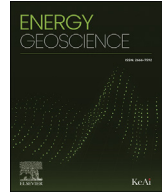




Contents lists available at ScienceDirect

Energy Geoscience

journal homepage: [www.keaipublishing.com/en/journals/energy-geoscience](http://www.keaipublishing.com/en/journals/energy-geoscience)

# Extraction of fractures in shale CT images using improved U-Net

Xiang Wu<sup>a</sup>, Fei Wang<sup>a,\*</sup>, Xiaoqiu Zhang<sup>b</sup>, Bohua Han<sup>c</sup>, Qianru Liu<sup>c</sup>, Yonghao Zhang<sup>c</sup>

<sup>a</sup> College of Geology Engineering and Geomatics, Chang'an University, Xi'an, Shaanxi, 710054, China

<sup>b</sup> Qinghai Oilfield Exploration and Development Research Institute, Dunhuang, Gansu, 736202, China

<sup>c</sup> China Petroleum Logging Co. Ltd., Xi'an, Shaanxi, 710077, China

## ARTICLE INFO

### Article history:

Received 31 December 2022

Received in revised form

9 February 2023

Accepted 20 April 2023

### Keywords:

CT slices

Fracture segmentation

Shale

U-Net

Deep learning

## ABSTRACT

Accurate extraction of pores and fractures is a prerequisite for constructing digital rocks for physical property simulation and microstructural response analysis. However, fractures in CT images are similar in grayscale to the rock matrix, and traditional algorithms have difficulty to achieve accurate segmentation results. In this study, a dataset containing multiscale fracture information was constructed, and a U-Net semantic segmentation model with a scSE attention mechanism was used to classify shale CT images at the pixel level and compare the results with traditional methods. The results showed that the CLAHE algorithm effectively removed noise and enhanced the fracture information in the dark parts, which is beneficial for further fracture extraction. The Canny edge detection algorithm had significant false positives and failed to recognize the internal information of the fractures. The Otsu algorithm only extracted fractures with a significant difference from the background and was not sensitive enough for fine fractures. The MEF algorithm enhanced the edge information of the fractures and was also sensitive to fine fractures, but it overestimated the aperture of the fractures. The U-Net was able to identify almost all fractures with good continuity, with an MIou and Recall of 0.80 and 0.82, respectively. As the image resolution increases, more fine fracture information can be extracted.

© 2023 Sinopec Petroleum Exploration and Production Research Institute. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

In low-permeability and low-porosity reservoirs, natural and artificial fractures are the main channels for fluid flow and mass transport. In addition, the presence of fractures also significantly affects the acoustic and mechanical properties of rocks, such as the dynamic and static modulus (Stoll et al., 2019; Liu et al., 2022). Accurate extraction and quantification of fractures are crucial for studying the fluid flow mechanisms in fractured reservoirs and the mechanical properties of the rocks. With the advancement of 3D computed tomography (CT) and computer performance, digital rock physics (DRP) provides an efficient method for studying rock microstructure. DRP digitizes the rock's pores, fractures, and

mineral composition, and then numerically simulates various physical processes in the digital object to obtain the rock's microscopic response. It has been used to study the micro-scale damage mechanism of rocks, fluid-solid coupling effects, and carbon dioxide capture and storage (Blunt et al., 2013; Ramandi et al., 2017; Zhu et al., 2020).

As a critical step in the digital rock physics workflow, the quality of image segmentation affects the subsequent numerical simulation results and analysis of rock microstructure properties (Iassonov et al., 2009; Wang et al., 2022). Image segmentation refers to the classification of pixels in CT images into components such as pores, fractures, and matrix. Ideally, the gray-scale histogram of rock images has multiple peaks representing different rock components. However, the reality is more complex. First, the aperture of fractures in CT images is only a few pixels, which is greatly limited by spatial resolution. In addition, noise and artifacts caused by ray hardening and partial volume effects during scanning also greatly affect the quality of CT images. It can be seen that accurate fracture segmentation in CT images is still a challenging task (Sheppard et al., 2004; Karimpouli et al., 2020; Lee et al., 2021; Lu et al., 2022).

\* Corresponding author.

E-mail address: [wangfei@chd.edu.cn](mailto:wangfei@chd.edu.cn) (F. Wang).



Production and Hosting by Elsevier on behalf of KeAi

For images with poor quality, appropriate filtering is usually used to remove noise and enhance the morphological information of fracture edges. Nonlinear filtering methods are mainly used in CT images (Lee et al., 2021), including Anisotropic Diffusion Filter, Unsharp Masking filter, Non-Local Mean filter and Membrane Enhancement Filter, etc. These methods can highlight the fracture information while limiting noise and artifacts (Perona and Malik, 1990; Buades et al., 2005; Martinez-Sanchez et al., 2014).

There are several common methods for image segmentation, such as thresholding, edge detection, region growing, and hybrid methods (Karimpouli et al., 2020; Wang et al., 2020; Lee et al., 2021). Karpyn et al. (2007) used different thresholds based on the grayscale histogram to segment saturated Berea sandstone with vertical fractures into oil, water, and matrix phases. However, when the differences between fractures and background are not significant, i.e., when the fracture features are not sufficiently distinct, global thresholding has poor segmentation performance. Edge detection algorithms extract boundary features by identifying pixels with significant grayscale changes, i.e., by using the grayscale gradient. Based on this idea, researchers use Canny, Sobel, and Laplacian operators to extract different edge features. Edge detection can only detect the edges of fractures and matrix, and the interior of fractures needs to be completed through a series of transformations such as dilation. The edge detection method is also sensitive to the boundaries between noise and non-fracture features, so this method often over-segments (Lee et al., 2021; Tang et al., 2021). Despite the recent developments of improved edge detection operators based on different principles can optimize segmentation results (Wang et al., 2020), their performances are still unstable and their applicability is limited. Region-growing based image segmentation methods including the watershed algorithm proposed by Beucher and Meyer (1993) and the active contour method introduced by Kass et al. (2004). Taylor et al. (2015) can accurately segment the pore space structure in sandstone. Lee et al. (2021) used improved Chan-Vese (Chan and Vese, 2001) active contour methods to segment the CT images of marble rock cores based on local threshold segmentation to obtain rough masks. This method is ineffective for recognizing micro-fractures in regions near the edges of the image, where the aperture of the fractures may be overestimated due to beam hardening.

