Adaptive Multi-Scale Feature Fusion Based Residual U-net for Fracture Segmentation in Coal Rock Images

Fengli Lu
  China university of Ming and Technology(Beijing)

Chengcai Fu
  China Universtity of Ming and Technology(BeiJing)

Guoying Zhang (✉ zhangguoying1101@163.com)
  https://orcid.org/0000-0002-6674-1632

Jie Shi
  China University of Ming and Technology(BeiJing)

Research

Adaptive Multi-Scale Feature Fusion Based Residual U-net for Fractures Segmentation in Coal Rock Images

Fengli Lu, Chengcai Fu, Guoying Zhang*, and Jie Shi

China University of Mining and Technology (Beijing), Ding No.11 Xueyuan Road, Haidian District, Beijing, China, 100083

*Correspondence: zhangguoying1101@163.com; Tel.: +86-137-1755-9517

Abstract:

Accurate segmentation of fractures in coal rock CT images is important for safe production and the development of coalbed methane. However, the coal rock fractures formed through natural geological evolution, which are complex, low contrast and different scales. Furthermore, there is no published data set of coal rock. In this paper, we proposed adaptive multi-scale feature fusion based residual U-uet (AMSFFR-U-uet) for fracture segmentation in coal rock CT images. The dilated residual blocks (DResBlock) with dilated ratio (1,2,3) are embedded into encoding branch of the U-uet structure, which can improve the ability of extract feature of network and capture different scales fractures. Furthermore, feature maps of different sizes in the encoding branch are concatenated by adaptive multi-scale feature fusion (AMSFF) module. And AMSFF can not only capture different scales fractures but also improve the restoration of spatial information. To alleviate the lack of coal rock fractures training data, we applied a set of comprehensive data augmentation operations to increase the diversity of training samples. Our network, U-net and Res-U-net are tested on our test set of coal rock CT images with five different region coal rock samples. The experimental results show that our proposed approach improve the average Dice coefficient by 2.9%, the average precision by 7.2% and the average Recall by 9.1%, respectively. Therefore, AMSFFR-U-net can achieve better segmentation results of coal rock fractures, and has stronger generalization ability and robustness.

Keywords: adaptive multi-scale feature fusion, U-net, fracture segmentation in coal rock CT image, dilation convolutions, residual U-net.

1. Introduction

The coal rock fracture network are the adsorption and seepage channels for gas and groundwater. The initiation and expansion of fractures are likely to cause the leakage of gas and groundwater when the underground mining activities cause the stress response of coal rock. Therefore, the accurately expressing the characteristics of fracture network are great significance for the safe production in coal mines, the coal seam fracturing and gas extraction. Industrial CT (Computed Tomography) technology is a nondestructive scanning method to detect the internal structure of coal, thereby we currently research coal rock fractures by coal rock CT images. However, coal rock CT images have the characteristics of high-density noise, weak coal rock fractures and different size fractures. The large-scale fractures through the entire section of coal rock CT image, while small-scale fractures have only 1-5 pixels. Further, coal rock CT image samples are lacking. Therefore, it is a challenging task for fracture segmentation in coal rock CT images.

The main digital image processing techniques based on traditional segmentation methods, such as threshold-based segmentation [1-2], edge detection [3-4], region growing [5] and segmentation based on watershed [6] etc. These traditional methods need to design the hand-craft feature representation for a new target segmentation [7]. However, the shape and location of coal rock fractures are unfixed, which make it difficult to define the coal rock fractures shape manually. Furthermore, the grayscale values of weak and tiny fractures are close to the grayscale values of the surrounding coal matrix. At present, the existing segmentation methods based on grayscale and texture features are very difficult to extract

weak fractures. To summarize, the traditional segmentation methods are not robust enough for the segmentation of weak and tiny fracture. It [8] proposed a coal rock fracture segmentation method based on the consistency of contour evolution and gradient direction consistency. However, the segmentation results of are still poor at the weak boundaries where the gradient change is not obvious. Therefore, the traditional segmentation method is not suitable for the fracture of coal rock.

In recent years, CNN (Convolutional Neural Network) are widely applied in the image segmentation because of its powerful feature extraction ability. U-net [9] was proposed to segment cell boundaries on limited data set, which achieved great success. It is a typical encoder-decoder structure, and are widely used in the field of image segmentation, especially in the medical field [10-12]. However, the simple feature fusion method will lead to weak feature extraction ability of U-net, so original U-net is easy to lose weak and tiny fractures. The issue is usually solved by increasing the size of the receptive field to capture context information and improving feature extraction ability of network. On one hand, PSPNet [13], Atrous Spatial Pyramid Pooling [14], Spatial pyramid pooling [15] and GRANet [16], these approaches leveraged spatial pyramid and dilated convolution to capture information. Although having achieved higher segmentation results, these methods cause too much computation burden. On other hand, RefineNet [17], G-FRNet [18] and FDNET [19], these methods reused feature maps of earlier layers in the encoder module to capture multi-scale feature and improve the restoration of spatial information, which obtained accurate segmentation results. However, these methods are not only complex in design but also require

3

large training data sets. And it may cause grid artifacts [21]. However, the train data of coal rock CT images is lacking.

However, no literature method so far coped with fractures of coal rock CT images based on multi-scale fusion method. Here, according to the characteristics of coal rock fractures, we proposed a novel network based on the strengths of the above methods and U-net, namely adaptive feature fusion based residual U-net for fractures segmentation in coal rock CT images (AMSFFR-U-net). The main contributions of our work are as followings:

1) We developed a novel network to segment coal rock fractures based on U-net structure.

2) We proposed an adaptive feature fusion module (AMSFF). The different scales of feature maps in encoding branch are fused by AMSFF to capture context information and reduce the loss of spatial information, which can improve the ability of multi-scale feature extraction and spatial information restoration.

3) To prevent over-fitting, we use a set of comprehensive data augmentation methods to increase the diversity of training samples.

The remaining content of the paper is organized as：In Section 2, we described the methodology of our proposed approach in detail. In Section 3, we introduced the data sets creation, data pre-processing and the data augmentation. In section 4, we reported the experimental results and comparisons. And we conclude our paper and discuss the future directions in Section 5.

2. Methods

As a particularly successful semantic segmentation structure, U-net [9] is widely used in

