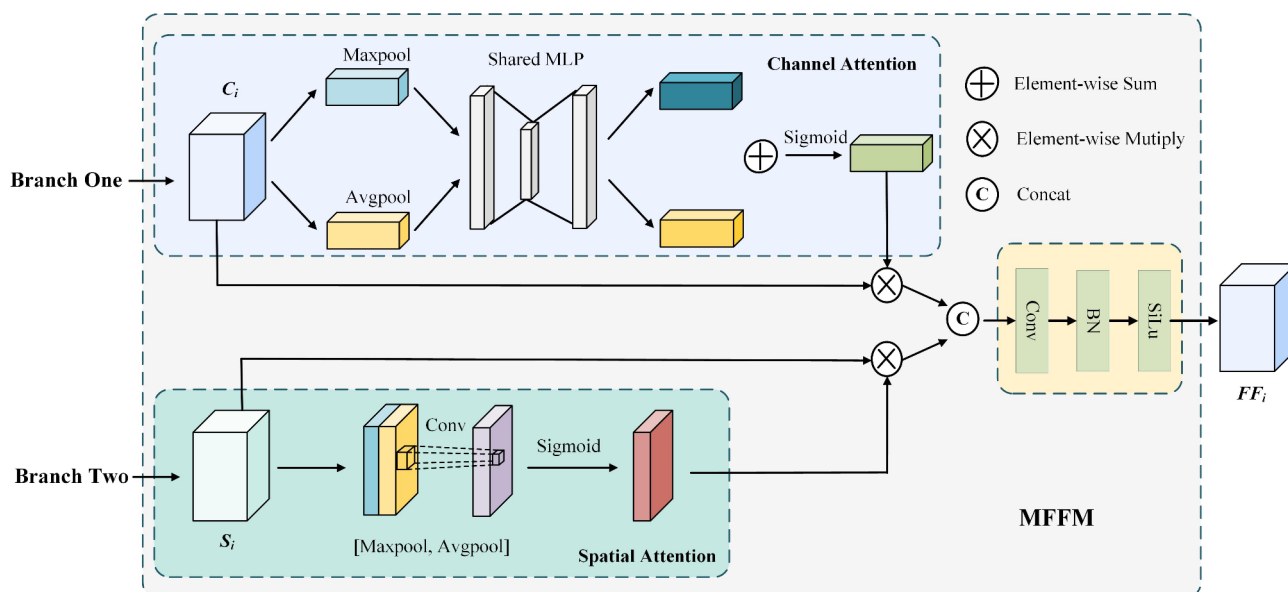


Fig. 5. MFFM network structure.

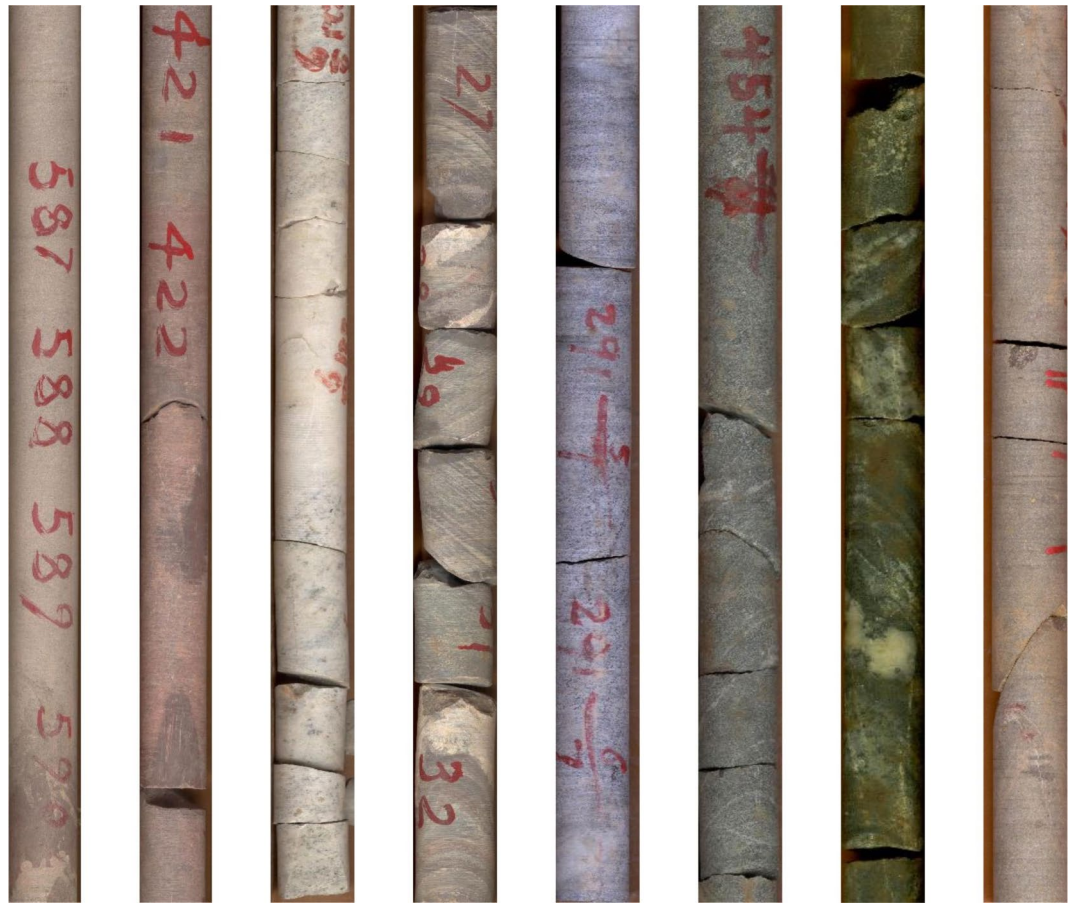


Fig. 6. Rock core image.