In view of the difficulty in accurately segmenting CT images, other hybrid algorithms have been proposed. Sheppard et al. (2004) used anisotropic diffusion and unsharp mask to preprocess CT images and proposed the Converging Active Contours (CAC) algorithm, which combined the watershed with active contour methods. Using parallel computing, the core data of 2000<sup>3</sup> voxels can be processed quickly and the pore structure and paleontological fossil structure in sandstone can be accurately extracted. Ramandi et al. (2017) proposed an improved CAC algorithm to segment each component of the coal images. The algorithm is divided into four stages: firstly, the resolvable structure is segmented, including fractures and pores; Secondly, the pores and fractures with sub-resolution are extracted; thirdly, the organic mineral regions of coal that are not penetrated by X-ray attenuation in the coal structure are extracted; finally, the inorganic mineral phases in the high-density area are segmented. The expansion and corrosion methods in morphology operation were used to distinguish the sub-resolution pores and fractures. Voorn et al. (2013) proposed Multiscale Hessian fracture Filtering (MHF) based on an image-based Hessian matrix. After applying a Gaussian filter to enhance fracture features, the structure information obtained from the Hessian matrix can be easily and accurately segmented. Deng et al. (2016) introduced the Technique of Iterative Local Thresholding (TILT) based on MHF to segment the carbonate core CT images of the Bedford Formation. After obtaining a rough mask of the

fracture using 3D Otsu global threshold segmentation and MHF, the local threshold method is applied to repeatedly reduce the rough fracture to a finer shape until its boundary reaches the fracture profile. Finally, the fracture aperture, porosity, roughness, and other morphological parameters were calculated quantitatively. Li and Zhang (2019) built a fracture contour evolution model according to the variation law of coal rock fracture morphology in three-dimensional space, and obtained the fine fracture structure by using the proposed 3D filtering method and gradient direction consistency model.

The aforementioned methods rely on setting feature extraction methods for different images and are susceptible to the choice of parameters and image quality. As a result, they can only be applied to specific types of images, lacking a certain degree of generalizability. Consequently, it is necessary to develop and validate a new model that can accurately detect rock fractures using advanced features, even in the presence of a significant amount of noise (Byun et al., 2021; Lu et al., 2022). In recent years, convolutional neural networks (CNN) have been widely applied in various fields as they are highly capable of extracting features from data. CNN has been used for tasks such as object detection, image classification and segmentation, and super-resolution reconstruction. One of the advantages of using CNNs is that the trained model can be applied to other similar data sets, making it suitable for transfer learning. In the field of earth science, CNNs have been used for tasks such as seismic impedance inversion, total organic carbon (TOC) content prediction, pore identification and multiphase flow prediction (Karimpouli et al., 2020; Asante-Okyerere et al., 2021; Wang et al., 2022; Niu et al., 2023). Researchers have also applied CNNs to fracture identification, including the segmentation of fractures in coal-rock CT images, electrical imaging well log images, and field outcrop images (Karimpouli et al., 2020; Byun et al., 2021; Lu et al., 2022).

U-Net is a semantic segmentation model based on convolutional neural networks (CNN) that was initially used for medical image processing. It is a classic encoder-decoder structure that uses skip connections to fuse low-level and high-level feature maps in order to obtain image features at different scales. The network is able to achieve good prediction results even with a small amount of training data (Ronneberger et al., 2015; Wang et al., 2022). Lei et al. (2021) used a U-Net with added attention modules to accurately segment inertinite and pyrite in coal. Lu et al. (2022) proposed an adaptive multi-scale feature fusion model based on U-Net, which aggregates semantic information at multiple levels and accurately identifies fractures in coal-rock CT images. Reinhardt et al. (2022) compared the results of threshold segmentation, clustering analysis, watershed segmentation, U-Net, and random forest algorithms for segmenting sandstone with a single fracture, and found that both the random forest algorithm and U-Net were able to identify almost all fracture pixels, but the former performed better and had higher training efficiency when dealing with small amounts of data.

Despite fracture extraction methods are well documented, the use of deep learning methods for extracting fractures in rocks mainly focuses on outcrop rocks and coal samples, with few reports on CT fracture segmentation of shale. To address the challenge of accurately extracting fractures in shale CT images, this study has built a CT image dataset that includes fractures with multiple scales, covering fractures of various shapes and apertures. Using a modified U-Net, we have conducted pixel-level classification of the images and analyzed the impact of image resolution and brightness on the segmentation results obtained by the model. We have also compared and contrasted the segmentation results from the U-Net with those from other traditional algorithms.

## 2. Material and methods

### 2.1. CT image dataset production

Shale CT images were collected from samples of the Longmaxi and Wufeng formations in the Sichuan Basin, resulting in data from 10 rock samples. Each sample had approximately 1000 slices, with a resolution of approximately 14  $\mu\text{m}$  per pixel. As the imaging scan produced highly similar slices with small vertical intervals between them, every 200<sup>th</sup> slice was selected from the original dataset, resulting in a total of 60 original slices. The built-in Image Labeler toolbox of Matlab is used to label the original image at pixel level, Figure 1 shows labeling process for a typical image which consists of different fracture apertures.

As illustrated in Fig. 2(a) and Fig. 2(b), the original shale CT images have low contrast, making it difficult to discern the edge features of fractures. To improve the brightness and contrast of the images while suppressing noise, we apply the contrast-limited adaptive histogram equalization (CLAHE) method as a pre-processing step (Zuiderveld, 1994). The algorithm is implemented in Opencv and consists of three steps. First, the image is divided into multiple blocks, and tile-Grid-Size is set to  $8 \times 8$  in the algorithm. Then calculate the cumulative gray distribution of each block, and adjust the gray value of each pixel in the original image according to the contrast limit threshold (clip-Limit is set to 2 in the algorithm). The resulting images are shown in Fig. 2(c) and (d). Figure 3 illustrates the change in grayscale values before and after applying the CLAHE method, which demonstrates that the low grayscale values in fracture regions are enhanced. In general, a larger dataset allows for more information to be learned by the model, improving its generalization ability. To further expand the dataset, we employ image augmentation techniques such as over-laying and elastic distortion, resulting in a total of 200 images.

To enable the model to learn the characteristics of fractures at various scales and apertures and consider computer performance, the original images ( $2048 \times 2048$  pixels) transformed into  $512 \times 512$  pixels through three paths. Firstly, the original image is directly compressed to  $512 \times 512$  pixels. Secondly, dividing the original image into 4 equal parts along each direction and getting 16 sub-images of  $512 \times 512$  pixels. Thirdly, dividing the image into 2 equal parts along each direction, resulting in 4 sub-images of  $1024 \times 1024$  pixels, and then compressing them to  $512 \times 512$  pixels. This process resulted in a multi-scale fracture dataset, as illustrated in Fig. 4. Through these operations, the dataset is expanded to 4200 images, which are split into a training set and validation set in a 9:1 ratio for training.

