International Journal of Geo-Information

MDPI

Article

A Novel Rock Mass Discontinuity Detection Approach with CNNs and Multi-View Image Augmentation

Ilyas Yalcin [1,2], Recep Can [1], Candan Gokceoglu [3] and Sultan Kocaman [4,*]

[1] Graduate School of Science and Engineering, Hacettepe University, 06800 Beytepe Ankara, Türkiye; ilyas.yalcin@hacettepe.edu.tr (I.Y.); recepcan@hacettepe.edu.tr (R.C.)
[2] Başkent OSB Technical Sciences Vocational School, Hacettepe University, 06909 Sincan Ankara, Türkiye
[3] Department of Geological Engineering, Hacettepe University, 06800 Beytepe Ankara, Türkiye; cgokce@hacettepe.edu.tr
[4] Department of Geomatics Engineering, Hacettepe University, 06800 Beytepe Ankara, Türkiye
* Correspondence: sultankocaman@hacettepe.edu.tr

Abstract: Discontinuity is a key element used by geoscientists and civil engineers to characterize rock masses. The traditional approach to detecting and measuring rock discontinuity relies on fieldwork, which poses dangers to human life. Photogrammetric pattern recognition and 3D measurement techniques offer new possibilities without direct contact with rock masses. This study proposes a new approach to detect discontinuities using close-range photogrammetric techniques and convolutional neural networks (CNNs) trained on a small amount of data. Investigations were conducted on basalts in Bala, Ankara, Türkiye. A total of 34 multi-view images were collected with a remotely piloted aircraft system (RPAS), and discontinuity lines were manually delineated on a point cloud generated from these images. The lines were back-projected onto the raw images to increase the amount of data, a process we call multi-view (3D) augmentation. We further evaluated radiometric and geometric augmentation methods, the contribution of multi-view augmentation to the proposed model, and the transfer learning performance of six different CNN architectures. The highest performance was achieved with U-Net + SE-ResNeXt-50 with an F1-score of 90.6%. The CNN model trained from scratch with local features also yielded a similar F1-score (91.7%), which is the highest performance reported in the literature.

Keywords: rock mass classification systems; close-range photogrammetry; deep learning; transfer learning; domain adaptation; data augmentation

check for updates

Citation: Yalcin, I.; Can, R.; Gokceoglu, C.; Kocaman, S. A Novel Rock Mass Discontinuity Detection Approach with CNNs and Multi-View Image Augmentation. *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 185. https://doi.org/10.3390/ijgi13060185

Academic Editors: Adrianos Retalis, Dimitrios D. Alexakis and Wolfgang Kainz

Received: 7 April 2024
Revised: 24 May 2024
Accepted: 28 May 2024
Published: 31 May 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Rock slope failures are generally controlled by discontinuities and their orientations [1]. Detection of rock discontinuities is required in various fields such as tunnel, gallery, and deep shaft construction in underground mining operations, rock falls, the classification and stabilization of rock masses, etc. In addition, Türkiye is prone to rockfall with hundreds of incidents per year [2], such as when numerous fatalities occurred on February 6, 2023, during the Kahramanmaras earthquakes (7.7 Mw and 7.6 Mw) that occurred due to rockfall in mountainous areas [3]. Well-known rock mass classification systems involve the recent versions of the Rock Mass Rating (RMR) system introduced by Bieniawski [4], the Geological Strength Index (GSI) established by Hoek and Brown [5], and the Q system developed by Barton [6]. The parameters defined in these systems aim to characterize rock mass discontinuities, which include spacing, persistence, roughness, aperture, infill, orientation of discontinuities, and the number of discontinuity sets [7].

