Shahrood University of Technology

Journal of Mining and Environment (JME)

Journal homepage: www.jme.shahroodut.ac.ir

Iranian Society of Mining Engineering (IRSME)

Identification of Rock Discontinuities with the use of the UAV Image Processing and Deep Learning

Mostafa Rahimiyan[1], Mohammad Ataei[1]*, Reza Kakaie[1], and Hossein Khosravi[2]

1. Faculty of Mining, Petroleum & Geophysics Eng., Shahrood University of Technology, Shahrood, Iran
2. Faculty of Electrical Engineering, Shahrood University of Technology, Shahrood, Iran

Article Info

Received 27 September 2025
Received in Revised form 7 November 2025
Accepted 24 November 2026
Published online 24 November 2026

DOI: 10.22044/jme.2025.16898.3316

Keywords

Rock Discontinuities
UAV
Image Processing
Deep Learning

Abstract

The identification of rock discontinuities is a critical factor in mining and construction projects. Traditional methods for conducting this task are often difficult, time-consuming, pose risks to human safety, and lead to incomplete evaluations. The introduction of unmanned aerial vehicles (UAVs) has transformed this process, allowing coverage of the entire area in a short time without endangering personnel. The aim of this paper is to employ deep learning using the Python programming language to develop and train a neural network based on the UNET++ architecture to automatically identify rock surface discontinuities from UAV-captured imagery. It also addresses challenges associated with supervised learning, particularly overfitting, by implementing data augmentation techniques and reducing model parameters by approximately 6%. Consequently, the pixel-wise precision criterion improved significantly from 53.27% to 75.6%. This work stands out from other studies by focusing specifically on UAV imagery for geological assessments, employing a dual strategy to overcome overfitting, and demonstrating effective performance despite limited training data. The results show that the model is capable of accurately identifying rock discontinuities and is a suitable method for the mining and construction industries.

1. Introduction

Prior to the initiation of any mining or civil engineering project involving earthworks, it is necssary to carry out comprehensive geotechnical studies in order to evaluate ground conditions and identify critical factors influencing resource estimation and project feasibility. These investigations, while time-consuming and costly, are indispensable. Neglecting them may lead to severe consequences, including budget overruns, schedule delays, redesign the project, or even altered the project economically unfeasible [1]. Discontinuity characterization is one of the most critical components in both pre-construction and operational phases, forming the basis for key design parameters in geotechnical and mining projects. Conventional methods such as linear sampling and window mapping are typically recommended for collecting geological discontinuity data in rock masses. However, these techniques present significant challenges: they are labor-intensive, potentially hazardous, and their results may vary depending on the expert's knowledge [2]. Consequently, they are rarely implemented systematically in mining operations.

The development of photogrammetric drone technology has revolutionized discontinuity studies by facilitating rapid aerial image acquisition across all ground, including unreachable areas. This approach generates large datasets, needs to apply artificial intelligence for data analysis. The integration of AI and UAV technology offers a precise, rapid, repeatable, and scalable solution while eliminating the risks and limitations associated with manual data collection. Drones not only accelerate aerial surveys but also provide safe access to rugged or hazardous terrains.

Historically, numerous efforts have been made to develop automated systems for rock discontinuity identification. McCarter applied photogrammetry for geotechnical data acquisition for the first time [3]. Ressini used the numerical methods to identify

Corresponding author: ataei@shahroodut.ac.ir (M. Ataei)

discontinuities in bivariate functions [4]. Lemy employed the Canny edge detection algorithm followed by dilation filters to connect adjacent edges, with final classification performed by artificial neural networks [5]. Knapen advanced automation by employing terrestrial laser scanners (TLS) for discontinuity mapping. After point cloud acquisition and triangulation, surface normals were computed and used as discontinuity poles. These normals were then processed through a fuzzy clustering system to classify discontinuity sets. However, TLS systems lack optical capabilities, preventing discontinuity-type discrimination [6].

With development of digital imaging, image processing techniques such as Wiener filtering and Canny edge detection became a standard method for discontinuity analysis [7]. A common preprocessing step (grayscale conversion) reduces color image dimensionality but compromises accuracy. To address this, Zizi Pi proposed a color gradient distribution (CGD) method that exploits chromatic contrast at discontinuity boundaries. Comparative analyses demonstrated CGD's superiority over grayscale/binary methods in both accuracy and computational efficiency [8]. In this paper, an image processing algorithm was developed to automatically identify rock discontinuities in drone-captured imagery from mining sites using deep learning method (implemented in python).

2. Evolution of UAV Photogrammetry

The application of drones in geomatics traces back to Prizbilla and Wester-Ebbinghaus [9]. Until 2008, data collection relied on terrestrial, manned aircraft, or satellite systems. Subsequent years saw the proliferation of UAV-based close-range photogrammetry [10, 11].

Vasuki (2013) utilized an Oktokopter to image a 200 × 250 m area from 35 m altitude. After orthomosaic generation, geological features were identified via edge detection using phase-matching algorithms. Following noise suppression, pixels with values above a set threshold were classified as edges and grouped into lines. Finally, discontinuity planes were calculated using the Fisher-Boles algorithm [12].

Kong et al. combined UAV photogrammetry with structure-from-motion (SfM) to create high-fidelity 3D rock mass models, demonstrating UAVs' superiority over traditional mapping in hazardous environments [13]. Pulla et al. further advanced this paradigm by integrating UAV-derived 3D models with machine learning for automated discontinuity classification in the Acoculco Caldera, showcasing the transformative potential of UAV-AI synergy in geotechnical applications [14].

3. Recent Developments in UAV–AI Integration

More recent advances have focused on integrating UAV photogrammetry with deep learning for automatic detection and segmentation of geological discontinuities. Kumar developed an automatic system for surface characterization in coal mining environments using satellite-based pixel-level classification, demonstrating the capability of neural networks for geological mapping in data-limited conditions [15]. Yalcin et al. applied transfer learning to UAV imagery for rock mass discontinuity identification, showing that pretrained CNNs can effectively mitigate data scarcity when site-specific fine-tuning is performed [16]. Huashan et al. proposed an enhanced RDC-UNet incorporating residual and attention modules, achieving higher mean Intersection over Union (IoU) scores (~70%) on mining-area datasets and showing that architectural refinement can improve boundary delineation [17]. Similarly, Li et al. integrated edge-feature fusion and multi-level upsampling into a DeepLabv3+ backbone (EMNet), improving segmentation accuracy on UAV-based remote-sensing images [18].

Wang et al. developed an automated CNN framework for detecting mining-induced fissures from UAV photogrammetry, emphasizing operational applicability in large-scale field conditions [19]. Multi-view image augmentation has also been introduced to enhance robustness and reduce overfitting, as demonstrated by Yalcin et al. , who showed that geometrically diversified UAV datasets significantly improve discontinuity detection accuracy [20]. Collectively, these studies highlight the rapid progress of UAV–AI integration in geotechnical applications, while also revealing remaining challenges in data diversity, generalization, and boundary precision—issues that the present study addresses through an optimized UNet++ architecture and targeted data-augmentation strategy under data-scarce conditions.

4. Operations Performed

In this paper, a model was developed to identify automatically rock mass discontinuities in images, including drone-captured imagery by using a supervised machine learning. Figure 1 shows the process of creating and utilizing this model.

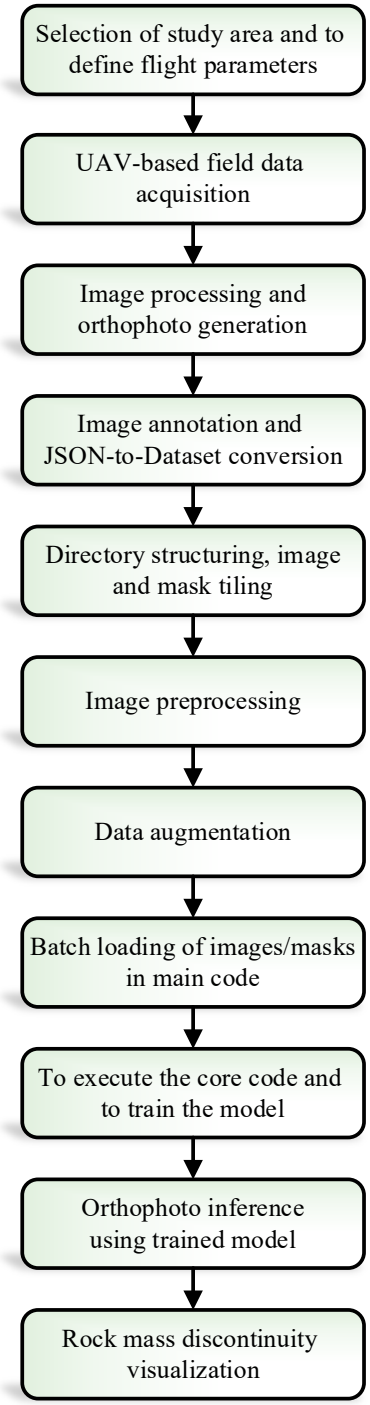


Figure 1. The process of developing a model for automatic detection of rock surface discontinuities in images

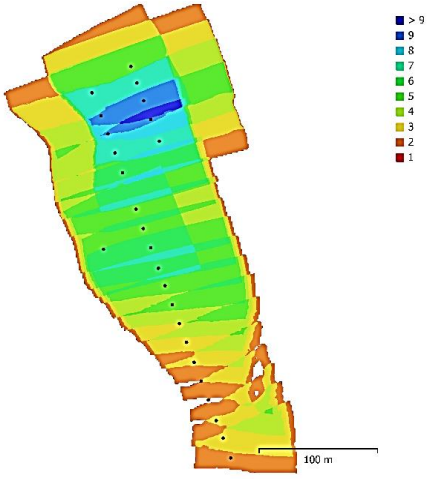
5. Database Preparation

The images for this study were collected using a Phantom 4 Pro drone from two locations. The rocks in both areas are Cretaceous limestones. Both sites are located in Semnan province in the northen part of Iran. The first area is near Shahroud city, and the second area is located in northern Semnan (Figure 2).