### 2.2. U-Net semantic segmentation

#### 2.2.1. U-Net integrated with scSE

The U-Net semantic segmentation model is a classical encoder-decoder structure, as shown in Fig. 5 (Ronneberger et al., 2015;

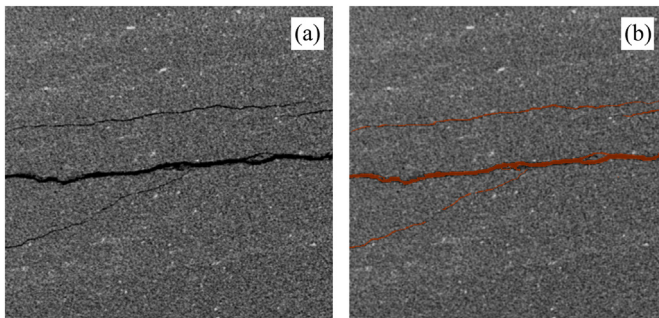


Fig. 1. Image annotation. (a) Original image; (b) Annotated image.

Wang et al., 2022). The left side is the encoder, i.e. the image feature extraction part, which primarily consists of convolutional layers and down-sampling layers. In this study, we employ the classic VGG16 network as the feature extraction network, with 5 down-samplings, resulting in a final feature map size of 1/32 of the original image. The right side of the decoder structure is used to restore the original resolution of the feature map and primarily consists of up-sampling layers (bilinear interpolation is used in this study) and skip connection layers. The skip connection layers combine semantic information from different network levels through concatenation, effectively capturing the texture and edge features of the image.

In addition, a scSE (Spatial and Channel Squeeze & Excitation) attention mechanism module is added after each convolutional block in the decoding stage, as shown in Fig. 6. The scSE module is a parallel combination of channel attention mechanism modules and spatial attention mechanism modules. By using compression and excitation, it assigns different weights to the spatial and channel features of the image, recalibrating the feature map to focus on more effective image information, thereby improving the accuracy of the classification results (Roy et al., 2018).

#### 2.2.2. Mixed loss function

The loss function is used to calculate the error between the manually annotated ground truth and the predicted values output by the model. It plays a decisive role in the training process of the model, and the network can optimize the weight coefficients of the model through the change of the loss function value. Due to the extremely imbalanced number of samples of fracture pixels and matrix pixels, and considering the problem of pixel-level classification for fracture segmentation, the sum of Focal Loss (Lin et al., 2017) and Tversky Loss (Hashemi et al., 2018) is selected as the mixed loss function to evaluate the model's classification effect on each pixel. Focal Loss adds balancing factors and probability factors to the cross-entropy loss function, effectively addressing the issue of imbalanced positive and negative sample numbers and the challenge of training some unidentifiable samples. Tversky Loss is a measure of the similarity between the prediction results and the true label set, and the commonly used Dice coefficient is a special form of it. In order to penalize false negatives (i.e. fracture pixels that are misclassified as background),  $\alpha$  and  $\beta$  in Tversky Loss are set to 0.3 and 0.7, respectively.

$$\text{Loss} = \omega \times \text{Floss} + (1 - \omega) \times \text{Tloss} \quad (1)$$

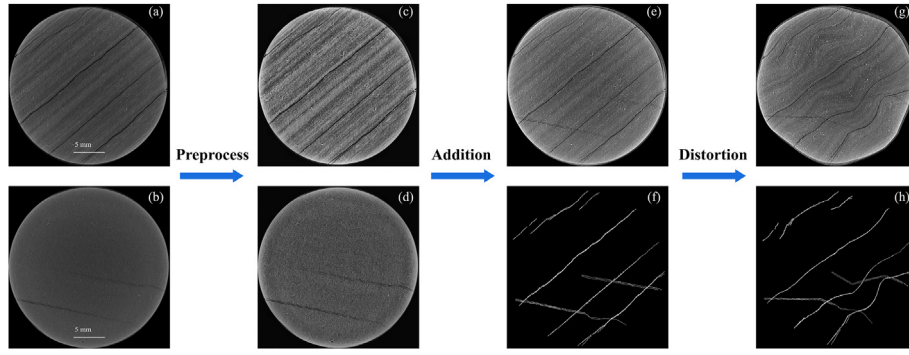
$$\text{Tloss}(\alpha, \beta) = \frac{X \cap Y}{X \cup Y + \alpha(X + Y) + \beta(X - Y)} \quad (2)$$

$$\text{Floss}(p_t) = -a_t(1 - p_t)^\gamma \log(p_t) \quad (3)$$

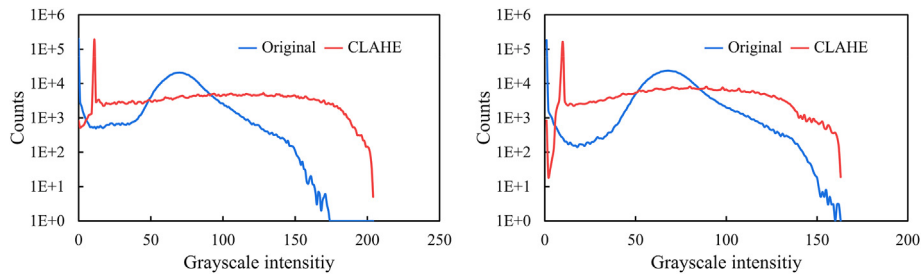
Where  $\omega$  is the coefficient of the two loss functions and is set to 0.5.  $X$  denotes the label value and  $Y$  denotes the predicted value. The probabilities of different categories are balanced by the balancing factor, and  $\gamma$  is the focus parameter.