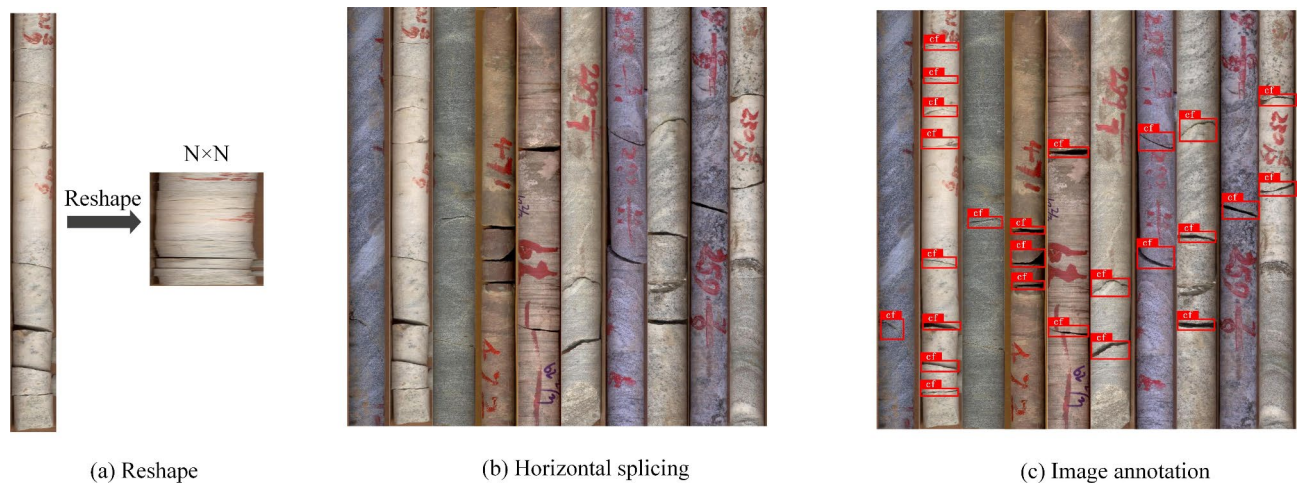


Fig. 7. Data set construction method.

YOLOX¹⁹: an anchor-free frame design and a dynamic label assignment strategy are adopted, which simplifies the training and inference process and improves the detection accuracy. In addition, a decoupled detection header is introduced to optimize the classification and regression tasks independently, which further improves the model performance.

YOLOv5: Compared to the predetermined anchor points of YOLOv4, YOLOv5 adopts an adaptive anchor point design, which makes the anchor points closer to the distribution and morphology of the actual objects.

YOLOv8 (Varghese et al., 2024): the C3 module of YOLOv5 was replaced by the C2f module with richer gradient streams, and the classification and detection heads were separated in the decoupled head.

YOLOv5-ECASA (YOLOv5 enhanced by ECASA attention module) (Alzubaidi et al., 2022): the ECASA attention is used to improve the YOLOv5 algorithm, which realizes the intelligent identification of rock core fissures.

Evaluation indicators

The evaluation metrics selected in this paper include Recall, Precision, F1, and mean average precision (mAP). Recall indicates the proportion of actually existing fissures that are correctly detected. Precision indicates the proportion of detected fissures that are actual existing fissures. The calculation formula is shown below.

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

where TP (True Positives) denotes the number of fissures correctly identified as fissures; FP (False Positives) denotes the number of backgrounds incorrectly identified as fissures. FN (False Negatives) denotes the number of fissures incorrectly identified as backgrounds.

F1 is the reconciled mean of precision and recall. mAP denotes the mean of precision at different recall levels.

$$F1_score = \frac{2Precision \cdot Recall}{Precision + Recall} \tag{9}$$

$$mAP = \int_0^1 Precision(Recall) d(Recall) \tag{10}$$

mAP@0.5 is the mean value of the average accuracy calculated when the Intersection over Union (IoU) is set to 0.5. IoU represents the degree of overlap between the predicted and real frames, which is calculated as the ratio of the intersection area to the concatenation area. of the predicted box to the real box. mAP@0.5:0.95 is the mean value of mAP calculated at different IoUs (from 0.5 to 0.95 with a step size of 0.05). mAP@0.5 can reflect the model's identification performance in general application scenarios. mAP@0.5:0.95 takes into account the model's performance under multiple IoU thresholds, which comprehensively reflects the detection capability of the model.