Figure 2. Location of the survey sites within the study region.

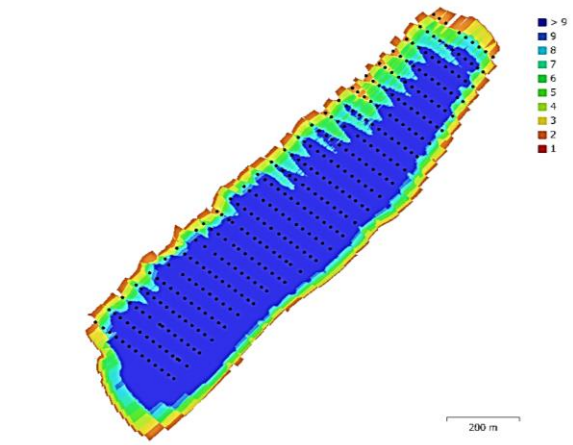
The surveys were conducted without using LiDAR technology, therefore the images were not obtained with uniform ground resolution. However, after image analysis using Metashape software, the ground resolution was calculated to be 2.21 cm/pixel for the first survey and 3.18 cm/pixel for the second survey (Figure 3 and 4 respectively). This measurement represents the average dimensions shown by each pixel on the ground.



Number of images:	27	Camera stations:	25
Flying altitude:	85.5 m	Tie points:	30,040
Ground resolution:	2.21 cm/pix	Projections:	86,679
Coverage area:	0.0495 km²	Reprojection error:	0.476 pix

Camera Model	Resolution	Focal Length	Pixel Size	Precalibrated
FC6310 (8.8mm)	5472 x 3648	8.8 mm	2.41 x 2.41 µm	No

Figure 3. Camera locations, image overlap and other specifications for the first-phase aerial survey imagery.



Number of images:	329	Camera stations:	329
Flying altitude:	115 m	Tie points:	6,487,170
Ground resolution:	3.18 cm/pix	Projections:	21,915,763
Coverage area:	0.525 km²	Reprojection error:	0.72 pix

Camera Model	Resolution	Focal Length	Pixel Size	Precalibrated
FC6310 (8.8mm)	5472 x 3648	8.8 mm	2.41 x 2.41 µm	No

Figure 4. Camera locations, image overlap and technical specifications for second-phase aerial survey imagery

The images show rock surfaces, predominantly exposed outcrops with joints and fractures of varying orientations and dimensions. These images were captured using a DJI Phantom 4 Pro UAV. Camera specifications and image acquisition parameters have presented in Table 1.

The smallest identifiable object dimensions in the images are calculated using Equation (1):

$$Sz = Npx \times \mathrm{GSD} \quad (1)$$

Where:

Sz - Minimum detectable feature size (cm)

Npx - Constant value (2-5) representing the number of pixels required for object identification

GSD - Ground Sampling Distance (cm/pixel) [21].

Based on this equation and average GSD values, the expected detectable discontinuity thickness ranges are:

- 4-11 cm for Phase 1 imagery
- 6-16 cm for Phase 2 imagery

6. Image Processing and Model Development

The process of developing the image processing model using supervised learning is described below. The first step in the supervised learning involves creating a suitable database or dataset. These data must be divided into three groups. training dataset, test dataset and validation dataset [22].

6.1. Data Preparation Process

a) In the initial stage, a random selection of UAV-captured images was manually annotated one-by-one to label them (Figure 5).

b) After labeling and format adjustment, the images and their corresponding masks (Figure 6) were saved to the computer's hard drive as the input database for the neural network.

c) The images in the database were divided as training, validation, and test sets.

Table 1. Camera Specifications and Image Acquisition Parameters

Parameter	Value
Camera Make / Model	DJI FC6310
Image Resolution	5472 × 3648 pixels (≈ 20 MP)
Focal Length	8.8 mm (35 mm equivalent: 24.0 mm)
Aperture (F Number)	f/6.3
Shutter Speed	1/640 s
ISO	100
Color Space	sRGB
White Balance	Auto
Metering Mode	Center-weighted average
Image Format	JPEG
Capture Date/Time	2023-06-10 13:48:43
GPS Coordinates	35°45′31.18″ N, 53°22′34.06″ E
GPS Altitude	2119.6 m above sea level
Relative Altitude	+119.80 m
Camera Pitch / Yaw / Roll	-90.00° / -54.20° / +0.00°
Drone Pitch / Yaw / Roll	-7.50° / -53.60° / -6.40°

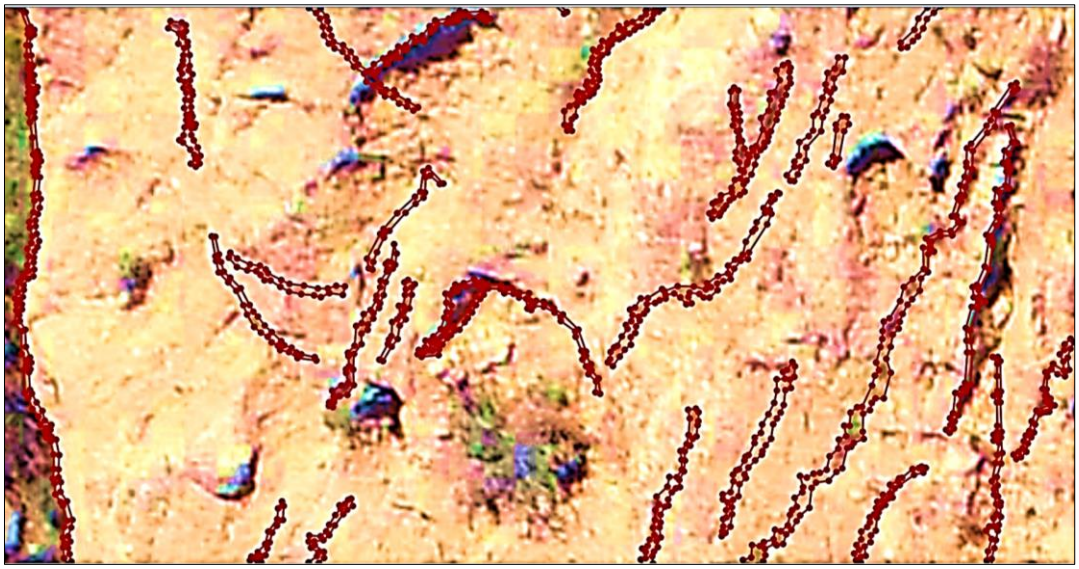
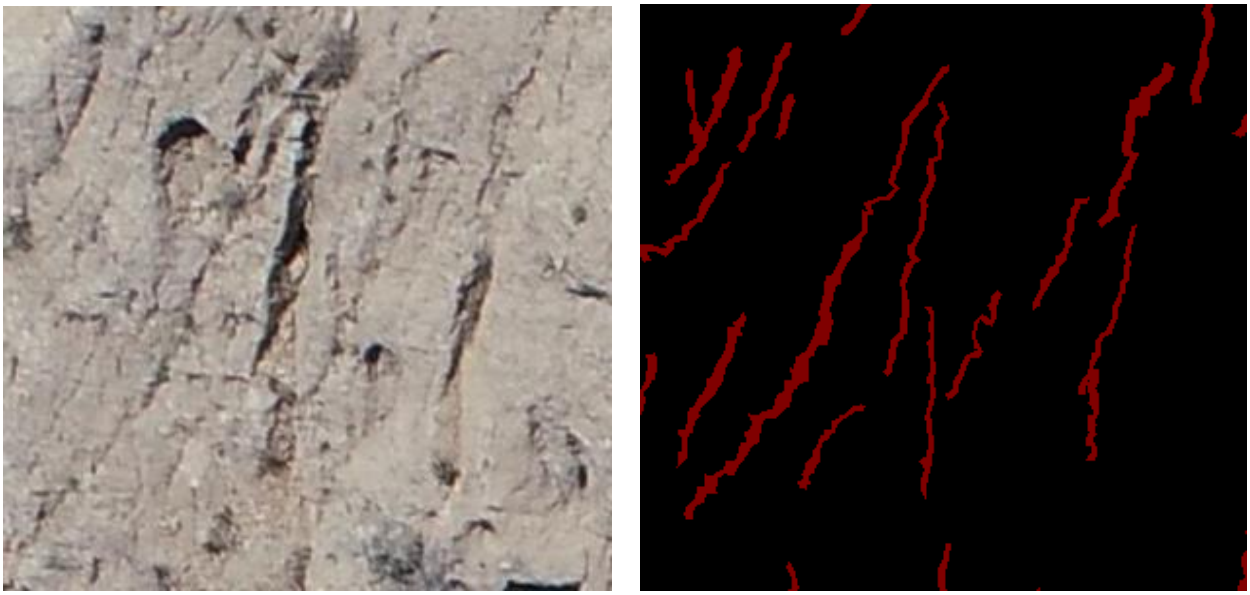