the field of image segmentation. It successfully achieved accurate segmentation results on 30 cell training samples. Moreover, there are some similarities between cell boundaries and coal rock fractures. And there is a lack of data on coal rock fractures. Therefore, this paper attempts to segment coal rock fractures based on U-net structure. However, coal rock fractures are complex compared to cell boundaries, such as weak and tiny fractures in coal rock. However, the original U-net will omit the weak and tiny coal rock fractures due to its simple feature fusion method. Moreover, the original U-net use VGG network, while feature extraction capability of VGG is weaker than ResNet. The residual learning framework can prevent the gradient disappearing and improve the performance of training when network is deepen [20]. Inspired by this, a novel network of fractures segmentation in coal rock CT images based on U-Net was proposed, which to strengthen the feature extraction ability of weak and tiny fractures. And the precise structure is shown in Fig.1.
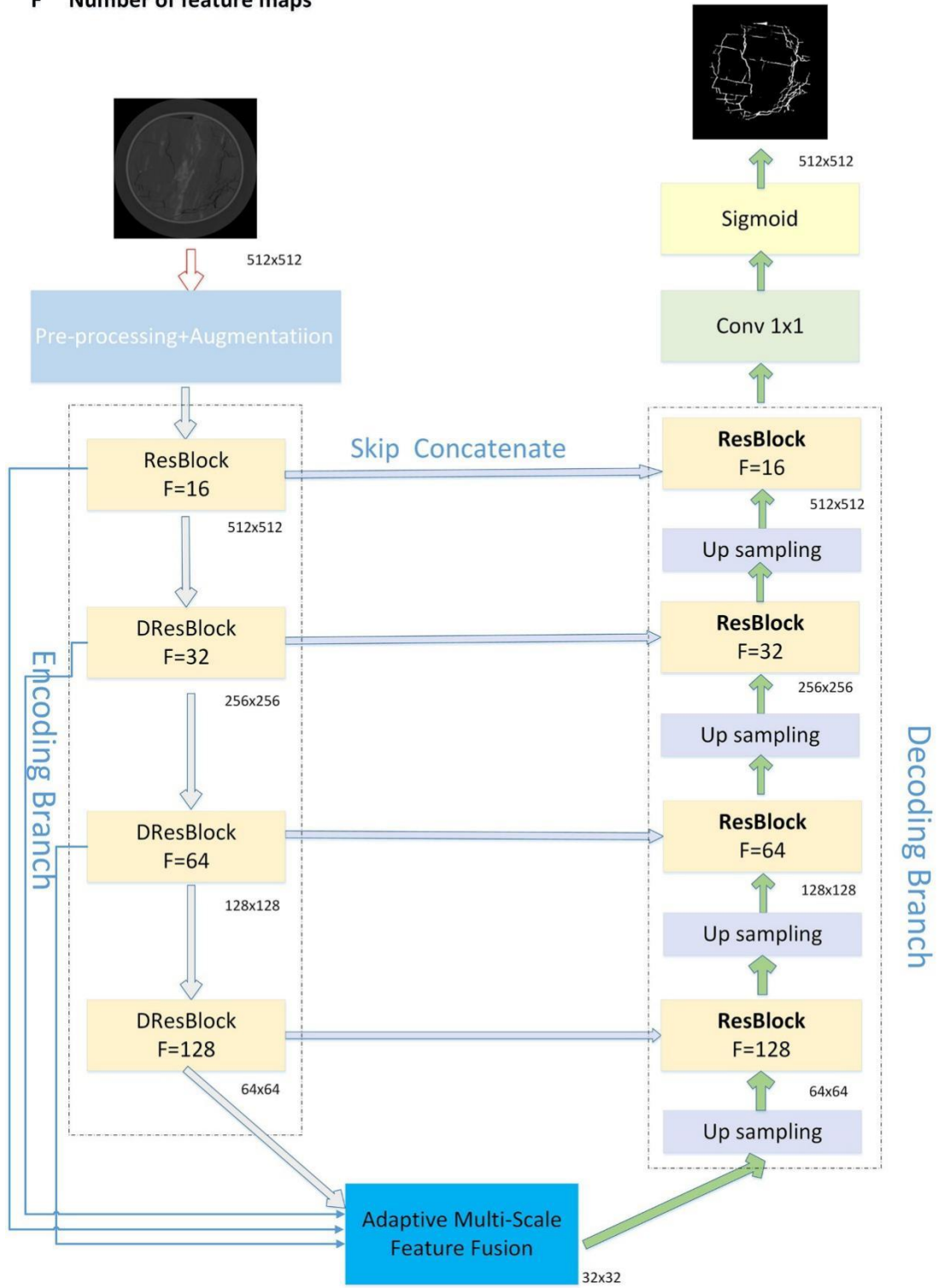
Fig.1. Architecture of our proposed, which comprises encoding branch, AMSFF and decoding branch. 1) Encoding branch consisted of one Residual Block (ResBlock) with a dilation rate of 1 and three Dilated Residual Block (DResBlock) with dilation ratios (1, 2, 3). 2) AMSFF adaptive fuse the different scales feature maps of all cell blocks from encoding branch. 3) Decoding branch consisted of four ResBlock and 4 up-sampling layers. 4)

The number below each rectangle represents the size of the feature map, such as 256×256.

The AMSFFR-U-Net has three main parts: encoding branch, adaptive feature fusion module (AMSFF) and decoding branch. Fig.1 shows the overall structure. In the architecture, each convolution layers except the last output layer is composed of batch normalized (BN), exponential linear unit function (ELU) and 3×3 convolution.

In the coal rock fracture images, the grayscale values of the weak and tiny fracture boundaries are close to the grayscale values of the coal matrix. It is easy to result in mis-segmentation when global context information is not fully understood. To reduce the mis-segmentation, it can make the network better understand the global context information by using dilated convolution to increase the receptive field usually. However, it is may cause grid artifacts by using the convolution of the same dilation rate many times on the feature images of the same resolution [21]. Hybird dilation convolution (HDC) [21] was proposed to solve the issue. Inspired by this work, this paper used three dilated convolution layers to construct DResBlock without increasing numbers of parameter and the dilated ratio D of three Conv 3×3 in the DResBlock is defined as 1,2,3, respectively. It can better capture global contextual fracture feature to reduce the loss of information of weak and tiny coal rock fractures. The structure of dilated residual block is depicted in Fig.2.
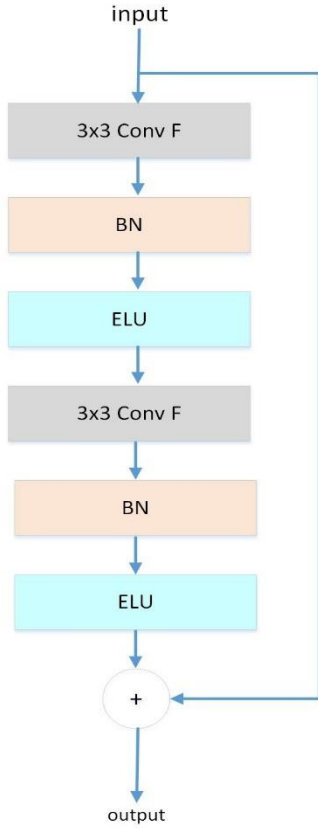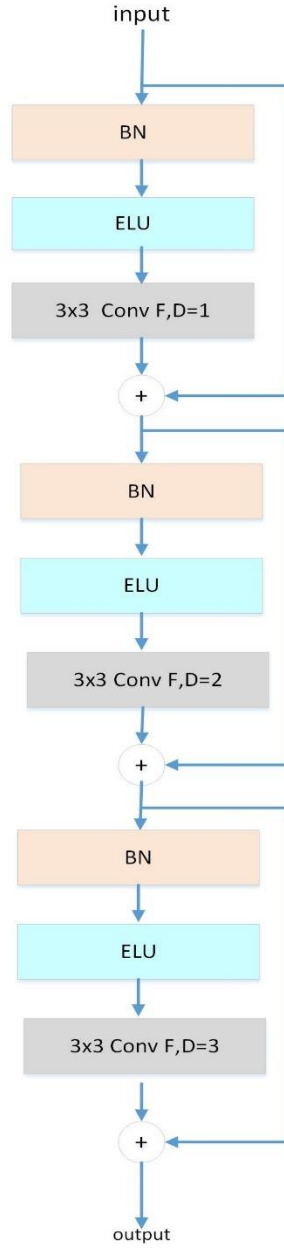
Residual Block(ResBlock)

Dilated Residual Block(DResBlock)

Fig.2 Illustration of the structure of dilated residual block. (a)The residual block, D=1. (b) The structure of dilated residual block, and the dilated ratio D of different convolutional layers is 1,2,3, respectively.

Furthermore, AMSFF module fuse the different scales feature maps from encoding branch, which is a largely supplement to DResBlock to capture multi-scale coal rock fractures.

Encoding branch: Encoding branch mainly focuses on learning fractures feature and

reducing the resolution of coal rock CT images gradually. And it consisted of one ResBlock

with a dilation rate of 1 and three DResBlock with dilation rates of 1, 2 and 3. However, Coal

rock images have not only sparse targets but also fractures of different scales, such as small

fractures with only 1-5 pixels. And the more obvious the fractures are, the closer the grayscale

values are to 0, while the coal matrix is similar to weak tiny fractures and have larger gray

values. Pooling operation is easy to cause the loss spatial information and feature information

of the weak and tiny fractures. Therefore, each DResBlock use a stride of 2 in the first layer to

reduce the size of feature instead of pooling operation.

