



## Automatic fracture characterization in CT images of rocks using an ensemble deep learning approach

Chuyen Pham <sup>a,b</sup>, Li Zhuang <sup>a,b</sup>, Sun Yeom <sup>b</sup>, Hyu-Soung Shin <sup>a,b,\*</sup>

<sup>a</sup> University of Science and Technology, Daejeon, 34113, Republic of Korea

<sup>b</sup> Korea Institute of Civil Engineering and Building Technology, Goyang, 10223, Republic of Korea



### ARTICLE INFO

**Keywords:**

Rock fracture  
Computed tomography  
Deep learning  
Rotated bounding box

### ABSTRACT

The presence of fractures in a rock mass can have a substantial influence on its mechanical and hydraulic properties. For many years, computed tomography (CT) scan has been effectively utilized to investigate the internal structures of rock. However, quantitative characterization of fracture based on CT data remains a challenging endeavor due to the inevitable blurry appearance of fractures in CT images, complexity in fracture patterns and the heterogeneity of host rock. In this study, a deep learning-based method is presented for automatically and accurately segmenting rock fractures in CT images, with special consideration for directional ambiguous fractures embedded in heterogeneous backgrounds. Our method involves two stages: (1) fracture detection in the form of minimal bounding boxes using the Faster R-CNN deep learning algorithm, and (2) fracture segmentation inside the detected bounding boxes using the U-Net deep learning algorithm. The detection stage aims to establish spatial constraints for the segmentation process, thereby significantly minimizing undesirable noise and artifacts existing in the background from consideration and consequently improving the segmentation result. In addition, we also develop a system that can automatically extract fracture properties including orientation, length, aperture, and wall roughness of fractures based on detection and segmentation results from our method. Experiments on CT images of fractured coarse-grained granite, fine-grained sandstone, and shale core samples were carried out in an attempt to confirm the applicability of our approach. The comparison between fracture segmentation results and the manually-annotated ground truths shows that our approach can achieve a competitive dice score of up to 0.942 and also outperform other state-of-the-art deep learning approaches including Mask R-CNN and U-Net alone.

### 1. Introduction

Fractures are defined as the structural discontinuities formed in a rock mass. As brittle in nature, most rock formations in Earth's crust are always fractured to some extent.<sup>1</sup> The research on rock fractures, both natural and man-made, has early sparked a lot of interest due to their effects on the mechanical and hydraulic properties of the host rock.<sup>2,3</sup> The existence of fractures in a rock mass can have either positive or negative implications for rock engineering related applications. For CO<sub>2</sub> storage and nuclear waste disposal, fractures are perhaps the most important factor to consider since they allow upward flow of CO<sub>2</sub> and other disposals from storage reservoirs such that compromising long-term storage capability and safety.<sup>4</sup> On the other hand, many commercial petroleum and geothermal reservoirs are located in deep fractured rock formations. In either case, fractures not only serve as

additional storage space, but also provide high-permeable channels for the transport of hydrocarbons or high temperature water to producing wells.<sup>5</sup> Accordingly, an accurate and thorough assessment of fracture properties like aperture, wall surface roughness and orientation are therefore critical for understanding rock properties and rock engineering behavior.

In practice, the most common way to characterize rock fracture is to observe them at outcrop and on core samples retrieved from the target formation. Recently, the introduction of computed tomography (CT) into geological sample characterization has undoubtedly offered geoscientists a new effective tool for analyzing internal rock structures such as pore, mineral, and fracture.<sup>6-11</sup> The greatest advantage of this technique is that it allows looking inside the rock samples without destructing them, which is a must-have attribute for characterizing the fracture network in rocks. Johns et al.<sup>12</sup> were among the first to use CT

\* Corresponding author. University of Science and Technology, Daejeon, 34113, Republic of Korea.  
E-mail address: [hyushin@kict.re.kr](mailto:hyushin@kict.re.kr) (H.-S. Shin).

images to characterize fractures in crystalline rock cores. Several other early studies were carried out on imaging and quantifying fractures in rocks as well as other solid samples using the CT technique.<sup>13–15</sup>

Quantitative analysis of fractures in CT images typically requires partitioning the fracture voxels from the others (i.e., minerals, pores), which is referred to as the segmentation process. In general, segmentation of CT images can be done manually or automatically. However, while the manual approach is error-prone and time-consuming, the previously established automatic methods share a number of limitations that inhibit their applications in fracture segmentation.<sup>16</sup> In the beginning, thresholding methods (i.e., Otsu's or Yen's algorithm) and region-growth method (i.e., watershed algorithm) were widely employed for CT image segmentation.<sup>17,18</sup> Since these methods rely on identifying a single grayscale value to partition fractures from the background, they severely suffer from imaging resolution limits, complicated fracture patterns, and also the occurrence of artifacts in CT images. The difficulties become more evident when attempting to measure extremely fine fractures in large heterogeneous rock samples.<sup>19</sup> To partly deal with this problem, ridge filters such as Hessian and Frangi filters are commonly adopted to enhance the contrast of curve-like structures (i.e., fractures) to the rest.<sup>20,21</sup> The effectiveness of this approach, however, is restricted since ridge filters also enhance the boundary of minerals in the background and perform poorly in the case of fractures with varying aperture. As a result, the large open fractures may be distinguished from the background, but with the consequence that some of the minor fractures were lost as well as false positive detections (i.e., pixels are not fracture, however, they may still be categorized as fracture pixel).

Other notable methods for CT image segmentation include unsupervised (i.e., k-means and fuzzy c-means) and supervised machine learning techniques (i.e., support vector machine, random forest and artificial neural network). The recent works of Chauhan et al.<sup>22</sup> and Guntoro et al.<sup>23</sup> review in detail the performance of machine learning methods on rock CT image segmentation. Although these studies indicate that the machine learning methods are obviously better than other conventional methods relying solely on grayscale, the segmentation result was found to be highly affected by the feature vector selection scheme, as well as sensitive to the grayscale contrast between phases.

For the past few years, deep learning has arisen as a powerful tool for image data processing tasks including object detection and segmentation. Given a sufficient amount of training data, deep learning algorithms are clearly proved to outperform the aforementioned methods. Several attempts have been made to apply deep learning to identify cracks or fractures in CT images and other imaging modalities. Lee et al.<sup>24</sup> and Karimpouli et al.<sup>25</sup> proposed different deep learning architectures for automatically segmenting fracture or cleat in CT images of rocks. Deep learning was utilized by Chen et al.<sup>26</sup> and Byun et al.<sup>27</sup> to trace cracks on rock outcrop images. Many studies tried to extract cracks on roadways and concrete constructions using deep learning.<sup>28,29</sup> Although prior researches have demonstrated the promise of using deep learning at addressing relatively clear cracks or fractures segmentation in images, they have yet to be tackled in the heterogeneous problem of images and vagueness of fractures. Encoder-decoder based deep learning architecture such as U-Net and SegNet is perhaps the most widely used algorithm for the task of crack and fracture segmentation.<sup>30,31</sup> Despite simple architecture and efficiency, those algorithms may have some difficulties in distinguishing two distinct objects that have the same intensity values. For this reason, prior knowledge such as shape or location can help to further improve U-Net performance. Another well-known deep learning segmentation algorithm is Mask R-CNN.<sup>32</sup> Despite of considering as the most effective way to obtain instance segmentation, many latest researches in biomedical image segmentation indicate that the performance of Mask R-CNN on pixel-wise segmentation is worse than U-Net.

In summary, the literature reviews presented above show that segmenting fractures in CT images and extracting fracture information are

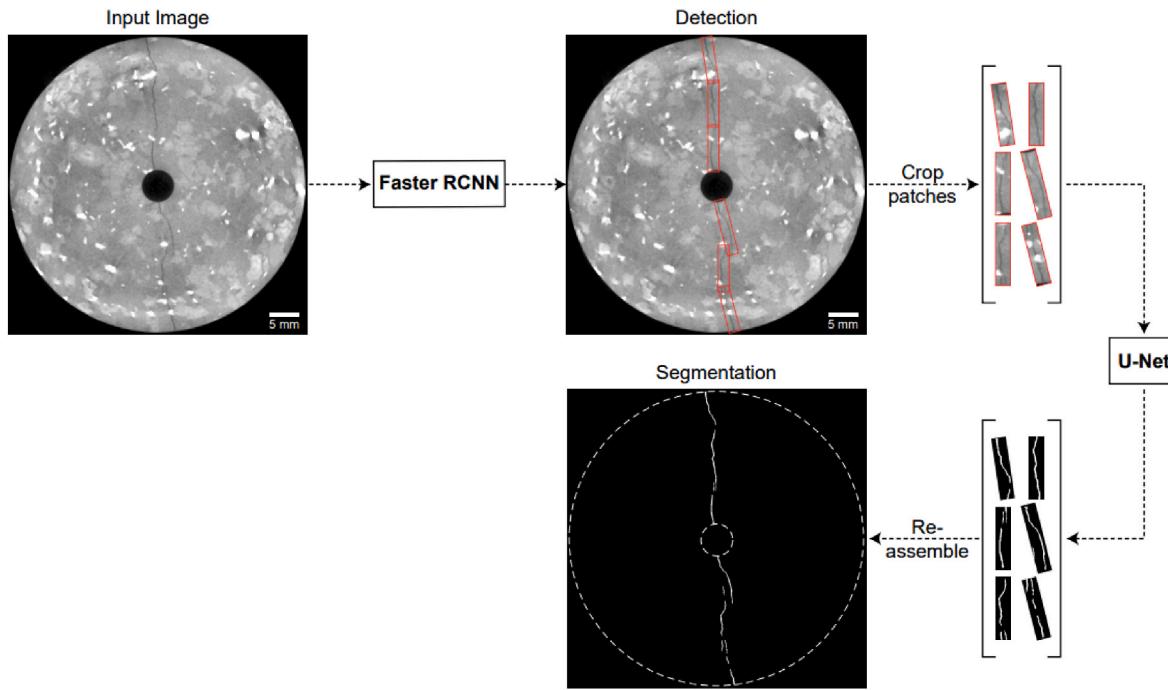
challenging tasks that still have to be addressed. In this study, we aim to develop an approach for automatic pixel-wise segmentation of fractures in CT images giving particular attention to dealing with ambiguous fractures and heterogeneity in the background. The novelty in our proposed approach is the combination of two deep learning algorithms i.e., Faster R-CNN<sup>33</sup> and U-Net<sup>30</sup> to detect and segment fractures in CT images. The fundamental concept is to use the bounding box to provide prior spatial information to guide the segmentation process and also remove as much background from segmentation consideration as possible. The idea of attention in deep learning is akin to this approach. Based on this, a novel additional implementation was conducted by introducing a new concept of rotated bounding box in fracture detection. Comparing to the typically-used bounding box (i.e., horizontal bounding box), rotated bounding box shows greater adaptability to fit with the tortuous fracture, such that minimizing the ratio of background within the bounding box. Accordingly, our approach is able to address the challenges of segmenting very fine fractures with highly variable aperture as well as the complexity of background. It is also worth noting that our objective is to automatically detect and segment the fracture with efficiency and accuracy demonstrated in respect of the manually-created ground truth rather than to examine the real aperture of the fracture. As a result, the accuracy of segmentation is only evaluated based on ground truth. In the following sections, we provide an overview of the detection and segmentation algorithms as well as describe the datasets and comprehensive experiments to demonstrate the applicability of our approach.

