Advances in Geo-Energy Research

Vol. 15, No. 2, p. 129-142, 2025

Original article

Intelligent identification of coal fractures using an improved U-shaped network

Dengke Wang[1,2], Lu Li[1], Hongtu Zhang[1]*, Shengwei Li[3], Fukai Zhang[4], Yuling Xia[4]

[1]*Henan Provincial Key Laboratory of Gas Geology and Gas Control, Henan Polytechnic University, Jiaozuo 454000, P. R. China*

[2]*State Key Laboratory of Coal Mine Disaster Dynamics and Control, Chongqing University, Chongqing 400044, P. R. China*

[3]*Faculty of Engineering and Natural Sciences, Tampere University, POB 589, Tampere 33014, Finland*

[4]*School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454000, P. R. China*

Keywords:
Fracture identification
U-shaped network
convolutional neural network
gas extraction
coal mine safety

Cited as:
Wang, D., Li, L., Zhang, H., Li, S., Zhang, F., Xia, Y. Intelligent identification of coal fractures using an improved U-shaped network. Advances in Geo-Energy Research, 2025, 15(2): 129-142.
https://doi.org/10.46690/ager.2025.02.05

Abstract:
To address the challenges of coal fracture image recognition, including interference from gangue and multiscale fractures, a multiscale coal fracture segmentation network model to significantly enhance the recognition of coal fracture structures is proposed. The model significantly enhances the recognition of fracture structures based on a U-shaped network architecture and the incorporation of several advanced techniques, including transfer learning, depthwise separable atrous convolutions, and residual modules. Transfer learning, by leveraging pretrained visual geometry group 16-layer network weights, bolsters the feature extraction capabilities of an encoder. Simultaneously, the integration of depthwise separable atrous convolutions and residual modules optimizes a decoder, thereby improving segmentation accuracy and the robust recognition of fractures within images. Experimental results based on qualitative and quantitative data showed that the proposed model surpassed traditional convolutional neural networks, demonstrating proficiency in identifying multiscale fractures in complex coal images. The model was applied to the identification of fractures in roadway surrounding rock boreholes. By extracting fractures from borehole imaging videos and planar diagrams, and conducting cross-validation, the study precisely delineated the fracture distribution. Additionally, to improve coal seam gas extraction efficiency, the grouting and sealing range for cross-layer extraction boreholes was determined.

1. Introduction

With deep mining becoming increasingly common in coal mines, compound dynamic disasters due to coal seam gas have greatly increased, seriously threatening the safety and efficiency of deep mining and the coordinated development of mining environments (He and Wang, 2023). Coal comprises various defects, including fractures and pores, which inevitably influence the occurrence and migration of coal seam gas (Liu et al., 2021). Meanwhile, dynamic disasters in mines triggered by the instability and failure of coal structures may change and expand pores and fractures within coal (Wang et al., 2024). Therefore, the identification of coal fractures has great practical significance for studying the laws governing the migration of coal seam gas and the stability of the surrounding rock.

Fractures are critical sites for gas leakage and play a significant role in gas-related disasters. In recent years, various detection technologies have been developed to analyze pore and fracture networks in coal at different measurement scales. These include methods such as mercury intrusion porosimetry (De Castro et al., 2020), low-pressure nitrogen adsorption

Yandy Scientific Press

*Corresponding author.
E-mail address: wdk@hpu.edu.cn (D. Wang); 13934802132@163.com (L. Li); zhanghongtu@hpu.edu.cn (H. Zhang); lishengwei_2020@163.com (S. Li); zhangfukai@hpu.edu.cn (F. Zhang); xyl@hpu.edu.cn (Y. Xia).

Received November 1, 2024; revised December 2, 2024; accepted December 25, 2024; available online December 30, 2024.

(Zhao et al., 2019), nuclear magnetic resonance (Liu et al., 2020), X-ray computed tomography (CT) (Li et al., 2020), scanning electron microscopy (Roslin et al., 2019), focused ion beam-scanning electron microscopy (Fang et al., 2019), and acoustic emission and ultrasonic techniques (Shi et al., 2019). Among these technologies, many scholars have favored CT technology, which allows sample destruction to be avoided and macroscopic and microscopic measurement results to be obtained. Li et al. (2012) used low-field nuclear magnetic resonance and industrial CT scanning to quantitatively characterize the physical properties of coal with different structures, uncovering relationships between the coal structures and their physical properties. Hao et al. (2020) applied industrial CT scanning technology to analyze the extension direction, development degree, and connectivity of primary cracks in coal under different loads from a microscopic perspective. Wang et al. (2022) conducted laboratory simulations of hydraulic fracturing in coal and utilized industrial CT scanning combined with digital volume correlation to quantify the spatial distribution, structural changes, and expansion of fractures. Wu et al. (2022) examined the microscopic damage evolution and fracture characteristics of coal under uniaxial and triaxial loading conditions using industrial CT scanning. In these studies, the researchers have employed industrial CT technology to provide comprehensive insights into the internal structures of coal, the evolution of fractures within it, and coal's dynamic responses to external forces. Corresponding image data have also been used as analytical objects for digital image processing, offering significant experimental support and theoretical guidance for coalbed methane exploitation and coal mine gas control.

Advances in digital image processing technology have led to the widespread application of image segmentation algorithms, including threshold segmentation, edge detection, and wavelet transform, to coal image analysis. Liu and Cai (2008) applied digital image processing techniques to crack analysis, utilizing a multiscale wavelet transform method to detect image edges and quantitatively assess cracks. Liu et al. (2009) applied a least squares support vector machine to predict the gray values of pixel neighborhoods within CT images. They developed an edge detection technique that integrates gradient and zero-crossing information. Bai and Liu (2009) integrated image processing techniques into rock fracture expansion analysis using a wavelet transform algorithm to detect edges in CT images and to quantitatively measure fractures. Wang et al. (2016) applied Wiener filtering to denoise CT images of coal mass fractures, performed threshold segmentation on a digital terrain model, and reconstructed coal pore structures in three dimensions. Mohebbi et al. (2017) employed the Hough transform to detect structural surface traces in rock masses, while Leng et al. (2021) used a canny operator for edge detection of structural surface traces to enhance the precision of image processing. Lou et al. (2023) employed empirical wavelet transform to filter and denoise electromagnetic radiation waveforms induced by coal fractures, thereby enhancing the accuracy and effectiveness of coal fracture identification. These studies demonstrate the extensive application of digital image processing technology to the segmentation and analysis of coal images. Processing CT images facilitates a more comprehensive understanding of the physical properties and fracture structures of coal, providing accurate data inputs for subsequent computer vision and deep-learning models. Despite these advances, certain limitations remain. Further research is needed to enhance the real-time performance, adaptability, and versatility of such algorithms for different coal types and to promote the broader application of digital image processing technology.

Computer vision technology has been increasingly used to detect fractures in coal roadway heading faces. Long et al. (2015) proposed a fully convolutional network model in which traditional fully connected layers were substituted with convolutional layers. This architecture leveraged skip connections to combine global and local features, enabling pixel-wise classification. Ronneberger et al. (2015) proposed a U-shaped network (U-Net) comprising an encoder-decoder structure to allow the semantic information extracted from images to be effectively learned and high-quality image segmentation to be achieved. The researchers improved the segmentation accuracy of the model by combining shallow and deep features via skip connections. Chaurasia and Culurciello (2017) proposed a semantic segmentation network based on deep learning that comprised an encoder-decoder structure. In the encoder section, the researchers incorporated techniques such as residual connections and depthwise separable convolutions to enhance the network's capacity to extract and process image features. Roy et al. (2019) introduced a parallel spatial and channel squeeze-and-excitation module, demonstrating that combined residual modules and attention mechanisms enhanced the focus on and extraction of feature information. Xie et al. (2019) applied a fully connected layer for individual pixel classification following convolutional neural network processing to achieve surface crack segmentation in coal subjected to impact loading from a Hopkinson pressure bar. Zhang et al. (2020) enhanced U-Net architecture by incorporating inception modules and dense connections, thereby significantly expanding the receptive field to extract detailed features of coal fractures. Lu et al. (2020) developed a convolutional neural network that fused parallel multiscale features to alleviate weak boundary fracture loss and improve the accuracy of coal mass fracture segmentation. Meanwhile, Zheng et al. (2022) achieved intelligent segmentation of coal fracture CT images using a deep learning U-Net model and completed three-dimensional reconstruction based on the segmentation results. Yan et al. (2023) proposed a coal fracture recognition network model based on an improved deeplab version 3 plus model. They employed a lightweight mobilenet version 2 module as the backbone feature extraction network for this model, which reduced the model parameters. These studies highlight the diverse applications of deep learning techniques to the intelligent recognition of coal fractures. However, challenges persist regarding the robustness, versatility, and adaptability of these models in dealing with small sample sizes, indicating a need for further research.