Figure 5. Labeled image showing discontinuity annotations



(a) **(b)**

Figure 6. a) Sample image; b) Corresponding mask

6.2. Network Architecture

This study employed the UNet++ architecture, a variant of CNN architectures. UNet++ was first introduced in 2018 by Zongwei Zhou et al. [23]. According to the UNet architecture, UNet++ features a U-shaped structure comprising three main components:

- Encoder (Contraction path)
- Decoder (Expansion path)
- Bottleneck layer

The encoder compresses input images, reducing their spatial dimensions while increasing depth. The decoder then reconstructs the images to their original dimensions [24]. UNet++ begins with an encoder backbone followed by the decoder network. The key distinction of UNet++ from UNet lies in its redesigned skip connections [23].

6.3. Implementation Details

The network architecture comprises: (a) an encoder (Figure 7) where input images (320×320×3) undergo five spatial reductions through ResNet34 blocks (E0:7×7 conv→MaxPool→E1-E4: residual blocks), producing feature maps at 160×160 (64ch), 80×80 (64ch), 40×40 (128ch), 20×20 (256ch), and 10×10 (512ch); and (b) a decoder (Figure 8) with four upsampling blocks (D0-D3) that concatenate encoder features through dual 3×3 convolutions (160→80→40→20 ch) before final expansion to 320×320×10.

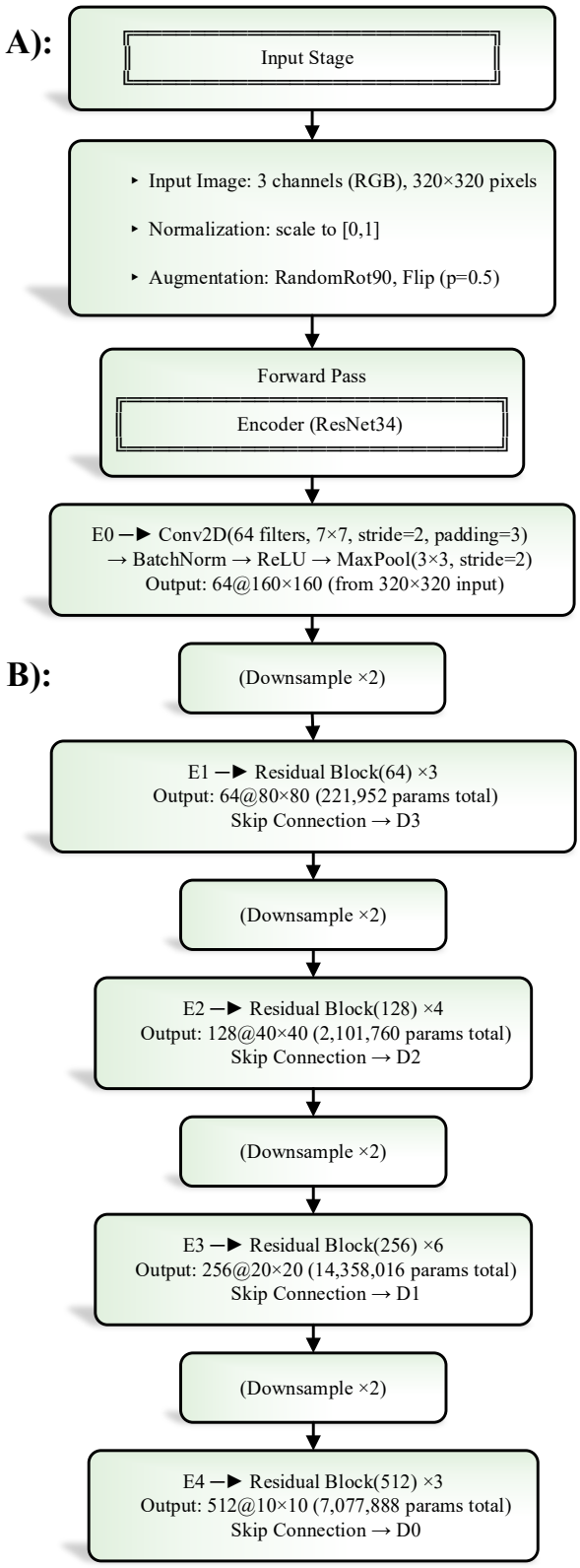


Figure 7. Encoder pathway architecture - The ResNet34-based encoder processes 320×320×3 input through successive downsampling blocks. (a) Initial block produces (b) Four residual stages

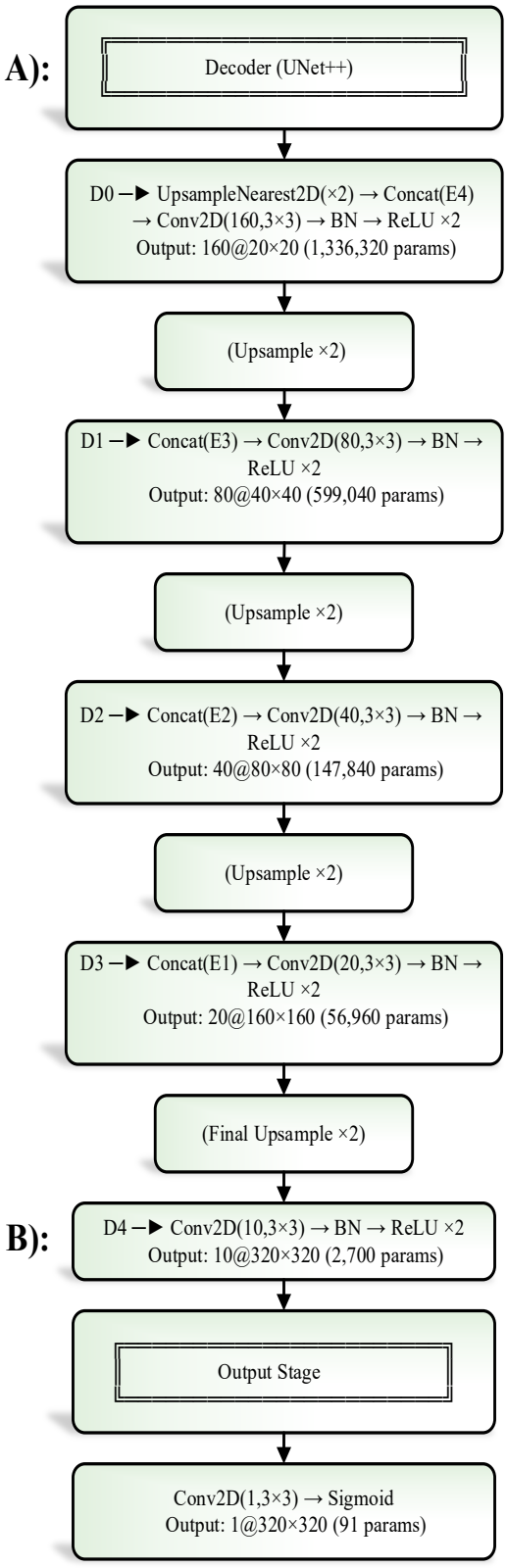


Figure 8. UNet++ decoder pathway - The decoder performs progressive upsampling with skip connections: a) D0-D3 blocks upsample and concatenate encoder features (E4-E1 respectively) through dual; b) Final D4 block expands to 320×320×10 before 1×1 convolution with sigmoid activation function.

7. Key steps and operations

The training and evaluation process begin with dataset preparation, which involved loading both the training and validation datasets and utilizing a DataLoader to manage data batches efficiently for each phase. The model architecture was based on the U-Net++ framework, implemented through a dedicated class. This configuration included a ResNet encoder with a depth of five layers, a series of progressively smaller decoder channels, a binary segmentation activation function, and a specific loss function for performance optimization.