## 2. Methodology

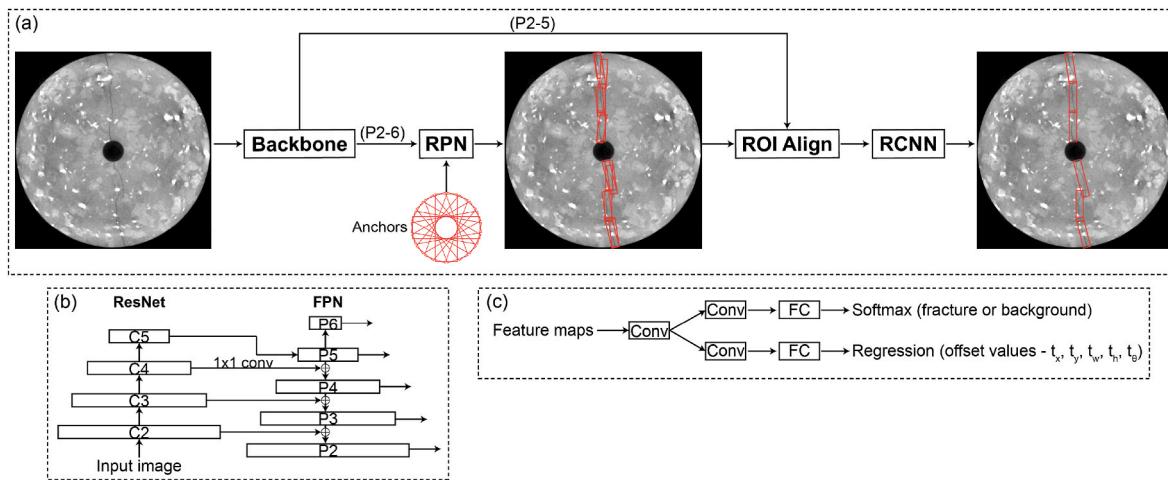
Our proposed approach is composed of two sequential stages: (1) modified Faster R-CNN to propose bounding boxes restricting the areas where fracture exists, and (2) U-Net to segment the fracture within the bounding box. The complete pipeline for getting fracture segmentation is shown in Fig. 1. Moreover, for the purpose of comparison, we also implement fracture segmentation using Mask R-CNN and U-Net alone. The following sections will explain each component in more depth.

### 2.1. Modified faster R-CNN with rotated bounding boxes

Faster R-CNN, a well-known region-based object detection algorithm, was first published by Ren et al.<sup>33</sup> Although several different models including YOLO and SSD have been developed with the aim of improving detection performance<sup>34,35</sup>, Faster R-CNN still remains as an outperforming algorithm in terms of detection accuracy. Regardless of the complexity in architecture, Faster R-CNN can be generally divided into three major components: backbone network, region proposal network (RPN) and region-based convolutional neural network (R-CNN) (Fig. 2). The backbone network refers to a deep convolutional neural network, which is used to convert the input images into multi-scale feature maps by successively convolving them with filters following by other operations (i.e., pooling and activation). The feature maps are higher-level representations of an image, in which important features that are difficult to spot with human eyes can be identified.<sup>36</sup> Following Lin et al.<sup>37</sup>, we employ the backbone network as a combination of a deep convolutional neural network (ResNet) and feature pyramid network (FPN) in this study. About RPN and R-CNN, they both compose several CNN layers and fully-connected layers as the output. Their operations are almost identical including: (1) predicting bounding boxes where fracture is likely to exist and (2) refining the coordinate of those boxes to better match with the fracture pattern. The sole difference is that R-CNN employs a specific number of partial regions in the feature maps corresponding to proposal boxes (i.e., RPN's output) for final bounding box prediction, whereas RPN uses the whole feature maps extracted from the input image. Since the architecture of general Faster R-CNN has been exhaustively explored in publications, we will only discuss the modifications within each component that are required for rotated bounding



**Fig. 1.** Illustration of fracture segmentation pipeline: Fracture detection using Faster R-CNN algorithm with rotated bounding boxes at the beginning stage followed by the second stage of fracture segmentation in detected bounding boxes using U-Net.



**Fig. 2.** (a) Rotated Faster R-CNN architecture, (b) backbone and (c) RPN and R-CNN.

box prediction and the key enhancement made in this study.

Before diving into the details of Faster R-CNN, it is beneficial to enumerate the concept of rotated bounding box and anchor box. The use of rotated bounding box for object detection has been introduced in the past few years primarily for scene text detection and ships, airplanes and other vehicles detection in aerial images.<sup>38,39</sup> By introducing a new parameter angle ( $\theta$ ), this type of bounding box can be rotated to any direction such that fitting better to the object's shape makes it unique and fully contained inside the bounding box. The angle parameter is defined as the angle formed between the positive vertical axis and the longer side of the bounding box ranging from  $-90^\circ$  to  $90^\circ$ . The other parameters for the representation of rotated bounding box include center point coordinate ( $x_c, y_c$ ), width ( $w$ ), and height ( $h$ ). It is also worth noting that the width of a bounding box is always less than its height. In line with rotated bounding box, anchor boxes are a collection of pre-defined rotated bounding boxes (i.e., with several specific scales, aspect ratios and angles), which are uniformly distributed over the

image. Those anchor boxes play a role as references for generating proposal boxes during the RPN stage.

RPN network will learn to identify the anchor boxes containing fracture and then produce target regression coefficients to transform those anchor boxes into better proposal bounding boxes. For that purpose, the RPN network is designed with a convolution layer followed by two sibling fully-connected (FC) branches known as classifier and regressor. During training, an anchor box can be labeled as positive, negative, or neutral corresponding to foreground (i.e., fracture-contained) boxes, background boxes and ambiguous boxes based on its overlap with ground truth boxes and preset thresholds (i.e., IoU – Intersection over Union).<sup>33</sup> Consequently, the classifier branch will learn how to predict anchor box label by using the Softmax activation function.<sup>33</sup> The output of this branch will be objectiveness scores, which reflect how confident the network is that the anchor box contains fracture. At the same time, one or more than one positive anchor boxes will be assigned to each specific ground truth box. The offset value between

assigned anchor boxes and the ground truth box will be calculated as the equations below. The first four equations present the scale-invariant translation between center coordinates and the log-space height-width shift. Additionally, the fifth equation presents the angle rotation in radians between orientations. The regressor branch will eventually learn to predict those offset values. Afterward, the predicted offset values will be used to transform an anchor box into the final predicted proposals via equations, which are simply the reverses of the above-mentioned equations.

$$t_x = (x - x_a) / w_a \quad (1)$$

$$t_y = (y - y_a) / w_a \quad (2)$$

$$t_w = \log(w / w_a) \quad (3)$$

$$t_h = \log(h / h_a) \quad (4)$$

$$t_\theta = (\theta - \theta_a) \times \frac{\pi}{180^\circ} \quad (5)$$

where  $x_a$ ,  $y_a$ ,  $w_a$ ,  $h_a$  and  $\theta_a$  represent the x-y center coordinate, width, height and orientation of anchor box, respectively, and  $x$ ,  $y$ ,  $w$ ,  $h$  and  $\theta$  represent the x-y center coordinate, width, height and orientation of the ground truth box, respectively. By examining the result from the first implementation, we recognized that an anchor box, which has a large different angle to ground truth box, can be assigned as positive anchor since its overlap with ground truth box is higher than assignment threshold (i.e.,  $\text{IoU} > 0.5$ ). To solve this issue, we modified the conventional IoU by multiplying it with an index referred to as angle weight to take into account for angle difference and then used the angle-weighted IoU for anchor assignment. This implementation helps to decay IoU respecting to the angle difference with Gaussian penalty function as following equations:

$$\text{IoU}_{\text{angle weighted}} = \text{weight} \times \text{IoU}_{\text{normal}} \quad (6)$$

$$\text{weight} = 1 - e^{-\alpha \times \left(1 - \frac{\Delta\theta}{90}\right)^2} \quad (7)$$

where  $\Delta\theta$  represents the acute angle between two oriented bounding boxes ranging from  $0^\circ$  to  $90^\circ$  and  $\alpha$  is empirical constant, which was set to 0.1 in this study. As mentioned earlier, a ground truth box can be matched up with one or more than one anchor boxes that lead to the prediction may compose of overlapping boxes. To remove redundant boxes, the proposal boxes will be passed through a filter algorithm known as rotated non-maximum suppression (rNMS). rNMS will keep or remove boxes according to their objectiveness scores and IoU index, such that if a bounding box overlaps considerably with a higher objectiveness scores bounding box (i.e.,  $\text{IoU} > 0.7$ ), it will be removed. After non-maximum suppression, the list of proposal bounding boxes is eventually clipped, leaving only up to 512 bounding boxes with highest objectiveness scores.

Although RPN has a bounding box regression branch, it only gives coarse bounding boxes that may contain fractures. The primary reason is due to the fact that RPN has to use high-variance anchor boxes to predict the potential boxes.<sup>40</sup> Thus, a second stage R-CNN is needed to further refine the bounding boxes. In this stage, proposal boxes will now serve as reference boxes and also the basis to crop feature maps for the input of this stage. Since the size of the proposal boxes is different, cropped feature maps cannot be directly input into R-CNN. Region of Interest alignment (ROI Align) is therefore carried out to extract fixed-size feature maps. The remaining operations within this stage are similar to RPN as described above. The outputs of R-CNN will be bounding box list along with their objectiveness scores, which is again filtered using rNMS with an IoU threshold of 0.5 to get the final predicted bounding box list.

For further improving the detection result, Cascade R-CNN, an extension of Faster R-CNN with two additional stages for bounding box refinement<sup>41</sup>, is also investigated in this study. The two additional stages of Cascade R-CNN are similar to R-CNN discussed above, but with different IoU thresholds for bounding box assignment, which are 0.6 and 0.7 corresponding the first and second stages, respectively. In this way, the bounding boxes are progressively refined through multiple stages, resulting in the more accurate final prediction of bounding boxes.

Faster R-CNN is optimized using a multi-task loss function, which is a summation of classification losses and box regression losses (Eq. (8)). For each region of interest, the loss function is calculated as follows:

$$L_{\text{total}} = L_{\text{rpn\_cls}} + L_{\text{rpn\_reg}} + L_{\text{rcnn\_cls}} + L_{\text{rcnn\_reg}} \quad (8)$$

$$L_{\text{cls}} = \frac{-1}{N_{\text{cls}}} \sum_{i \in \{0,1\}} p_i \log(u_i) \quad (9)$$

$$L_{\text{reg}} = \frac{1}{N_{\text{reg}}} \sum_{i \in \{x,y,w,h,\theta\}} \text{smooth}_{L1}(t_i^u - v_i) \quad (10)$$

$$\text{smooth}_{L1}(x) = \begin{cases} 0.5x^2, & \text{if } |x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases} \quad (11)$$

where  $L_{\text{cls}}$  and  $L_{\text{reg}}$  represent the classification loss and bounding box regression loss in each stage (i.e., RPN and R-CNN). The classification loss is basically a categorical cross-entropy loss that computes the difference between the predicted probability vector ( $u$ ) and the ground truth one-hot vector ( $p$ ) (Eq. (9)). The bounding box regression loss for all rotated bounding box representations adopted the  $\text{smooth}_{L1}$  loss function (Eq. (10)), which is defined in Girshick.<sup>42</sup> In short,  $\text{smooth}_{L1}$  loss function, also known as Huber loss function, is a combination of mean square error loss and mean absolute error loss in which the square term will be used if absolute error falls below 1 and absolute term otherwise (Eq. (11)).