Engineering rock mass classifications take into consideration the most important geological aspects affecting rock mass so as to rate its quality and form the backbone of the empirical design approach for engineering structures such as tunnels, slopes, etc., and they are also commonly employed in rock mechanics applications [8]. Among the

rock mass classification systems, the widely used RMR [4], GSI [5], and Q [6] systems employ the discontinuity characteristics of rock masses. When calculating basic RMR, the spacing of discontinuities and condition of discontinuities are the two main parameters. The condition of the discontinuities contains five different discontinuity features, i.e., persistence, aperture, roughness, infilling, and weathering [4]. In addition, a correction considering the orientation of discontinuities is then applied and the final RMR is obtained. The Q system [6] is a numerical assessment of rock mass quality using six parameters, and three of these parameters are related to discontinuities, such as the number of joint sets, the roughness of the most favorable discontinuity, and the degree of alteration or filling along the weakest discontinuity. Finally, the GSI system [5] employs two main parameters: rock mass structure and the quality of the discontinuity. In addition, the rock mass structure is controlled by the number and orientation of discontinuities. As can be seen from this short assessment, the rock mass behavior is directly controlled by the characteristics of discontinuities.

Scan-line [9] and window mapping [10] are the conventional methods used to collect data on the discontinuity properties of rock surfaces. In the scan-line method, discontinuities, appropriately visible as lines, are determined on discontinuity surfaces, and the measurements are taken at the intersections. In the window mapping method, measurements are carried out in a rectangular area to reduce orientation bias [11]. In both approaches, the measurements are carried out using a compass and a clinometer [12], which can be limited or dangerous to use due to rockfalls or inaccessibility to the discontinuities. When measuring the orientation of rock discontinuities, the compass must be set on the discontinuities. However, they can be inaccessible depending on the elevation of the discontinuity with respect to the ground. In addition, unstable rock masses may cause injuries during measurements, thus posing life-threatening risks. Likewise, scan-line studies also potentially involve such risks. To avoid these challenges in fieldwork, rock discontinuities can be automatically detected from images, but this approach also has several difficulties. The first and foremost one is irregular shapes, which require different approaches to many other object (or feature) extraction or segmentation tasks, such as building or road detection. Furthermore, most rock surfaces have non-uniform shapes and colors sourced from coating or erosion. Shadows and visibility issues also increase the level of difficulty. Although some discontinuities are detected as lines with regular or irregular shapes, others appear as surfaces based on the observer's viewpoint or occlusions.

In recent years, geospatial technologies such as Light Detection and Ranging (LiDAR) and aerial and mobile close-range photogrammetry have been used to measure discontinuities as they allow for offline precise measurements of large regions after a short field campaign. LiDAR sensors produce high-precision 3D point cloud data [13] and reliable results on discontinuity sets [14]. However, since a study area may not be fully scanned by a LiDAR sensor, it is necessary to establish multiple measurement stations on unfavorable rock mass terrains or to adapt multisensory data to acquire a complete scene. Remotely piloted aircraft systems (RPASs) are used to capture optical images from different altitudes without any station setup. The Structure from Motion (SfM) technique also enables the production of a 3D model without the requirement of a metric camera. The SfM method can calculate camera calibration parameters, image perspective center positions, and rotations in model space by extracting key points from overlapping source imagery [15]. Optical imaging from an RPAS platform offers a cost advantage over LiDAR, and it has been used for rock discontinuity measurement by several scholars such as Salvini et al. [16], Kong et al. [17], Wang et al. [18], Song et al. [19], etc.

Most rock mass discontinuity determination methods from 3D point clouds are based on surface extraction algorithms, such as the Discontinuity Set Extractor (DSE) developed by Riquelme et al. [20], or a combination of methods, as proposed by Liu et al. [21]. The point cloud data size is a major limitation as it can increase the computational cost. Ozturk et al. [22], Chen and Jiang [23], and An et al. [24] determined rock discontinuities using mobile phone images, which is also a practical approach depending on the size of the

region to be imaged and the rock mass parameters. All studies mentioned above rely on point clouds for detecting rock mass discontinuities. The methods, their strengths, and weaknesses are briefly summarized in Table 1. However, producing dense point clouds is a time-consuming process and can mainly be performed in the office using desktop or cloud-based tools. Depending on the size of the study site, the data size also increases [18], leading to unnecessary data production. Additionally, point cloud processing and the detection of discontinuities may require a high level of expertise and computational skills.