Adaptive Multi-Scale Feature Fusion: However, in the encoder-decoder structure, much

spatial and feature information is lost because continuously reduce the sizes of the feature

maps in encoding branch, especially weak and tiny fractures. To address the issue, this paper

proposed adaptive feature fusion module to fuse the different scales feature maps of all from

encoding branch. AMSFF and DResBlock are complementary to each other in functionality.

On one hand, AMSFF can extract multi-scale fractures by fusion different scales feature maps.

On the other hand, AMSFF can significantly improve the ability of the restoration of spatial

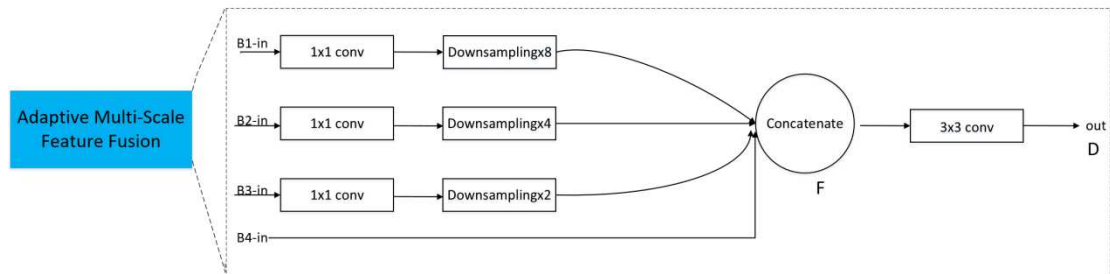information for decoding branch. AMSFF as shown in Fig.3.

Fig.3 Structure of AMSFF: 1) The input feature maps (B_1, B_2, B_3) undergo a 1×1 convolution layer to reduce

parameters calculation, firstly; then they reduce the dimension by convolutions with strides 2 to obtain the same

size. 2) The input feature map B_4 directly connected because of its dimensions are the same as the feature map F.

3) The merged feature map F undergo a 3×3 convolution as the input of decoding branch.

Here, AMSFF as transformation from the encoding branch to the decoding branch. We expressed the output information of all layers of the encoding branch as B_1, B_2, B_3, B_4, respectively. The feature information after AMSFF fusion produces rich Multi-scale semantic information. However, the dimensions and sizes of the feature information (B_1, B_2, B_3, B_4) are different. Therefore we must ensure that B_1, B_2, B_3, B_4 to be fused has the same spatial resolution and number of channels. Obviously, B_4 can be concatenated directly as its dimensions are the same as the feature map F to output. In order to reduce parameters calculation to increase processing speed, the input feature maps (B_1, B_2, B_3) first undergo a 1×1 convolution layer, then the dimension was reduced by convolutions with stride 2 to obtain the same size. Here, 1×1 convolution and 3×3 convolution both composed of BN, ELU and convolution operation. All the resulting feature maps $B_n^{'}$ (n=1, 2, 3, 4) and fusion results F are described as follows:

$$B_n^{'} = T_n(B_n), n = 1,2,3,4 \tag{1}$$

$$F = [B_1^{'}, B_2^{'}, B_3^{'}, B_4^{'}] \tag{2}$$

Here, T_n represents the different number of convolutions layers are performed by different feature map B_n. Finally, the merged feature map F undergo a $3\times3\times128$ convolution with stride 2 to reduce the size of feature maps to 32×32, which is as the input of up-sampling in the decoding branch.

Decoding branch: Decoding branch comprises four up-sampling and four ResBlock with the standard convolution. The "skip-concatenation" concatenated with the feature maps from the

corresponding encoding branch. Decoding branch gradually up-sample the feature map F output by the AMSFF and the feature maps from the corresponding encoding branch with up-sampling of with 2×2, and final restores to the same resolution as its input original image. Finally, a 1×1 convolution and activation function to obtain an accurate pixel-level binary coal rock fracture segmentation results.

Algorithm: To complement the previous description, we describe the whole algorithm flow chart is presented in is Algorithm 1.

Input: image sets I_i

Output: the Predicted results of coal rock fractures

1. For $i = 1$: number(I_i)

2.　CLAHE to pro-processing, generated datasets I_p.

3.　a set of comprehensive data augmentation methods to data n, generated datasets I_j.

4. end for

5. For j=1 : number(I_j)

6.　d1=ResBlock(I_1,16)

7.　d2=DResBlock(d1,32, stride=2)

8.　d3=DResBlock(d2,64, stride=2)

9.　d4=DResBlock(d4,128, stride=2)

10.　AMSFF=AMSFF. concat (d1,d2,d3,d4)

11.　U1=upsample(AMSFF).concat(U1,d4)

12.　C1=ResBlock(U1,128)

13.　U2=upsample(U1).concat(U2,d3)

14.　C2=ResBlock(U2,64)

15.　U3=upsample(U2).concat(U3,d2)

16.　C3=ResBlock(U3,32)

17.　U4=upsample(U3).concat(U4,d1)

18.　C4=ResBlock(U4,16)

19. End for

Algorithm 1: The coal rock fractures segmentation based on AMSFFR-U-Net.

Loss Function: The coal rock CT images belong to the imbalance of samples i.e large background proportion and small target proportion. Dice coefficient loss [22] is effective when data is unbalanced. Here, we use the Dice coefficient loss as the segmentation loss function of. Dice coefficient represents the similarity between the ground value and the predicted value. And the higher the similarity, the more accurate the model segmentation. It is defined as follows:

$$L_{DSC} = 1 - \frac{2\sum_{i=1}^{N} X_i * Y_i}{\sum_{i=1}^{N} X_i + \sum_{i=1}^{N} Y_i} \tag{3}$$

Where, X_i and Y_i represent the ground truth and predicted value respectively. N denotes pixels of coal rock CT images training samples. Furthermore, due to the limited samples of coal rock fractures, we construct our loss function with the $L2$ regularization to avoid over-fitting, it is as shown in equation (6):

$$L_{loss} = L_{DSC} + L2 \tag{4}$$

3. DATA

3.1 Datasets Creation

The experimental coal rock CT images are collected from State Key Laboratory Coal Resources and Safe Mining, China University of Mining & Technology (Beijing). And annotating images by staff of State Key Laboratory Coal Resources and Safe Mining, China University of Mining & Technology (Beijing) who have been engaged in coal rock fracture labeling for many years, which ensures the authenticity and reliability of the annotation images of coal rock fractures. In this experiment, these coal rock samples from five different regions. 1000 frames coal rock CT images were obtained from each sample. Due to the

fractures of adjacent frames are very similar, one frame is extracted every 40 frames on each

sample for annotating to prevent over-fitting. A total of 120 frames rock CT images obtained

from 5 different coal rock samples. 100 frames were selected randomly as training set. And

the remaining 20 frames were used as test set.

3.2 Data Pre-processing

As coal rock is a natural body, coal rock CT images have the characteristics of low

contrast and uneven gray. And it is difficult to identify fractures without pre-processing

operation. Here, we use globally CLAHE (contrast limited adaptive histogram equalization)

to improve the contrast of the original images, maintain global information and enhance the

contrast of the edges of coal rock fractures. Fig.4 shows an example of data pre-processing.
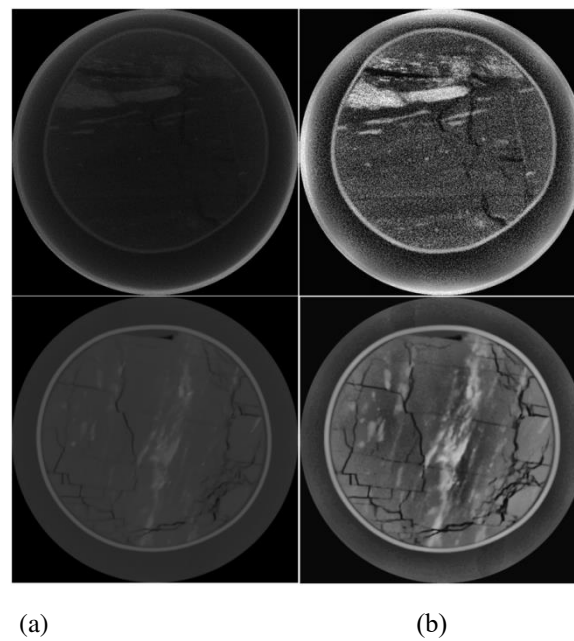

(a)   (b)

Fig.4 The result image after: (a) original image. (b) after globally CLAHE operation.