## 2.2. U-Net

U-Net is a deep learning-based semantic segmentation algorithm primarily presented by Ronneberger et al.<sup>30</sup> for segmenting 2D biomedical images. A typical U-Net architecture consists of a contraction path and a symmetric expanding path connecting via skip connections. The contraction path, also known as an encoder, comprises several stacks of convolution layer following by ReLU activation and max-pooling operation to downscale the image. The contraction path, like the backbone in Faster R-CNN, helps extract the feature maps from input image and learn the context of the image. Conversely, the expanding path, as indicated by its name, progressively upscales the feature map back to its original size by using a stack of transpose-convolution (so-called deconvolution) layers. The sigmoid activation is used after the last convolutional layer for pixel-wise binary classification. The breakthrough in U-Net is the introduction of skip connections, which is a bridge to directly propagate the feature maps from the contraction path to expanding path. These skip connections play an important role to recover spatial information lost during downscaling, making U-Net suitable for fine-grained details semantic segmentation. To date, a variety of U-Net-type segmentation algorithms have been created and have become a standard in the segmentation of biomedical as well as other complex image segmentation such as satellite or aerial images. In this study, we investigate using a variant of U-Net so-called the Residual Attention U-Net in order to obtain better segmentation results. For this variant, the traditional contraction block will be replaced by residual block (i.e., a component of ResNet-34) and attention gates are added to each level of skip connection.<sup>43</sup> While the residual blocks allow the network to learn more complicated patterns, the attention gates allow the network to focus on specific areas of the image where an interesting object can be found rather than the full

image. The accuracy of the segmentation result has improved greatly as a result of these two improvements. Fig. 3 depicts the U-Net architecture used in this study.

To deal with the class imbalance problem (i.e., the number of uninterested background pixels is considerably higher than the number of fracture pixels), a weighted loss function ( $L_{total}$ ), which is a combination of dice loss ( $L_{dice\_loss}$ ) and binary cross-entropy loss ( $L_{cross\_entrop\_loss}$ ), was used for training the network (Eq. (12)). While the binary cross-entropy loss treats background and fracture pixels equally, the Dice loss function, also known as region-based loss, is the harmonic mean of Precision (i.e., the possibility that the model to not predict a background pixel as a fracture pixel) and Recall (i.e., the possibility that the model to find all fracture pixels). Thus, it equally weighs false positive and false negative predictions for addressing class imbalance (Eq. (13)).<sup>44</sup>

$$L_{total} = L_{dice\_loss} + L_{cross\_entrop\_loss} \quad (12)$$

$$L_{dice\_loss} = 1 - \text{Dice\_coefficient} = 1 - \frac{2}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}} = 1 - \frac{2}{\frac{TP+FP}{TP} + \frac{TP+FN}{TP}} \quad (13)$$

where  $TP$ ,  $FP$  and  $FN$  are true positive, false positive and false negative areas, respectively.

### 3. Validation study

#### 3.1. Data preparation

We validated our algorithms using the CT images of different rock samples including two granites (Pocheon granite and Gonghe granite), gray sandstone and Mancos shale. Pocheon granite and gray sandstone were artificially fractured in hydraulic fracturing tests in laboratory<sup>45,46</sup>, while Gonghe granite contains an inclined natural fracture.<sup>47</sup> CT images of shale samples fractured in the triaxial test were collected from open-source.<sup>48</sup> The CT scanning of granites and sandstone was performed by using a multi-tube industrial X-ray CT equipment at the Korea Institute of Civil Engineering and Building Technology. Each scan provides a stack of 1024 image slices with a size of 1024x1024. In general, CT images are visualized in grayscale with value ranging from 0 to 255

representing the X-ray attenuation coefficient of the rock's components. Accordingly, high density minerals appear brighter than pores and fractures, which were filled by air under dry condition. Among the samples, granite samples are shown to be highly heterogeneous with multiple mineral components that are irregularly distributed. While quartz and feldspar have almost similar ranges of density, micas (particularly biotite) have greater density resulting in high intensity regions in the images, which highly influence surrounding materials and present a significant challenge for fracture segmentation when it passes through this material. This effect is quite pronounced in Gonghe granite, which has a greater biotite concentration (13.4%) than that in Pocheon granite (2.6%). For sandstone and shale samples, despite showing relatively high homogeneity, the fracture pattern is more complex with crossings and branches. The spatial resolutions of all the CT images used in this study are ranging from 50  $\mu\text{m}$  to 75  $\mu\text{m}$ . Prior to further image processing, CT images were subjected to a non-local means filter and contrast limited adaptive histogram equalization to remove noise and increase phase contrast.

Fig. 4(a) shows an example of a CT image of Pocheon granite sample with bi-wing hydraulic fractures emanating from the wellbore. Three major mineral components are clearly observed in the image, in which white, light gray, and dark gray shades colors represent biotite, feldspar and quartz clusters, respectively. The fractures appear in the image as dark and tortuous line structures with thickness varied along their length. The thickness of fractures observed on CT image is affected by CT values of their surrounding matrix as a result of the partial volume effect.<sup>12</sup> Since pixels at the fracture's boundary can be made up of both air (i.e., a fracture filled with air) and solid matrix, the CT number of these pixels is a function of both attenuation coefficients of the two phases and their volume percentage. In addition, diffraction of the X-ray beam induced by sharp edges such as open fractures, on the other hand, results in a lower CT number in voxels surrounding the fracture. As a result of the two phenomena, CT values near the fracture boundary gradually alter, making the exact fracture boundary being obscure. Several calibration methods for estimating the true aperture of fracture from CT images have been introduced including missing attenuation, peak height or inverse point-spread function.<sup>12,15,49</sup> Fig. 4(b) shows the corresponding grayscale intensity histogram of the CT image in Fig. 4(a), which appears as a unimodal and almost normal distribution. This is

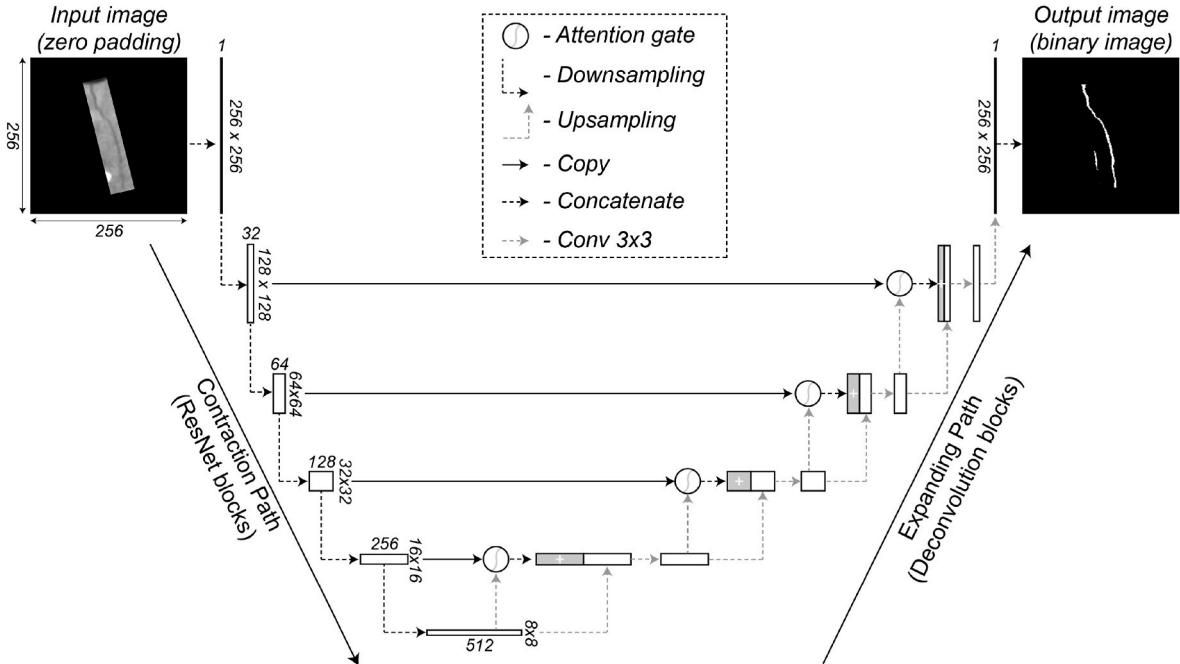
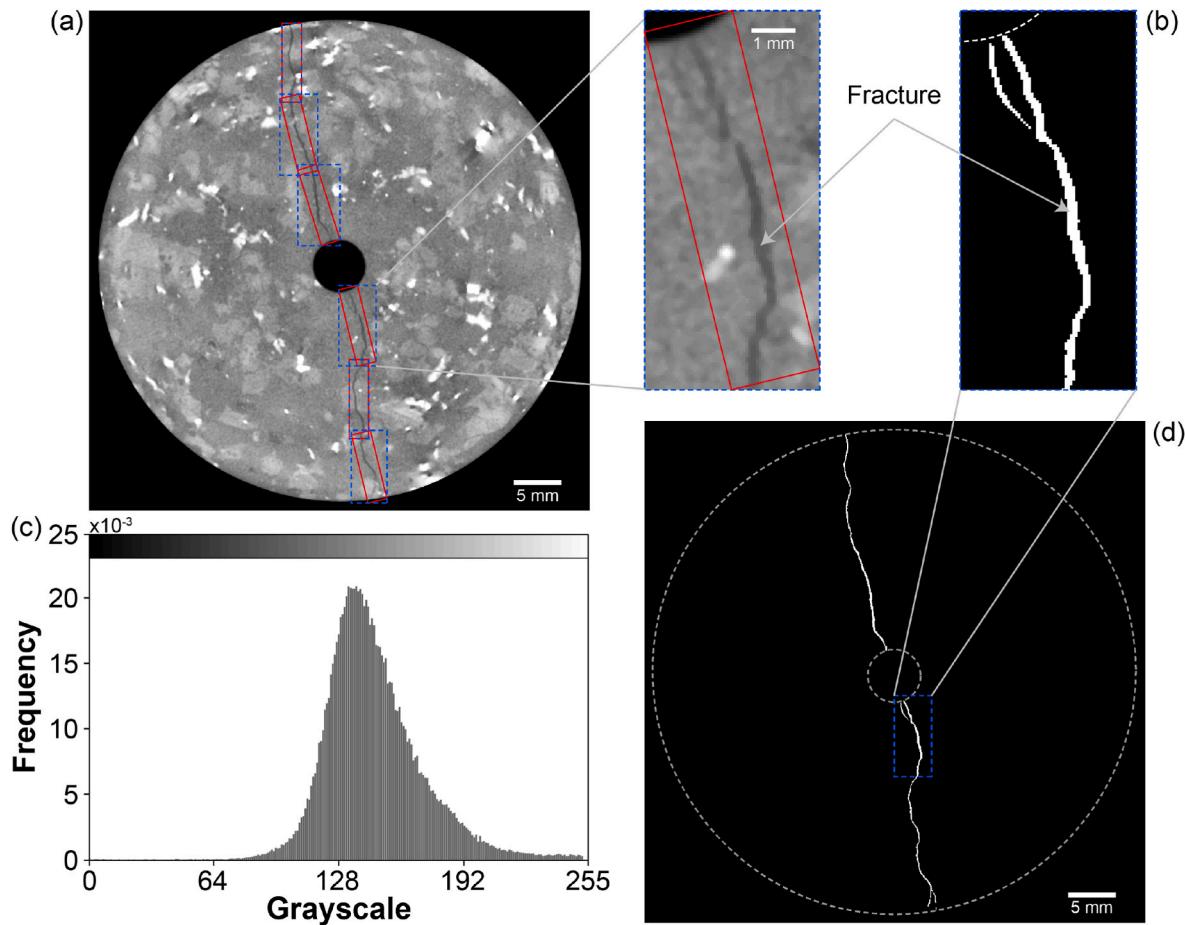


Fig. 3. The architecture of the Residual Attention U-Net.