Traditional methods for identifying coal fractures have proved highly successful, particularly for simple images in which fractures exhibit clear edges and high contrast with the

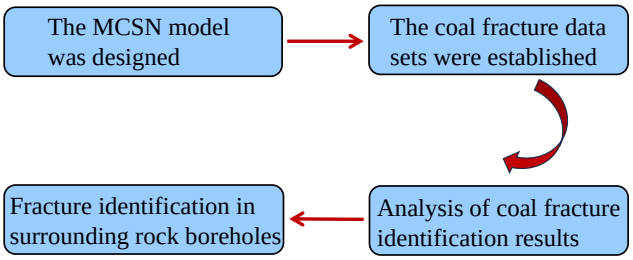

Fig. 1. Flowchart of the overall research procedure.

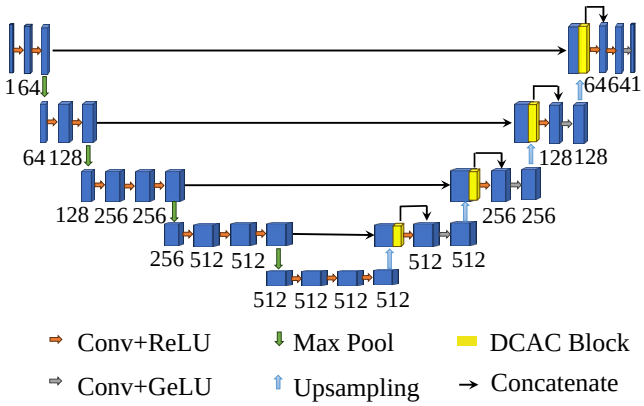

Fig. 2. The architecture of the MCSN model.

background. The algorithms they incorporate generally perform well under such conditions. However, several challenges remain. The robustness of existing models in noisy environments is inadequate, which often compromises crack detection accuracy. Furthermore, the complexity of multiscale fractures limits some models' ability to effectively integrate feature information across different scales, thereby impacting overall recognition performance. Hence, it is crucial to conduct research on deep learning algorithms specifically for coal fractures. Such research would facilitate the automated detection of fractures, eliminating the need for traditional labor-intensive manual methods while better accommodating the complex and variable structures of coal fractures.

For engineering applications, accurate and efficient identification of coal fractures is essential for ensuring coal mine safety, underground engineering stability, and disaster prevention in mining operations. Multiscale fusion strategies, improved depthwise separable atrous convolution (DCAC) modules and residual modules can be incorporated into the design of multiscale coal fracture segmentation networks (MCSN). Compared to existing technologies, the MCSN model offers enhanced accuracy, robustness, and practicality, making them particularly well-suited for handling complex, multiscale coal fracture images.

2. U-Net network and overall framework

A schematic diagram of U-Net (a convolutional neural network for image segmentation), named for its distinctive U-shaped architecture. The network model primarily comprises two key components: a symmetrically arranged encoder and a decoder. The encoder consists of five downsampling modules, each incorporating a 3 × 3 convolution kernel, a rectified linear unit (ReLU) activation function, and a 2 × 2 max pooling layer. Each pooling operation reduces the feature map size by half. The decoder comprises five upsampling modules, each featuring a 3 × 3 convolution kernel and a 2 × 2 deconvolution layer. These modules progressively restore the resolution of the feature map as the decoding process unfolds. Additionally, U-Net employs skip connections between the encoder and decoder to facilitate the integration of shallow and deep features. Finally, a 1 × 1 convolution kernel carries out classification and produces the final semantic segmentation output.

Based on U-Net, to accomplish more accurate image segmentation, the research procedure shown in Fig. 1 is followed, which comprises the following four steps:

1) Design an MCSN model for the identification of coal fractures.
2) Construct a dataset of coal fracture images.
3) Perform qualitative and quantitative analyses of the coal fracture image extraction results.
4) Apply the MCSN model to identify fractures in the surrounding rock boreholes of roadways.

3. The MCSN network model

This section first presents an overview of the MCSN model architecture and provides detailed sequential descriptions of the multiscale fusion strategy, the DCAC module, and the residual module.

3.1 Overall network architecture

The MCSN model developed is based on U-Net network architecture and uses an encoder-decoder structure combined with skip connections to segment complete crack images from complex coal mass images. The overall structure of the model is illustrated in Fig. 2.

The encoder progressively reduced image resolution while extracting high-level features. Combined convolutional and pooling layers were employed for downsampling to capture features at different scales. A visual geometry group 16-layer (VGG16) network was adopted as the backbone for feature extraction and a 3 × 3 convolution kernel was employed in the five convolutional layers, with kernels per layer set at 64, 128, 256, 512, and 512, respectively (Alshammari, 2022). A ReLU activation function was applied each convolutional layer to enhance the network's nonlinear representation capacity. The VGG16 backbone reduced the input image size from 512 × 512 pixels with 3 channels to 32 × 32 pixels with 512 channels, resulting in five initial feature layers that effectively concentrated the features of coal fractures.

The decoder restores the high-level features extracted by the encoder to the original image dimensions through upsampling and skip connection operations to produce a binary segmentation map of coal fractures. During upsampling, neighboring interpolation expands the feature map size, which accelerates the training of the segmentation model. An improved feature extraction network with a redesigned convolutional module was developed to achieve spatially cascaded depthwise separable atrous convolutions, and to mitigate the network de-

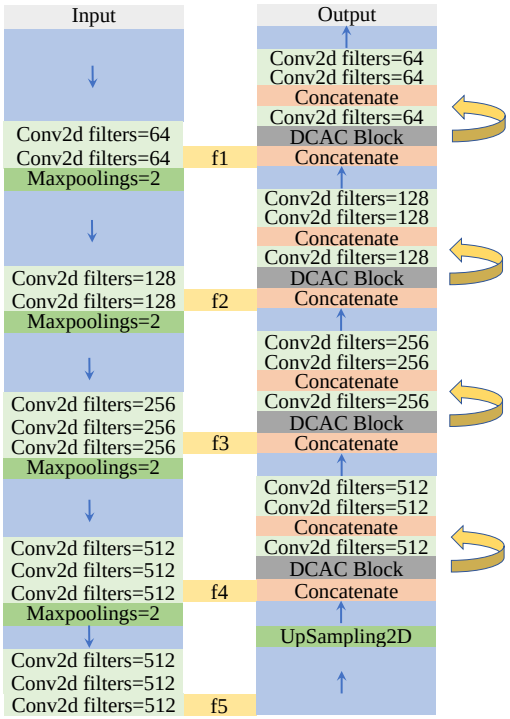

Fig. 3. Multiscale fusion architecture diagram.

gradation that typically occurs with increased depth, residual connections were incorporated into the new design.

Shallow features capture more detailed spatial and positional information, and deep features encapsulate richer semantic content. By fusing five preliminary effective feature layers, the model recovered spatial details in the decoder while accurately understanding semantic information. Through skip connections, this fusion mechanism enabled the model to better address complex scenarios involving shape variations, scale changes, and occlusions, thereby enhancing segmentation accuracy and robustness.

3.2 Multiscale fusion process

The concept of multiscale fusion underpinning the MCSN model refers to the integration of feature information from different scales into the encoder and decoder, which facilitates the comprehensive use of local and global contextual information for accurate semantic segmentation prediction. This fusion is achieved primarily through two methods: 1) employing multiscale receptive fields in the network structure modules to capture feature information at varying scales simultaneously and 2) introducing a multiscale feature fusion mechanism into the MCSN, as illustrated in Fig. 3.