It can be seen from Fig.4, the contrast between the fractures and the coal matrix in the coal

rock images is significantly improved after globally CLAHE operation.

3.3 Data Augmentation

At present, there is no published data set of coal rock CT images. However, manual annotation of coal rock fracture images requires a lot of time and cost. In order to overcome the limitation of coal rock CT images, we perform data augmentation operations on the original coal rock CT images and the corresponding annotation images to increase the diversity of training samples, to avoid over-fitting and to improve the generalization of the model. In the paper, we applied a set of comprehensive data augmentation methods summarized in Table1.

Table 1 Summary of implementation data augmentation methods.

B represents the factor of brightness; α, σ control the degree of the elastic distortion; β represents the weight of the image superposition.

Methods	Value Range
Random rotation	±30º
Horizontal-flip	50%
Width-shift-range	20%
Height-shift-range	20%
Brightness	B $\in$ (0.5 , 1)
Image superposition	β=0.5
Elastic deformation	α= 1000, σ=40

Basic data augmentation methods: random rotation, horizontal-flip, width-shift-range and contrast factor. Owing to the characteristic of sparse target of coal rock fracture images, these basic data augmentation methods are not sufficient to increase the diversity of fracture samples. We further adopted the methods of image superposition and elastic transformation, which can generate large and reasonable training data. The mathematical representation of image superposition is the two images are superposed together according to a certain linear relationship while keeping the image size unchanged. The mathematical principle is as

follows：

$$g(x) = \beta f_0(x) + (1-\beta) f_1(x) \qquad (5)$$

Here, β is the coefficient of linear superposition of two images. Then, the elastic deformation operation is performed on the superimposed image [23]. σ is a Gaussian of standard deviation, and it convolved with the displacement field. α is a scaling factor controls the intensity of the deformation. First, the displacement field A is obtained by the original displacement field convolved with σ. Then, the elastic deformation displacement field is obtained by the displacement field a multiplying by α. Finally, the elastic deformation displacement field is applied to the image after affine transformation to generate the elastic deformation images. The data augmentation process of a coal-rock image is shown in Fig.5.
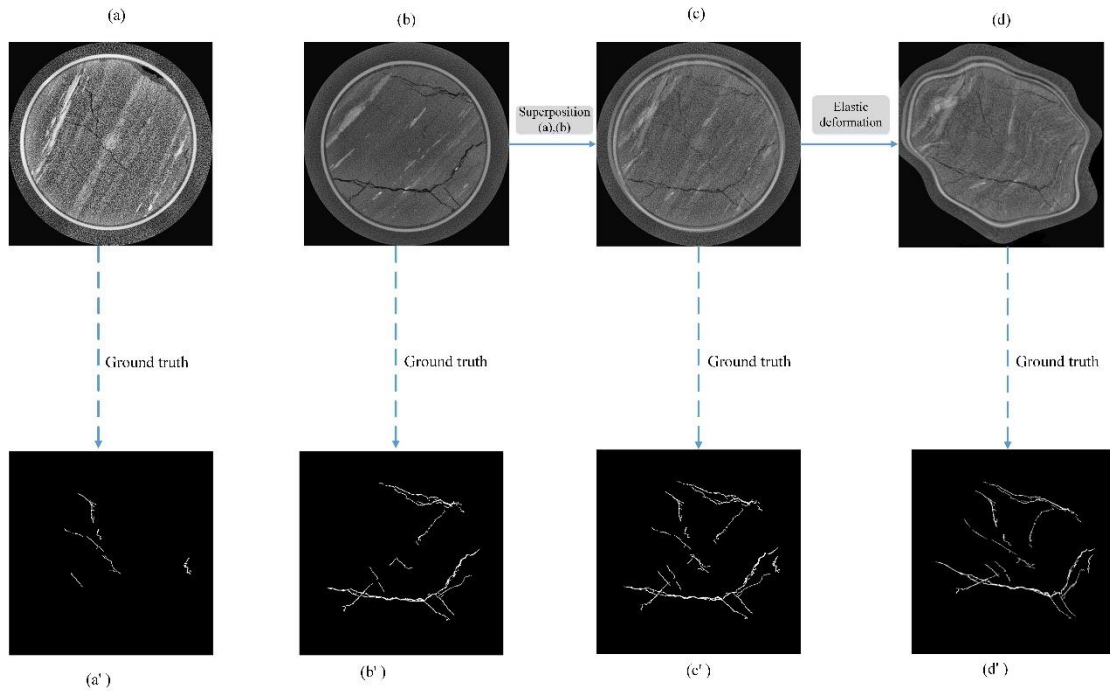

Fig.5 The process of data augmentation. (a), (b) The original coal rock CT images. (a′), (b′) The ground truth of (a) and (b). (c) The result of superimposing (a) on (b). (c′) The ground truth of (c). (d) The result of (c) after elastic deformation. (d′) The ground truth of (d).

4. Experiments

4.1 Implementation Details

Our proposed model was implemented on Keras framework and was optimized by Adam (Adaptive moment estimation) optimizer. The network was trained on 2G GPU and 128G memory of Ubuntu 16.04. Data augmentation is used during training to alleviate the problem of lacking data samples (Detailed in 3.3), and 8000 samples are generated as the training samples. We employed the "poly" learning rate policy with the initial rate is multiplied by $(1 - \frac{iter}{max\_iter})^{power}$, and the initial learning rate 0.0001, with power=0.9. And the batch size is 4. This paper use 512×512 images to train our network on the training samples. And the obtained model is used to segment coal rock fractures on our test set.

4.2 Results Analysis and Discussion

Our test set consists of coal and rock samples from five different regions (Detailed in 3.1). To verify performance and generalization of U-net, Res-U-net and AMSFFR-U-net, this paper tested three networks on our test set. And the fracture segmentation results are shown in Fig.6.
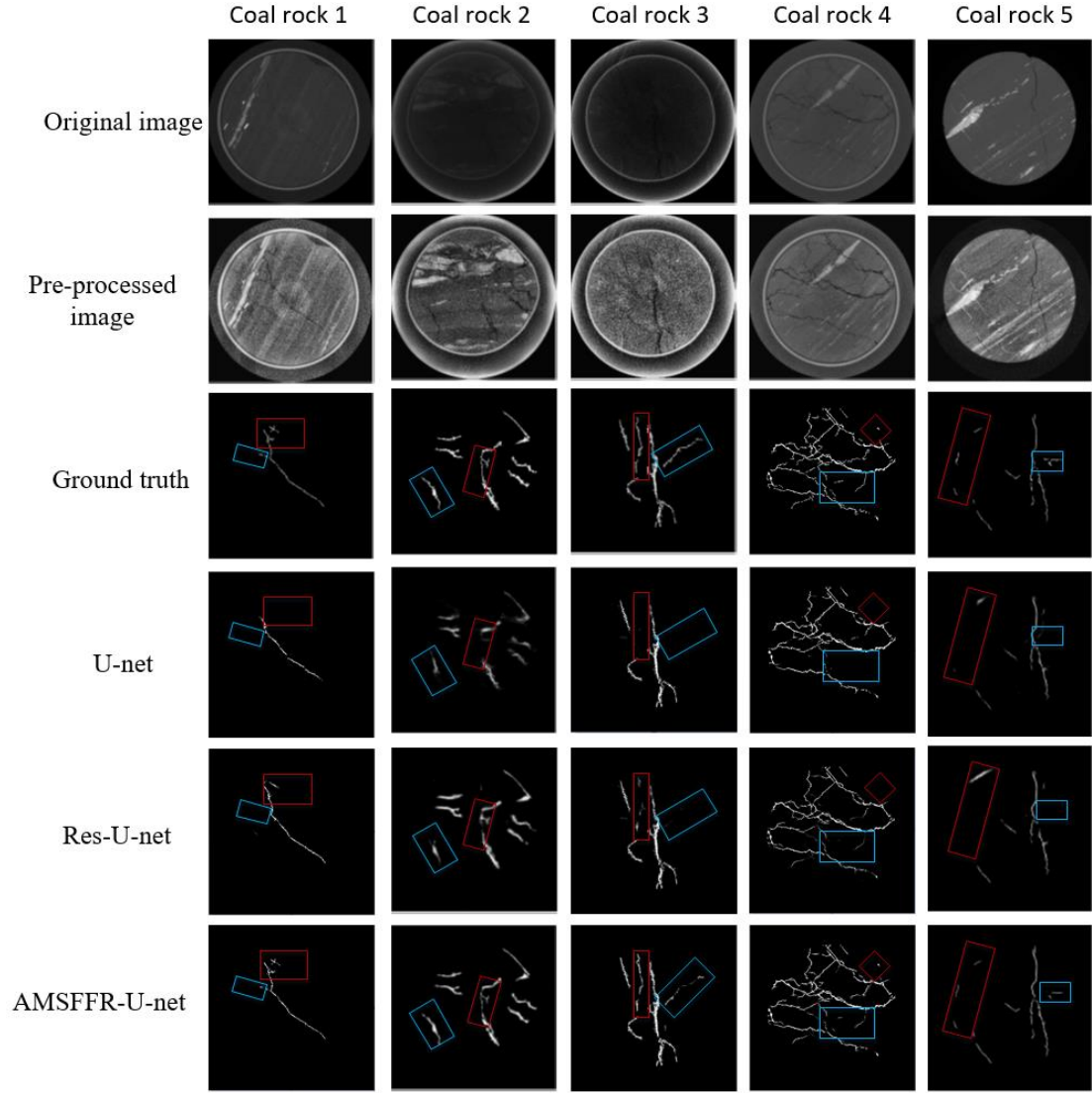
Fig.6 The fracture segmentation results of three networks on five different region of coal rock samples.