**Fig. 4.** An example of input image: (a) CT image with horizontal ground truth bounding boxes (blue) and rotated ground truth bounding boxes (red), (b) zoom-in images within a bounding box, (c) Corresponding normalized histogram of the CT image, (d) Corresponding ground truth segmentation of the CT image. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

mainly attributed to two facts: (1) the total area of fracture in the image is relatively small in comparison to that of the uninterested background and (2) fractures have the same gray-scale intensity values as objects in background. As a result, simply assigning a global grayscale threshold is impossible to separate fractures from background.

For fracture detection dataset, 800 images were randomly selected from the previous sets. After that, a graphical image annotation tool (i.e., roLabelImg - <https://github.com/cvgict/roLabelImg>), is employed to annotate fracture with rotated bounding boxes in those CT images. To avoid excessively large uninterested background included in the bounding boxes, we annotate fractures with a sequence of overlapping rotated boxes, each of which tightly encompasses a portion of the fracture. The scale of the bounding boxes is varied from 20 to 30 pixels in width and the aspect ratio is ranging from 1/3 to 1/5, which are closely matched with the scale and aspect ratios of the anchor boxes. By annotating the bounding boxes in accordance with the anchor boxes, the model can learn to associate specific anchor boxes with particular detected bounding box sizes and aspect ratios, therefore, improving the model's performance. Fig. 4 depicts how the ground truth bounding boxes were annotated. The red and blue rectangles in Fig. 4(a) correspond to rotated bounding box and regular bounding box, respectively. Out of 800 images, 600 (75%), 100 (12.5%) and 100 (12.5%) images were randomly selected as training, validation and test dataset, respectively.

To create segmentation dataset, we must classify individual pixels in the CT image into 2 classes (i.e., fracture and background). Since this is a challenging and time-consuming operation, only a portion of the above-mentioned 800 CT images was chosen for fracture segmentation

annotation. The ground truth annotation for segmentation was done automatically with a plugin of the Fiji package namely Trainable Weka Segmentation<sup>50</sup> and then carefully amended for final segmentation ground truth. It is worth mentioning that exact labeling of the fracture width is impractical due to the fact that the fracture boundary is ambiguous as a result of partial volume and other effects. Accordingly, the ground truth mask may not reflect the real aperture of the fracture. Totally, 450 images were labeled. We keep out the same 100 images as in fracture detection dataset for test set and used the remaining for training and validation with the ratio of 8:2, respectively. By taking ground truth bounding boxes as the basis, 2431 patches were cropped out from 450 images. The number of images used for training, validation and testing for each model (i.e., detection model and segmentation model) is summarized in Table 1.

### 3.2. Deep learning experiments and results

In general, the algorithms presented in this study were implemented in Python using Numpy, OpenCV and deep learning library –

**Table 1**  
Dataset summary.

Dataset	Train set (image)	Validation set (image)	Test set (image)	Total (image)
Detection	600	100	100	800
Segmentation	280 (1965 patches)	70 (466 patches)	100 (753 patches)	450 (3184 patches)

TensorFlow.<sup>51</sup> We carried out experiments on a 4-core Intel Core i7-6700K @ 4.00 GHz processor with NVIDIA Geforce RTX 3090 GPU with memory of 24 GB. Two models i.e., detection model (Faster R-CNN) and segmentation model (U-Net) have been trained and validated separately and then be assembled for obtaining fracture segmentation result seamlessly. For each model, several alternative architectures have been investigated by changing backbone network and using cascade network in Faster R-CNN model and using attention gate in U-Net model. All other hyper-parameters such as anchor setting, batch size and number of training epoch, etc., were kept the same for comparison purpose. The detailed configuration of each model is described below.

For training and evaluating Faster R-CNN algorithm, all the images will be resized to a relatively lower resolution of  $640 \times 640$  to obtain the same size image for the input of the model as well as speed up the training and evaluation. In the inference stage, the predicted bounding box will be reverted to the original size of the image. During training, several different data augmentation methods (i.e., rotation, flip, brightness and contrast shift and sharpness) are also applied to training set for generating more diverse CT images for training. This strategy plays an important role to prevent an overfitting problem in deep learning and also improve generalization performance of the model. Since there were no available pre-trained weights for rock CT images, the training was performed from scratch by initializing each layer's weights using initialization method described in He et al.<sup>52</sup> The optimizer for training the model over 100 epochs is stochastic gradient descent (SGD), which has an initial learning rate of 0.001, momentum of 0.9, and weight decay of 0.0005. For hyperparameters of anchor box, five anchor scales of (16, 24, 32, 40 and 48) and five anchor strides of (4, 8, 16, 32 and 64) were chosen to train the model. Each anchor scale and stride correspond to a feature map level of FPN (i.e., P2, P3, P4, P5 and P6). For rotated Faster R-CNN, anchor box orientations are set to be distributed evenly over  $0^\circ$ – $90^\circ$  with  $15^\circ$  increment in combination with aspect ratios of (2, 1/2, 3, 1/3, 5, 1/5) such that provide  $360^\circ$  span anchor boxes. The other hyperparameters of the model were kept the same as the original implementation of Faster R-CNN.<sup>33</sup>

The mean average precision (mAP) with IoU thresholds of 0.5 was used for quantitative evaluations of our proposal method.<sup>53</sup> Since fractures were annotated with a group of consecutive boxes instead of a single large box, predicted bounding box group is sometimes misaligned with the ground truth owing to the difference in individual box size and number of boxes, but still catches the same detection area. Therefore, a new reasonable box matching rule is required to prevent underestimating new algorithm operation. In addition to IoU, we also include the orientation of the predicted bounding box and the number of ground truth bounding boxes that overlap it, as criterion for determining whether the predicted box is true positive or false positive detection. Accordingly, there are two cases that a predicted bounding box can be defined as true positive prediction: (1) the predicted bounding box has IoU higher than 0.5 and (2) the bounding box overlaps with more than 2 ground truth bounding boxes and the different angle between orientation of this bounding box and the ground truth boxes is less than  $10^\circ$ .

In the context of segmentation, an image will be cropped into small patches regarding to the ground truth bounding boxes as mentioned before. Each patch is then padded with zero padding into the pre-defined pixel size as a requirement of the same size image for training. During training, different augmentation methods including rotation, elastic transformation and contrast shift are also randomly applied to each patch to generate more diverse training data. We trained the model with 100 epochs by using Adam optimizer for required gradient descent optimization with an initial learning rate of 0.001. Since the output of the network is the probability map of a single patch, the probability maps of those cropped patches will be assembled at the inference stage to acquire the probability map for the entire image. From the merged probability map, the binary image of fracture is obtained by applying a threshold of 0.5 to the probability map.

For quantitative analysis of the segmentation results, we use the

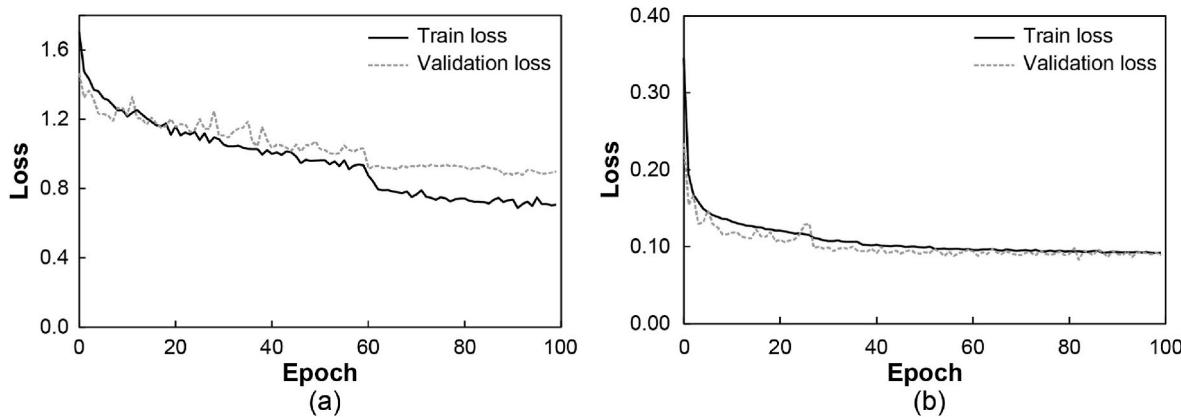
well-known Dice coefficient as a metric to evaluate the final segmentation result. The Dice coefficient evaluates the degree of overlap between the predicted and the reference segmentation masks. Accordingly, the Dice score is ranging from 0 to 1, with higher values signifying better segmentation result.

To ensure that the models were properly trained, we first look at the learning curves. Overall, both the training and validation loss curves exhibit diminishing trends as the training progresses, indicating that all models work well on unseen data (Fig. 5). The training curves of the rotated Faster R-CNN (Fig. 5(a)) reveal that the training and validation losses begin to converge after around 80 epochs. As a result, more training is unnecessary because it will not improve the model's performance. The model achieves its best mAP of approximately 89.7% on the test dataset. Both rotated and horizontal Faster R-CNN approaches show similar performance in terms of mAP. For training and inference speed, the conventional horizontal bounding box approach is certainly faster than the rotated bounding box approach adopted in this study. This is due to two facts that the number of anchor boxes in the case of horizontal bounding box approach is much less than that in rotated bounding box and IoU estimation in rotated bounding box is far more complicated than the case of horizontal bounding box, such that requiring increased computing time. However, the rotated bounding box approach can achieve accurate detection with less background area as fitting better to elongate fractures, which is an important factor to improve the segmentation result (Fig. 6(c) and d). Furthermore, different backbone networks also have an impact on model performance and training time. Better performance is achieved by using a deeper backbone network. In this work, we used ResNet-101 as the backbone since it dramatically improves detection accuracy while only requiring a minor increase in training time. For cascade Faster R-CNN, if the training time is disregarded, this network may be used to slightly improve detection as well as segmentation results.

The segmentation model, like the detection models, performs well on the test dataset, as indicated by the convergence of the validation and training loss curves (Fig. 5(b)). The highest Dice score on fracture segmentation was achieved when combining the U-Net model with the rotated Faster R-CNN model. We further study the performance of the method through comparison with two baseline models, i.e., Mask R-CNN (instance segmentation) and U-Net alone (semantic segmentation). As we can see in Fig. 6 and Table 2, our ensemble approaches (i.e., horizontal Faster R-CNN + U-Net and rotated Faster R-CNN + U-Net) can achieve Dice scores higher than 0.9 in all cases with the best case achieving 0.942 using cascade rotated Faster R-CNN. All ensemble approaches outperformed U-Net and Mask R-CNN with average improvements of approximately 8% and 26%, respectively. However, the ensemble approaches required roughly 2–3 times and 1.5–2 times longer than U-Net and Mask R-CNN, respectively, in inference. Among two ensemble approaches, the use of rotated Faster R-CNN results in a better Dice score. We can see that the rotated bounding box effectively eliminate the background from segmentation consideration while preserving detection accuracy, resulting in lower false positive and false negative rate in segmentation, and thus overtaking the horizontal bounding box on the same training dataset. In conclusion, the combination of rotated Faster R-CNN and U-Net yields the best result.