Experimental results

The evaluation metrics results of each model are shown in Table 1. The F1, mAP@0.5 and mAP@0.5:0.95 of IDA-RCF are 93.09%, 94.44%, and 84.61%, respectively, which are the highest among all the models, indicating that IDA-RCF has a better identification accuracy of rock core fissures. Compared with DETR, the F1, mAP@0.5 and mAP@0.5:0.95 of IDA-RCF is improved by 4.62%, 4.82%, and 9.41%, respectively. DETR is built based on Transformer, which has good global feature capability but lacks in extracting the local features of the fissure, while two-branch feature extraction of the IDA-RCF takes advantage of both CNN in local feature extraction and SA in global feature extraction, thus achieving better identification accuracy. Among the four types of models built based on the YOLO framework (YOLOX, YOLOv5, YOLOv5-ECASA, YOLOv8), YOLOv5-ECASA has the best evaluation metrics, which is due to the fact that the model introduces the ECASA attention, which is able to filter the redundant information better for the fissure images with complex background interference and help the network put more attention to the fissure region. However, YOLOv5-ECASA is constructed entirely by convolutional neural networks. Due to the limitation of the localization of the convolutional kernel, it is insufficient in capturing the global contextual information, thus the accuracy is lower than that of IDA-RCF. In addition, IDA-RCF has the smallest difference between mAP@0.5:0.95 and mAP@0.5, which indicates that IDA-RCF is more broadly applicable, which can be used in more project scenarios that have different requirements on the detection accuracy.

Model	Precision (%)	Recall (%)	F1 (%)	mAP@0.5 (%)	mAP@0.5:0.95
IDA-RCF	92.57	93.62	93.09	94.44	84.61
DETR	88.16	88.79	88.47	89.62	75.20
Faster-RCNN	87.20	87.84	87.52	89.03	76.86
EfficientDet	88.26	89.34	88.80	90.16	79.65
YOLOX	87.63	88.55	88.09	89.04	74.24
YOLOv5	86.52	87.10	86.81	87.56	73.32
YOLOv5-ECASA	89.14	90.27	89.70	90.67	79.23
YOLOv8	88.31	89.57	88.94	89.58	78.02

Table 1. Results of model performance metrics.

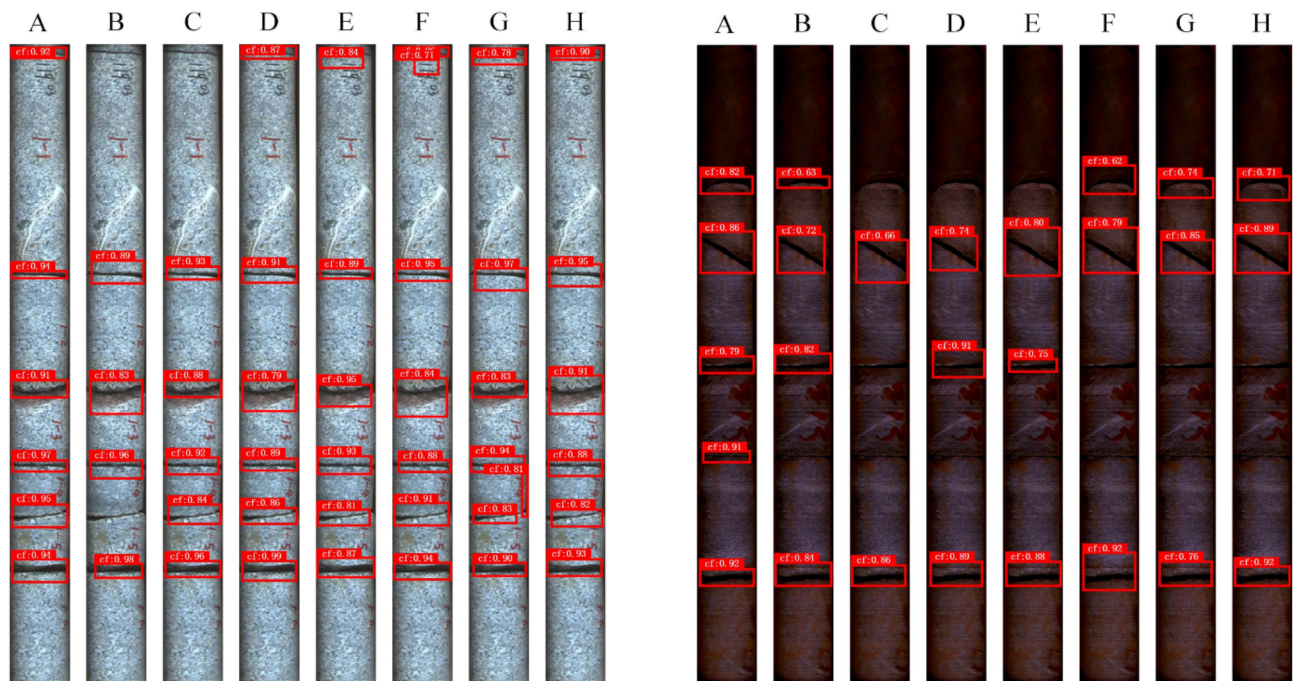
The different geological conditions are mainly reflected in the color differences of rock cores. Figure 8a and b show the fissure identification results for rock cores in different geologic conditions. IDA-RCF, DETR, Faster-RCNN, EfficientDet, YOLOX, YOLOv5, YOLOv5-ECASA, and YOLOv8 are denoted as A, B, C, D, E, F, G, H in the figure. For a more objective comparison, the shown images are similar in the integrity and the background complexity. When the color of the rock is light (Fig. 8a), the models are able to accurately identify the coarser fissures. For the thin fissures, DETR and Faster-RCNN showed missed recognition. When the core color is dark (Fig. 8b), the overall confidence of each model appears to decrease, which is due to the fact that the color of fissure region is nearly black, a darker background will make it difficult for the models to differentiate the demarcation area between the background and the fissure. There are five fractures in Fig. 8b. IDA-RCF correctly identified all the fissures, while DETR, Faster-RCNN, EfficientDet, YOLOX, YOLOv5, YOLOv5-ECASA, and YOLOv8 identified 4, 2, 2, 3, 3, 3, 3, respectively. The above results show that IDA-RCF has better geological applicability in identifying the fissures under different core colors.