Here, we use two rectangles of different colors to mark two representative fracture locations. We can see from Fig.6 the fracture segmentation results of AMSFFR-U-net are closer to ground truth compared with the other two methods. Especially, the weak and tiny fractures segmentation results of AMSFFR-U-net on the column 3 and the column 5 are better than the other two networks. However, the coal rock fractures are segmented by Res-U-net and U-net, which is missing the weak and small-scale fractures, and the boundaries of weak and tiny fractures are blurred. As can be seen from the fourth row of Fig.6, U-net can only segment some large-scale fractures with obvious boundaries while it lost a lot of weak and

tiny fractures. Although the segmentation results of Res-U-net are better than U-net, the phenomenon of missing some small-scale fractures is still serious, which can be seen from the fifth row 5. The coal rock fracture network obtained by the AMSFFR-U-net has high accuracy. However, to further improve the accuracy of the segmentation results, we will expand the training samples to enhance the feature extraction ability of the AMSFFR-U-net.

In Fig.7, we compare the P-R(Precision-Recall) curve of AMSSFFR-U-net with U-net and Res-U-net. Obviously, the area under the P-R curve of the other two methods is smaller than our proposed method.
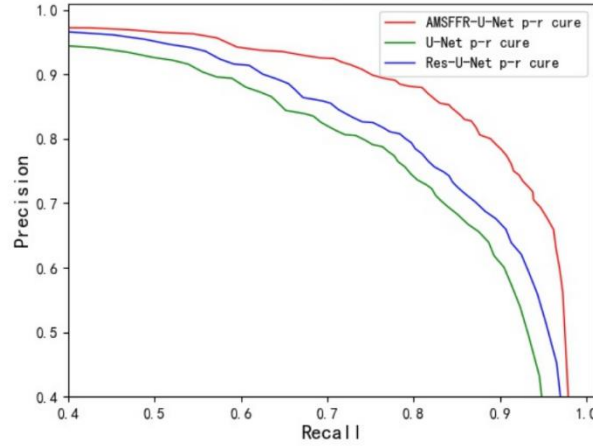

Fig.7 Precision-Recall curves for various methods on our test set of coal rock CT images.

4.2 Evaluation Metrics

To quantitatively evaluate the performance of the AMSFFR-U-Net, we use DSC (Dice similarity coefficient), Precision, Recall and Accuracy as evaluation indicators, which commonly used evaluation indicators for semantic segmentation, the detailed description are as follows:

$$DSC = \frac{2TP}{(FP + 2TP + FN)} \tag{6}$$

$$Pr\,ecision = \frac{TP}{TP + FP} \tag{7}$$

$$Re\,call = \frac{TP}{TP + FN} \tag{8}$$

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \tag{9}$$

Here, 1) TP: the numbers of correctly segmented fracture pixels; 2) TN: the numbers of correctly segmented background pixels; 3) FN: the number of pixels that the fracture was incorrectly segmented into the background; 4) FP: the number of background pixels are predicted as fracture pixels incorrectly.

The four evaluation indicators mentioned above, the higher value, the better the segmentation result. And U-net, Res-U-net and AMSFFR-U-Net were trained on our training set. Table 2 shows the average value of four evaluation indicators for segmentation results on our test set.

Table 2 The evaluation for segmentation results.

Metrics / Methods	DSC	Precision	Recall	Accuracy
U-net	0.874	0.814	0.823	0.936
Res-U-net	0.896	0.825	0.814	0.960
AMSSFFRU-net	0.925	0.897	0.905	0.981

It can be seen from Table 2 that the value of the four evaluation indicators of U-NET is the lowest, due to U-net is relatively weak in feature extraction and feature fusion. And Res-U-net uses residual block in the U-net structure to enhance the ability of feature extraction, so the value of the four evaluation indicators of U-NET are higher than U-net. Furthermore, as can be seen from Table 2, this paper proposed approach improve the average Dice coefficient by 2.9%, the average precision by 7.2% and the average Recall by 9.1%, respectively. That's enough to prove that the AMSSFFR-U-net achieves the best segmentation results than the other two methods.

5. Conclusions

In this paper, we proposed AMSFFR-U-net for automatically segment coal rock fractures,

which achieved very good performance. DResBlock of the different dilated ratio (1, 2, 3) are used in the encoding branch to extract different scale fractures. At the same time, there are residual connections in the DResBlock structure to improve the feature extract ability of the network. Moreover, AMSFF module adaptively fused the feature maps from different layers of encoding branch, which is a largely supplement to DResBlock to capture multi-scale coal rock fractures. In addition, AMSFF can improve the restoration of spatial information as well. Furthermore, we applied a set of comprehensive data augmentation operations to increase the diversity of training samples to prevent over-fitting during training. The experimental results show that AMSFFR-U-net outperforms the state-of-art algorithms, such as U-net and Res-U-net in the fractures segmentation tasks in coal rock CT images.

It is great significance for mastering the evolution rule of the fractures in the coal rock under mining stress and the coal rock seepage calculation. However, as future developments, we will consider the 3D spatial information and the connectivity of fracture among the coal rock CT images. And modeling the evolution rule of coal rock fractures feature will be investigated as well.

Abbreviations
CT: Computed Tomography; CNN: Convolutional neural network; AMSFFRU-uet: adaptive multi-scale feature fusion based residual U-uet; AMSFF: adaptive multi-scale feature fusion; BN: batch normalized; ELU: exponential linear unit function; DResBlock: dilated residual block; Adam：Adaptive moment estimation.
Availability of data and material
Our data is not shared temporarily. But, if you have any requirement, please feel free to contact us.
Competing interests
The authors declare no conflicts of interest.

Authors' contributions

References

[1] Wang G, Tse P W, Yuan M, Automatic internal crack detection from a sequence of infrared images with triple-threshold Canny edge detector [J]. Measurement Science & Technology,Volume 29, Number 2 (2018). doi:10.1088/1361-6501/aa9857.