#### 4. Method applications in rock property characterization

This section presents the results of identification and characterization of fracture in two sets of CT images of Gonghe granite using the optimal method, i.e., rotated Faster R-CNN assembled with U-Net. Based on the fracture detection and segmentation results, fracture geometry information, such as the distributions of fracture orientation and length, fracture aperture, fracture wall roughness, etc., is quantitatively assessable. Furthermore, we study how segmentation result impacts estimations of fracture aperture and fracture surface roughness. This analysis illustrates the significance of the quantitative discrepancies in



**Fig. 5.** Learning curves: (a) rotated Faster R-CNN and (b) U-Net.

fracture geometry estimations, which could be critical, especially in fluid flow modeling.

#### 4.1. Fracture orientation and length

Fracture orientation can be defined in terms of strike direction and dip angle. For the first set of CT data, the fracture was artificially created along the intrinsic cleavage (i.e., rift plane) of Pocheon granite in hydraulic fracturing. The strike direction and length of fracture can be obtained from the segmentation result on each CT slice. In the case of a single relatively straight fracture existing in the image, the fracture strike coincides with the orientation of the fitting line. For the case of two or more sets of fractures with more complicated patterns such as branching or crossing, it becomes trickier since we have to split fractures into various sets based on orientation and intersecting points.<sup>54</sup> However, with the combination of detected the rotated bounding boxes and segmentation result of fractures, the fracture orientation and length can be quickly obtained. In short, the orientation of the rotated bounding box generally can represent the orientation of fracture in the box. Even when the bounding box is close to junction points, this approach is highly reliable as it only depicts the dominant trend of fracture in the bounding box. Also, this method can show local variations in fracture orientation as a fracture is not sampled along with a fixed orientation. We illustrate fracture orientation distribution measured on a CT image in Fig. 7 using a rose diagram, where its radius represents the length of the fracture. The dominant trend of the fracture in this slice is NW-SE and the lesser concentrations trend ENE-WSW, which is similar to what we observed in the CT image.

#### 4.2. Fracture aperture

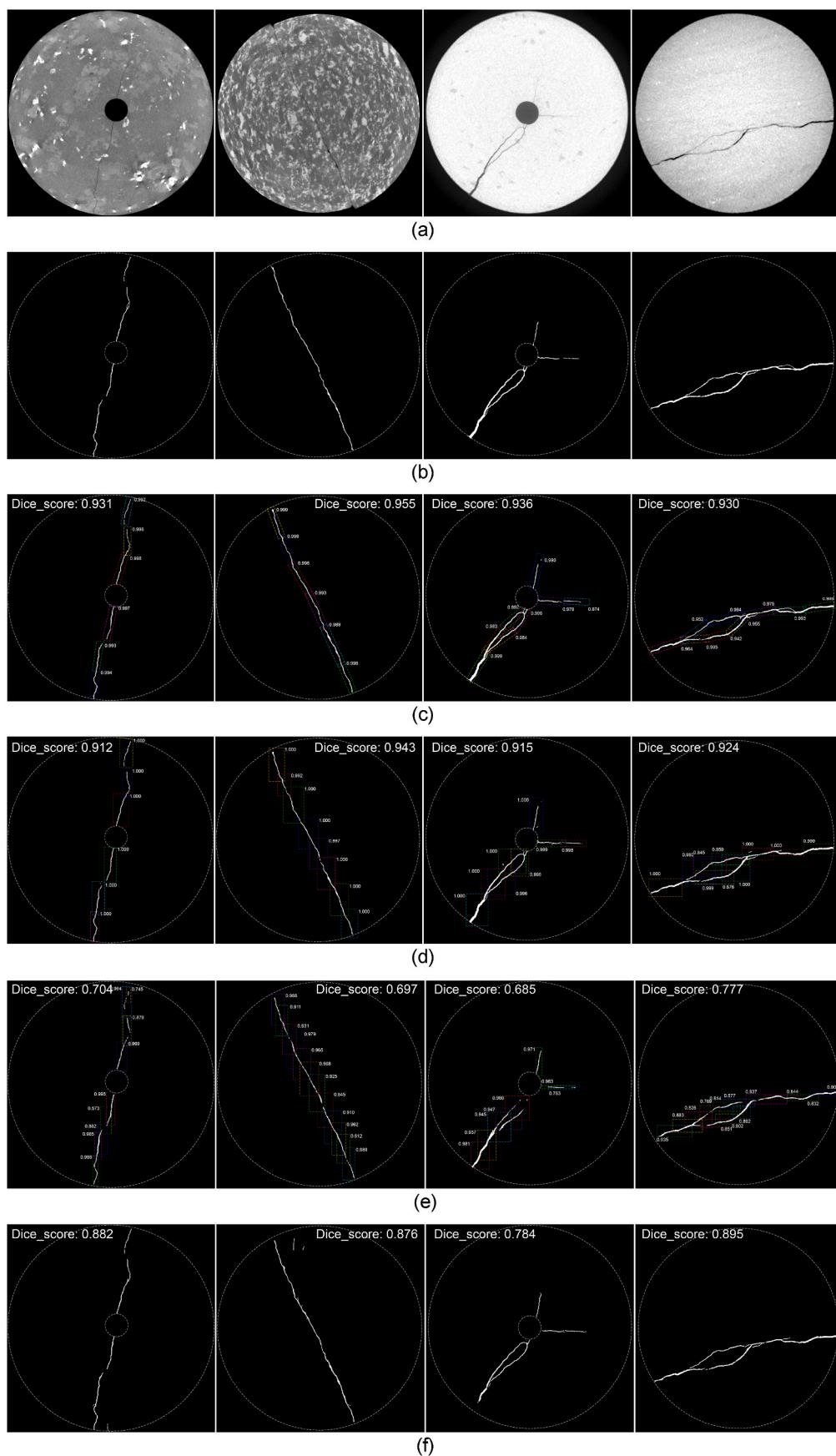
Fracture aperture is defined as the perpendicular distance between two fracture walls. This is a crucial parameter, especially for fluid flow in fractures.<sup>3,55–57</sup> In this study, we use the method given by Bear<sup>58</sup> to estimate fracture aperture from fracture segmentation results. Accordingly, the aperture at a point on the centerline of fracture is considered to be the diameter of a circle drawn around that point and expanded in size until the circle contacts both fracture walls. However, as previously mentioned, the resolution of CT images as well as the presence of artifacts may have a significant influence on the characterization of fracture apertures. As a result, the fracture aperture from the CT images is frequently calibrated with the aperture obtained from higher resolution measurements such as Scanning Electron Microscopy (SEM)<sup>18</sup> and spacers with known apertures.<sup>49</sup> In this study, an SEM image was taken at a plane in the middle of the sample to examine if there was a difference in fracture aperture obtained from those two images. Since the rock sample still appear intact as the fracture does not completely cross the sample and the process of creating thin section for SEM scanning

involves gluing the sample and then polishing, the integrity of the fracture is maintained and the thin section still accurately represents the original state of the rock sample. By avoiding disturbances to the fracture during the thin section preparation, we can obtain reliable and accurate observations of the fractured granite sample using SEM. Since the resolution of the SEM scan is much higher than CT scan, fracture aperture estimated from SEM scan can be considered as the true aperture of fracture, which can be used for validating the aperture estimated by our deep-learning approach on CT images. For consistency and direct comparison between two image modalities, the corresponding CT image is registered against SEM image. Fig. 8(a) and (b) show SEM image and corresponding CT slice, respectively. At every calibration point marked on image, the aperture is obtained.

While measurements from SEM image represent the distribution of fracture apertures ranging from 45 to 98  $\mu\text{m}$ , the CT image segmentation result shows fracture apertures ranging from 59 to 260  $\mu\text{m}$ . This indicates that CT image segmentation overestimates fracture aperture, which is due to two facts. First, the pixel-wise segmentation process cannot take sub-pixel aperture into account. For example, a fracture with an aperture slightly less or larger than a pixel may only be predicted as one pixel. Secondly, due to the two partial volume effects discussed earlier, the ground truth for training may not reflect the real aperture. The plot of CT apertures vs SEM apertures is presented in Fig. 8(d). We can see that a single fracture aperture value determined by CT image segmentation can correlate to a wide range of aperture determined by SEM image. However, a linear trend line can still support a moderately positive correlation between CT aperture and average SEM aperture with a regression coefficient  $R^2$  of 0.8803. The estimated apertures from CT data are 1.1–2.7 times greater with apertures ranging from 50 to 100  $\mu\text{m}$  in high-resolution SEM images. It is reasonable to infer that the gap will decrease if the fracture aperture becomes higher, as the partial volume effects are weakened. Furthermore, despite certain errors in fracture aperture estimation, CT scan is still widely employed for assessing fracture in rock since it is a non-destructive technique that can provide 3D internal microstructure of rock mass, whereas SEM scan require a cross-section sample, which might be relatively more expensive and time-consuming than CT scanning and only provides information about the surface of sample. Further research is needed to address the discrepancy between aperture obtained from CT image deep learning model segmentation and from SEM image.

#### 4.3. Fracture surface roughness

Another set of Gonghe granite CT data was used for assessing fracture surface roughness in Fig. 9(a). Using the ensemble approach presented in this study, a 3D model of fracture was created from CT data in Fig. 9 (b). Two surfaces of fracture are then extracted from the 3D model. Since the surface of fracture is continuous, Inverse-distance Weighting



**Fig. 6.** Examples of segmentation results using different methods: (a) CT images, (b) Ground truth, (c) Rotated Faster R-CNN + U-Net, (d) Horizontal Faster R-CNN + U-Net, (e) Mask R-CNN and (f) U-Net alone.