The decoder fused five preliminary effective feature layers, each of which contains information at different scales. The fusion process began with 2 × 2 upsampling with a stride of 2 on the initial feature layers extracted by the encoder. These layers were then concatenated to perform the first feature fusion of the multiscale information. The shallow feature images, which retained detailed information about the original image at different scales, were stacked to preserve self-similarity. Subsequently, the fused feature maps were input into DCAC modules, where convolution-based feature maps were stacked to reintegrate the multiscale information. The atrous convolution captures structural information from various receptive fields, and self-similarity effectively enhanced feature representation.

3.3 Depthwise separable atrous convolution module

Coal fractures have diverse, complex shapes, typically with a small number of fractures, and fracture pixels comprise only a small portion of the entire coal image. To address the issue of data class imbalance and enhance the accuracy and efficiency of fracture recognition, a spatially cascaded DCAC module was designed for the feature reconstruction process. This module integrates atrous convolution with depthwise separable convolution (Wang et al., 2019; Dai et al., 2023). Traditional convolution operations focus primarily on local neighborhood information, which may fail to capture the fine features of small, sparse fractures. In contrast, atrous convolution is a specialized operation that expands a receptive field without increasing the number of parameters or computational loads, thereby improving the model's ability to capture long-range contextual information. The introduction of a dilation rate allows the convolution kernel to cover a broader input area within the receptive field. The dilation rate is defined as:

$$y(m,n)=\sum_{i=1}^{M}\sum_{j=1}^{N}u(m+r\times i,n+r\times j)w(i,j) \tag{1}$$

where $y(m,n)$ denotes the output feature map obtained via atrous convolution; m and n represent the horizontal and vertical coordinates of the output feature map, respectively; and M and N represent the length and width of the convoluted image, respectively. u refers to the input image; r is the dilation rate, which controls the spacing within the convolution; w represents the convolution kernel; and i and j are the indices of the convolution kernel.

Atrous convolution with varying dilation rates is illustrated in Fig. 4, using a 3 × 3 convolution kernel as an example. The red dots indicate the positions at which the convolution kernel operates on the input, while the blue area represents the original receptive field of the input. At a dilation rate of 1, there is no gap between the convolution kernel elements, which corresponds to a conventional convolution operation. This is ideal for detecting subtle fracture structures, such as microscopic cracks or slight texture variations, ensuring that even small fracture features are not overlooked by the model. At a dilation rate of 2, the spacing between convolution kernels increases, allowing the model to capture the relationships between local features and their surrounding contexts. This facilitates the integration of broader contextual information, making it particularly useful for analyzing prominent macroscopic fracture features. Similarly, at a dilation rate of 4, the model can capture a wider range of contextual information. This dilation rate enhances the model's ability to discern broader spatial relationships, allowing large-scale fractures or other structural features to be identified.

Coal has small and large macrocracks that range widely

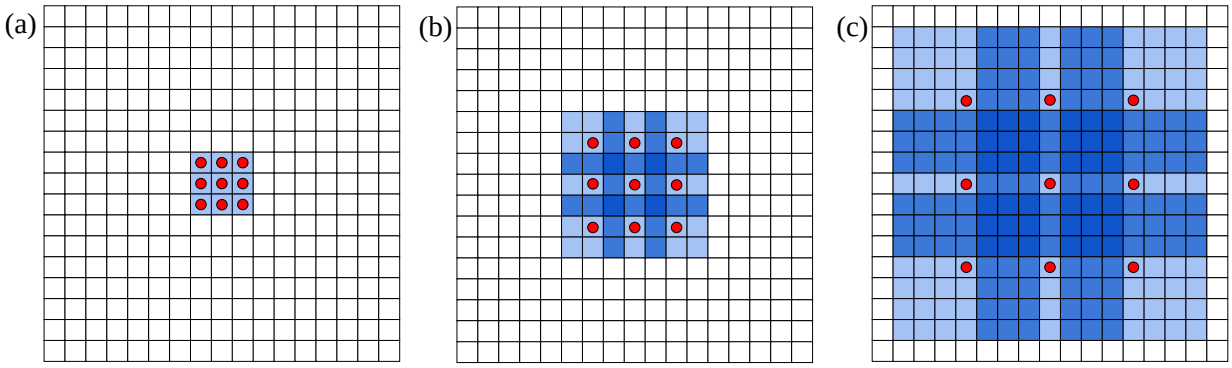

Fig. 4. Receptive field at different dilation rates. (a) dilation rate of 1, (b) dilation rate of 2, and (c) dilation rate of 4.

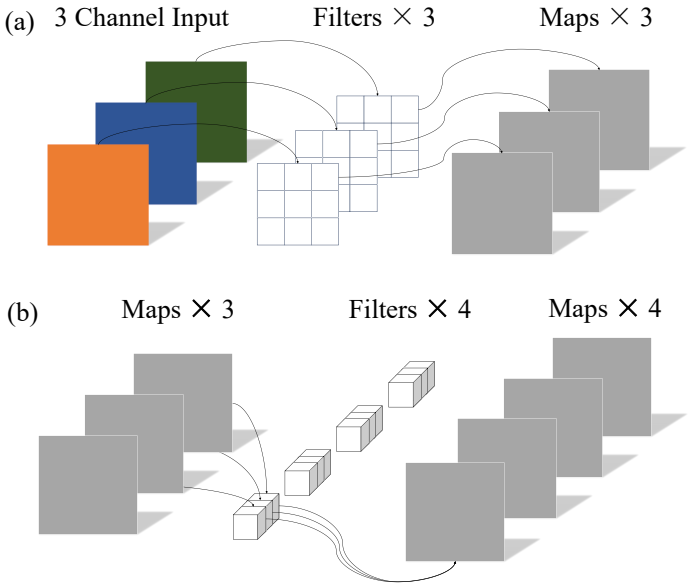

Fig. 5. Depthwise separable convolutions. (a) depthwise convolution and (b) pointwise convolution.

in sizes. Thus, during the identification process, crack structures must be considered across multiple scales. However, indiscriminately increasing the dilation rate may result in the loss of critical details or the blurring of local features within an image. In fracture segmentation, maintaining a balance between local detail and macroscopic structure is crucial for achieving optimal performance. To address this challenge and preserve local perceptual capabilities while capturing as many fine details as possible, dilated convolutions with dilation rates of 1, 2, and 4 were applied in parallel. This strategy captured more dimensional feature information and enhance the accuracy of crack segmentation.

Depthwise separable convolution was regarded as another type of convolution operation. It decomposes standard convolution into two distinct steps: depthwise convolution and pointwise convolution. In depthwise convolution, the convolution was applied independently to each input channel, whereas in pointwise convolution, a 1 × 1 convolution kernel was used to convolve the output generated by the depthwise convolution step. This two-step process effectively combined the feature maps produced by depthwise convolution, mapping them into the final feature space, as shown in Fig. 5.

By integrating the spatial cascade of dilated convolutions into depthwise separable convolutions, the MCSN model can better capture relationships and semantic information between objects, particularly when dealing with small structures, such as cracks. For each input channel of the feature map, three convolutions of the depthwise convolution module were applied separately. The resulting output feature maps were then concatenated along the channel dimension, creating a larger feature map that was passed to the pointwise convolution layer. Finally, the features are mapped to the desired feature space. This method enabled each dilation rate to effectively capture contextual information at different scales, while pointwise convolution combines the output feature maps and adjusts the number of output channels. The multiscale fusion of depthwise separable atrous convolutions significantly enhanced the model's ability to segment complex details, such as edges, textures, and small structures.

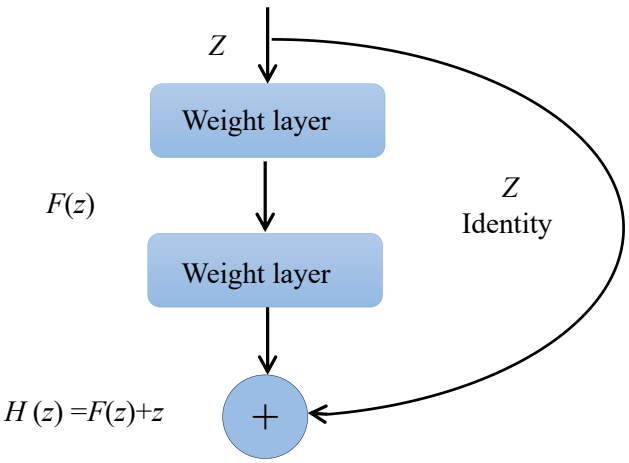

Fig. 6. Residual structure.