Figure 9 shows the fissure identification results with different integrity of rock core. The fragmentation of the cores in Fig. 9a–d increases sequentially. The more fragmented the core is, the more number of fissures and the richer the morphology, which makes the identification more difficult. For the core with good integrity in Fig. 9a, all models accurately identified the fissure locations. In Fig. 9b, EfficientDet, YOLOv5, and YOLOv8 showed incomplete identification of the fissure area. Faster-RCNN, EfficientDet, YOLOv5-ECASA incorrectly identified one fissure into two rectangular boxes. In Fig. 9c and d, the missed identification of fissures in the comparison models is obviously increased, and the phenomenon that the predicted rectangular box area is obviously larger than the actual fissure area occurs. Comprehensively analyzing the identification results of IDA-RCF, although a small amount of missed identification occurs with the increase of fragmentation, it is still the least among all the models, in addition, the rectangular box can well reflect the position occupied by the fissures in the core, which can provide the basis for the subsequent accurate calculation of the fissure rate in the rock core.

Figure 10 shows the identification results for the complex background case. In Fig. 10a, there is a lot of dark interference in the background, Faster-RCNN, EfficientDet, YOLOX, YOLOv5 and YOLOv8 incorrectly identify the dark material as fissure. In Fig. 10b, combination of low light brightness, small fissure widths, and interference in the background causes great difficulty in identification. EfficientDet and YOLOX did not identify any fissures, Faster-RCNN, EfficientDet, YOLOX, YOLOv5 and YOLOv8 showed varying degrees of misidentification. IDA-RCF accurately identifies all fissures without incorrect identification, indicating that the proposed algorithm has good resistance to environmental interference, which is suitable for complex and variable conditions.

Ablation experiment

The effect of Deformable convolution, Two-branch feature extraction network, and MFFM adopted in this paper are tested through ablation experiment.



(a) Fissure identification results for light rock color (b) Fissure identification results for dark rock color

Fig. 8. Results of fissure identification for different geological conditions.