**Table 2**  
Fracture segmentation performance using different models.

Models		mAP@.5	Dice scores	Inference time (sec/img)
Mask R-CNN with ResNet-101 backbone		0.886	0.739	1.38
U-Net		–	0.864	0.97
Faster R-CNN + U-Net	Horizontal + ResNet-50	0.872	–	–
	Horizontal + ResNet-101	0.908	0.927	2.15
	Horizontal + ResNet-101 + Cascade	0.913	0.933	2.39
	Rotated + ResNet-50	0.865	–	–
	Rotated + ResNet-101	0.897	0.932	2.42
	Rotated + ResNet-101 + Cascade	0.905	0.942	2.73

Interpolation (IDW) is then used for filling in the gaps where our approach failed to detect the fractures (i.e., fractures totally under resolution of CT scan). This technique computes average values at the gaps by weighting those from neighboring points regarding to their distances.<sup>59</sup> Fig. 9(c) and (d) show the upper fracture surface obtained from blue light 3D scanning and CT data, respectively. Fracture roughness is quantified based on the deviation of a fracture surface profile from its best-fit plane. Li and Zhang<sup>60</sup> reviews a variety of 2D and 3D methods and parameters exist to measure roughness. In this study, we quantified the roughness of the fracture surface using the Joint Roughness Coefficient (JRC), which has been widely adopted as 1D dimensionless metric for quantifying natural fracture surface topography.<sup>61,62</sup> As the JRC values varying from fracture to fracture and with the sample size, many investigations show that the root mean square of the first derivative of the fracture surface profile (Z2) shows the best correlation with JRC.<sup>63</sup> Following the ISRM instruction<sup>64</sup>, JRC of a 3D single fracture is estimated based on the average value of 2D cross-section profiles along fracture surface. These 2D cross-section profiles were derived from fracture surface point clouds using IDW algorithm. In all, 106 profiles were retrieved with each line spaced 1 mm apart. The Z2 values are calculated based on the local slopes of each profile with intervals between measured data points equal to 0.5 mm (Eq. (14)). We have also compared the JRC values estimated by using two different cross-section line spacings of 0.5 mm and 1.5 mm, and found that the influence of this value on the JRC value is insignificant. The relationship between JRC and Z2 can be represented in Eq. (15). Although the JRC values vary slightly across profiles, as shown in Fig. 10(b), the predicted JRC of the fracture surface is  $10.83 \pm 2.24$ , which is comparable to the JRC values

( $10.12 \pm 2.18$ ) obtained from the blue light 3D scanning surface. This is quite obvious given that the spatial resolutions of CT scan (approximately  $60 \mu\text{m}$ ) and blue light 3D scanning ( $50 \mu\text{m}$ ) are comparable. However, the benefit of CT scanning over the blue light 3D scanning is that it can be utilized for closed fractures while the latter can only be used for open fractures. Closed fractures (joints) are common in deep rock core samples, and quantifying their characteristics is vital for determining rock properties as well as estimating in-situ stress.

$$Z_2 = \sqrt{\frac{1}{(N-1)(\Delta x)^2} \sum_{i=1}^{N-1} (z_{i+1} - z_i)^2} \quad (14)$$

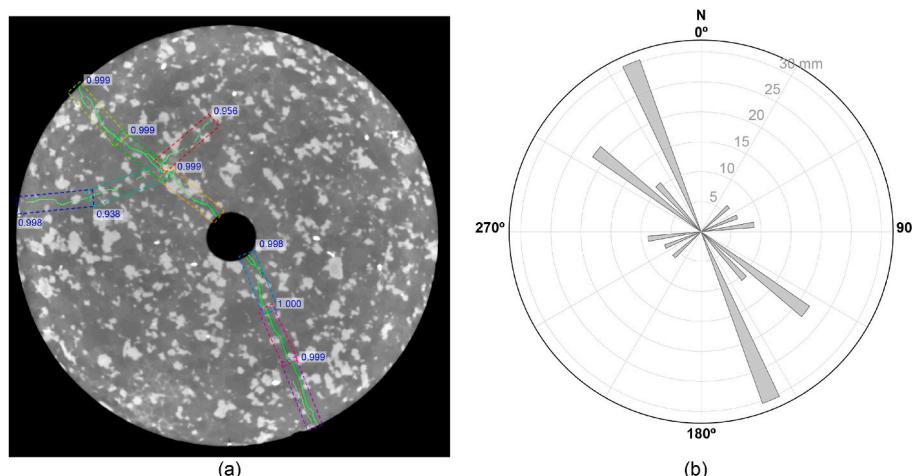
$$JRC = 61.79(Z_2) - 3.47 \quad (\Delta x = 0.5 \text{ mm}) \quad (15)$$

where  $\Delta x$  is sampling interval,  $z$  is the elevation at a sampling point, and  $N$  is the number of sampling points along a cross-section line.

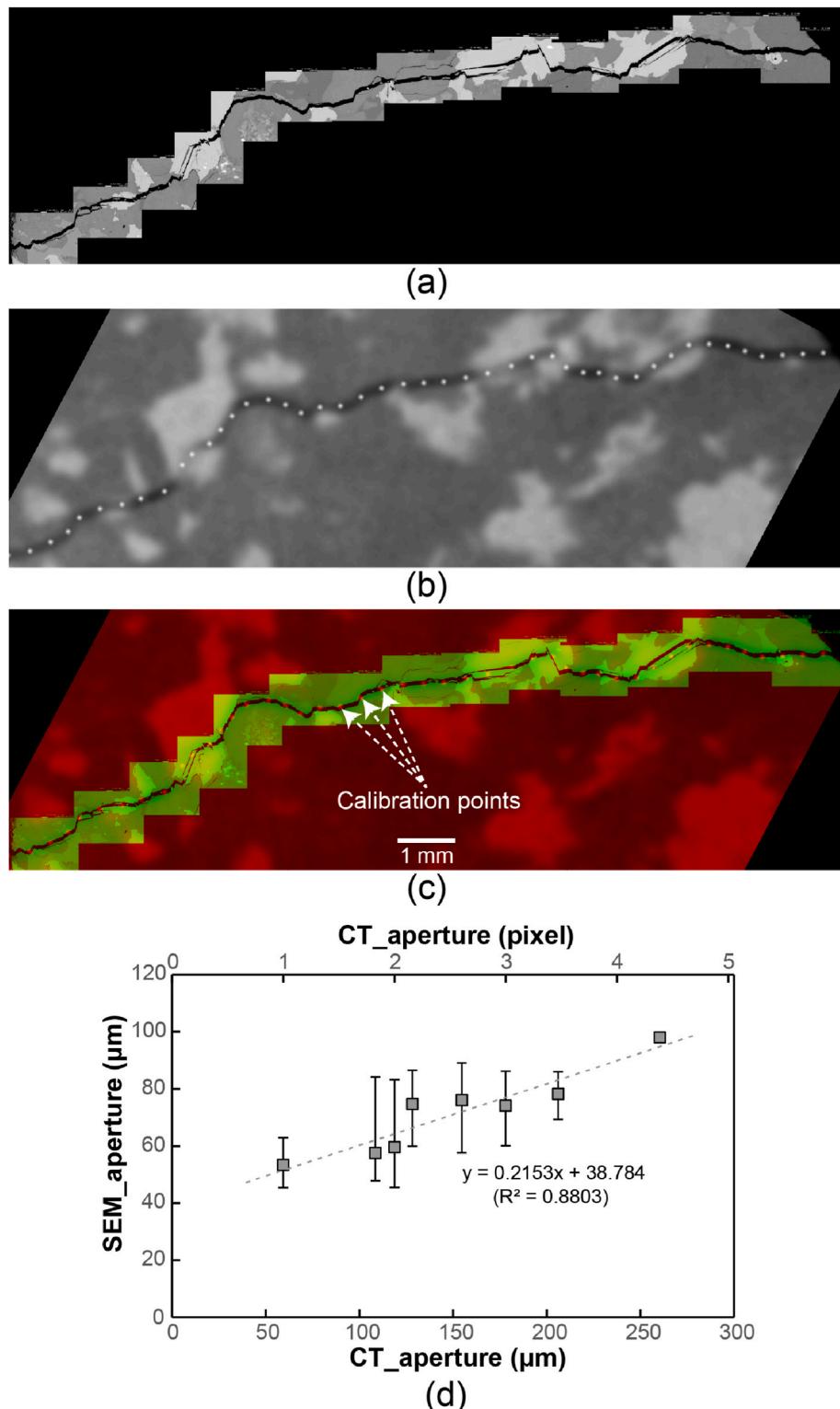
## 5. Discussions

From the previous results, we have shown that our newly developed algorithm can produce reasonable results on fracture segmentation regardless of heterogeneity in CT image and high complexity in fracture patterns. Beyond that our assemble approach has a few essential advantages over the others. Regarding the detection performance, the use of the rotated bounding box can effectively eliminate the background in bounding box, thereby setting a more robust basement for fracture segmentation. It is also worth noting that information on fracture length and orientation can be extracted from a rotated bounding box without the need for a further segmentation stage. To some extent, this approach could be generalized to other problems such as cracks or fractures detection in buildings, roads, tunnels and other structures, where not require pixel-by-pixel sophisticated segmentation, but rather considering the overall pattern such as length and orientation.

The segmentation results and the quantitative fracture characterizations suggest two findings. First, our proposed approach has shown good performance on CT images of the four different types of rock, achieving highly complete fracture segmentations with low false positive errors. Except for severe cases where fracture was completely invisible on CT scan, our approach can identify the majority of fractures in CT images, even when high-density biotite minerals causing artifacts in CT images are crossing the fractures. It can be observed that the proposed method achieves better Dice score by approximately 8%–26% improvement compared with that in a single model (i.e., U-Net alone and Mask R-CNN). This improvement might be related to addressing the



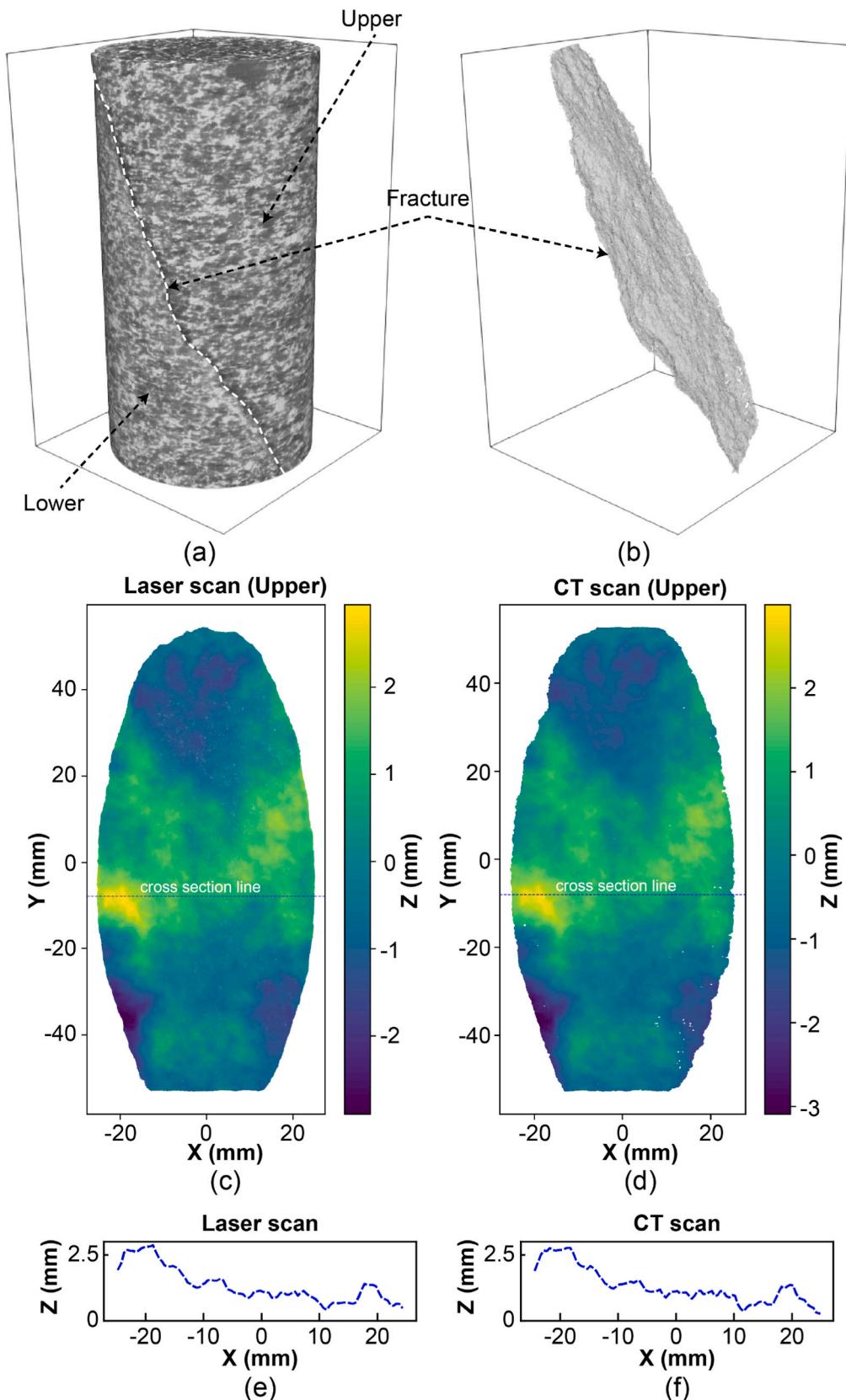
**Fig. 7.** (a) CT image and fracture network (green) on Gonghe granite obtained by using rotated Faster R-CNN + U-Net, (b) Rose diagram of the orientation and length of the detected fractures in the CT image. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)