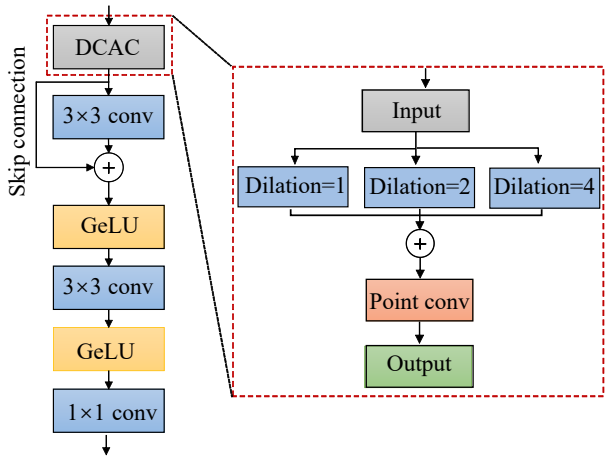

Fig. 7. Multiscale feature extraction network.

3.4 Residual module

Residual networks were employed to address the issues of gradient vanishing and explosion that arise as model depth increases (He et al., 2016). The residual module is formally defined as:

$$y = F(z,\{W_i\}) + z \tag{2}$$

where z represents the input of the residual module, y represents the output of the residual module, and $F(z,W_i)$ is the residual map to be learned by the network.

Unlike traditional convolutional neural network architectures, the residual neural network employs "shortcut" connections-also referred to as skip connections-as illustrated in Fig. 6. This modification changes the input to the activation function from the traditional network output $H(z) = F(z)$ to $H(z) = F(z) + z$. This structural enhancement can improve the model's expressive capability, with the activation function being the ReLU.

The network was constructed based on residual learning, incorporating a newly designed convolutional component. The original convolution operation was replaced with two newly designed 3 × 3 standard convolution modules, as depicted in Fig. 7. The entire module facilitates identity mapping through shortcut connections by forming residual connections. The activation function employed is the Gaussian error linear unit (GeLU) (Minhyeok, 2023)-a high-performance activation function for neural networks. The formula is:

$$xP(X \le x) = x\emptyset(x) \tag{3}$$

where $\emptyset(x)$ is the cumulative distribution of Gaussian normal distribution. $P(X \le x)$ denotes the probability that the stochastic variable X is no greater than x, where x symbolizes the value fed into the activation function. The complete form of Gaussian normal distribution is:

$$xP(X \le x) = x\int_{-\infty}^{x} \frac{e^{\frac{-(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} dX \tag{4}$$

where μ and σ represent the mean and standard deviation of the normal distribution, respectively.

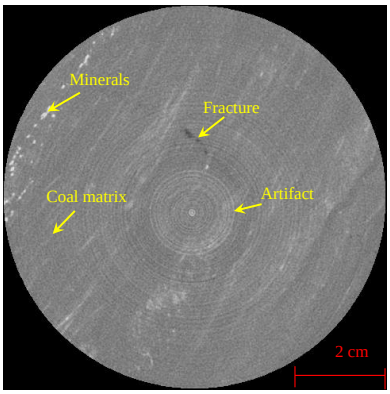

Fig. 8. CT scanning image.

This approach facilitates the direct propagation of gradients during backpropagation, thereby enhancing the model's training efficiency and alleviating the vanishing gradient problem.

4. Experimental description and data

4.1 Sample preparation and experimental equipment

Coal blocks were collected from the main mining seam of the Zhaogu No. 2 Mine, located in Xinxiang City, Henan Province, China. The large coal blocks were processed into cylindrical specimens, each with a 25-mm diameter and 50-mm height. The industrial CT rock scanning system proposed by Wang et al. (2021) was employed to scan four coal samples.

4.2 Scan results

The coal samples were scanned layer by layer from top to bottom, using representative slice images of the coal fractures, as shown in Fig. 8. According to the principles of industrial CT scanning, different gray values correspond to varying material densities in the scanned images. Typically, fractures have the lowest density and thus appear as black regions in the images. The density of a coal matrix is higher than that of the fractures it contains, but lower than that of the minerals; therefore, the coal matrix appears as a dark gray area in the images. Miner-

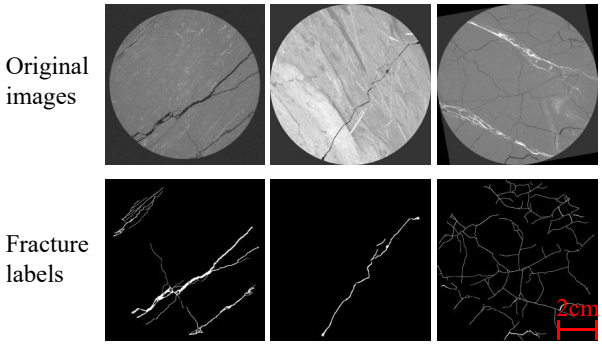

Fig. 9. Distinguishing fractures in CT slices.

als, being the densest components, are represented by white or light gray areas. However, due to the limitations of CT scanning technology, ring-like artifacts may occur during imaging, and the heterogeneity of the coal mass introduces several key challenges: 1) the gray values of fractures may resemble those of the coal matrix, 2) the gray values of fractures may be close to that of gangue, and 3) the topological structure of fracture development is complex.

4.3 Dataset production

Datasets are essential for deep learning, directly influencing the performance and generalization ability of the MCSN model. Given the absence of an open-source CT segmentation image dataset for coal fractures, a custom dataset was manually created comprising 400 images derived from CT slices of four coal samples. Specifically, 100 CT slices from each coal sample were selected, which resulted in 400 images, including both tensile and shear fractures. These fractures varied in size, shape, and orientation, reflecting the inherent complexity and heterogeneity of coal masses.

The image processing software ImageJ was employed to annotate the CT slice images at the pixel level and to identify the location, shape, and size of the fractures. During the annotation process, fractures and backgrounds were categorized into two classes using distinct colors: fracture regions marked as white areas with a grayscale value of 225 and nonfracture regions marked as black areas with a grayscale value of 0, as illustrated in Fig. 9.

To augment the dataset and enhance the model's generalization and robustness, data augmentation techniques were applied to the original images, including spatial transformations such as fixed-angle rotation (ranging from -30° to 30°) and horizontal flipping. Pixel transformations by adding salt-and-pepper noise and Gaussian blur were also applied to simulate different image acquisition conditions and environments. The augmented dataset, comprising 3,200 images (including the original data), was divided into a test set (320 images), a validation set (640 images), and a training set (2,240 images) at a 1:2:7 ratio. Stratified sampling was employed to ensured that fracture images of the coal samples were evenly distributed across the training, validation, and test sets according to fracture type. This approach provided a sufficient volume of training data to enable the model to better learn and generalize complex fracture structures.

4.4 Test environment and parameter settings

A computing platform running the Windows 10 operating system was used as the test environment. The platform was equipped with an Intel(R) Xeon(R) Gold 6226R CPU @ 2.90 GHz processor, 192 GB of memory, and an NVIDIA Quadro RTX 5000 graphics card. Python was employed as the programming language and pytorch as the deep learning framework.

To train the MCSN model, a transfer learning technique based on frozen layers was used, and the hyperparameters were configured as follows: a batch size of 6, a learning rate of 0.0001, and a momentum of 0.9. Additionally, the model was trained for 100 epochs to ensure convergence during the training process.

4.5 Loss function

Dice loss is commonly used as a loss function for image segmentation, especially for tasks involving class imbalances. Unlike other loss functions, such as cross-entropy loss, dice loss effectively handles imbalanced segmentation problems because it emphasizes object classes rather than the background. In the case of coal fracture extraction, where the background occupies most of the image and fractures are comparatively small, dice loss helps reduce the background's influence on the loss function, leading to more accurate fracture segmentation.

The similarity between segmentation results and the ground truth was measured using the dice coefficient, which is a widely used evaluation metric employed to quantify the overlap between the predicted and actual regions (Wang et al., 2020). Derived from the dice coefficient, dice loss was served as the loss function for segmentation tasks. A dice loss value of 0 signifies perfect overlap between the predicted and true labels, while a value of 1 indicates no overlap; thus, a value close to 1 indicates a high degree of overlap and more accurate segmentation. Low dice loss values reflect a better alignment between a model's predictions and the ground truth. The dice loss can be calculated by:

$$\text{dice} = \frac{2\left|X \cap Y\right|}{\left|X\right| + \left|Y\right|} \tag{5}$$

$$\text{dl} = 1 - \frac{2\left|X \cap Y\right|}{\left|X\right| + \left|Y\right|} \tag{6}$$

where dl stands for dice loss, and X and Y represent the true results and predicted results of the task, respectively.