#### 2.2.3. Hyperparameter settings

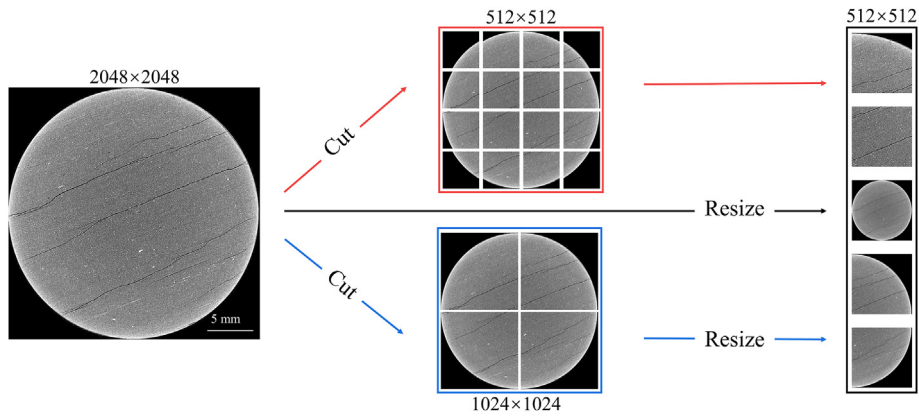
Refining the performance of a CNN requires persistent tuning of its network parameters to achieve increasingly precise predictions. To streamline the training process, it is typical to manually define hyperparameters. The number of iterations, or epochs, refers to the total number of times the model is updated during the training process. In the present study, we set the number of epochs to 50 for efficiency. The batch size refers to the number of CT images that are input to the model for calculation at once. In order to consider



**Fig. 2.** Images preprocessing and enhancement methods. (a) and (b) Original CT images; (c) and (d) Preprocessed image; (e) Superimposed image; (g) Distorted image; (f) and (h) Labels.



**Fig. 3.** Gray histogram before and after preprocessing.



**Fig. 4.** Multi-scale data set production process.

computational performance, the batch size was set to 2 in this experiment. The optimizer is used to compute the gradient of the loss function with respect to the network parameters, and updates the weights of the network nodes in the direction of the gradient descent in order to minimize the loss. The learning rate determines the magnitude of the update to the network parameters as they move in the direction of the gradient descent. After multiple trials, we selected the Adam optimizer for this experiment, with an initial learning rate of 0.004 and an exponential decay method for updating the learning rate, with a decay coefficient of 0.95.

### 3. Results and discussion

#### 3.1. U-Net training process

The training process of the model is shown in Fig. 7. The size of

the loss steadily decreases during the training process, converging in fewer than 10 epochs and the final validation set loss is slightly larger than the training set loss, without overfitting. In addition, the model's accuracy was measured using the recall and mean intersection-over-union (Mlou). Mlou is the mean of the intersection and union of the sets of true and predicted values for each class; recall is the probability of positive samples being predicted as positive among the actual positive samples, and a higher recall means that more fracture pixels can be predicted. The calculation methods for Mlou and recall are shown in Equations (4) and (5), respectively. Consistent with the change in the loss function, the model achieved high Mlou and sensitivity in the first few epochs of training, but at around the 24th epoch, recall reached a maximum and Mlou reached a minimum, which may be the result of the model over-segmenting. The model's final Mlou is stable at 0.80, and the final recall is stable at 0.82.



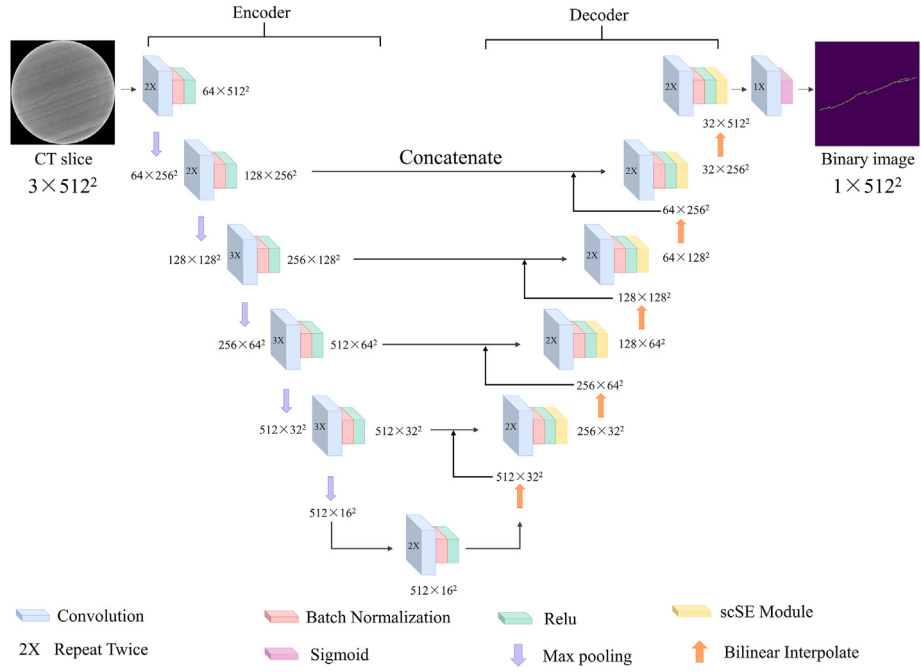


Fig. 5. U-Net network structure (feature extraction network being Vgg16).

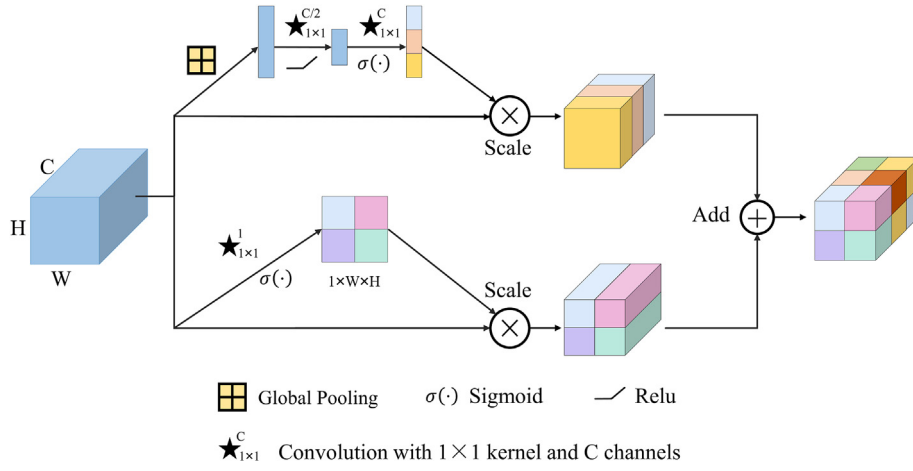


Fig. 6. scSE module.