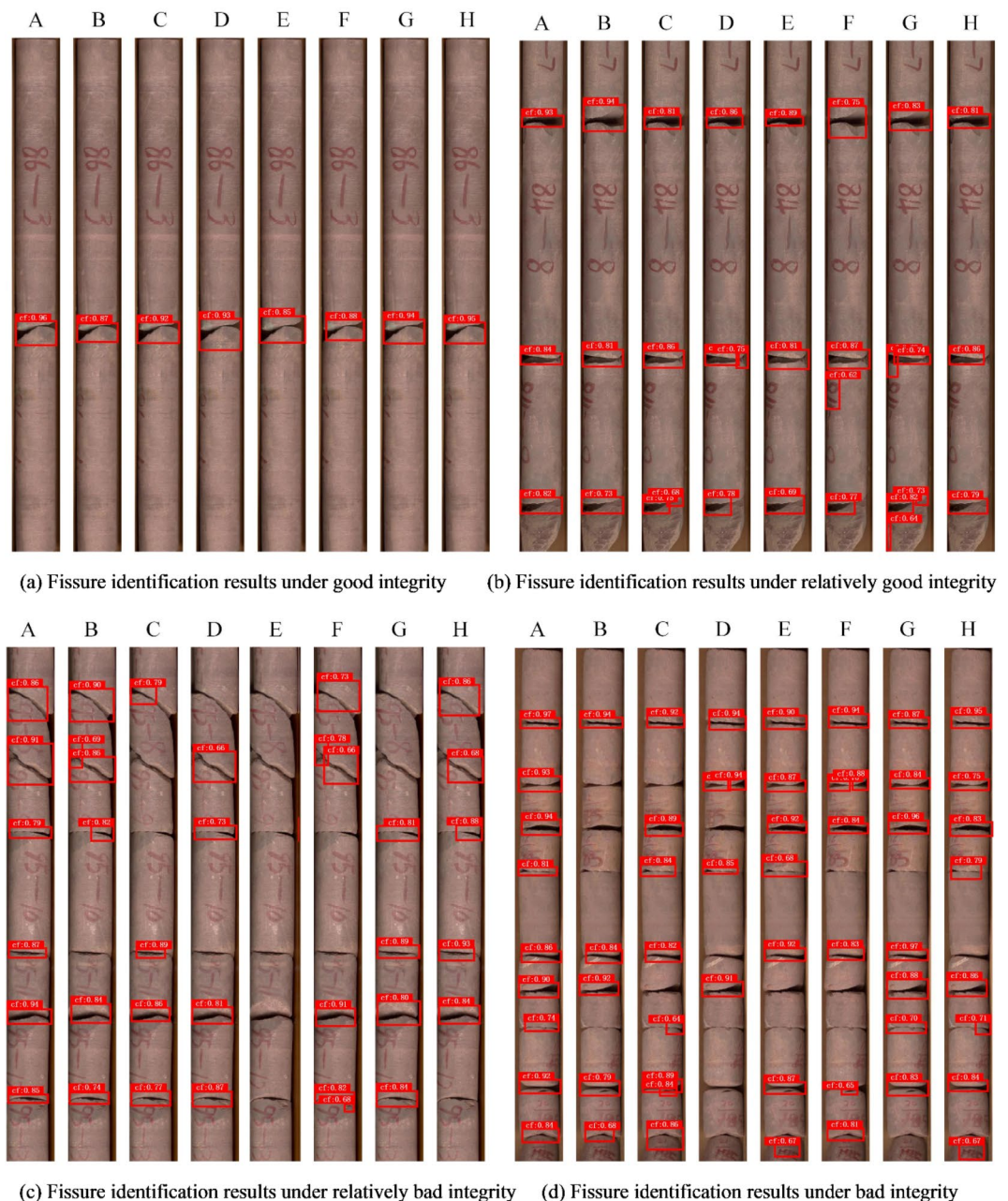


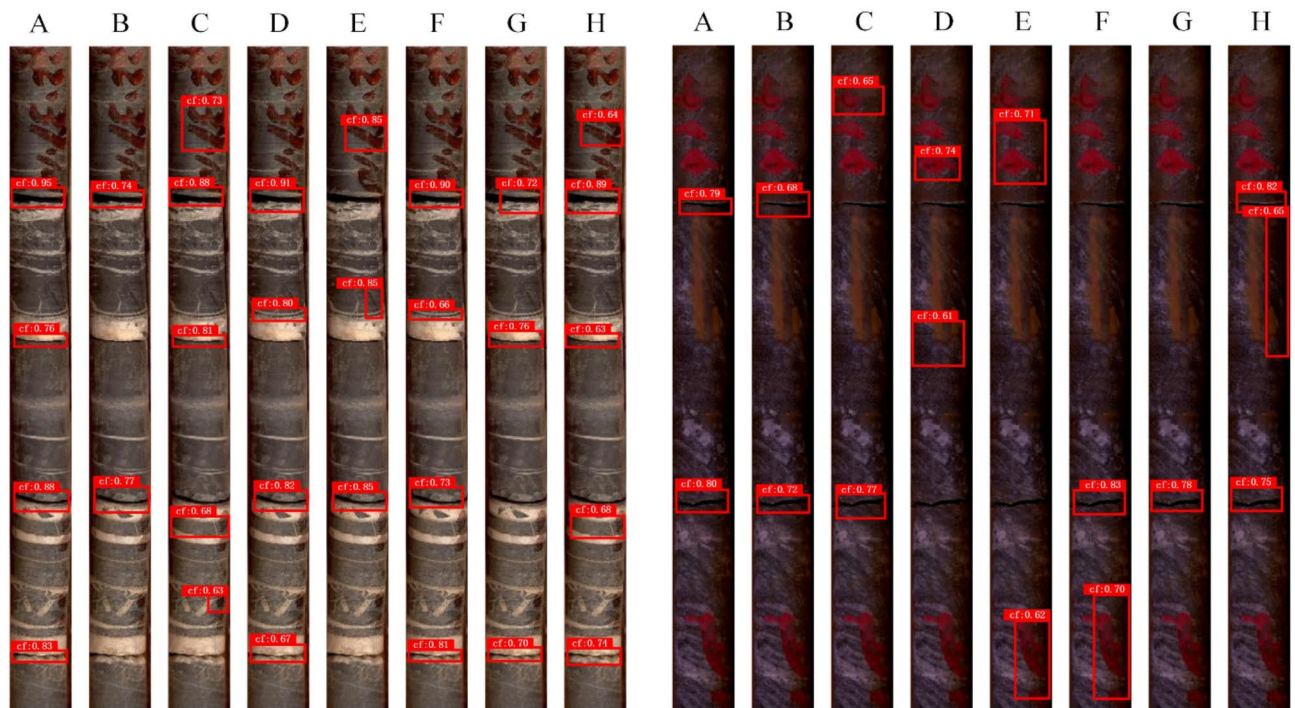
Fig. 9. Results of fissure identification under different integrity.

The effect of deformable Convolution

The deformable convolution in IDA-RCF is replaced with the traditional convolution to construct the comparison model. The F1, mAP@0.5 and mAP@0.5:0.95 of this model are 91.35%, 92.36%, and 80.22%, respectively, which are 1.74%, 2.08%, and 4.39% lower compared with IDA-RCF. Analyzing Fig. 11 (The model with deformable convolution are denoted as A1, the model with traditional convolution are denoted as A2), the model with traditional convolution correctly identifies the transverse fissures with relatively regular shapes, but the identification of curved and diagonal fissures is unsatisfactory. The model with deformable convolution accurately localizes the fissures in different directions. The above results show that the deformable convolution has better ability to deal with geometric deformation and different scales of fissures.

The effect of two-branch feature extraction network

Branch one and branch two are respectively used as the backbone to construct the comparison model, the test results are shown in Table 2; Fig. 12 (The model with two-branch feature extraction network are denoted as B1, the model with branch one are denoted as B2, the model with Branch two are denoted as B3). The F1, mAP@0.5 and mAP@0.5:0.95 of the model with two-branch feature extraction network are improved by 3.1%, 3.6%, and 4.14% compared with the model with branch one, and 2.27%, 2.69%, and 3.29% compared with the



(a) Fissure identification results with dark interference (b) Fissure identification results with combinatorial interference

Fig. 10. Results of fissure identification in complex background.

model with branch two, which indicates that the combination of the advantages of CNN and SA can effectively improve the identification accuracy of fissures.

The effect of MFFM

To investigate the effectiveness of MFFM in fusing global features from branch one and local features from branch two, the models with and without MFFM are compared. The results show that the F1, mAP@0.5 and mAP@0.5:0.95 of the model with MFFM is improved by 1.27%, 1.15%, and 2.08% compared to that of the model without MFFM, respectively. Figure 13 (The model with MFFM one are denoted as C1, The model with MFFM one are denoted as C2) shows that the model with MFFM has better anti-interference ability for complex backgrounds and the ability to capture the detailed fissure features. The model without MFFM appeared to incorrectly identify the background as a fissure several times. In addition, for fissures with small widths, the model without MFFM had missed identifications and incomplete identifications.