4.6 Evaluation indicators

The segmentation of coal fractures involves pixel classification, with each pixel classified as either a part of the background or a fracture. Confusion matrices are widely used evaluation tools for such classification and for assessing the segmentation accuracy and performance of a model during the prediction process.The matrix was derived from four key classification outcomes-true positive, true negative, false positive, and false negative-and used to calculate several

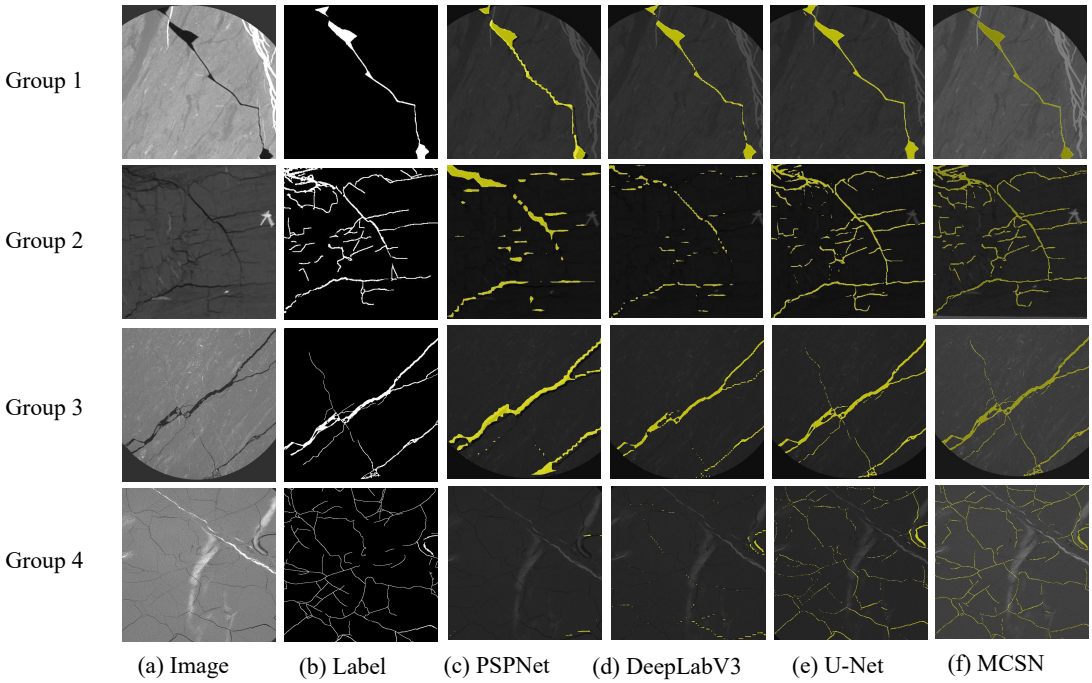

Fig. 10. Visual comparison of the different network models.

commonly used evaluation metrics for image segmentation, including recall, precision, mean pixel accuracy (MPA), and mean intersection over union (MIoU). The specific formulas for these metrics are:

$$\mathrm{Re} = \frac{P_{ii}}{P_{ii} + P_{ij}} \tag{7}$$

$$\mathrm{Pre} = \frac{P_{ii}}{P_{ii} + P_{ji}} \tag{8}$$

$$\mathrm{MPA} = \frac{1}{k+1} \sum_{i=0}^{k} \frac{P_{ii}}{\sum_{j=0}^{k} P_{ij}} \tag{9}$$

$$\mathrm{MIoU} = \frac{1}{k+1} \sum_{i=0}^{k} \frac{P_{ii}}{\sum_{j=0}^{k} P_{ij} + \sum_{j=0}^{k} P_{ji} - P_{ii}} \tag{10}$$

where Re stands for recall, Pre stands for precision, k is the number of foreground classes, and i and j are the pixel categories. P_{ii} represents a true positive outcome, which is the number of actual fracture pixels correctly classified as fractures; P_{ij} is a false negative outcome, which represents the number of fracture pixels classified as nonfracture; and P_{ji} represents a false positive outcome (i.e., the number of nonfracture pixels incorrectly classified as fractures).

5. Results and discussion

5.1 Qualitative analysis of the model segmentation results

Four representative image groups were selected to conduct segmentation comparison experiments using a pyramid scene parsing network (PSPNet), deeplab version 3 model (DeepLabV3), U-Net model, and MCSN model, which are all benchmark models widely used for image segmentation. Their performance and technological maturity have been verified, so they provide reliable comparison benchmarks for the effectiveness of new models. These experiments revealed the performance of each algorithm in segmenting fractures under varying conditions, as shown in Fig. 10.

For the first group of coal CT fracture images, the primary challenge lay in the substantial grayscale gradients of the fractures, necessitating the model to handle pixel-level variations with a high degree of precision. Traditional network models struggled with images that have significant grayscale gradients, resulting in blurred edges and insufficient local contrast. Consequently, regions that should been continuous appeared disjointed. In contrast, the MCSN model effectively addressed these issues, enhancing the detection of boundary information and improving the models' ability to capture fine details.

In the second group of coal CT fracture images, the complex topology of the fractures posed challenges. In regions with intricate fracture patterns and numerous fine textures, traditional network models exhibited varying degrees of misdetection. Specifically, for some groups, many fine spiderweb-like structures were mistakenly identified as blocky fractures, leading to the merging of separate fractures. This issue stems from inadequate segmentation performance, which prevents models from reliably preserving fracture shapes and structures. In contrast, the proposed MCSN model correctly identified most fractures, with only a few minor exceptions.

The third group of coal CT fracture images presented the challenge of identifying fine cracks, since the fractures' grayscale values closely resembled those of the coal matrix. Although traditional network models captured the complex features of images, they struggled to accurately detect very small or subtle fissures. The proposed MCSN model captured such fine details, enabling cracks with pixel values nearly indi-

Table 1. Comparative evaluation metrics for different neural network models.

Methods	MIoU (%)	Re (%)	Pre (%)	MPA (%)
PSPNet	68.52	66.04	66.54	66.04
DeepLabV3	75.49	78.10	82.87%	78.10
U-Net	79.19	83.48	89.65	83.49
MCSN	87.69	92.75	91.55	92.75

stinguishable from the background to be identified and improving the precision of fracture detection.

In the fourth group of coal CT fracture images, the primary challenge was interference caused by gangue and minerals. The high gloss and reflectivity of these materials often obscures the edges or fine details of fractures, preventing accurate fracture identification. When the boundaries between fractures and gangue or minerals become blurred, traditional network models misclassified fractures as part of the gangue or minerals, or even failed to detect the fractures altogether, resulting in poor segmentation performance. However, the proposed MCSN model effectively mitigated the interference caused by the heterogeneity of coal, leading to enhanced fracture detection accuracy.

In conclusion, the MCSN model consistently outperformed the other models in terms of fracture identification, regardless of the simplicity or complexity of the fracture structures in the images. It exhibited superior resilience to noise interference, effectively preserving both the trajectory and structural integrity of the fractures.

5.2 Quantitative analysis of the model segmentation results

The MCSN model was compared with three other representative models. The first model (PSPNet) is a deep convolutional neural network developed for semantic image segmentation (Zhao et al., 2017). It is primarily based on a pyramid pooling module that captures contextual information across multiple scales, which enhances the model's ability to interpret and accurately segment images. The second model (DeepLabV3) is a deep learning framework for semantic segmentation (Chen et al., 2018). It combines dilated convolutions, an encoder-decoder architecture, and a spatial pyramid pooling module obtain accurate boundary and contextual information from images. The third model (U-Net) is a convolutional neural network designed specifically for image segmentation tasks (Liu and Wang, 2022). Featuring an encoder-decoder structure with skip connections, U-Net excels at capturing multiscale features, resulting in high-quality segmentation outcomes.

The transfer learning and training methods described previously were applied to these deep learning models and the models were validated using the enhanced dataset. Table 1 shows a comparison of the evaluation metrics for each model.