Table 1. A brief overview of selected studies from the literature.

	Data Type	Study Goal	Method	Weakness
Salvini et al. [16]	3D models and point clouds derived from RPAS and Terrestrial Laser Scanner (TLS)	Measuring surface roughness	Roughness measurements were compared manually with the traditional method	Manual calculation for surface roughness measurements
Kong et al. [17]	3D point clouds derived from RPAS data	Discontinuity detection with scan-line and dip/dip direction calculations	Normal vector calculation and cluster analysis for points	Difficulties in detecting discontinuities such as linear traces and cracks
Wang et al. [18]	3D models and point clouds derived from RPAS data	Identification of discontinuities and calculation of rockfall potential	Point cluster analysis and image processing	Time consumption depending on octree levels
Song et al. [19]	3D models and point clouds derived from RPAS data	Measuring surface roughness and 3D laser scanner comparison with RPAS	Cloud-to-cloud comparison with M3C2 algorithm	Ambient lighting and equipment sensitivity
Ozturk et al. [22]	3D models and point clouds derived from mobile phone data	Detection of discontinuity and kinematic analysis	Plane fitting and normal vector calculation method	Discontinuities are detected manually
Chen and Jiang [23]	3D models and point clouds derived from mobile phone data	Discontinuity detection with dip/dip direction calculation	Coordinate transformation and normal vector calculation	Discontinuities are detected manually
An et al. [24]	3D models and point clouds derived from mobile phone data	Measuring surface roughness	Image reconstruction and mesh model generation	Ambient lighting and requires large working area

On the other hand, conventional edge detection algorithms such as Canny or Sobel filters, can also be used to detect rock discontinuities from images. However, due to the nature of the problem, the identified discontinuities contain a high level of noise sourced from topographic variations and the presence of different textures on rock surfaces [25]. Convolutional neural network (CNN) architectures have been widely used for various image processing, feature extraction, and object segmentation tasks (e.g., see Qiu et al. [26] and Yalcin et al. [27–30]), including edge detection. However, they require large amounts of data and computational resources for model training. Yalcin et al. [27,28] detected rock discontinuities on orthophotos using a CNN, and they emphasized that blurring was observed at the discontinuities due to the inferior quality of the digital surface model (DSM), especially at locations with poor illumination and shadow. On the other hand, detecting discontinuities in geometrically unprocessed (raw) images requires the use of multiple images with different viewing angles to obtain the 3D position information and ensure model completeness. However, when multiple images are used, image measurement uncertainty is introduced to the data as the manual interpretation of rock discontinuity in images is difficult and delineating the same lines in multiple images can be highly challenging. Since a large amount of training data may not be obtainable in many sites,

the development of novel augmentation approaches is needed for learning from a small amount of data.

In order to overcome these issues, in this study, we propose a novel multi-view image (3D) augmentation method to introduce variation to learning data using perspective imaging geometry, and we also propose applying a transfer learning approach to meet the data reduction requirement. Transfer learning applies knowledge from previous tasks to new ones, reducing the need for extensive training data [31]. Domain adaptation, a specific type of transfer learning, occurs when related tasks have different data distributions [32]. We investigated the domain adaptation of CNNs pre-trained with crack data in our process as these approaches have recently proven to be successful in increasing model performance with a small amount of data being used for fine tuning (Yalcin et al. [27,28,30]). A practical application was carried out on basalts in Bala town of Ankara using data acquired with an RPAS. The ground truth was manually delineated as line vectors on the 3D model, which was produced in Agisoft Metashape. The discontinuity lines defined in object space were first back-projected to the image space and then to pixels using orientation parameters obtained from the bundle block adjustment. The CNN evaluated here was trained with labels in multi-view images, and the effectiveness of the multi-view augmentation was assessed accordingly. In addition, a transfer learning approach was tested by comparing two different CNN architectures and three different backbones with domain adaptation. The data, methods, and the results are explained and discussed in the following section.