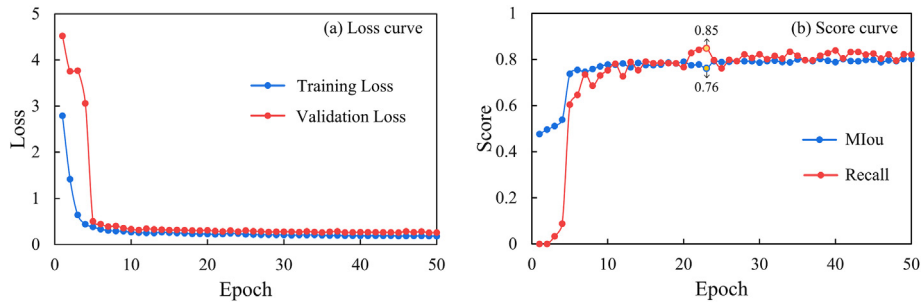


Fig. 7. Loss value and precision change during training. (a) Loss curve; (b) Score curve.

$$MIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{TP}{FN + FP + TP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

Where  $TP$ ,  $TN$ ,  $FP$  and  $FN$  respectively represent true positive, true negative, false positive and false negative.

### 3.2. Factors affecting U-Net segmentation results

Figure 8 shows the prediction results of the U-Net at different resolutions and after preprocessing. The yellow is the fracture recognized by the model. For the two original images, it can be seen that the inside of the image is dark and the outside is brighter, especially for original image 2, which is the result of beam hardening during imaging. The model's recognition of fractures in the original image is poor. Although increasing the resolution can extract more fracture pixels, the aperture of the recognized fractures is larger and does not match the actual fracture features. However, the model can accurately extract the fractures in the preprocessed image. When the resolution is increased, it can be seen that the aperture of the fractures recognized by the U-Net is more distinct and more in line with the real image. In addition, from the marked red rectangle, it can be seen that larger sizes provide more details of microfractures, which the model can also accurately extract. However, from the marked blue rectangle, it can be found that as the resolution increases, the model's sensitivity to some wide and rough fractures is not high, which may be due to the pixels occupied by the fracture aperture exceeding the range of the

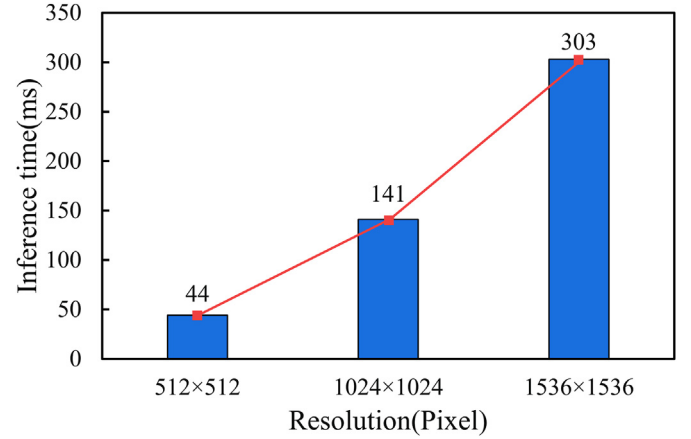


Fig. 9. Inference time at different resolutions.

U-Net model. Figure 9 shows the time required for the model to infer a single image of different sizes. It can be seen that as the size of the image input to the model increases, the inference time for a single image increases significantly. Consequently, it is necessary to strike a balance between resolution and computational efficiency.

### 3.3. Analysis of segmentation results by different methods

Figure 10 and Figure 11 show the fractures extracted by the U-Net, the Membrane Enhancement Filter (MEF), Otsu, and Canny edge detection, including two-dimensional slices and three-dimensional views. The three-dimensional view shows the result

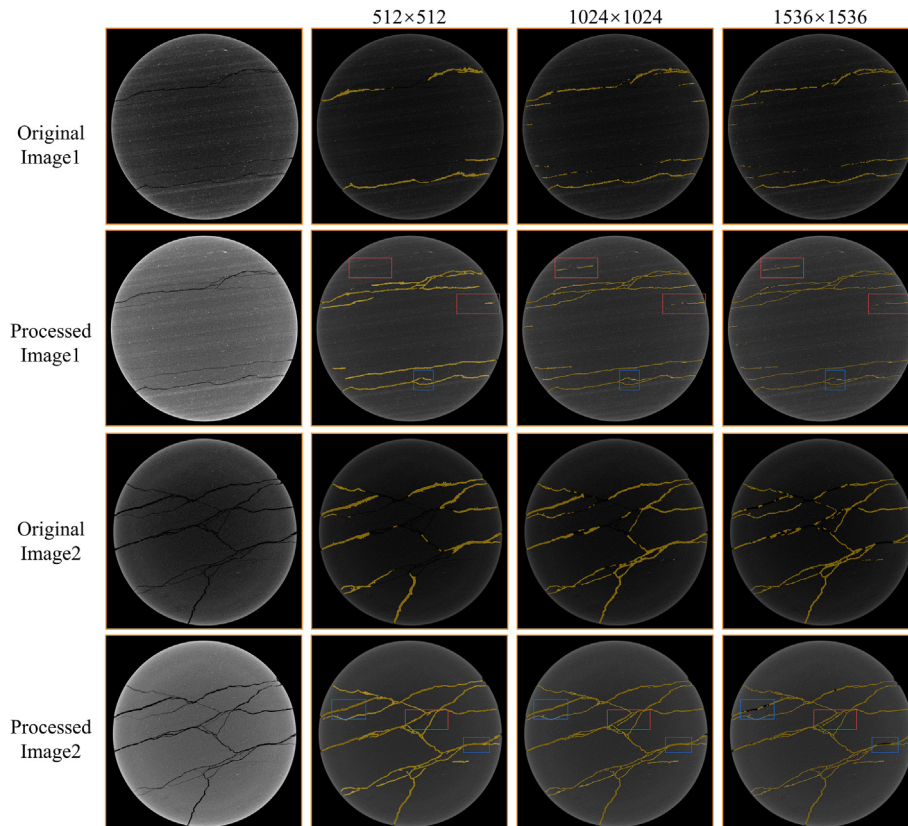


Fig. 8. The U-Net model outputs for different processing modes.