Rock core integrity evaluation experiments

The degree of RCI was evaluated using the fissure rate, which is the average number of fissures per unit length (1 m). The relationship between fissure rate and the degree of RCI is shown in Table 3. In this paper, the fissure rate is automatically calculated based on the fissure detection results of IDA-RCF. The calculation process is shown in Fig. 14. Sixteen rock core images are selected for RCI evaluation experiments. The results are shown in Table 4. The relative error between the fissure rate calculated by IDA-RCF and manual statistics is only 4.38%, and the prediction accuracy of RCI degree is 93.8%. The above results indicates that the proposed method can accurately evaluate the degree of RCI, which can greatly reduce the manpower and time consumed by manual evaluation means if it is applied to the practical projects.

Conclusion

In this paper, an intelligent detection algorithm for rock core fissures based on deep learning, named IDA-RCF, is proposed by combining Deformable convolution, Two-branch feature extraction network, and Multi-level feature fusion network, which realizes the automatic evaluation of rock core integrity in accordance with the fissure identification results. The main conclusions are as follows:

- (1) The F1, mAP@0.5 and mAP@0.5:0.95 of IDA-RCF are 93.09%, 94.44%, and 84.61%, respectively, which has higher fissure identification accuracy compared with other commonly used networks. In addition, IDA-RCF has the smallest difference between mAP@0.5:0.95 and mAP@0.5, which indicates that IDA-RCF is more broadly applicable, which can be used in more project scenarios that have different requirements on the detection accuracy.

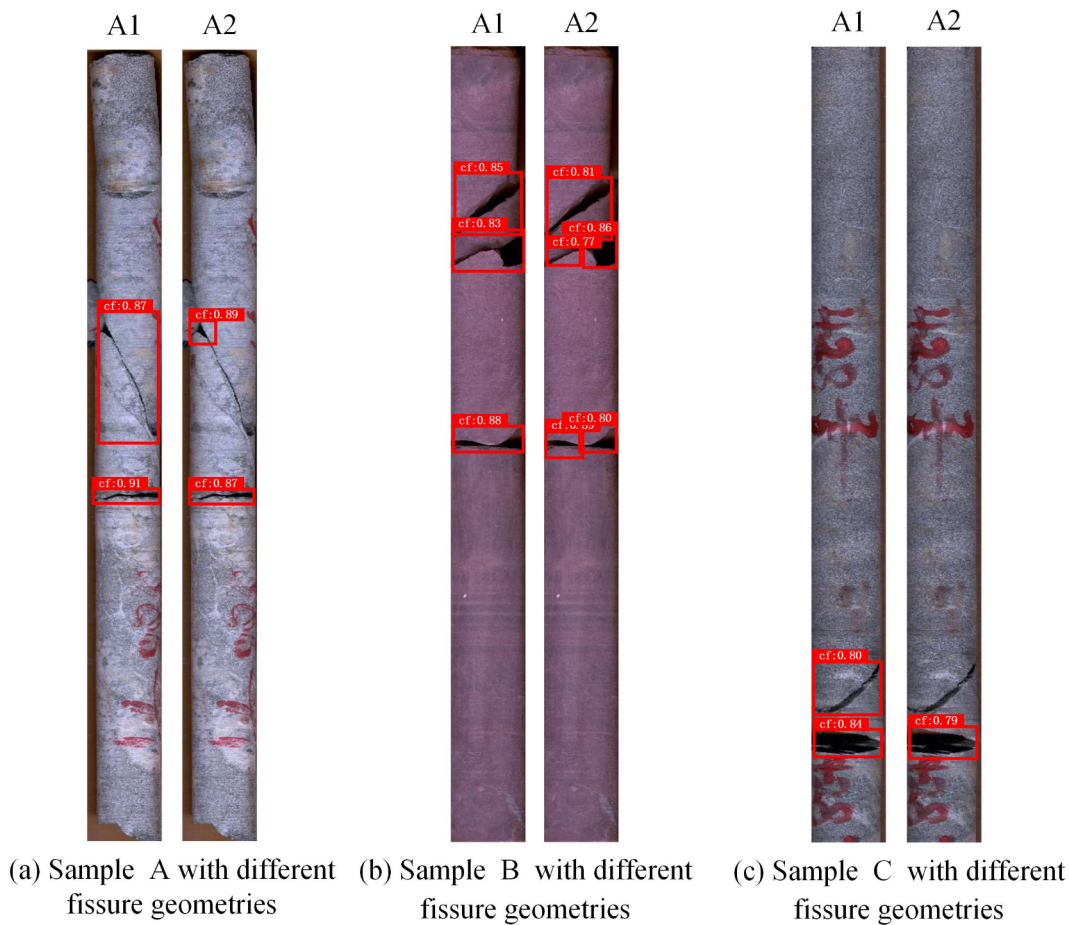


Fig. 11. Identification results of models with deformable convolution and traditional convolution.

Model	Precision (%)	Recall (%)	F1 (%)	mAP@0.5 (%)	mAP@0.5:0.95
With branch one	89.63	90.36	89.99	90.84	80.47
With branch two	90.37	91.28	90.82	91.75	81.32
With two-branch feature extraction network	92.57	93.62	93.09	94.44	84.61

Table 2. Model evaluation indicators with different branch.