**Fig. 8.** Fracture aperture analysis data and results of typical Gonghe granite sample (a) SEM image, (b) CT image corresponding to the same locations in (a), (c) Overlapped image (CT - red channel and SEM – green channel), and (d) Correlation between aperture measurements from CT and SEM images. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

issue of class imbalance, which occurs when the number of background pixels exceeds the number of fracture pixels, as well as model generalization due to the lack of training data. Although some background remains in bounding box, the vast majority of it has been eliminated by segmentation consideration such as mitigating class imbalance problem. Furthermore, cropping image patches by predicted bounding boxes can

help to improve model generalization. Because the aperture and grayscale intensity of a fracture vary greatly throughout its length, single models (i.e., Mask R-CNN and U-Net alone) tend to focus on identifying clear open parts rather than unclear parts. Cropping patches and feeding them to model independently allows the model to consider different fracture patterns equally. This helps our model to be better generalized



**Fig. 9.** Analysis data and results of a Gonghe granite sample. (a) 3D volume rendering of CT data, (b) 3D fracture derived from CT data using rotated Faster R-CNN + U-Net approach, (c) Upper fracture surface from laser scan, (d) Upper fracture surface from CT data using rotated Faster R-CNN + U-Net, and (e, f) Examples of two corresponding cross-section profiles derived from two surfaces.

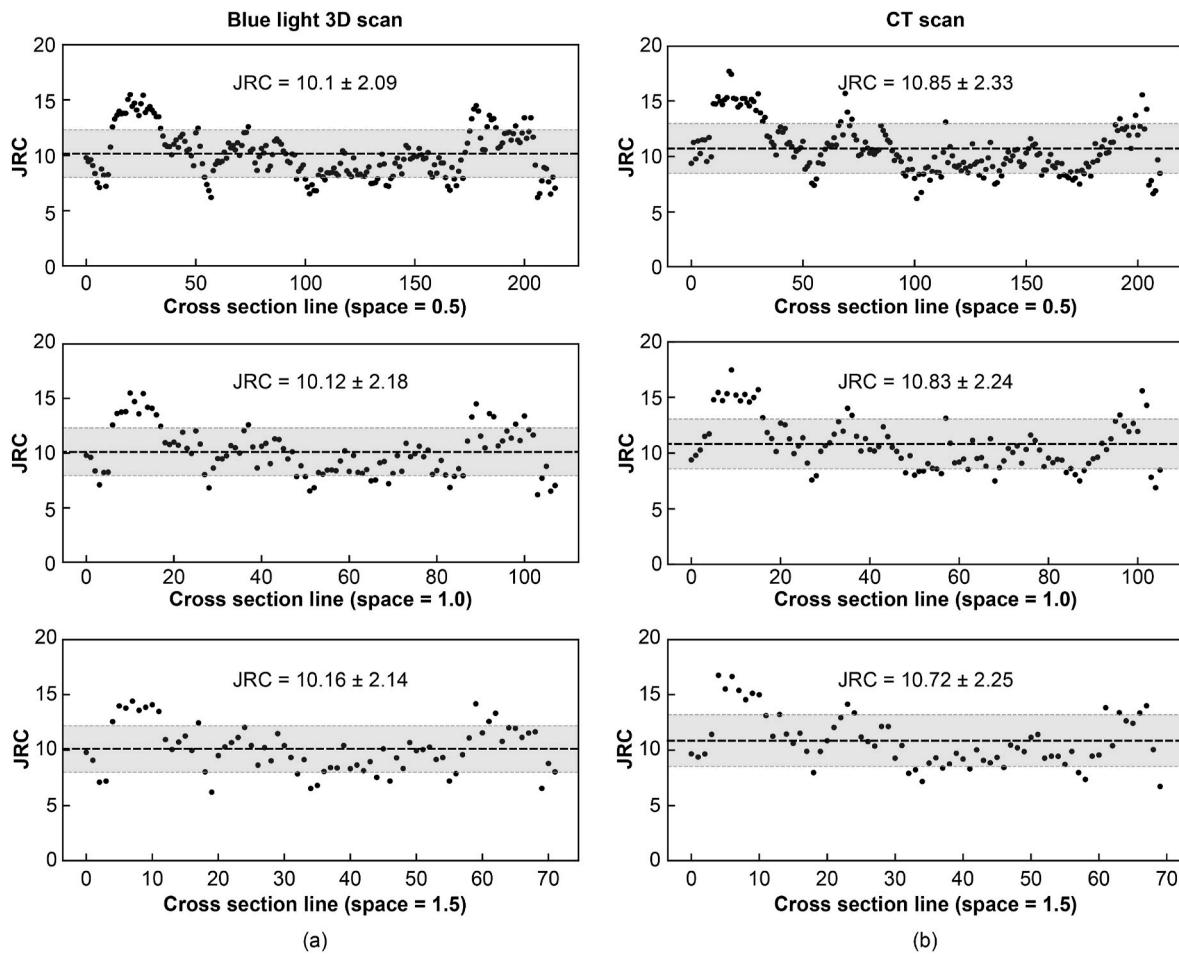


Fig. 10. JRC values estimated from fracture surface point cloud data extracted from (a) blue light 3D scan and (b) CT scan.

since the target object of interest to learn is not always the same throughout the training data. Second, deep learning model is data-driven approach and its performance heavily relies on the size of training dataset. In this study, the problem of data scarcity is unavoidable since the segmentation ground truth annotation process is too difficult and time-consuming. However, by cropping a single annotated image into multiple patches to conduct training, we can avoid expensively annotating large datasets. Therefore, with the same cost of annotation, our method outperforms Mask R-CNN and U-Net, which use the fully-annotated images for training rather than cropping into patches, resulting in less training data and the models prone to overfitting. Furthermore, for U-Net and Mask R-CNN training, squeezing images to a smaller size is unavoidable due to computing resource limitation as well as the requirement of same size image. While the resize operation causes the loss of information on the details of small and thin objects like fracture, cropping patch strategy guarantees that the whole context of the input image is preserved, which is critical for improving segmentation performance and reducing usage of the computational resources.

Although the accuracy of fracture segmentation using the proposed method has been demonstrated against ground truth, the segmentation results still need to be carefully calibrated concerning true aperture of the fracture. It would be the focus of future research to improve the accuracy of fracture segmentation within bounding boxes. We expect that increasing the quantity and quality of data used for training would improve the segmentation result. Additionally, a manner for calibrating fracture apertures precisely will be further investigated to achieve higher accuracy of automatic aperture measurement.

## 6. Conclusions

In this paper, we presented an ensemble deep learning approach to address several inherent issues in the extraction of fracture information from CT images of rocks. The approach was developed by combining two independent deep learning algorithms in function: Faster R-CNN for fracture detection and U-Net for fracture segmentation. Experiments with distinct CT images in four different types of rocks show that the ensemble deep-learning approach proposed in this study represents advances in three regards. First, rather than considering the whole image, the segmentation process is taken in operation only within the detected bounding box, which considerably alleviates the disturbance problem of complex background of fracture and reduces resource waste in computation. Particularly employing the rotated bounding box is more beneficial. Second, the images are cropped into smaller patches according to bounding boxes for training U-Net segmentation model, virtually mitigating the class imbalance problem caused by the relatively large numbers of background pixels in comparison to fracture pixels while also improving the model generalization by increasing the diversity of the input image. Finally, despite certain increases in computational complexity as well as training and inference time, our proposed approach outperformed other deep learning algorithms (i.e., Mask R-CNN and U-Net alone) with the same cost of ground truth annotation, making it a superior solution to dealing with the problem of a limited dataset. The proposed approach is applied and further validated through characterization of several engineering properties of rock fractures including orientation, length, aperture, and wall surface roughness. The results show that our proposed approach can effectively and accurately extract the fracture geometrical information from CT images, and the

measurements on fracture aperture and surface roughness are comparable with observations from other methods (i.e., SEM and blue light 3D scanning).

## Authorship statement

**Chuyen Pham:** Method development and implementation, result analysis, writing original manuscript. **Li Zhuang:** Scientific advice, manuscript review and editing. **Sun Yeom:** Data acquisition and pre-processing. **Hyu-Soung Shin:** Research planning and supervision, scientific advice, result analysis, manuscript review and editing.

## 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.

## Data availability

Data will be made available on request.

## Acknowledgments

This research was supported by the project “Development of environmental simulator and advanced construction technologies over TRL6 in extreme conditions” and the International Collaboration Program (Project No. 20230361-001) funded by KICT.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ijrmms.2023.105531>.