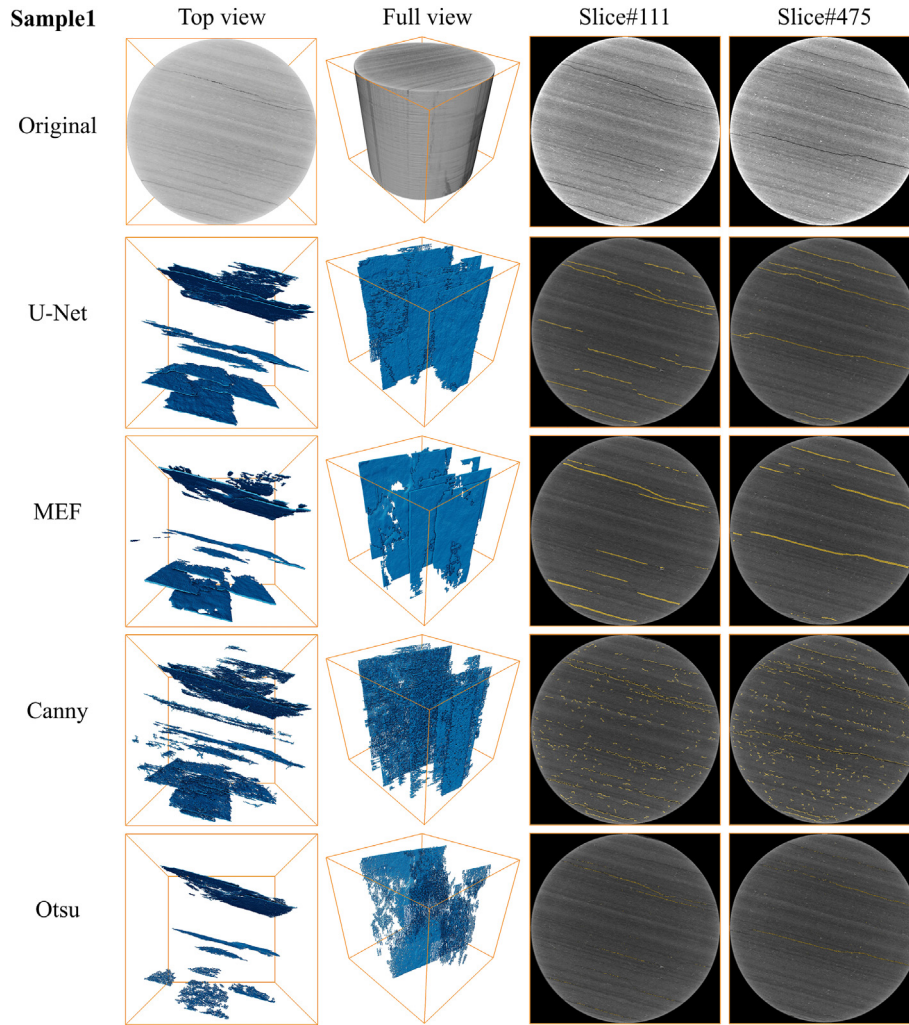


Fig. 10. Segmentation results of Sample 1.

of removing islands in the three-dimensional fracture body with a threshold of 300 voxels after initial segmentation. The two-dimensional slice shows the segmentation result before removing islands, where the yellow lines are the extracted fractures. MEF is an edge enhancement algorithm proposed by for cell structures, which propagates local information in three-dimensional space based on tensor voting algorithms and enhances the structural information of voxels belonging to the same structure (Martinez-Sanchez et al., 2014). This method is implemented in Avizo, where the feature scale and tensor voting scale are set to 1 and 10 pixels, respectively. Canny edge detection consists of four steps. First, Gaussian filtering is performed with a Gaussian filter kernel of 3 pixels. Then, Sobel filtering is used to obtain gradient information. Next, non-maximal suppression is performed along the gradient direction, and finally, the edges are connected based on the given upper bound of 100 and lower bound of 1. Otsu is a global threshold segmentation method known as the maximum interclass variance method. Its principle is that the larger the variance between the foreground and background segments obtained by the adaptive threshold size, the greater the difference between the two parts, and the more likely they are to be two different components. Figure 12 shows the core porosity measured by helium method and fracture volume fraction calculated by different methods. The porosity by helium method was measured by UltraPore-400 in

accordance with API RP 40:1998, recommended practices for core analysis, IDT. It should be noted that helium porosity measures the entire core (4 cm–5 cm in height), whereas X-ray CT scans only measure the central core at 1 cm in height due to imaging method limitations.

According to the segmentation results, the Otsu algorithm can only recognize fractures that are more obvious and is not sensitive to fractures with small gray scale differences with the background (matrix), or even if some microfractures are recognized, but their continuity is very poor, they may be considered islands on the three-dimensional body and removed.

From the two-dimensional slice, it can be seen that the Canny edge detection algorithm is able to roughly identify the position of the fracture, but it can only segment the boundary between the matrix and the fracture, and further filling is needed for the interior of the fracture. In addition, this method will excessively segment and mistakenly detect other areas with significant changes in gray scale gradient as fractures. However, from the three-dimensional perspective, it can be seen that these false detections can be removed after removing the isolated islands. As can be seen in Fig. 12, the final fracture porosity is significantly different from the laboratory measurement, mainly due to the inability to extract the interior of the fracture.