During the forward pass, the input images were processed by the model, and the outputs were compared with the ground truth masks using the defined loss function. Training and backpropagation followed, where the optimizer updated the model weights based on the computed gradients, while training and validation losses were recorded for each epoch.

For visualization and reporting, training and validation losses were plotted to facilitate performance analysis, and the model checkpoints were saved at the end of every epoch. The overall workflow for model training and evaluation is illustrated in Figure 9. After the model development and initial setup which included code implementation, debugging, and finalization, various outputs were generated by running the model on the dataset.

8. Initial Training Phase

The model was first trained without data augmentation to establish baseline performance. Training and validation loss curves, obtained using the UNet++ architecture, are presented in Figures 10 and 11, demonstrating the model's convergence behavior.

As can be seen from Figures 11 and 12 indicates that the training process initially converges, but after sixteen epochs, no significant progress is observed. The decrease in the loss value in the training curve, despite the lack of improvement and reduction in the validation loss, suggests overfitting of the network weights. This not only fails to enhance the model performance but also impairs its semantic recognition capability.

The primary causes of overfitting are insufficient training data and model complexity. The first step to address this issue is increasing the dataset size. However, since generating labeled training data (database construction) is highly time-consuming and expensive. Deep neural networks typically require at least around a thousand images for training [25]. Accordingly, reducing the number of images below this threshold leads to decreased accuracy.

To tackle this challenge, data augmentation techniques were employed. Given the nature of the data in this study, horizontal flip, rotation, and vertical flip were selected among various augmentation methods.

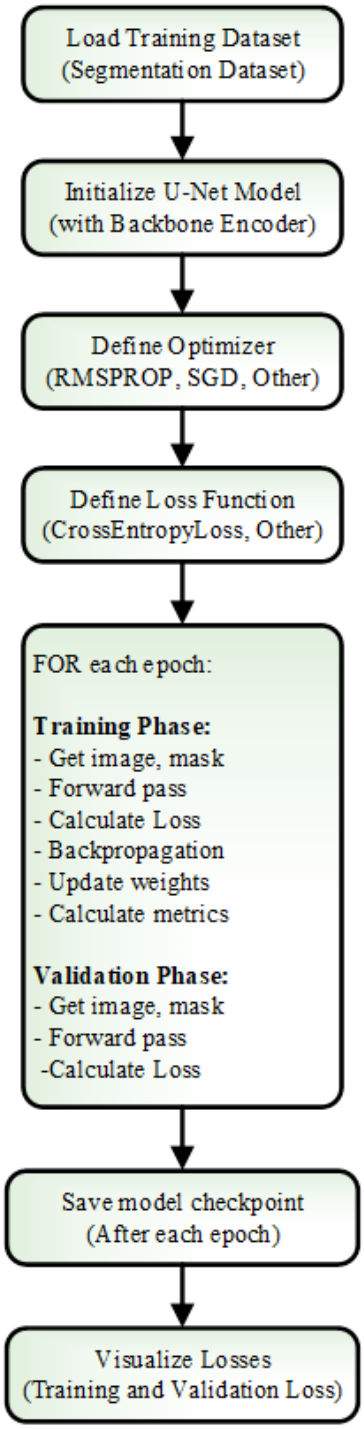


Figure 9. Model Training and Evaluation Flow

Implementing data augmentation significantly improved model performance and reduced network error. The network performance in training and validation modes without data augmentation was observed in Figures 11 and 12. The minimum loss value during training (without augmentation) was 0.302 at epoch 40, while in validation, it was 0.207 at epoch 16.

Figures 12 and 13 illustrate the network performance when using data augmentation. According to these pictures, over 40 training epochs (the same number as without augmentation), the minimum loss value during training increased from 0.302 to 1.19, and in validation, it improved from 0.207 to 0.15.

As previously mentioned, due to overfitting in the non-augmented training process, the relatively low loss value (0.302 vs. 1.19) is not meaningful. Notably, the reduction in validation loss when using augmentation (0.15 vs. 0.207) confirms the effectiveness of this approach.

Due to the observed performance improvement after data augmentation, the model training was extended for an additional 60 epochs (Figures 14 and 15).

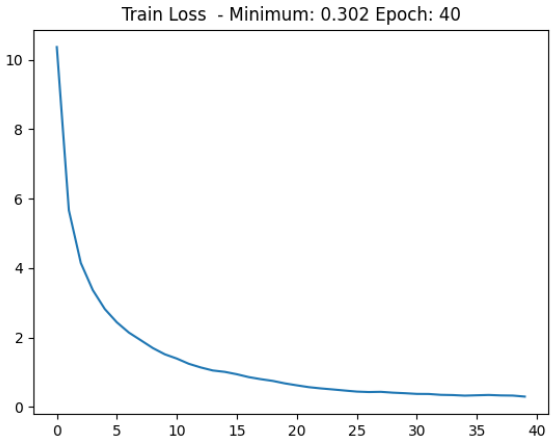


Figure 10. Training Loss Progression

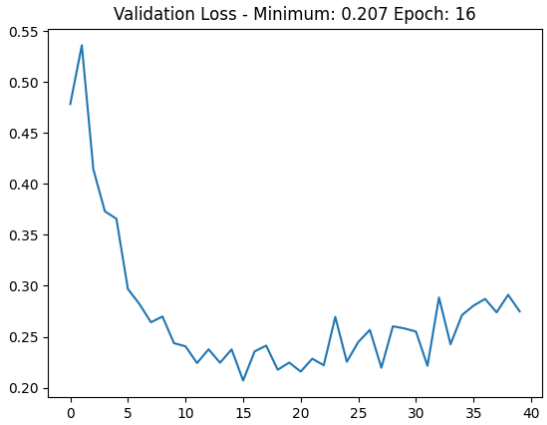


Figure 11. Validation Loss Progression

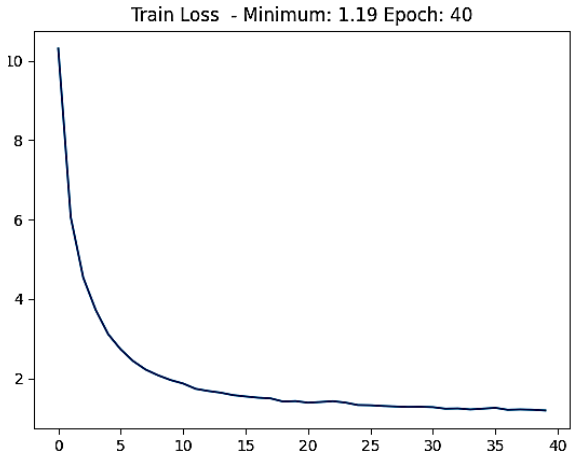


Figure 12. Network performance during the training process when using data augmentation techniques

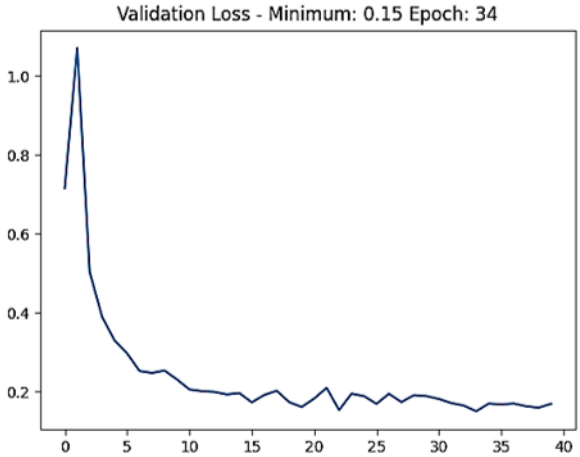


Figure 13. Network performance during the validation process when using data augmentation techniques

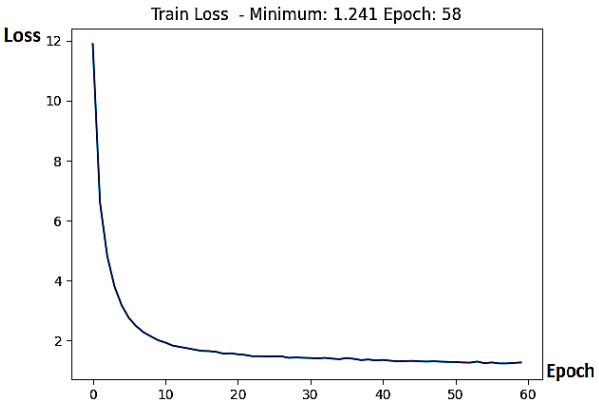


Figure 14: Network performance across 60 training epochs with data augmentation techniques applied

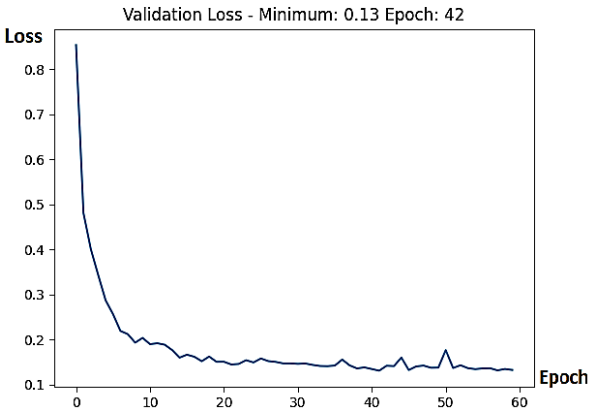


Figure 15. Network validation performance during 60 training epochs with data augmentation techniques

In the next phase, the model, after being trained for 60 epochs, is loaded and experimentally evaluated to assess its performance and quality prior to further steps.For this purpose, the corresponding code (initial inference) was developed. This code randomly selects an image from the test dataset, processes it after loading the model, and displays the resulting output mask.