[2] J. Than Chia Ming, N. M. Noor, O. M. Rijal, R. M. Kassim and A. Yunus, Enhanced au tomatic lung segmentation using graph cut for Interstitial Lung Disease. 2014 IEEE Conferenc e on Biomedical Engineering and Sciences (IECBES), Kuala Lumpur, pp. 17-21, 2014, doi:10. 1109/iecbes.2014.7047479.

[3] Li S, Cao Y,Cai H, Automatic pavement-crack detection and segmentation based on steerable matched filtering and an active contour model. Journal of Computing in Civil Engineering(2017), vol. 31, no. 5.

[4] R. Kaur, M. Juneja, and A. K. Mandal, A hybrid edge-based technique for segmentation of renal lesions in CT images. Multimedia Tools and Applications, vol. 78, no. 10, pp. 1291 7–12937, Aug. 2018. doi:10.1007/s11042-018-6421-7.

[5] P.R.F. Pedro et al., Novel and powerful 3d adaptive crisp active contour method applied in the segmentation of ct lung images. Medical Image Analysis (2017), 35, 503-516.

[6] A. Tareef et al., Multi-pass fast watershed for accurate segmentation of overlapping cervical cells. IEEE Trans. Med. Imaging 37, 2044–2059 (2018).doi: 10.1109/TMI.2018.2815013.

[7] X. Zhang et al., A cognitively-inspired system architecture for the Mengshi cognitive vehicle. Cognit. Comput. 12, 140–149 (2019).

[8] Zhiwei Li, Guoying Zhang, Fracture Segmentation Method Based on Contour Evolution and Gradient Direction Consistency in Sequence of Coal Rock CT Images. Mathematical Problems in Engineering(2019).

[9] O. Ronneberger et al., U-Net: convolutional networks for biomedical image segmentation.Lect. Notes Comput. Sci. 234–241 (2015). doi: 10.1007/978-3-319-24574-4_28.

[10] A. Fabijańska, Segmentation of corneal endothelium images using a U-Net-based convolut ional neural network. Artificial Intelligence in Medicine, vol. 88, pp. 1–13, Jun(2018). doi:10. 1016/j.artmed.2018.04.004.

[11] Jiang, Y., Wang, F. , Gao, J. , & Cao, S, Multi-path recurrent u-net segmentation of retinal fundus image. Applied Sciences (2020), 10(11), 3777. doi: 10.3390/app10113777.

[12] Krishna D S, Renukanand P K, Sreedhar B K, Giridhar S, Liang Z, Shamira P, Drunet: a dilated-residual U-net deep learning network to segment optic nerve head tissues in optical coherence tomography images. Biomedical Optics Express (2018), 9(7), 3244-.

[13] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, Pyramid Scene Parsing Network. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jul. 2017. doi:10.1109/cvpr.2017.660.

[14] Chen L C, Papandreou G, Kokkinos I, et al, DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE transactions on pattern analysis and machine intelligence (2018), 40(4):834. doi:10.1109/TPAMI.2017.2699184.

[15] He, Kaiming, et al, Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Transactions on Pattern Analysis & Machine Intelligence 37.9(2014):1904-16. doi:10.1007/978-3-319-10578-9_23.

[16] Z. Feng, H. Yong, and S. Xukun, GRANet: Global Refinement Atrous Convolutional Neural Network for Semantic Scene Segmentation. 2018 25th IEEE International Conference on Image Processing (ICIP), Oct. 2018. doi:10.1109/icip.2018.8451636.

[17] Lin, G.; Milan, A.; Shen, C.; and Reid, I, Refinenet: Multi-path refinement networks with identity mappings for high-resolution semantic segmentation. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jul. 2017. doi:10.1109/cvpr.2017.549.

[18] M. A. Islam, M. Rochan, N. D. B. Bruce and Y. Wang, Gated Feedback Refinement Network for Dense Image Labeling. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 4877-4885, doi: 10.1109/CVPR.2017.518..

[19] M. Zhen, J. Wang, L. Zhou, T. Fang, and L. Quan, Learning Fully Dense Neural Networks for Image Semantic Segmentation. Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, pp. 9283–9290, Jul. 2019. doi:10.1609/aaai.v33i01.33019283.

[20] K. He, X. Zhang, S. Ren and J. Sun, Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 770-778, doi:10.1109/CVPR.2016.90.

[21] Wang, Panqu, Pengfei Chen, Ye Yuan, Ding Liu, Zehua Huang, Xiaodi Hou, and Garrison Cottrell, Understanding Convolution for Semantic Segmentation. 2018 IEEE Winter Conference on Applications of Computer Vision (WACV) (March 2018). doi:10.1109/wacv.2018.00163.

[22] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. Jorge Cardoso, Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations. Lecture Notes in Computer Science, pp. 240–248, 2017. doi:10.1007/978-3-319-67558-9_28.

[23] S. Schaefer et al., Image deformation using moving least squares. ACM Trans. Graphics2 5, 533-540 (2006).
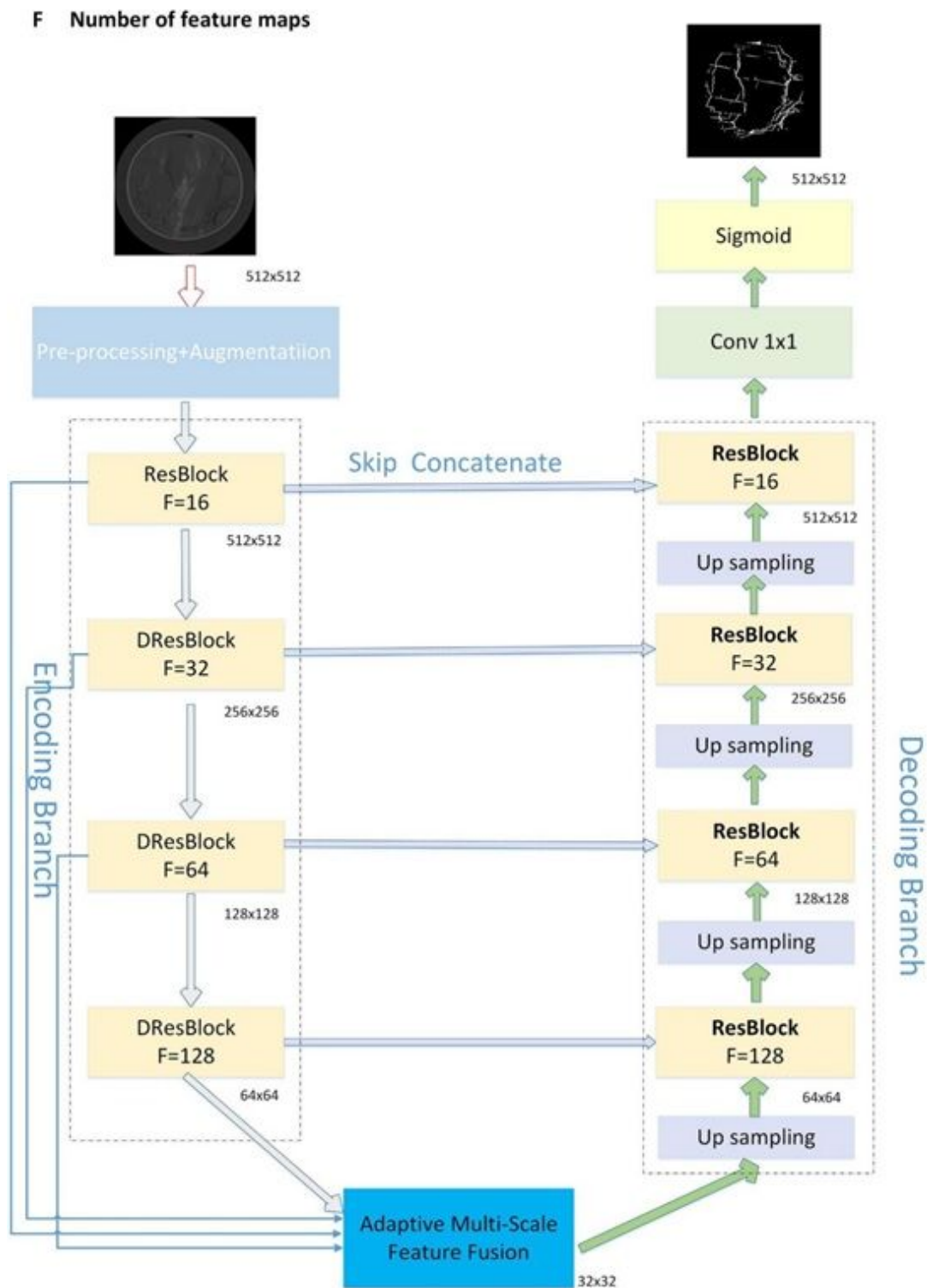
Figures