The MEF algorithm demonstrates an advantage in its ability to

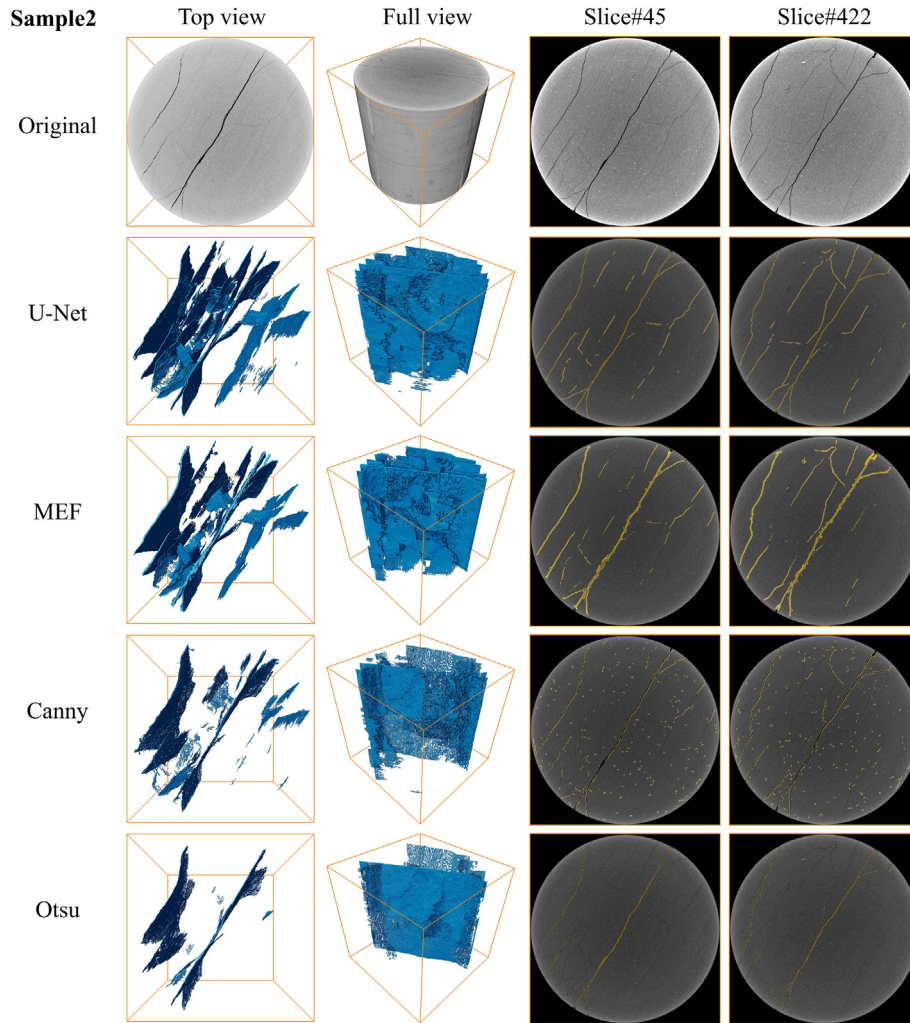


Fig. 11. Segmentation results of Sample 2.

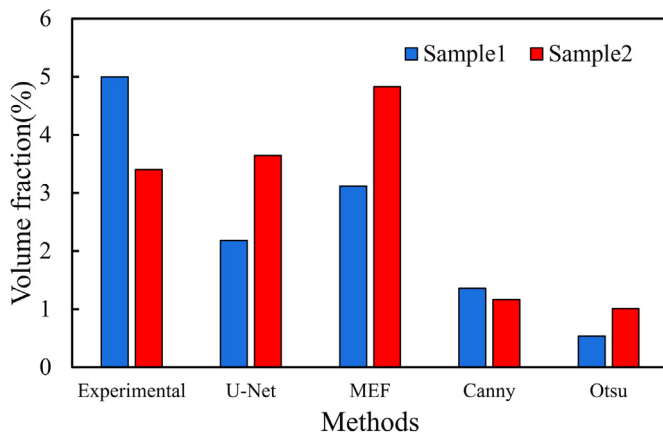


Fig. 12. Volume fraction of fractures obtained by different methods.

reinforce information about similar components in three-dimensional space, enabling the identification of a majority of visible fractures and a portion of micro-fractures with good continuity. However, analysis of two-dimensional slice information reveals a significant overestimation of fracture aperture by this

method, leading to a large discrepancy between the calculated fracture porosity and the measured value in the laboratory.

The U-Net semantic segmentation model, through the utilization of skip connection layers, is able to effectively combine information from multiple scales within the input images, resulting in segmentation outputs that accurately capture the natural features of rock fractures. Comparison of the calculated fracture porosity values obtained using this model with those measured in the laboratory reveals a high degree of proximity. Analysis of the model's performance on Sample 1, as demonstrated through the prediction results for 2D slices, indicates that the U-Net is able to accurately identify a majority of the fractures present, with particularly strong performance on Slice #475 where it is able to extract micro-fractures that were not captured by other methods. In the case of Sample 2, which exhibits more complex fracture shapes, the U-Net model demonstrates a high level of accuracy in identifying the majority of fractures. However, instances of under-detection do occur for some highly challenging-to-identify fractures, leading to lower connectivity for some fractures.

#### 4. Conclusions

- (1) The CLAHE algorithm can effectively remove noise and enhance the visibility of dark fracture features, thereby



facilitating more accurate fracture extraction. The multi-scale nature of the dataset enhances the model's generalization ability.

- (2) The U-Net demonstrates strong performance in integrating semantic information at various levels to produce accurate fracture segmentation results. As the resolution of the input image increases, the network is able to identify more micro-fractures. However, when the size of the image becomes too large, the aperture of the fractures may exceed the model's recognition range, resulting in missed detections.
- (3) In this study, methods such as Otsu, Canny, and MEF were evaluated for their ability to segment fractures in shale CT images. It was found that these methods require manual parameter tuning to achieve relatively satisfactory results. Specifically, the Canny edge detection algorithm tended to exhibit over-segmentation, and was limited in its ability to only identify the boundary between fractures and matrix due to the nature of its algorithm. The Otsu algorithm was found to be less sensitive to fine fractures, only able to accurately extract those with clear differences from the background. The MEF method was able to extract most fractures, including some micro-fractures, but tended to overestimate the aperture of the fractures.

Overall, the results suggest that the U-Net is the most effective and accurate method for extracting fractures in shale CT images. However, there is still potential for improvement. The original dataset used in this study only includes data from 10 core samples, and it may be beneficial to construct a more representative dataset by incorporating shale CT images with a diverse range of characteristics. Additionally, the current model processes 2D slices individually, ignoring vertical connectivity information. It is possible that a 3D semantic segmentation model may produce more satisfactory results in this application.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgements

This work was supported by the Natural Science Basis Research Plan in Shaanxi Province of China (Program No. 2022JM-147).