Although the PSPNet, DeepLabV3, and U-Net models generally demonstrated strong semantic segmentation performance, the performance was suboptimal for the coal CT fracture dataset, as shown in Table 1. The PSPNet model, which utilizes a pyramid pooling module to capture multiscale contextual information, however, due to the abundance of fine texture and edge features in coal CT images, the global scene information was not fully captured or leveraged by the PSPNet model, thereby affecting the accuracy of semantic segmentation. The DeepLabV3 model, based on a deep residual network encoder and incorporating complex operations, such as dilated convolutions and multiscale attention mechanisms, exhibited a highly complex hierarchical structure, which led to overfitting and higher training costs. In addition, expanding the receptive field without pooling reduces image resolution, which caused information loss and negatively affect segmentation accuracy. Compared to more complex architectures, the U-Net model, which was originally designed for medical imaging, has fewer parameters and thus reduces overfitting. However, its reduced complexity limited its learning capability, making it challenging to accurately capture small or subtle features in images. Consequently, there is room to considerably improve segmentation accuracy.

Table 2. Single-image processing times for different network models.

Methods	PSPNet	DeepLabV3	U-Net	MCSN
Time (ms)	130.00	271.61	164.28	182.99

In contrast, the MCSN model consistently outperformed the other models across various evaluation metrics, achieving 87.69% MIoU, 92.75% Re, 91.55% Pre, and 92.75% MPA. This model demonstrated exceptional overall performance and robust generalization for the coal CT image dataset.

The image processing times for each model were evaluated using the previously defined test dataset. The times required by the models to process single images are presented in Table 2. Of the deep learning models were compared in this work, the U-Net model exhibited the fastest processing speed for crack extraction, while the PSPNet and DeepLabV3 models demonstrated the slowest extraction speeds. The proposed MCSN model achieved a moderate processing speed, balancing performance and computational efficiency. Notably, the MCSN model delivered the best crack extraction performance without the longest processing time. This demonstrates that the MCSN model excels in terms of both segmentation accuracy and computational efficiency.

In summary, the proposed MCSN model offers several notable advantages: 1) its robust feature extraction capabilities and expansive receptive field facilitate the precise identification of image details and structures; 2) it effectively integrates multiscale feature information, facilitating the processing of variously sized target objects; and 3) the enhanced residual structure contributes to stable and efficient network training.

5.3 Ablation experiment

To evaluate the effectiveness of each module in coal fracture identification, ablation experiments were designed and the results are listed in Table 3.

Table 3. Results of the ablation experiments.

VGG16	DCAC	Residual module	MIoU (%)	Re (%)	Pre (%)	MPA (%)
×	×	×	79.19	83.48	89.65	83.49
✓	×	×	81.37	90.75	91.03	90.75
✓	✓	×	85.24	91.05	91.45	91.05
✓	✓	✓	87.69	92.75	91.55	92.75

Note: The symbol ✓ indicates a module's inclusion in the basic U-Net network model, while the symbol × indicates its exclusion.

The experimental results demonstrated that the VGG16 network, DCAC modules and residual modules all positively contributed to the model's performance. As shown in Table 3, the introduction of the VGG16 backbone network improved across all evaluation parameters. The VGG16 network's robust feature extraction capabilities enabled the model to better capture and understand complex data features. DCAC modules further improved the model's quantitative evaluation values. This improvement could be attributed to DCAC modules reducing the computational load while enhancing the model's capacity to extract multiscale features by performing separate computations in the spatial and channel dimensions. Additionally, atrous convolutions enabled the network to capture a broader range of features, thereby leading to a better understanding of the structure and context information of fractures. When the VGG16 network, DCAC modules, and residual modules were integrated, the model achieved optimal fracture identification performance. Skip connections allowed the model to learn hierarchical features, enabling the network to perceive cracks at multiple scales and extract them with greater accuracy. Ultimately, when these three enhancements were combined, the MCSN model outperformed the benchmark U-Net model, with improvements in MIoU, Re, Pre, and MPA of 8.5%, 9.27%, 1.9%, and 9.26%, respectively. These combined effects significantly contributed to the model's overall improved performance, thereby providing stronger support for practical applications.

Despite these advances, the MCSN model has some remaining limitations that should be addressed in future research. First, although depthwise separable and atrous convolutions reduced the model parameters and improved computational efficiency, the model still suffers from relatively long training times, particularly when processing large-scale datasets. Consequently, its real-time applications may be limited. Second, the model's accuracy decreases when used for extremely blurred field images, which adversely affects crack extraction performance. This degradation is primarily due to factors such as suboptimal lighting conditions, equipment limitations, and operational errors. Future researchers should focus on optimizing the model's structure to further reduce parameters and improve crack identification accuracy. These enhancements will increase the model's performance in practice and strengthen its adaptability and robustness in complex environments.

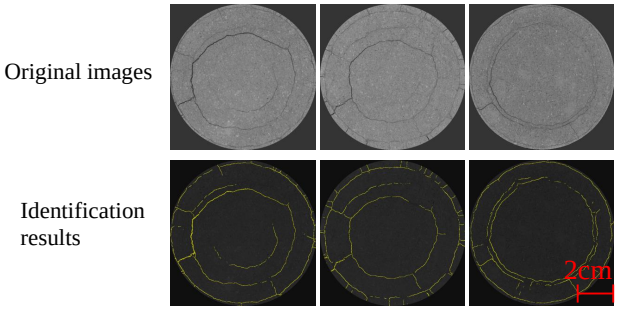

Fig. 11. Identification results for rock fractures.

6. Engineering applications

6.1 Ablation experiment

The rock samples were scanned using an industrial CT scanning system, which allowed the extraction of clear, representative images for use in the experimental testing of the proposed MCSN model. Some of the fracture recognition results are presented in Fig. 11.

Clearly, the MCSN model effectively captured and accurately classified fracture features in rock images, demonstrating its broad applicability. Moreover, the model achieved an MIoU of 86%, an Re of 91.41%, a Pre of 91.6%, and an MPA of 91.41% for the rock fracture images. This indicates that the model is suitable for analyzing the internal structures of coal seams and assessing other complex geological materials. Hence, the underlying principles of the MCSN model can be directly applied in engineering environments, such as to identify borehole fractures in the surrounding rocks of roadways.

6.2 The project profile

The borehole imaging was conducted in an engineering context in the bottom extraction roadway 13,161 of the Baiping Coal Mine in Henan Province, China. The primary coal extraction seam at this mine is the No. 2-1 coal seam, in which the lithological distribution of the 13,161 working face roof primarily consists of a large sandstone layer containing medium-grained sandstone. The mechanical parameters of this rock layer indicate high strength and good integrity, making it favorable for borehole drilling. The average thickness of the coal seam is 5.5 m, with inclination angles ranging from 7° to 27°. The maximum measured gas content is 15.8 m^3/t, and the maximum gas pressure was 1.70 MPa. There are no

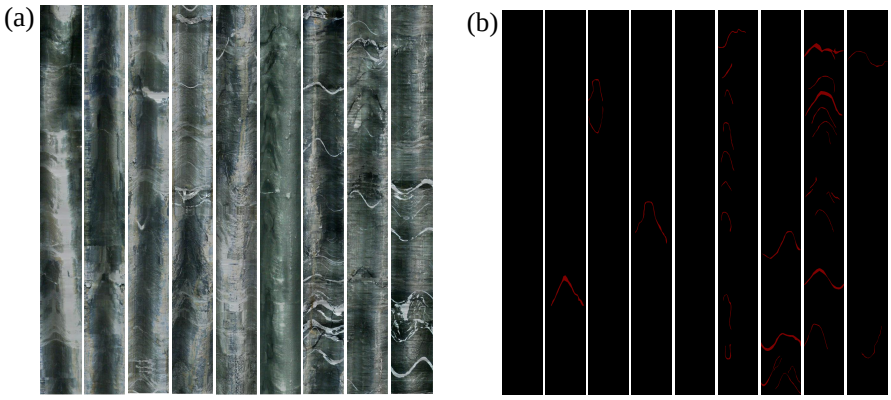

Fig. 12. Intelligent fracture identification results for borehole fractures (unfolded display). (a) unfolded display of borehole fractures in the surrounding rock and (b) fracture identification results for the MCSN model.