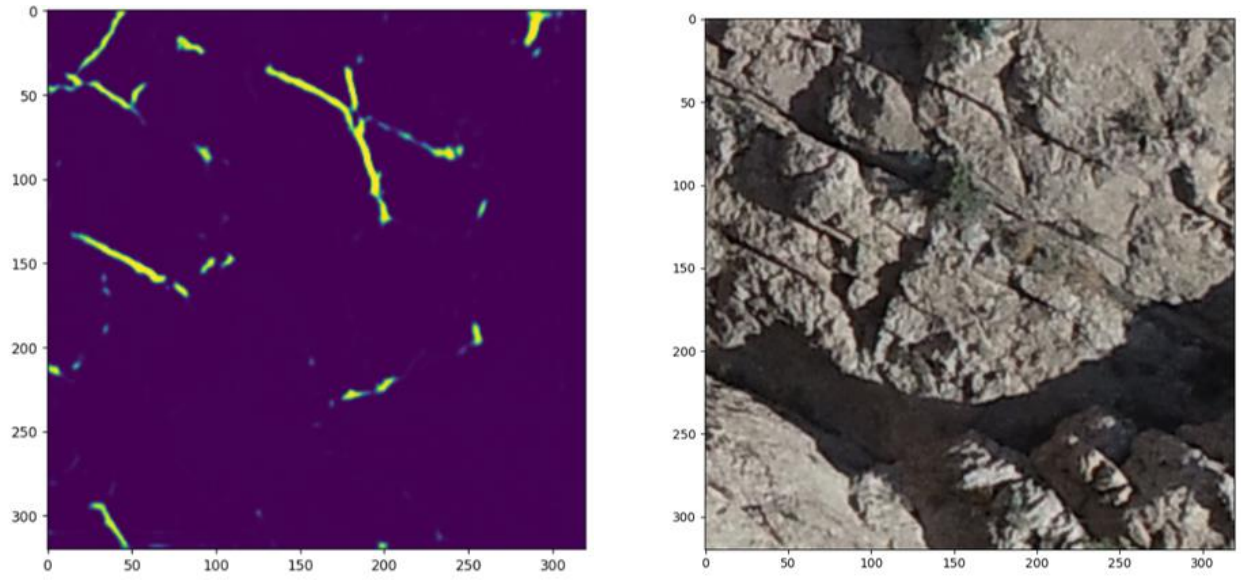
Figure 16 shows the performance quality of the model trained without data augmentation after 40 epochs, while Figure 17 demonstrates the performance of the model trained with data augmentation (also after 40 epochs) in detecting rock mass surface discontinuities on the same image.

Beyond the superior performance evident from visual inspection, quantitative metrics are essential for evaluating the model. To quantitatively assess the network performance, two metrics are considered: Pixel-wise Precision (PWP) and Intersection over Union (IoU), also referred to as the Jaccard Index. The formulation of pixel-wise accuracy is given in Equation (2), and the IoU is defined in Equation (3).

$$\text{Pixel-wise Precision} = \frac{\text{Number of Correctly Predicted Pixels}}{\text{Total Number of Labeled Pixels in the Ground Truth}} \quad (2)$$

$$\text{IOU} = \frac{|A \cap B|}{|A \cup B|} \quad (3)$$

The IoU is the most widely used evaluation metric in tasks such as object detection and tracking. Object detection consists of two sub-tasks: localization, which determines an object's position in an image, and classification, which assigns a class label to that object [26].



(a) (b)

Figure 16. Performance of the trained network (40 epochs) without data augmentation. a) Network prediction; b) Original image

As shown in Equation (3), the IoU metric is calculated as the ratio of the intersection area (overlap) to the union area between the predicted bounding box and the ground truth. Thus, IoU quantitatively measures how well a predicted bounding box aligns with the ground truth box.

When comparing model predictions against labeled masks (ground truth data) for calculating mean pixel-wise accuracy, both background pixels and predicted pixels are considered. For this reason, this metric is primarily used in classification tasks.

Since rock surface images typically do not contain equal or near-equal sample sizes of positive and negative classes, and one class (surface discontinuities) has fewer samples than the other, this type of dataset is considered "imbalanced." The number of pixels corresponding to rock surface discontinuities is usually significantly lower than background pixels. For example, in the mask shown in Figure 16, the labeled discontinuity pixels (3,927) are vastly outnumbered by background pixels (98,473). Therefore, in addition to the IoU metric, the PWP metric has been employed.

PWP metric, is particularly useful in image segmentation tasks, especially in contexts where it needes to evaluate how well the model detects the target class

relative to the total number of true instances, focusing solely on the positive class.

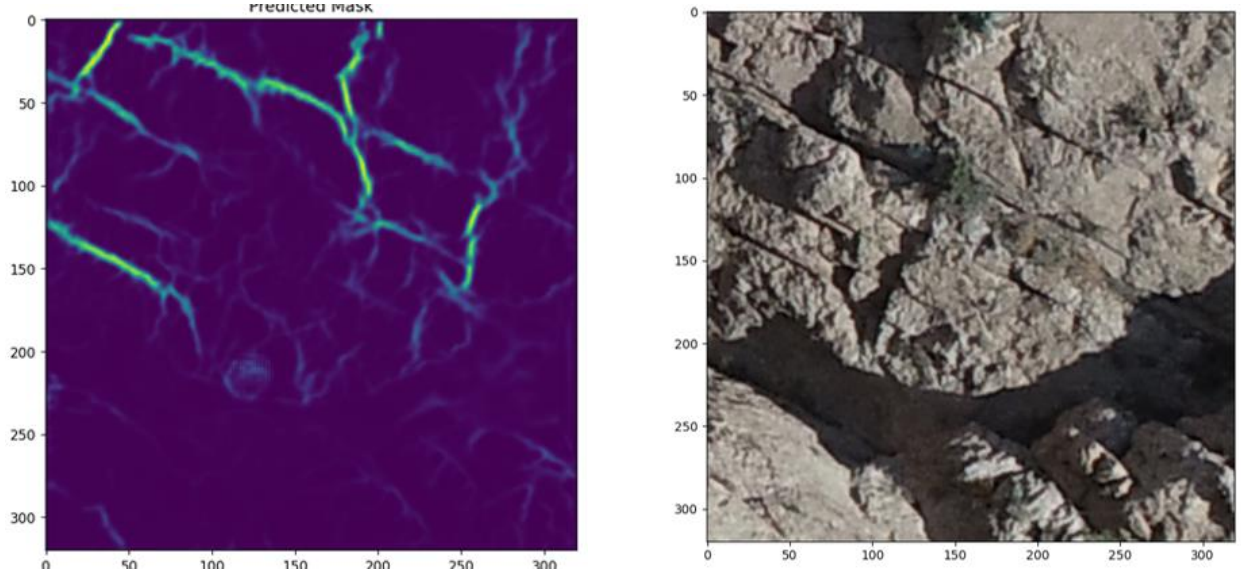
It is useful to evaluate model performance on imbalanced datasets, as it provides a clearer picture of how well the model identifies the target class compared to the ground truth

The code for the PWP metric operates as follows: It counts the pixels in the predicted output that exceed a user-defined threshold and match the corresponding pixels in the ground truth labeled. The ratio of this count to the total number of labeled pixels is then used as the score of model success.

Upon analyzing the computation methods of IoU and PWP, it appears that the PWP metric is more practical than IoU for evaluating model performance in image segmentation tasks. In addition to these metrics, visual inspection clearly demonstrates the model accuracy and output quality.