## References

- 1 Bishop AW. The strength of crustal materials. *Eng Geol.* 1974;8:139–153.
- 2 Jaeger JC. Friction of rocks and stability of rock slopes. *Geotechnique.* 1971;21: 97–134.
- 3 Zimmerman RW, Bodvarsson GS. Hydraulic conductivity of rock fractures. *Transport Porous Media.* 1996;23:1–30. <https://doi.org/10.1007/BF00145263>.
- 4 Zoback MD, Gorelick SM. Earthquake triggering and large-scale geologic storage of carbon dioxide. *Proc Natl Acad Sci USA.* 2012;109:10164–10168. <https://doi:10.1073/pnas.1202473109>.
- 5 Burchett TP. Carbonate rocks and petroleum reservoirs: a geological perspective from the industry. *Geological Society, London, Special Publications.* 2012;370:17–37.
- 6 Van Geet M, Swennen R, Wevers M. Towards 3D petrography: application of microfocus computed tomography in geological science. *Comput Geosci.* 2001;27: 1091–1099. [https://doi.org/10.1016/S0098-3004\(00\)00154-0](https://doi.org/10.1016/S0098-3004(00)00154-0).
- 7 Ketcham RA, Carlson WD. Acquisition, optimization and interpretation of X-ray computed tomographic imagery: applications to the geosciences. *Comput Geosci.* 2001;27:381–400. [https://doi.org/10.1016/S0098-3004\(00\)00116-3](https://doi.org/10.1016/S0098-3004(00)00116-3).
- 8 Andrà H, Combaret N, Dvorkin J, et al. Digital rock physics benchmarks—Part I: imaging and segmentation. *Comput Geosci.* 2013;50:25–32.
- 9 Wennberg OP, Renman L, Basquet R. Computed tomography scan imaging of natural open fractures in a porous rock: geometry and fluid flow. *Geophys Prospect.* 2009;57: 239–249. <https://doi.org/10.1111/j.1365-2478.2009.00784.x>.
- 10 Kyle JR, Ketcham RA. Application of high resolution X-ray computed tomography to mineral deposit origin, evaluation, and processing. *Ore Geol Rev.* 2015;65:821–839. <https://doi.org/10.1016/j.oregeorev.2014.09.034>.
- 11 Schmitt M, Halisch M, Müller C, Fernandes CP. Classification and quantification of pore shapes in sandstone reservoir rocks with 3-D X-ray microcomputed tomography. *Solid Earth.* 2016;7:285–300. <https://doi:10.5194/se-7-3441-2015>.
- 12 Johnn RA, Steude JS, Castanier LM, Roberts PV. Nondestructive measurements of fracture aperture in crystalline rock cores using X-ray computed tomography. *J Geophys Res.* 1993;98:1889–1900. <https://doi.org/10.1029/92JB02298>.
- 13 Keller A. High resolution, non-destructive measurement and characterization of fracture apertures. *Int J Rock Mech Min Sci.* 1998;35:1037–1050. [https://doi.org/10.1016/S0148-9062\(98\)00164-8](https://doi.org/10.1016/S0148-9062(98)00164-8).
- 14 Van Geet M, Swennen R. Quantitative 3D-fracture analysis by means of microfocus X-Ray Computer Tomography ( $\mu$ CT): an example from coal. *Geophys Res Lett.* 2001; 28(17):3333–3336.
- 15 Vandersteen K, Busselen B, Van Den Abeele K, Carmeliet J. Quantitative characterization of fracture apertures using microfocus computed tomography. *Geol Soc, London, Spec. Publ.* 2003;215:61–68. <https://doi.org/10.1144/GSL.SP.2003.215.01.06>.
- 16 Wildenschild D, Hopmans JW, Vaz CMP, Rivers ML, Rikard D, Christensen BSB. Using X-ray computed tomography in hydrology: systems, resolutions, and limitations. *J Hydrol.* 2002;267:285–297.
- 17 Iassonov P, Gebrenegus T, Tuller M. Segmentation of X-ray computed tomography images of porous materials: a crucial step for characterization and quantitative analysis of pore structures. *Water Resour Res.* 2009;45(9):1–12.
- 18 Ramandi HL, Mostaghimi P, Armstrong RT, Saadatfar M, Pinczewski WV. Porosity and permeability characterization of coal: a micro-computed tomography study. *Int J Coal Geol.* 2016;154:57–68.
- 19 McBeck J, Cordonnier B, Renard F. The influence of spatial resolution and noise on fracture network properties calculated from X-ray microtomography data. *Int J Rock Mech Min Sci.* 2021;147:104922. <https://doi:10.1016/j.ijrmms.2021.104922>.
- 20 Voorn M, Exner U, Rath A. Multiscale Hessian fracture filtering for the enhancement and segmentation of narrow fractures in 3D image data. *Comput Geosci.* 2013;57: 44–53. <https://doi.org/10.1016/j.cageo.2013.03.006>.
- 21 Deng H, Fitts JP, Peters CA. Quantifying fracture geometry with X-ray tomography: technique of iterative local thresholding (TILT) for 3D image segmentation. *Comput Geosci.* 2016;20(1):231–244.
- 22 Chauhan S, Rühaak W, Anbergen H, et al. Phase segmentation of X-ray computer tomography rock images using machine learning techniques: an accuracy and performance study. *Solid Earth.* 2016;7:1125–1139.
- 23 Guntoro PI, Ghorbani Y, Koch PH, Rosenkranz K. X-ray microcomputed tomography ( $\mu$ CT) for mineral characterization: a review of data analysis methods. *Minerals.* 2019;9(3):183. <https://doi.org/10.3390/min9030183>.
- 24 Lee D, Karadimitriou N, Ruf M, Steeb H. *Detecting Micro Fractures with X-Ray Computed Tomography.* 2021. arXiv: arXiv:2103.12821.
- 25 Karimpouli S, Tahmasebi P, Saenger EH. Coal cleat/fracture segmentation using convolutional neural networks. *Nat Resour Res.* 2020;29:1675–1685.
- 26 Chen JY, Zhou ML, Huang HW, Zhang DM, Peng ZC. Automated extraction and evaluation of fracture trace maps from rock tunnel face images via deep learning. *Int J Rock Mech Min Sci.* 2021;142, 104745. <https://doi.org/10.1016/j.ijrmms.2021.104745>.
- 27 Byun H, Kim J, Yoon D, Kang IS. A deep convolutional neural network for rock fracture image segmentation. *Earth Sci Inform.* 2021;14:1937–1951. <https://doi.org/10.1007/s12145-021-00650-1>.
- 28 Liu Y, Yeoh JKW. Robust pixel-wise concrete crack segmentation and properties retrieval using image patches. *Autom Constr.* 2021;123, 103535. <https://doi.org/10.1016/j.autcon.2020.103535>. ISSN 0926-5805.
- 29 König J, Jenkins M, Mannion M, Barrie P, Morrison G. Optimized deep encoder-decoder methods for crack segmentation. *Digit Signal Process.* 2021;108, 102907. <https://doi.org/10.1016/j.dsp.2020.102907>.
- 30 Ronneberger O, Fischer P, Brox T. In: *U-Net: Convolutional Networks for Biomedical Image Segmentation.* vol. 9351. MICCAI, Springer, LNCS; 2015:234–241.
- 31 Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans Pattern Anal Mach Intell.* 2017;39(12):2481–2495. <https://doi.org/10.1109/TPAMI.2016.2644615>.
- 32 He K, Gkioxari G, Dollár P, Girshick RB. Mask R-CNN. *CoRR.* 2017.
- 33 Ren S, He K, Girshick R, Sun J. Faster R-CNN. Towards real-time object detection with region proposal networks. In: *NIPS.* 2015.
- 34 Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: unified, real-time object detection. In: *CVPR.* 2016.
- 35 Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector. In: *European Conference on Computer Vision.* Springer; 2016:21–37. [https://doi.org/10.1007/978-3-319-46448-0\\_2](https://doi.org/10.1007/978-3-319-46448-0_2).
- 36 Goodfellow I, Bengio Y, Courville A. *Deep Learning.* MIT Press; 2016. <http://www.deeplearningbook.org>.
- 37 Lin T, Dollár P, Girshick RB, He K, Hariharan B, Belongie SJ. Feature pyramid networks for object detection. *CVPR.* 2017.
- 38 Jiang Y, Zhu X, Wang X, et al. R2CNN: Rotational Region CNN for Orientation Robust Scene Text Detection. 2017. arXiv preprint arXiv:1706.09579.
- 39 Ding J, Xue N, Long Y, Xia GS, Lu Q. Learning ROI transformer for oriented object detection in aerial images. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 2019:2849–2858.
- 40 Vu TX, Jang H, Pham TX, Yoo CD. Cascade RPN: delving into high-quality region proposal network with adaptive convolution. In: *Proc. NIPS.* 2019:1–11.
- 41 Cai Z, Vasconcelos N. In: *Cascade R-CNN: Delving into High Quality Object Detection.* CVPR; 2018.
- 42 Girshick R. Fast R-CNN. In: *Proceedings of the International Conference on Computer Vision IEEE.* 2015:1440–1448. <https://doi.org/10.1109/iccv.2015.169>.
- 43 Oktay O, Schlemper J, Le Folgoc L, et al. Attention U-Net: Learning where to Look for the Pancreas. 2018. arXiv [Preprint]. arXiv: 1804.03999.
- 44 Sudre CH, Li W, Vercauteren TK, Ourselin S, Cardoso MJ. Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support.* Cham: Springer; 2017:240–248.
- 45 Zhuang L, Kim KY, Jung SG, Diaz M, Min KB. Effect of water infiltration, injection rate and anisotropy on hydraulic fracturing behavior of granite. *Rock Mech Rock Eng.* 2019;52. <https://doi.org/10.1007/s00603-018-1431-3>.
- 46 Zhuang L, Kim K, Diaz M, Sun Y. Evaluation of water saturation effect on mechanical properties and hydraulic fracturing behavior of granite. *Int J Rock Mech Min Sci.* 2020;130, 104321. <https://doi.org/10.1016/j.ijrmms.2020.104321>.
- 47 Ji Y, Zhuang L, Wu W, Hofmann H, Zang A, Zimmermann G. Cyclic water injection potentially mitigates seismic risks by promoting slow and stable slip of a natural

- fracture in granite. *Rock Mech Rock Eng.* 2021;1–17. <https://doi.org/10.1007/s00603-021-02438-7>.
- 48 Luo J, Ramos M, Espinoza D. Fracture patterns in laminated Mancos shale. *Digital Rocks Portal*; 2019 [Online]. Available: <http://www.digitalrockspotal.org>.
- 49 Ketcham RA, Slottke DT, Sharp JM. Three-dimensional measurement of fractures in heterogeneous materials using high-resolution X-ray computed tomography. *Geosphere*. 2010;6(5):499–514. <https://doi.org/10.1130/GES00552.1>.
- 50 Arganda-Carreras I, Kaynig V, Rueden C, et al. Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification. *Bioinformatics*. 2017;33(15):2424–2426.
- 51 Abadi M, Agarwal A, Barham P, et al. TensorFlow: large-scale machine learning on heterogeneous systems. <http://tensorflow.org/>; 2015.
- 52 He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification. *IEEE Int. Conf. Comput. Vis.*. 2015; 1026–1034 <https://doi.org/10.1109/ICCV.2015.123>.
- 53 Lin TY, Maire M, Belongie S, et al. Microsoft COCO: common objects in context. <http://arxiv.org/abs/1405.0312>; May 2014.
- 54 Jing Y, Armstrong RT, Ramandi HL, Mostaghimi P. Coal cleat reconstruction using micro-computed tomography imaging. *Fuel*. 2016;181:286–299.
- 55 Snow DT. Rock fracture spacings, openings and porosities. *J. Soil Mech. Found Div, Proc. ASCE*. 1968;94:73–91.
- 56 Witherspoon PA. Investigation at Berkeley on fracture flow in rocks: from the parallel plate model to chaotic systems. In: Faybishenko B, Benson S, Witherspoon PA, eds. *Dynamics of Fluid in Fractured Rocks*. Washington, DC, USA: American Geo-physical Union; 2000:1–58.
- 57 Brush DJ, Thomson ER. Fluid flow in synthetic rough walled fracture: Navier–Stokes, Stokes, and local cubic law simulations. *Water Resour Res.* 2003;39(4):1085. <https://doi.org/10.1029/2002WR001346>.
- 58 Bear J. Modeling flow and contaminant transport in fractured rocks. In: Bear J, Tsang CF, de Marsily G, eds. *Flow and Contaminant Transport in Fractured Rock*. San Diego, Calif: Academic; 1993:1–37.
- 59 Isaaks EH, Srivastava RM. *An Introduction to Applied Geostatistics*. New York: Oxford University Press; 1990.
- 60 Li Y, Zhang Y. Quantitative estimation of joint roughness coefficient using statistical parameters. *Int J Rock Mech Min Sci.* 2015;77:27–35.
- 61 Barton N, Choubey V. The shear strength of rock joints in theory and practice. *Rock Mech Rock Eng.* 1977;10:1–54. <https://doi.org/10.1007/BF01261801>.
- 62 Barton N, Bandis S, Bakhtar K. Strength, deformation and conductivity coupling of rock fracture. *Int J Rock Mech Min Sci.* 1985;22:121–140.
- 63 Tse R, Cruden DM. Estimating joint roughness coefficients. *Int J Rock Mech Min Sci Geomech Abstracts*. 1979;16(5):303–307.
- 64 ISRM. Suggested Methods for rock characterization testing and monitoring. In: Brown ET, ed. *Commission on Testing Methods*. Oxford: Pergamon Press; 1981.