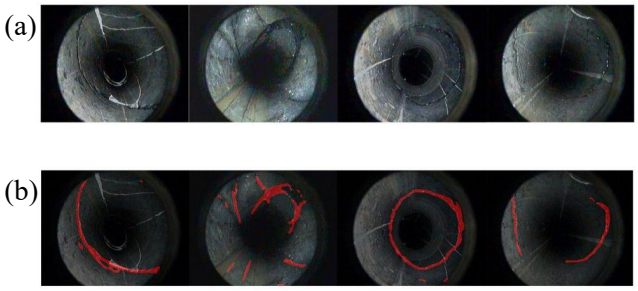

Fig. 13. Intelligent fracture identification results for borehole fractures (cross-sectional display). (a) cross-sectional display of borehole fractures in the surrounding rock and (b) fracture identification results for the MCSN model

risks of coal and gas outbursts in the main mining coal seam.

6.3 Fracture identification results for surrounding rock boreholes

Borehole imaging devices were used to collect visual data, and the MCSN model was applied to intelligently analyze the data and visualize fractures in the roadway's surrounding rock. The visual data consisted of borehole videos and planar unfolded images. By extracting fractures from these images and performing cross-validation, the accuracy and reliability of the analysis was significantly improved. As shown in Figs. 12 and 13, the model precisely classified each pixel into a fracture or nonfracture category. Fracture pixels were marked in red, while nonfracture pixels were marked in black, allowing fractures and nonfractures to be clearly distinguished. A few representative borehole images were selected for display. The fracture sizes varied, reflecting the complexity of underground fracture structures. The fractures showed no significant fillings, and the borehole walls maintained good integrity.

6.4 Determination of the grouting range and verification of gas extraction effectiveness

Based on the comprehensive analysis of borehole endoscopic results, with the issue of varying sealing lengths under different hole-opening positions and borehole inclination angles taken into account, an outer normal was constructed through the coal seam floor. The vertical distance from the opening position to the coal seam floor was used as a relative reference standard to describe the extent of the seal zone. According to the endoscopic results, the loosening ranges of the surrounding rock in the bottom gas extraction roadway (No. 13,161) were 3.5-4.7 m and 5.2-7.6 m, indicating the presence of two distinct fracture loosening zones within the surrounding rock. After further consideration, the overall loosening range of the surrounding rock was established as 3.5-7.6 m. Table 4 shows the recommended sealing ranges for specific borehole positions and inclinations, with loosening and sealing ranges calculated from the borehole opening position in the roadway.

Based on the identified fracture distribution ranges, the fracture zones in the surrounding rock boreholes were effectively sealed to minimize air leakage, thereby enhancing the gas extraction concentrations and purity. Field measurements of gas extraction concentrations were conducted in the bottom gas extraction roadway (No. 13,161) of the mine, where cross-layer boreholes had been drilled. In this roadway, one set of test boreholes and one set of reference boreholes were drilled, with each set comprising four boreholes. For the test group, the grouting range was determined based on the conclusions drawn from the fracture distribution analysis, while for the reference group, conventional practices for determining the grouting range were followed. Compared to the reference boreholes, the average gas concentration in the test group boreholes increased by 31% within the first 7 days, 21% within the first 15 days, and 25% within the first 30 days. Throughout the entire extraction period, the average gas concentration in the test group boreholes increased by 15%. A comparison of the gas concentrations between the reference and test groups is shown in Fig. 14. Compared to the reference boreholes, the average pure gas flow rate in the test group boreholes increased by 26% within the first 7 days, 14% within the first 15 days, and 20% within the first 30 days. Over the entire extraction period, the average pure gas flow rate increased by 12%. A comparison of pure gas flow rate between the reference and test groups is presented in Fig. 15. The results indicate that intelligent identification of surrounding rock borehole fractures, followed

Table 4. Sealing ranges for different borehole positions and inclinations.

Borehole inclination (°)	Distance from the center line (m)	Distance from the coal roof (m)	Distance from the right side (m)	Designed borehole length (m)	Surrounding rock loosening range (m)	Recommended sealing range (m)
5	2.5	2.5	0	38	11.8-15.9 17.5-21.6	24-28
15	2.5	2.5	0	29	8.4-15.8	17-29
30	2.5	2.5	0	22	6.0-8.4 9.0-11.4	15-22
45	2.5	2.5	0	18	5.1-9.7	11-18
60	0	0	2.5	22	5.7-8.0 8.9-11.4	14-22
75	0	0	2.5	25	4.5-8.6 10.2-12.4	15-25

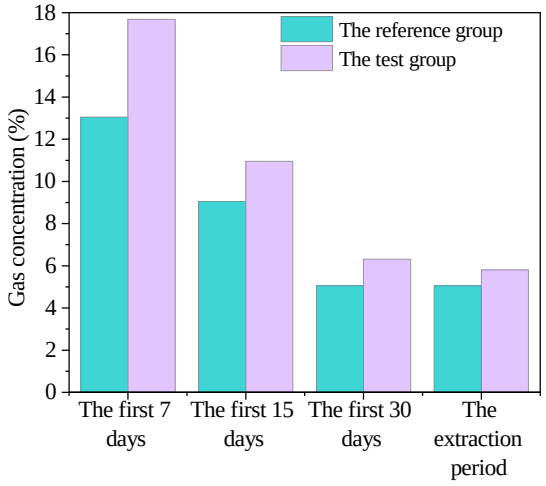

Fig. 14. Comparison of gas extraction concentrations.

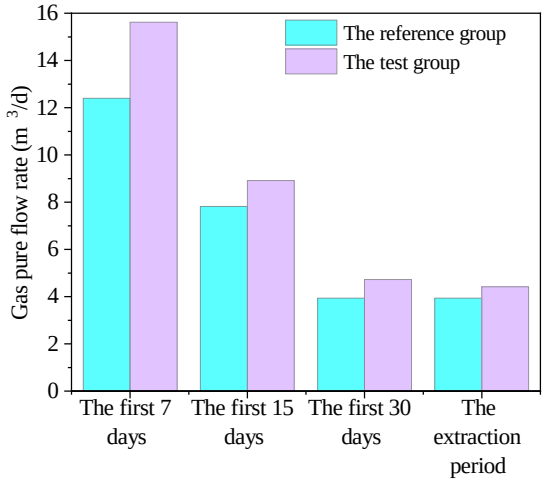

Fig. 15. Comparison of pure flow rates of gas extraction.

by grouting based on the identification results, significantly improved borehole sealing quality and enhanced the gas extraction concentration of the coal seam.

7. Conclusion

To achieve high-precision identification of coal fractures, a novel MCSN fracture segmentation model was designed, and its segmentation performance was evaluated qualitatively and quantitatively. The model was further applied to identify fractures in the surrounding rock boreholes. The main conclusions are as follows:

1) The VGG16 was adopted as the backbone feature extraction network, the MCSN model was built by integrating a DCAC module and a multiscale feature extraction network. This proposed MCSN model incorporated a multiscale fusion function, which enabled it to capture local and global contextual information for accurate semantic segmentation predictions. Additionally, it effectively mitigated the noise interference typically associated with traditional image segmentation methods during fracture extraction.
2) The MCSN model successfully addressed interference in fracture recognition caused by gangue, artifacts, and complex pore structures to achieve high-precision coal fracture identification. In terms of coal fracture identification, the evaluation MIoU, Re, Pre, and MPA metrics for the model were 87.69%, 92.75%, 91.55%, and 92.75%, respectively. For rock fracture identification, the corresponding evaluation metrics reached 86%, 91.41%, 91.6%, and 91.41%, respectively.
3) The MCSN model was utilized to conduct an on-site characterization of fractures in gas extraction boreholes. The distribution ranges and development degrees for surrounding rock cracks were accurately determined, which provided a reliable reference for the reasonable sealing length of gas extraction boreholes. Optimizing the sealing length using the MCSN model increased the average gas concentration in the extraction boreholes by 15% during the extraction period, while the average pure gas flow rate improved by 12%, thereby enhancing the overall

effectiveness of coal gas extraction.

Acknowledgements

This work was funded by the National Natural Science Foundation of China (Nos. 52174174 and 52274193), the Innovative Scientific Research Team of Henan Polytechnic University in China (No. T2022-1), and the Special Program for Basic Research of Key Scientific Research Projects of Colleges and Universities in Henan Province of China (No. 21zx004).