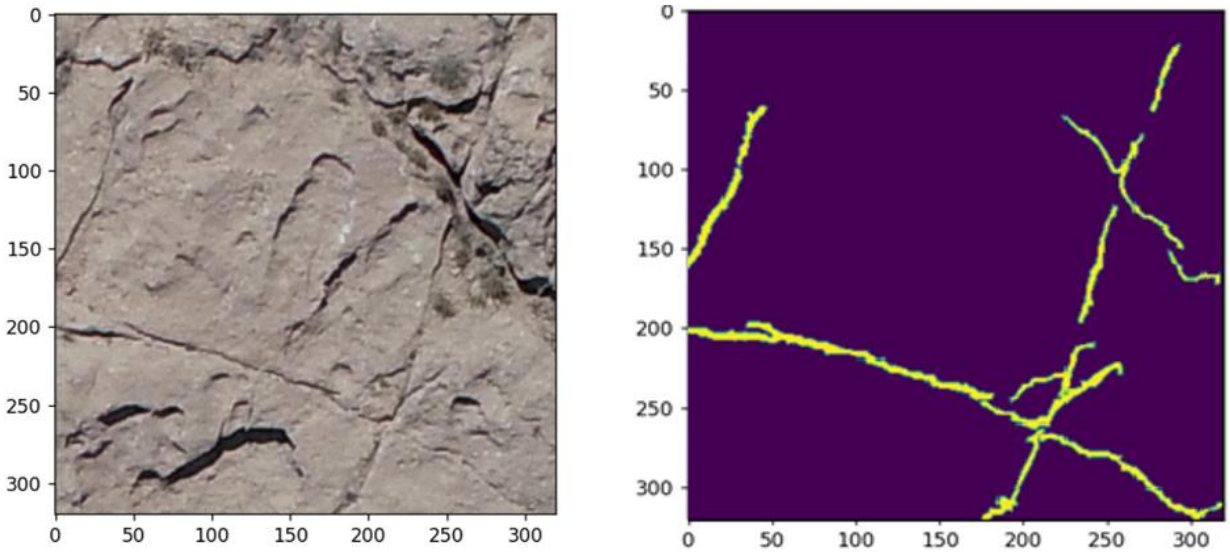
Figure 18 displays the selected test image and its corresponding ground truth mask. Meanwhile, Figure 19 compares the performance of:

- The trained model without data augmentation (left),
- The model trained with data augmentation over 60 epochs (right).



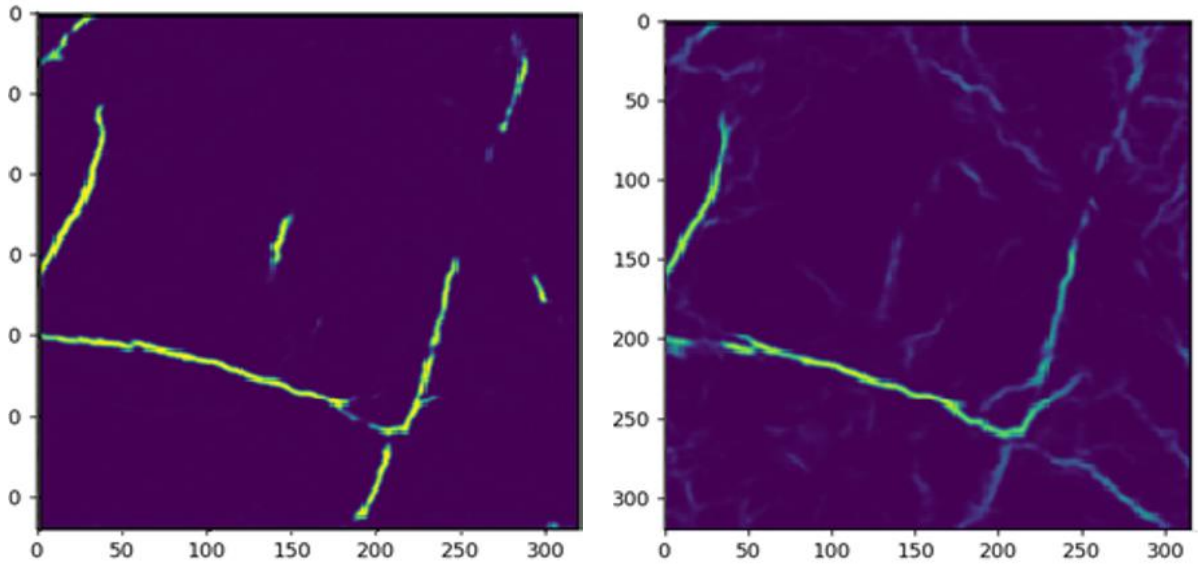
(a) **(b)**

Figure 17. Performance of the trained network (40 epochs) with data augmentation. a) Network prediction; b) Original image



(a) **(b)**

Figure 18. a)Selected test image for model evaluation; b) Corresponding ground truth mask



(a) (b)

Figure 19. Prediction comparison of models trained for 60 epochs - Without data augmentation (a) vs. With data augmentation (b)

Up to this point, efforts to improve model performance and address overfitting have focused on data augmentation. A second approach to mitigate or overcome overfitting involves reducing the number of trainable parameters. Accordingly, in the next step, we decreased the trainable parameters. Table 2 summarizes the model architecture with channel dimensions of 256, 128, 64, 32, and 16 for the feature maps in layers one through five of the Decoder section, respectively. By progressively reducing the dimensions of the feature maps, the model performance improved. The optimal performance was achieved with the feature map dimensions of 160, 80, 40, 20, and 10 in layers one through five of the Decoder, respectively. This reduction decreased the total network parameters from approximately 26 million to 24 million.

9. Optimizing Model Performance by Reducing Parameters

As demonstrated in Table 3, the summary of the model , reducing the number of parameters has improved the model accuracy according to the PWP metric which increased from 71% to 75%.

Table 2. Model Summary with Feature Map Dimensions [256, 128, 64, 32, 16]

Row no.	Component / Model Property	Value
1	Total Parameters	26,078,609
2	Trainable Parameters	26,078,609
3	Non-Trainable Parameters	0
4	Input Image Memory (MB)	1.17
5	Forward/Backward Pass Memory (MB)	1,289.45
6	Parameter Memory (MB)	99.48
7	Estimated Total Memory (MB)	1,390.11
8	PWP Metric Score (%)	71.02
9	IoU Metric Score (%)	27.74

Table 3. Model Summary with Feature Map Dimensions [160, 80, 40, 20, 10]

Row no.	Component / Model Property	Value
1	Total Parameters	24,517,103
2	Trainable Parameters	24,517,103
3	Non-Trainable Parameters	0
4	Input Image Memory (MB)	1.17
5	Forward/Backward Pass Memory (MB)	1,186.62
6	Parameter Memory (MB)	93.53
7	Estimated Total Memory (MB)	1,281.32
8	PWP Metric Score (%)	75.6
9	IoU Metric Score (%)	45.7

The predictions of the trained model after implementing both data augmentation and model parameter reduction are demonstrated in Figure 20. Visual inspection clearly shows superior discontinuity detection performance compared to the two previous model versions. A quantitative comparison of model performance across three configurations including baseline (without data augmentation), with data augmentation only, and the combined approach (data augmentation + parameter reduction) are presented in Table 4.

Table 4. Performance metrics comparison between models trained without and with data augmentation

Row	Model Development Approach	Metric	Score (%)
1	Without data augmentation	IOU	41.5
2	Without data augmentation	PWP	53.27
3	With data augmentation	IOU	44.8
4	With data augmentation	PWP	71.02
5	With data augmentation and model parameter reduction	IOU	45.56
6	With data augmentation and model parameter reduction	PWP	75.6

The trend of the final model's accuracy values, with the application of data augmentation and reduction of trainable network parameters, over 60 training epochs is presented in Table 5. In the next step, the final model obtained after 60 training epochs, incorporating the modifications of reducing the number of model parameters and applying data augmentation, was executed on a UAV image. As an example, the code was applied to a section of the UAV image shown in Figure 21. The result of this performance is presented in Figure 22, where the white lines indicate the discontinuities detected by the model. This model was generated after 60 training epochs. The total training time was two hours and thirty-five minutes. The specifications of the computer used for this processing are as follows:

- **GPU:** Nvidia GeForce 840M with 4 GB memory
- **CPU:** Intel® Core™ i7–5500U @ 2.40 GHz
- **Storage:** SanDisk Ultra II SSD
- **RAM:** 8.0 GB DDR3
- **Operating System:** Windows 10, 64-bit