Figure 1

Architecture of our proposed, which comprises encoding branch, AMSFF and decoding branch. 1) Encoding branch consisted of one Residual Block (ResBlock) with a dilation rate of 1 and three Dilated Residual Block (DResBlock) with dilation ratios (1, 2, 3). 2) AMSFF adaptive fuse the different scales

feature maps of all cell blocks from encoding branch. 3) Decoding branch consisted of four ResBlock and 4 up-sampling layers. 4) The number below each rectangle represents the size of the feature map, such as .
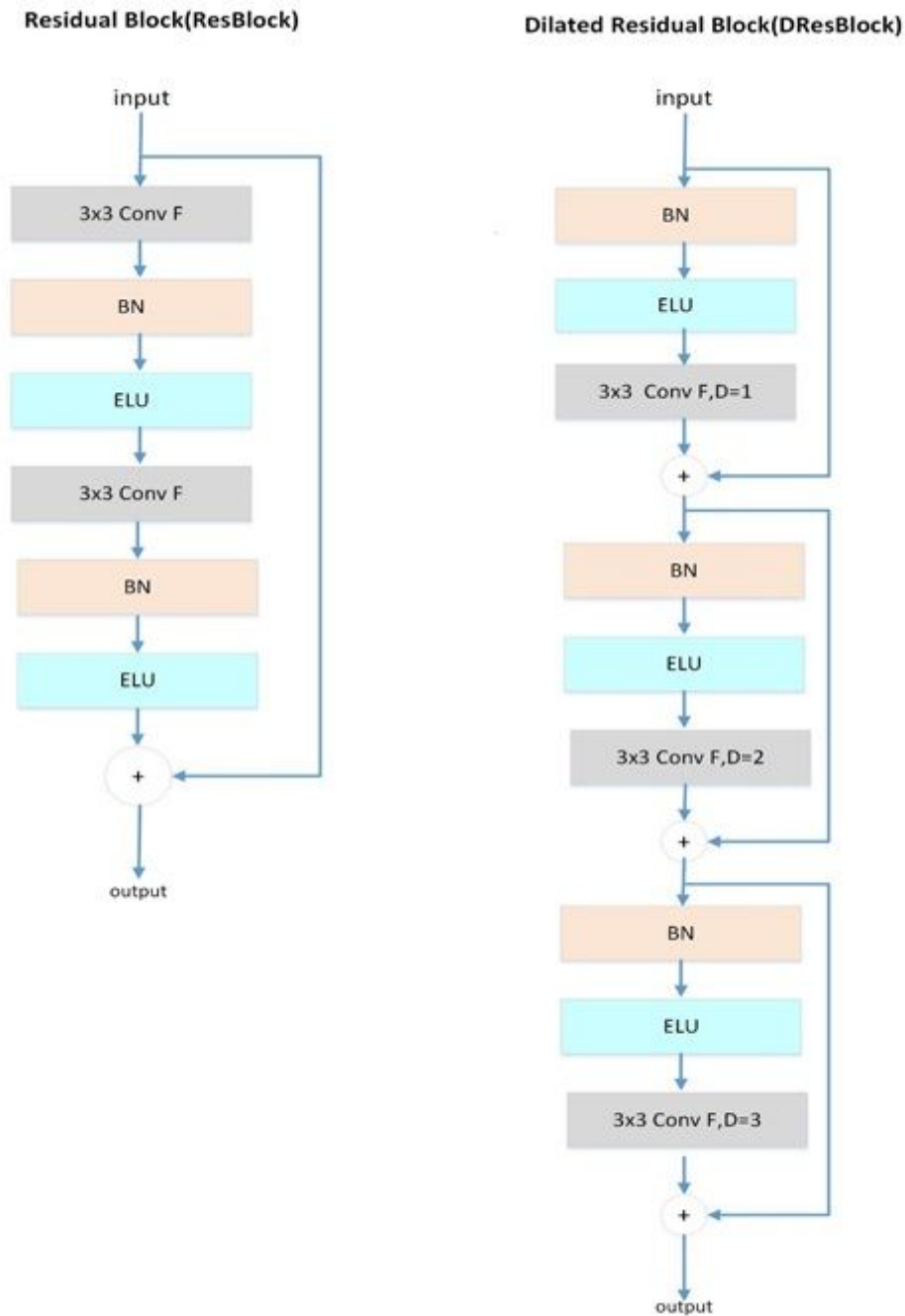
Residual Block(ResBlock)

input

3x3 Conv F

BN

ELU

3x3 Conv F

BN

ELU

+

output

Dilated Residual Block(DResBlock)

input

BN

ELU

3x3 Conv F,D=1

+

BN

ELU

3x3 Conv F,D=2

+

BN

ELU

3x3 Conv F,D=3

+

output

Figure 2

Illustration of the structure of dilated residual block. (a)The residual block, D=1. (b) The structure of dilated residual block, and the dilated ratio D of different convolutional layers is 1,2,3, respectively.
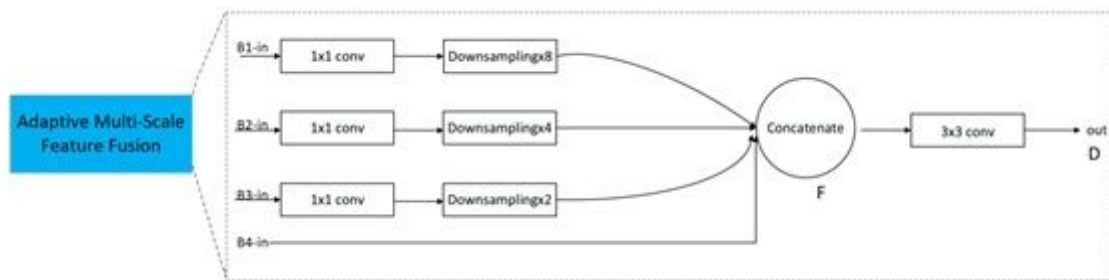
Figure 3

Structure of AMSFF: 1) The input feature maps (, ,) undergo a convolution layer to reduce parameters calculation, firstly; then they reduce the dimension by convolutions with strides 2 to obtain the same size. 2) The input feature map directly connected because of its dimensions are the same as the feature map . 3) The merged feature map undergo a convolution as the input of decoding branch.
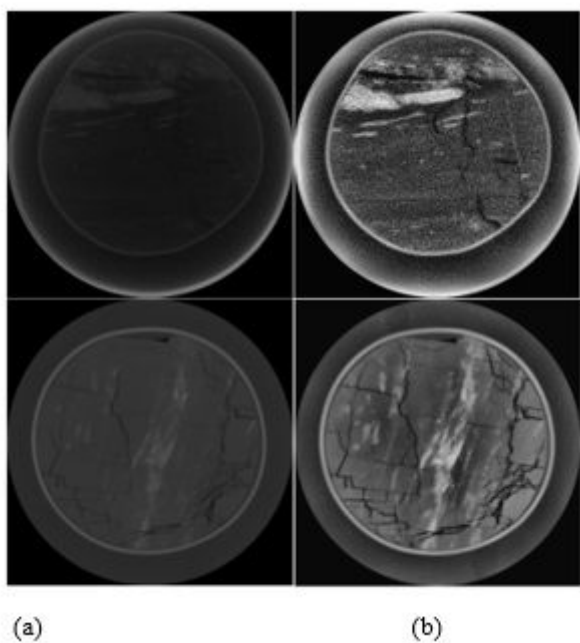