2. Materials and Methods

2.1. Study Area Characteristics

Bala town in Ankara Province (Türkiye) was selected as the study area due to its accessibility (60 km southeast of Ankara city center), as shown in Figure 1. A total of 18 lithological units exist in Bala, as explained in Figure A1 [33] and Table A1 in Appendix A. The working area, defined as the Evciler Volcanics formation of Bala, consists of basalts, which were preferred due to their smooth rock surfaces and strong features. Evciler Volcanics is described as consisting of white-colored tuffs, followed by red-gray-colored scoria and lapilli stones, and it is composed of olivine basalt lavas from top to bottom [33]. In addition, the formation is equivalent to the Pliocene-aged Bozdag Basalt. The lengths of rock blocks in the area range from 5 cm to 300 cm. The sampling area has an approximate width of 23 m and a length of 52 m.

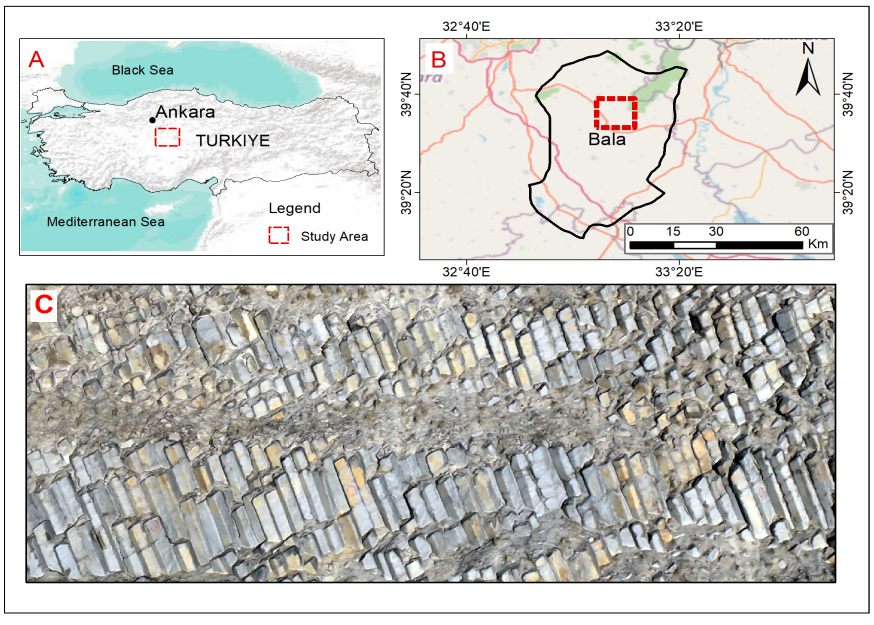


Figure 1. The location maps of the study area at (**A**) country, (**B**) city scales, and (**C**) a perspective view of the basalt columns.

2.2. The Overall Workflow

A workflow (Figure 2) consisting of four main steps was carried out in the detection of rock discontinuities: (i) photogrammetric processing (image acquisition, bundle block adjustment, and point cloud generation) (see the purple block in Figure 2 and refer to Section 2.3 for further details); (ii) training data preparation including augmentation (see the blue block in Figure 2 and refer to Section 2.4); (iii) training from scratch with local features (see Section 2.5.1), random initialization, and domain adaptation (see Section 2.5.2); and (iv) validation (see Section 2.6). The training and validation stages are highlighted with beige and green blocks in Figure 2. Besides the various data augmentation techniques, we investigated the performances of the CNN trained from scratch with a small amount of data and the transfer learning based on the images obtained from a crack dataset, which were adapted from local features (as explained below).