- (2) Compared with other models, IDA-RCF has better applicability to the task of fissure identification under different geology, rock core integrity, and complex backgrounds, which is mainly reflected in the fact with less cases of missing identification and incorrectly identifying the background as fissures.
- (3) Based on the detection results of IDA-RCF, the fissure rate is automatically calculated and used to predict the degree of rock core integrity. The relative error between the fissure rate calculated by IDA-RCF and manual statistics is only 4.38%, and the prediction accuracy of RCI degree is 93.8%, which indicates that the proposed method can accurately evaluate the degree of RCI.

Finally, it is noted that the evaluation of RCI with only the fissure rate is not comprehensive. In the future work, we plan to build a high-precision semantic segmentation algorithm for intelligent prediction of RQD (Rock Quality Designation), which is another key evaluation index reflecting the RCI. Then, combining the two indexes of RQD and fissure rate to evaluate the degree of RCI to further improve the prediction accuracy.



Fig. 12. Identification results of models with different branch.

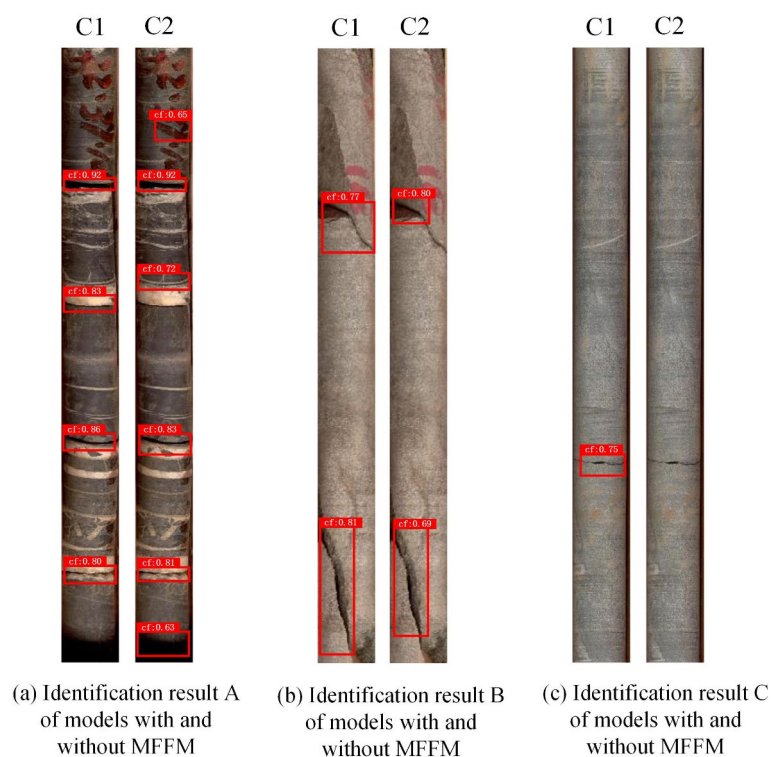


Fig. 13. Identification results of models with and without MFFM.

Fissure rate(/m)	0~5	5~10	10~25	25~50	>50
The degree of RCI	Intact (A)	Relatively intact (B)	Relatively broken (C)	Broken (D)	Extremely broken (E)

Table 3. Relationship between fissure rate and degree of rock core integrity.

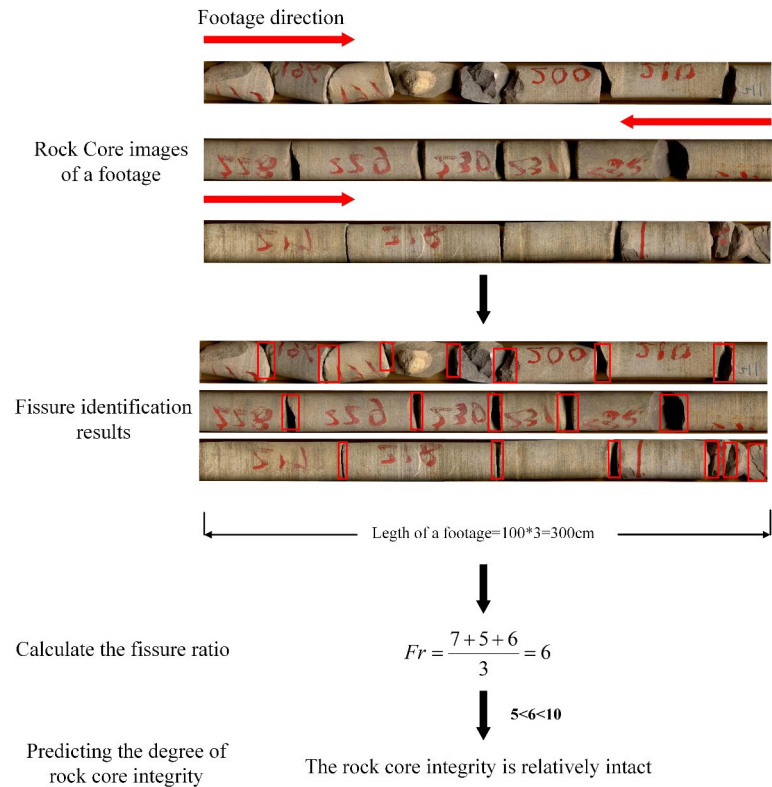


Fig. 14. Process for predicting the degree of RCI.