### References

- Asante-Okyere, S., Ziggah, Y.Y., Marfo, S.A., 2021. Improved total organic carbon convolutional neural network model based on mineralogy and geophysical well log data. *Unconventional Resour.* 1, 1–8.
- Beucher, S., Meyer, F., 1993. The morphological approach to segmentation: the watershed transformation. *Mathematical morphol. image processing* 18, 49.
- Blunt, M.J., Bijeljic, B., Dong, H., Gharbi, O., Iglauer, S., Mostaghimi, P., Paluszny, A., Pentland, C.H., 2013. Pore-scale imaging and modelling. *Adv. Water Resour.* 51, 197–216.
- Buades, A., Coll, B., Morel, J., 2005. A non-local algorithm for image denoising. 2005 IEEE Computer Soc. Conf. Computer Vision Pattern Recognition (CVPR'05) 2, 60–65.
- Byun, H., Kim, J., Yoon, D., Kang, I., Song, J., 2021. A deep convolutional neural network for rock fracture image segmentation. *Earth Sci. Inform.* 14, 1937–1951.
- Chan, T.F., Vese, L.A., 2001. Active contours without edges. *IEEE Trans. Image Process.* : publ. IEEE Signal Process. Soc. 10 (2), 266–277.
- Deng, H., Fitts, J.P., Peters, C.A., 2016. Quantifying fracture geometry with X-ray tomography: Technique of Iterative Local Thresholding (TILT) for 3D image segmentation. *Comput. Geosci.* 20 (1), 231–244.
- Hashemi, S.R., Mohseni Salehi, S.S., Erdoğan, D., Prabhu, S.P., Warfield, S., Gholipour, A., 2018. Asymmetric loss functions and deep densely-connected networks for highly-imbalanced medical image segmentation: application to multiple sclerosis lesion detection. *IEEE Access* 7, 1721–1735.
- Iassonov, P., Gebrenegus, T., Tuller, M., 2009. Segmentation of X-ray computed tomography images of porous materials: a crucial step for characterization and quantitative analysis of pore structures. *Water Resour. Res.* 45 (9).
- Karimpouli, S., Tahmesbi, P., Ramandi, H.L., 2020. A review of experimental and numerical modeling of digital coalbed methane: imaging, segmentation, fracture modeling and permeability prediction. *Int. J. Coal Geol.* 228, 103552.
- Karpyn, Z.T., Grader, A.S., Halleck, P.M., 2007. Visualization of fluid occupancy in a rough fracture using micro-tomography. *J. Colloid Interface Sci.* 307 (1), 181–187.
- Kass, M., Witkin, A.P., Terzopoulos, D., 2004. Snakes: active contour models. *Int. J. Comput. Vis.* 1, 321–331.
- Lee, D., Karadimitriou, N.K., Ruf, M., Steeb, H., 2021. Detecting micro fractures: a comprehensive comparison of conventional and machine-learning-based segmentation methods. *Solid Earth* 13 (9), 1475–1494.
- Lei, M., Rao, Z., Wang, H.D., Chen, Y.L., Yu, H., 2021. Maceral groups analysis of coal based on semantic segmentation of photomicrographs via the improved U-Net. *Fuel* 294 (2), 120475.
- Li, Z.W., Zhang, G.Y., 2019. Fracture Segmentation Method Based on Contour Evolution and Gradient Direction Consistency in Sequence of Coal Rock CT Images. *Mathematical Problems in Engineering*, 2019.
- Lin, T.Y., Goyal, P., Girshick, R.B., He, K.M., Dollár, P., 2017. Focal loss for dense object detection. *IEEE Trans. Pattern Anal. Mach. Intell.* 42, 318–327.
- Liu, Z.Y., Zhao, H.F., Shi, H.E., 2022. Experimental study on stress monitoring in fractured-vuggy carbonate reservoirs before and after fracturing. *J. Petrol. Sci. Eng.* 218, 110958.
- Lu, F.L., Fu, C.C., Zhang, G.Y., Shi, J., 2022. Adaptive Multi-Scale Feature Fusion based U-net for fracture segmentation in coal rock images. *J. Intell. Fuzzy Syst.* 42, 3761–3774.
- Martinez-Sanchez, A., Garcia, I., Asano, S., Lucic, V., Fernandez, J.J., 2014. Robust membrane detection based on tensor voting for electron tomography. *J. Struct. Biol.* 186 (1), 49–61.
- Niu, X., Zhang, J., Liu, J., 2023. Seismic impedance inversion in depth domain based on deep learning. *Unconventional Resour.* 3, 72–83, 2023.
- Perona, P., Malik, J., 1990. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* 12, 629–639.
- Ramandi, H.L., Mostaghimi, P., Armstrong, R.T., 2017. Digital rock analysis for accurate prediction of fractured media permeability. *J. Hydrol.* 554, 817–826.
- Reinhardt, M., Jacob, A., Sadeghnejad, S., Cappuccio, F., Arnold, P., Frank, S., Enzmann, F., Kersten, M., 2022. Benchmarking conventional and machine learning segmentation techniques for digital rock physics analysis of fractured rocks. *Environ. Earth Sci.* 81.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional Networks for Biomedical Image Segmentation. *ArXiv*, abs/1505.04597.
- Roy, A.G., Navab, N., Wachinger, C., 2018. Recalibrating fully convolutional networks with spatial and Channel “Squeeze and excitation” blocks. *IEEE Trans. Med. Imag.* 38, 540–549.
- Sheppard, A.P., Sok, R.M., Averdunk, H., 2004. Techniques for image enhancement and segmentation of tomographic images of porous materials. *Phys. Stat. Mech. Appl.* 339, 145–151.
- Stoll, M., Huber, F., Trumm, M., Enzmann, F., Meinel, D., Wenka, A., Schill, E., Schäfer, T., 2019. Experimental and numerical investigations on the effect of fracture geometry and fracture aperture distribution on flow and solute transport in natural fractures. *J. Contam. Hydrol.* 221, 82–97.
- Tang, Y.D., He, L., Xiao, H.G., Wang, R.H., Lu, W., Xu, T., 2021. Fracture extraction from smooth rock surfaces using depth image segmentation. *Rock Mech. Rock Eng.* 54, 3873–3889.
- Taylor, H.F., O'Sullivan, C., Sim, W.W., 2015. A new method to identify void constrictions in micro-CT images of sand. *Comput. Geotech.* 69, 279–290.
- Voorn, M., Exner, U., Rath, A., 2013. Multiscale Hessian fracture filtering for the enhancement and segmentation of narrow fractures in 3D image data. *Comput. Geosci.* 57, 44–53.
- Wang, H.S., Dalton, L.E., Fan, M., Guo, R., McClure, J., Crandall, D., Chen, C., 2022. Deep-learning-based workflow for boundary and small target segmentation in digital rock images using UNet++ and Ik-EBM. *J. Petrol. Sci. Eng.* 215, 110596.
- Wang, S.Q., Liang, S., Peng, F., 2020. Image edge detection algorithm based on fuzzy set. *J. Intell. Fuzzy Syst.* 38 (4), 3557–3566.
- Zhu, W., Zhao, L.C., Wang, C.C., Shan, R., 2020. Characterization of wave-induced pore fluid flow based on dynamic stress strain simulation on digital rocks. *Chin. J. Geophys.* 63 (6), 2386–2399.
- Zuiderveld, K., 1994. Contrast limited adaptive histogram equalization. *Graphics Gems* 474–485.