Conflict of interest

The authors declare no competing interest.

References

Alshammari, A. Construction of VGG16 convolution neural network (VGG16_CNN) classifier with NestNet-based segmentation paradigm for brain metastasis classification. Sensors, 2022, 22(20): 8076.

Bai, C., Liu, X. The edge detection technology of CT image for study the growth of rock crack. Paper Presented at 2009 ISECS International Colloquium on Computing, Communication, Control, and Management, Sanya, China, 8-9 August, 2009.

Chaurasia, A., Culurciello, E. Linknet: exploiting encoder representations for efficient semantic segmentation. Paper Presented at 2017 IEEE Visual Communications and Image Processing (VCIP), St. Petersburg, FL, USA, 10-13 December, 2017.

Chen, L., Papandreou, G., Kokkinos, L., et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.

De Castro, A. R., Agnaou, M., Ahmadi-Sénichault, A., et al. Numerical porosimetry: Evaluation and comparison of yield stress fluids method, mercury intrusion porosimetry and pore network modelling approaches. Computers & Chemical Engineering, 2020, 133: 106662.

Dai, Y., Li, C., Su, X., et al. Multi-scale depthwise separable convolution for semantic segmentation in street-road scenes. Remote Sensing, 2023, 15(10): 2649.

Fang, H., Sang, S., Liu, S., et al. Methodology of three-dimensional visualization and quantitative characterization of nanopores in coal by using FIB-SEM and its application with anthracite in Qinshui basin. Journal of Petroleum Science and Engineering, 2019, 182: 106285.

Hao, D., Tu, S., Zhang, C., et al. Quantitative characterization and three-dimensional reconstruction of bituminous coal fracture development under rock mechanics testing. Fuel, 2020, 267: 117280.

He, K., Zhang, X., Ren, S., et al. Deep Residual Learning for Image Recognition. Paper Presented at 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27-30 June, 2016.

He, M., Wang, Q. Rock dynamics in deep mining. International Journal of Mining Science and Technology, 2023, 33(9): 1065-1082.

Jain, P., Dubey, A., Saba, L., et al. Attention-based UNet deep learning model for plaque segmentation in carotid ultrasound for stroke risk stratification: An artificial intelligence paradigm. Journal of Cardiovascular Development and Disease, 2022, 9(10): 326.

Leng, B., Yang, H., Hou, G., et al. Rock mass trace line identification incorporated with grouping algorithm at tunnel faces. Tunnelling and Underground Space Technology, 2021, 110: 103810.

Liu, S., Li, X., Wang, D., et al. Experimental study on temperature response of different ranks of coal to liquid nitrogen soaking. Natural Resources Research, 2021, 32: 1467-1480.

Liu, Z., Liu, D., Cai, Y., et al. Application of nuclear magnetic resonance (NMR) in coalbed methane and shale reservoirs: A review. International Journal of Coal Geology, 2020, 218: 103261.

Li, Y., Cui, H., Zhang, P., et al. Three-dimensional visualization and quantitative characterization of coal fracture dynamic evolution under uniaxial and triaxial compression based on CT scanning. Fuel, 2020, 262: 116568.

Li, S., Tang, D., Xu, H., et al. Advanced characterization of physical properties of coals with different coal structures by nuclear magnetic resonance and X-ray computed tomography. Computers & Geosciences, 2012, 48: 220-227.

Liu, F., Wang, L. UNet-based model for crack detection integrating visual explanations. Construction and Building Materials, 2022, 322: 126265.

Liu, J., Jiang, Y., Zhao, Y., et al. Crack edge detection of coal CT images based on LS-SVM. Paper Presented at 2009 International Conference on Machine Learning and Cybernetics, Hebei, China, 12-15 July, 2009.

Liu, X., Cai, T. Research on the forecast model of coal mine roof based on image processing. Paper Presented at 2008 International Conference on Computer Science and Software Engineering, Wuhan, China, 12-14 December, 2008.

Long, J., Shelhamer, E., Darrell, T., et al. Fully convolutional networks for semantic segmentation. Paper Presented at 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 07-12 June, 2015.

Lou, Q., Jia, B., Wan, X., et al. Discrete characteristics of instantaneous frequency of EMR induced by coal and rock fracture. Measurement Science and Technology, 2023, 34(12): 125020.

Lu, F., Fu, C., Zhang, G., et al. Convolution neural network based on fusion parallel multiscale features for segmenting fractures in coal-rock images. Journal of Electronic Imaging, 2020, 29(2): 023008.

Mohebbi, M., Bafghi, A. F., Marji, M. F., et al. Rock mass structural data analysis using image processing tech-

niques (Case study: Choghart iron ore mine northern slopes). Journal of Mining and Environment, 2017, 8(1): 61-74.

Minhyeok, L. Mathematical analysis and performance evaluation of the GELU activation function in deep learning. Journal of Mathematics, 2023, 2023: 4229924.

Roslin, A., Pokrajac, D., Zhou, Y. Cleat structure analysis and permeability simulation of coal samples based on micro-computed tomography (micro-CT) and scan electron microscopy (SEM) technology. Fuel, 2019, 254: 115579.

Roy, A. G., Navab, N., Wachinger, C. Recalibrating fully convolutional networks with spatial and channel "squeeze and excitation" blocks. IEEE Transactions on Medical Imaging, 2019, 38(2): 540-549.

Ronneberger, O., Fischer, P., Brox, T. U-net: Convolutional networks for biomedical image segmentation. Paper Presented at International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5-9 October, 2015.

Shi, Q., Qin, Y., Zhou, B., et al. Porosity changes in bituminous and anthracite coal with ultrasonic treatment. Fuel, 2019, 255: 115739.

Wang, C., Liu, Y., Song, D., et al. Evaluation of bedding effect on the bursting liability of coal and coal-rock combination under different bedding dip angles. Advances in Geo-Energy Research, 2024, 11(1): 29-40.

Wang, D., Zeng, F., Wei, J., et al. Quantitative analysis of fracture dynamic evolution in coal subjected to uniaxial and triaxial compression loads based on industrial CT and fractal theory. Journal of Petroleum Science and Engineering, 2021, 196: 108051.

Wang, G., Yang, X., Zhang, X., et al. Numerical simulation of gas flow in pores and fissures of coal based on segmentation of DTM threshold. Chinese Journal of Rock Mechanics and Engineering, 2016, 35(1): 119-129. (in Chinese)

Wang, L., Wang, C., Sun, Z., et al. An improved dice loss for pneumothorax segmentation by mining the information of negative areas. IEEE Access, 2020, 8: 167939-167949.

Wang, X., Pan, J., Wang, K., et al. Fracture variation in high-rank coal induced by hydraulic fracturing using X-ray computer tomography and digital volume correlation. International Journal of Coal Geology, 2022, 252: 103942.

Wang, Y., Wang, G., Chen, C., et al. Multi-scale dilated convolution of convolutional neural network for image denoising. Multimedia Tools and Applications, 2019, 78: 19945-19960.

Wu, Y., Wang, D., Wei, J., et al. Damage constitutive model of gas-bearing coal using industrial CT scanning technology. Journal of Natural Gas Science and Engineering, 2022, 101: 104543.

Xie, B., Ai, D., Yang, Y. An automatic pixel-level crack identification method for coals experiencing SHPB impact tests. Journal of Geophysics and Engineering, 2019, 16(2): 297-308.

Yan, Z., Wang, H., Geng, Y., et al. Coal rock interface image recognition method based on improved DeeplabV3+and transfer learning. Coal Science and Technology, 2023, 51(S1): 429-439. (in Chinese)

Zhang, Z., Wu, C., Coleman, S., et al. DENSE-INception U-Net for medical image segmentation. Computer Methods and Programs in Biomedicine, 2020, 192: 105395.

Zheng, J., Qi, Z., Liu, J., et al. Segmentation of micro-cracks in fractured coal based on convolutional neural network. Journal of Mining Science and Technology, 2022, 7(6): 680-688. (in Chinese)

Zhao, H., Shi, J., Qi, X., et al. Pyramid sceneparsing network. Paper Presented at 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21-26 July, 2017.

Zhao, J., Qin, Y., Shen, J., et al. Effects of Pore Structures of Different Maceral Compositions on Methane Adsorption and Diffusion in Anthracite. Applied Sciences, 2019, 9(23): 5130.