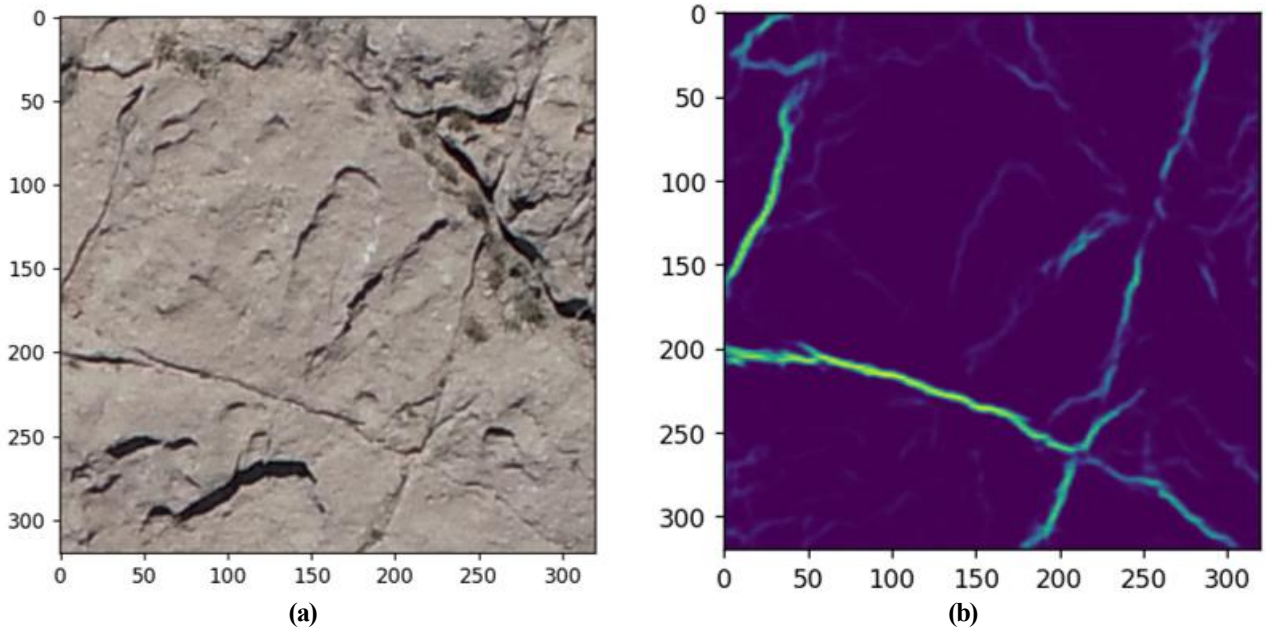


Figure 20. Performance comparison of the (a) original input image versus (b) optimized model after implementing both data augmentation and parameter reduction

10. Discussion

This study developed a Python-based deep learning workflow for detecting rock mass discontinuities from UAV imagery. Despite constraints such as suboptimal image resolution due to the lack of ground control points and a relatively small annotated dataset (~200 images), the optimized model, enhanced through data augmentation and parameter reduction, achieved a pixel-wise precision exceeding 75%, up from an initial 53%.

The results show that most visible discontinuities were successfully detected even with partial annotations, confirming the feasibility of automated UAV-based mapping that reduces the need for in-field inspections, particularly in inaccessible areas.

The study highlights that model performance depends strongly on dataset size, image quality, and geological variability. Although the dataset was limited, this constraint reflects the practical conditions of field-based geotechnical studies, where acquiring large labeled datasets is both costly and time-consuming. Nonetheless, reliable segmentation was achieved under data-scarce conditions, suggesting that targeted augmentation and optimization can ensure stable model performance. However, UAV-based imagery remains restricted to surface observations, and variations in imaging conditions may affect accuracy—underscoring the need for standardized data acquisition and photogrammetric control.

Table 5. Training Progression of Final Model Accuracy with Data Augmentation and Parameter Reduction

Epoch	Train Loss	Validation Loss	Epoch	Train Loss	Validation Loss	Epoch	Train Loss	Validation Loss
1/60	7.914	0.92	21/60	1.506	0.145	41/60	1.323	0.154
2/60	5.292	0.435	22/60	1.471	0.151	42/60	1.303	0.137
3/60	4.109	0.342	23/60	1.458	0.142	43/60	1.315	0.131
4/60	3.432	0.307	24/60	1.432	0.14	44/60	1.302	0.144
5/60	2.947	0.258	25/60	1.462	0.143	45/60	1.278	0.138
6/60	2.585	0.239	26/60	1.463	0.15	46/60	1.306	0.142
7/60	2.326	0.22	27/60	1.416	0.149	47/60	1.29	0.147
8/60	2.126	0.198	28/60	1.459	0.146	48/60	1.283	0.136
9/60	2.029	0.181	29/60	1.398	0.144	49/60	1.255	0.153
10/60	1.909	0.182	30/60	1.434	0.154	50/60	1.248	0.143
11/60	1.817	0.171	31/60	1.415	0.15	51/60	1.255	0.138
12/60	1.747	0.177	32/60	1.396	0.142	52/60	1.263	0.137
13/60	1.69	0.167	33/60	1.376	0.156	53/60	1.268	0.135
14/60	1.664	0.161	34/60	1.392	0.146	54/60	1.294	0.166
15/60	1.621	0.164	35/60	1.337	0.155	55/60	1.255	0.14
16/60	1.572	0.155	36/60	1.362	0.138	56/60	1.25	0.14
17/60	1.58	0.155	37/60	1.364	0.138	57/60	1.235	0.145
18/60	1.558	0.147	38/60	1.323	0.136	58/60	1.237	0.14
19/60	1.523	0.149	39/60	1.345	0.145	59/60	1.209	0.138
20/60	1.491	0.148	40/60	1.344	0.137	60/60	1.25	0.137



Figure 21. Sample UAV (drone) image as input to the final model

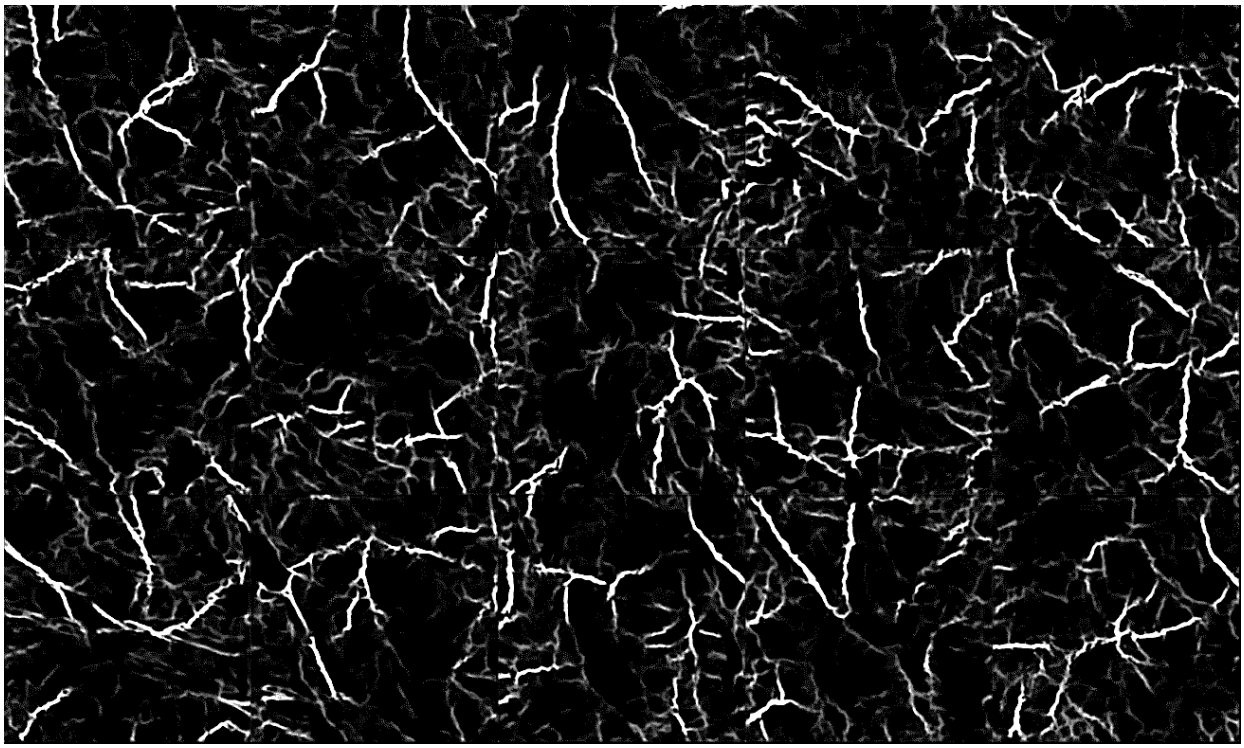


Figure 22. Final trained model's prediction on sample UAV (drone) image

11. Conclusions

The proposed deep learning framework demonstrated effective discontinuity detection from UAV imagery, achieving satisfactory segmentation accuracy despite limited training data. This outcome highlights the robustness and adaptability of the workflow for real-world geotechnical applications.

Future work should focus on improving generalization and scalability by expanding annotated datasets and comparing alternative CNN architectures. High-quality imagery with photogrammetric ground control, semi-supervised and active learning approaches to reduce manual labeling, and the use of multi-view or oblique UAV imagery can further enhance completeness. Additionally, transfer learning and the integration of complementary data such as multispectral or LiDAR imagery can improve accuracy. Establishing standardized UAV acquisition protocols and extending the approach to broader mining applications, with support from dedicated annotation teams, will enhance reproducibility, efficiency, and operational readiness.

References

[1]. H. Madani. (2009). Tunneling. Tehran: Amirkabir University of Technology.