(a)   (b)

Figure 4

The result image after: (a) original image. (b) after globally CLAHE operation.
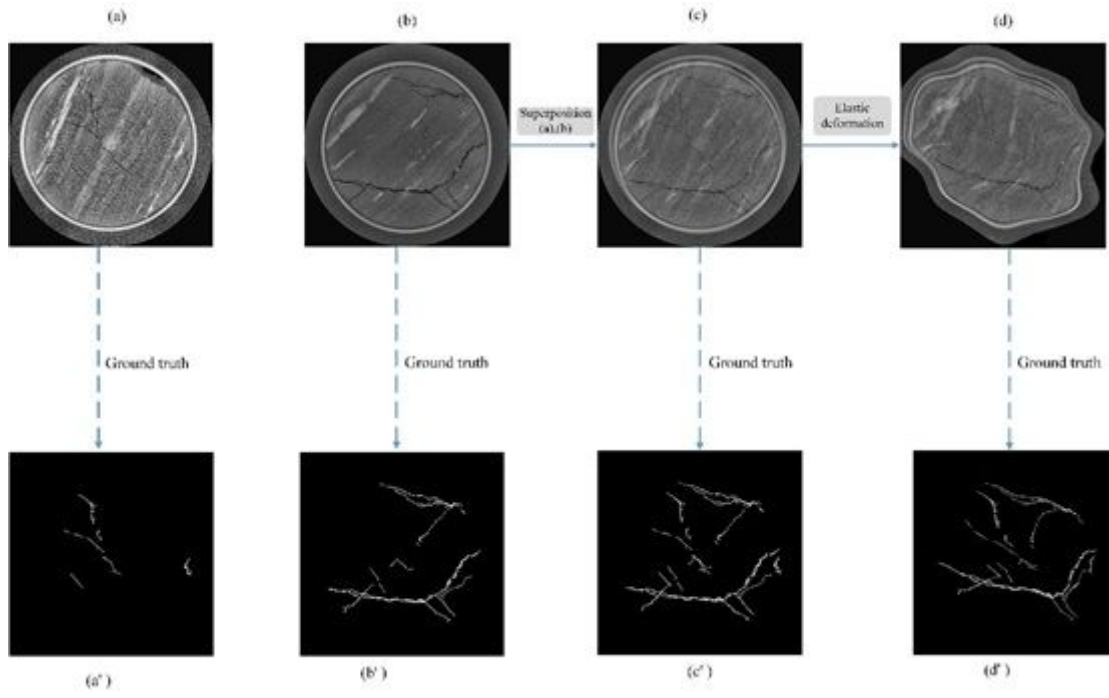
Figure 5

The process of data augmentation. (a), (b) The original coal rock CT images. (a▯), (b▯) The ground truth of (a) and (b). (c) The result of superimposing (a) on (b). (c▯) The ground truth of (c). (d) The result of (c) after elastic deformation. (d▯) The ground truth of (d).
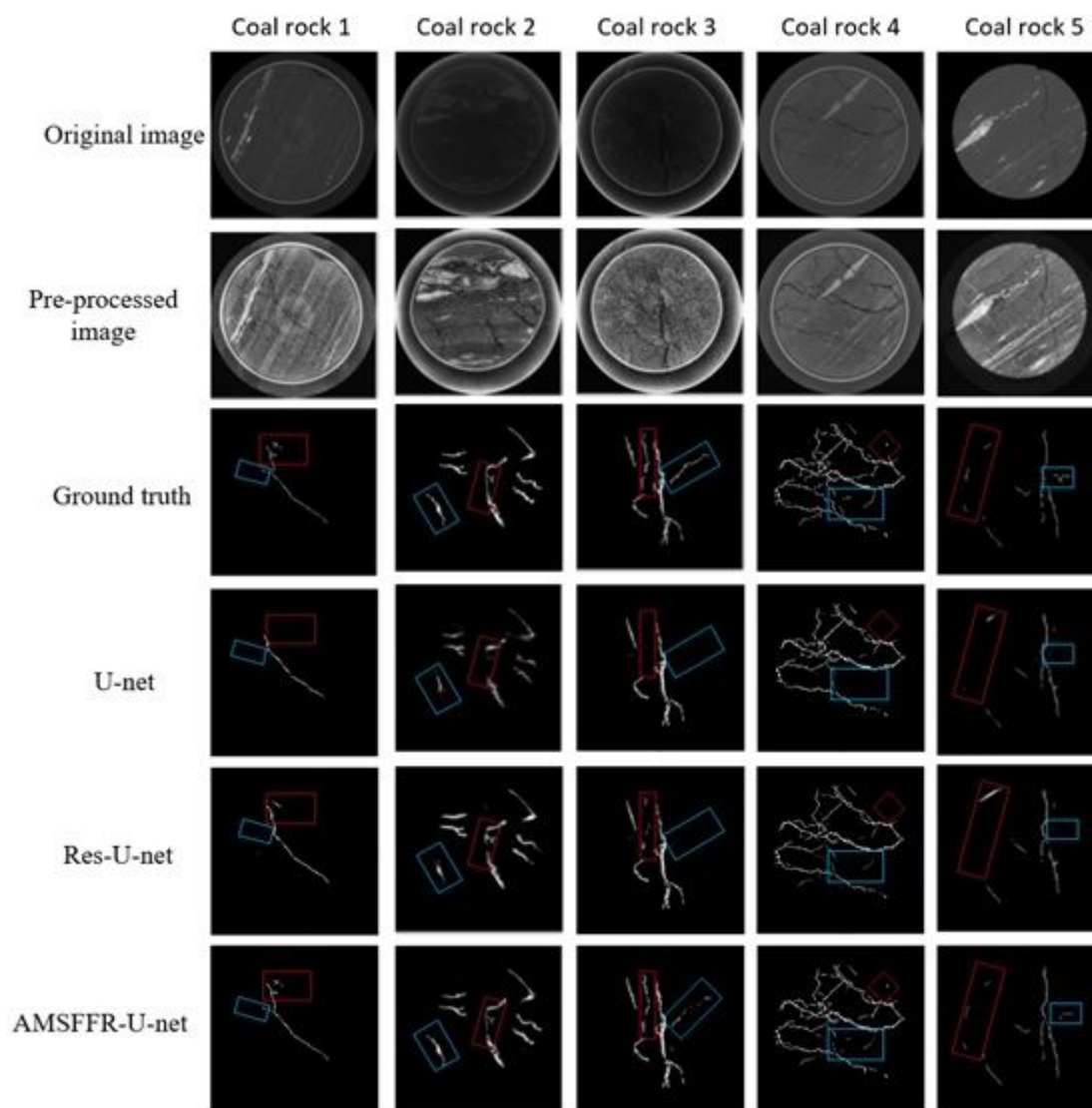
Figure 6

The fracture segmentation results of three networks on five different region of coal rock samples.

Figure 7
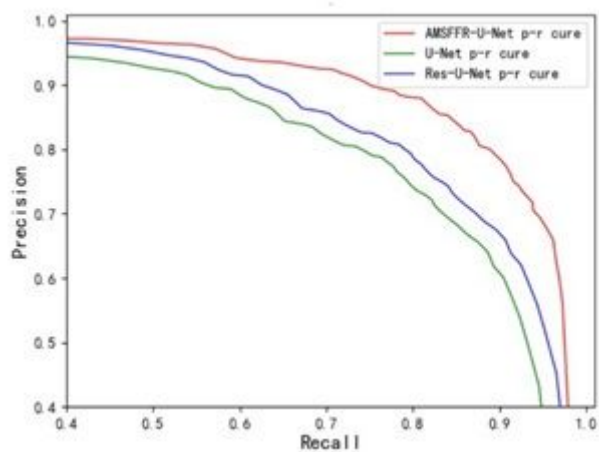
Precision-Recall curves for various methods on our test set of coal rock CT images.