Sample number	Fissure ratio from manual statistics (/m)	Predicted fissure ratio (/m)	Relative error (%)	Predicted integrity level	Real integrity level
1	3.33	3.33	0.00	A	A
2	3.25	3.50	7.69	A	A
3	19.00	17.50	7.89	C	C
4	2.33	2.33	0.00	A	A
5	4.00	4.33	8.33	A	A
6	5.75	5.25	8.70	B	B
7	3.33	3.33	0.00	A	A
8	16.50	15.50	6.06	C	C
9	26.00	24.00	7.69	C	D
10	4.00	4.00	0.00	A	A
11	4.50	4.25	5.56	A	A
12	7.00	7.00	0.00	B	B
13	7.00	7.00	0.00	B	B
14	3.25	3.50	7.69	A	A
15	6.33	5.67	10.53	B	B
16	4.00	4.00	0	A	A
Average value			4.38	–	–

Table 4. Predicted results of rock core integrity.

Data availability

The datasets generated and analysed during the current study are not publicly available due to commercial constraints but are available from the corresponding author on reasonable request.

Received: 19 December 2024; Accepted: 17 February 2025

Published online: 22 February 2025

References

1. Mahmoodzadeh, A. et al. Dynamic prediction models of rock quality designation in tunneling projects. *Transp. Geotechnics*. **27**, 100497 (2021).
2. Liu, T., Jiang, A., Zhang, Z., Liu, H. & Zheng, J. Estimation of the lengths of intact rock core pieces and the corresponding RQD considering the influence of joint roughness. *KSCE J. Civ. Eng.* **27** (6), 2689–2703 (2023).
3. Baraboshkin, E. E., Demidov, A. E., Orlov, D. M. & Koroteev, D. A. Core box image recognition and its improvement with a new augmentation technique. *Comput. Geosci.* **162**, 105099 (2022).
4. Hall, J., Ponzi, M., Gonfalanini, M. & Maletti, G. Automatic extraction and characterisation of geological features and textures from borehole images and core photographs. In *SPWLA Annual Logging Symposium* SPWLA-1996. SPWLA. (1996).
5. Saricam, T. & Ozturk, H. Estimation of RQD by digital image analysis using a shadow-based method. *Int. J. Rock Mech. Min. Sci.* **112**, 253–265 (2018).
6. Huyan, J., Ma, T., Li, W., Yang, H. & Xu, Z. Pixelwise asphalt concrete pavement crack detection via deep learning-based semantic segmentation method. *Struct. Control Health Monit.*, **29** (8), e2974. (2022).
7. Xiong, F. et al. Intelligent algorithm for rock core RQD based on object detection and image segmentation to suppress noise and vibration. *Adv. Civ. Eng.* **2024** (1), 3599911 (2024).
8. Alzubaidi, F., Mostaghimi, P., Swietojanski, P., Clark, S. R. & Armstrong, R. T. Automated lithology classification from drill core images using convolutional neural networks. *J. Petrol. Sci. Eng.* **197**, 107933 (2021).
9. Su, R. et al. A framework for RQD calculation based on deep learning. *Min. Metall. Explor.* **40** (5), 1567–1583 (2023).
10. Fu, D., Su, C. & Li, X. Automatic estimation of rock quality designation based on an improved YOLOv5. *Rock Mech. Rock Eng.* **57** (4), 3043–3061 (2024).
11. Saxena, N., Day-Stirrat, R. J., Hows, A. & Hofmann, R. Application of deep learning for semantic segmentation of sandstone thin sections. *Comput. Geosci.* **152**, 104778 (2021).
12. Zhou, Z., Zhang, J., Gong, C. & Wu, W. Automatic tunnel lining crack detection via deep learning with generative adversarial network-based data augmentation. *Undergr. Space*. **9**, 140–154 (2023).
13. Liu, H., Miao, X., Mertz, C., Xu, C. & Kong, H. Crackformer: Transformer network for fine-grained crack detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* 3783–3792 (2021).
14. Ma, D. et al. Automatic detection and counting system for pavement cracks based on PCGAN and YOLO-MF. *IEEE Trans. Intell. Transp. Syst.* **23** (11), 22166–22178 (2022).
15. Varghese, R. & Sambath, M. YOLOv8: A novel object detection algorithm with enhanced performance and robustness. In *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)* 1–6 (2024).
16. Carion, N. et al. End-to-end object detection with transformers. In *European Conference on Computer Vision* 213–229 (2020).
17. Ren, S., He, K., Girshick, R. & Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **39** (6), 1137–1149 (2016).
18. Tan, M., Pang, R. & Le, Q. V. Efficientdet: Scalable and efficient object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* 10781–10790 (2020).
19. Zheng, G., Songtao, L., Feng, W., Zeming, L. & Jian, S. YOLOX: Exceeding YOLO series in 2021. arxiv preprint [arXiv:2107.08430](https://arxiv.org/abs/2107.08430) (2021).

Author contributions

Z.H wrote the main manuscript text, H.M conducted data Curation and formal analysis, L.Y prepared Figs. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 and 14 and provided funding acquisition. All authors reviewed the manuscript.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research work was financially supported by Scientific Research Project of Hengyang City Guiding Plan (202323016888), Guangdong Provincial Department of Housing and Urban-Rural Development Science and Technology Innovation Plan Project (2023-K49-103119), Guangdong Lingnan Township Green Building Industrialization Engineering Technology Research Center (2023GLTG002), College Student Innovation and Entrepreneurship Training Program Project of Guangdong Province (S202411347096).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to L.Y.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025