[2]. Seyedahmad Mehrishal, Jineon Kim, Yulong Shao, Jae Joon Song. (2024). Artificial intelligence-aided semi-automatic joint trace detection from textured three dimensional models of rock mass. Institute of Rock and Soil Mechanics, Chinese Academy of Sciences.

[3]. M. McCarter. (1976). Application Of Plane Table Photogrammetry To Open Pit Mapping. pp. ARMA-76-0185, 25 August.

[4]. M. Rossini. (1998). 2D-Discontinuity Detection from Scattered Data. Computing, © *Springer-Verlag.*

[5]. F. Lemy and J. Hadjigeorgiou. (2003). Discontinuity trace map construction using photographs of rock exposures. *Pergamon,Elsevier Ltd*.

[6]. B. VAN KNAPEN and S. SIEFKO. (2006). Identification and characterisation of rock mass discontinuity sets using 3D laser scanning. The Geological Society of London.

[7]. M. Mohebbi, A. Yarahmadi Bafghi, M. Fatehi Marji and J. Gholamnejad. (2017). "Rock mass structural data analysis using image processing techniques (Case study: Choghart iron ore mine northern slopes)," *JME, Journal of Mining & Environment*.

[8]. Z. Pi, Z. Zhou, X. Li and S. Wang. (2021). Digital Image Processing Method for Characterization of Fractures, Fragments, and Particles of Soil/Rock-Like Materials. mathematics.

[9]. Przybilla, H.-J., Wester-Ebbinghaus. (1979). Bildflug mit ferngelenktem Kleinflugzeug. vol. 47.

[10]. I. Colomina, M. Blázquez, P. Molina, M.E. Parés and M. Wis. (2008). "TOWARDS A NEW PARADIGM FOR HIGH-RESOLUTION LOW-COST PHOTOGRAMMETRYAND REMOTE SENSING," vol. XXXVII Part B1.

[11]. H. Eisenbeiss. (2008). THE AUTONOMOUS MINI HELICOPTER: A POWERFUL PLATFORM FOR MOBILE MAPPING. Vols. CH-8093.

[12]. Y. Vasuki, E. J. Holden, P. Kovesi and S. Micklethwaite. (2013). A Geological Structure Mapping Tool using Photogrammetric Data. *23rd International Geophysical Conference and Exhibition, 11-14 August 2013 - Melbourne, Australia*.

[13]. Deheng Kong, Charalampos Saroglou, Faquan Wu, Peng Sha and Bo Li d. (2021). "Development and application of UAV-SfM photogrammetry for quantitative characterization of rock mass discontinuities," *International Journal of Rock Mechanics and Mining Sciences.*

[14]. Antonio Pola, Arturo Herrera-Díaz, Sergio Rogelio,Tinoco-Martínez, José Luis Macias. (2024). "Rock characterization, UAV photogrammetry and use of algorithms of machine learning as tools in mapping discontinuities and characterizing rock masses in Acoculco Caldera Complex," *Bulletin of Engineering Geology and the Environment, vol*. 83:260.

[15]. Kumar, A. (2024). A promising automatic system for studying of coal mine surfaces using Sentinel-2 data to assess a classification on a pixel-based pattern. *Journal of Mining and Environment, 15*(1), 41–54.

[16]. Yalcin, I., Can, R., Kocaman, S., & Gokceoglu, C. (2023). Rock mass discontinuity determination with transfer learning. *ISPRS Archives, XLVIII-M-1*, 609–616.

[17]. Huashan, L., et al. (2025). Identification of ground fissures in mining areas from UAV images based on RDC-UNet. *Elektronika ir Elektrotechnika, 31*(1), 15–22.

[18]. Li, Q., et al. (2023). Semantic segmentation of UAV remote sensing images based on edge feature fusing and multi-level upsampling integrated with Deeplabv3+ (EMNet). *Remote Sensing, 15*(3), 895–906.

[19]. Wang, K., et al. (2024). An automated approach for mapping mining-induced fissures using CNNs and UAV photogrammetry. *Remote Sensing, 16*(12), 2090.

[20]. Yalcin, I., Can, R., Kocaman, S., & Gokceoglu, C. (2024). A novel rock mass discontinuity detection approach with CNNs and multi-view image augmentation. *ISPRS International Journal of Geo-Information, 13*(6), 185.

[21]. Umberto Andriolo, Konstantinos Topouzelis, Tim H.M. van Emmerik, Apostolos Papakonstantinou, Jo˜ao Gama Monteiro, Atsuhiko Isobe, Mitsuko Hidaka h,Shin'ichiro Kako, Tomoya Kataoka, Gil Gonçalves. (2023). "Drones for litter monitoring on coasts and rivers: suitable flight altitude and image resolution," *elsevier*.

[22]. Muthuramalingam Sivakumar, Sudhaman Parthasarathy, Thiyagarajan Padmapriya. (2024). Trade-off between training and testing ratio in machine learning for medical image processing, *PeerJ computer science*.

[23]. Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh and Jianming Liang. (2018). UNet++: A Nested U-Net Architecture for Medical Image Segmentation. *Springer International Publishing*.

[24]. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *Springer International Publishing*.

[25]. Saleh Shahinfar, Paul Meek, Greg Falzona. (2020). How many images do I need" Understanding how sample size per class affects deep learning model performance metrics for balanced designs in autonomous wildlife monitoring. *Elsevier*.

[26]. Rezatofighi, Hamid and Tsoi, Nathan and Gwak, JunYoung and Sadeghian, Amir and Reid, Ian and Savarese, Silvio. (2019). "Generalized Intersection over Union," *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, no. https://giou.stanford.edu/.

[27]. K. X. C. Z. E. G. a. Y. L. Yilin Dong. (2023). "E-FPN: Evidential Feature PyramidNetwork for Ship Classification," Remote Sensing.

انجمن مهندسی معدن ایران
نشریه مهندسی معدن و محیط زیست
نشانی نشریه: www.jme.shahroodut.ac.ir
دانشگاه صنعتی شاهرود

شناسایی ناپیوستگی‌های سنگی با استفاده از پردازش تصاویر پهپادی (UAV) و یادگیری عمیق

مصطفی رحیمیان[۱]، محمد عطائی[۱*]، رضا کاکایی[۱] و حسین خسروی[۲]

۱. دانشکده مهندسی معدن، نفت و ژئوفیزیک، دانشگاه صنعتی شاهرود، شاهرود، ایران
۲. دانشکده مهندسی برق، دانشگاه صنعتی شاهرود، شاهرود، ایران

اطلاعات مقاله

تاریخ ارسال: ۲۰۲۵/۰۹/۲۷
تاریخ داوری: ۲۰۲۵/۱۱/۰۷
تاریخ پذیرش: ۲۰۲۵/۱۱/۲۴

DOI: 10.22044/jme.2025.16898.3316

کلمات کلیدی

ناپیوستگی‌های سنگی
پهپاد (UAV)
پردازش تصویر
یادگیری عمیق

چکیده

شناسایی ناپیوستگی‌های سنگی، عاملی حیاتی در پروژه‌های معدنی و عمرانی محسوب می‌شود. روش‌های سنتی برای انجام این کار اغلب دشوار و زمان‌بر هستند، ایمنی افراد را به خطر می‌اندازند و به ارزیابی‌های ناقص منجر می‌شوند. به‌کارگیری پهپادها (UAVs) این فرآیند را متحول کرده است؛ به‌طوری که می‌توان کل منطقه را در زمانی کوتاه و بدون به خطر انداختن کارکنان، مورد بررسی قرار داد. هدف این پژوهش، استفاده از یادگیری عمیق و زبان برنامه‌نویسی پایتون برای توسعه و آموزش یک شبکه عصبی مبتنی بر معماری UNET++ است تا ناپیوستگی‌های سطحی سنگ‌ها به‌طور خودکار از روی تصاویر ثبت‌شده توسط پهپادها شناسایی شوند. همچنین، با به‌کارگیری تکنیک‌های افزایش داده (Data Augmentation) و کاهش حدود ۶ درصدی پارامترهای مدل، چالش‌های مرتبط با یادگیری نظارت‌شده، به‌ویژه «بیش‌برازش» (Overfitting)، برطرف شده‌اند. در نتیجه، معیار دقت در سطح پیکسل (Pixel-wise precision) به‌طور قابل‌توجهی از ۵۳.۲۷ درصد به ۷۵.۶ درصد بهبود یافته است. وجه تمایز این پژوهش نسبت به سایر مطالعات، تمرکز ویژه بر تصاویر پهپادی برای ارزیابی‌های زمین‌شناختی، استفاده از راهکاری دوگانه برای مقابله با بیش‌برازش، و ارائه عملکردی کارآمد با وجود داده‌های آموزشی محدود است. نتایج نشان می‌دهند که این مدل قادر است ناپیوستگی‌های سنگی را با دقت بالا شناسایی کند و روشی مناسب برای صنایع معدنی و عمرانی به شمار می‌رود.

نویسنده مسئول مکاتبات: ataei@shahroodut.ac.ir (م. عطائی)