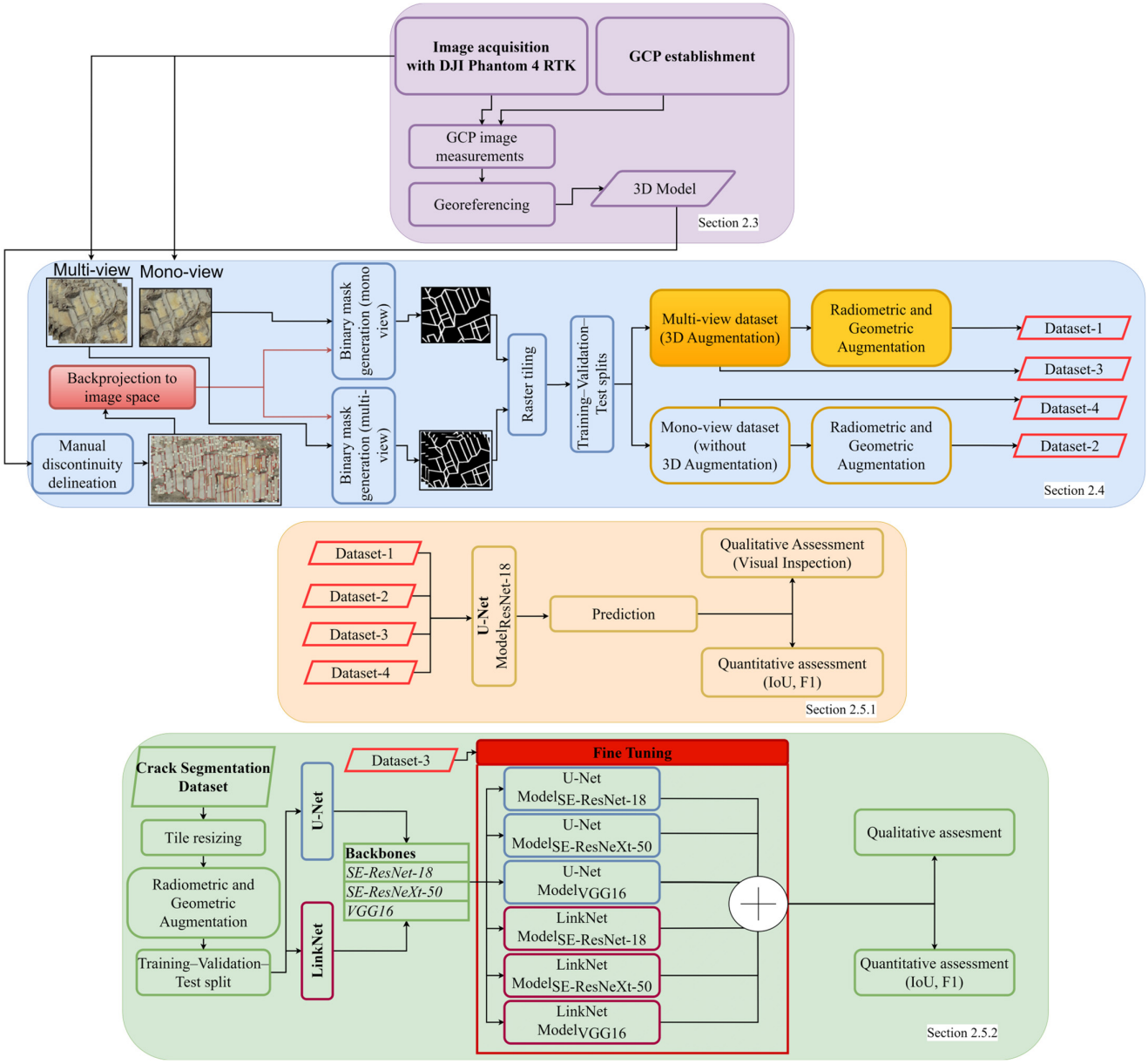


Figure 2. The overall workflow.

2.3. Photogrammetric Processing

The photogrammetric process involves a field campaign for Ground Control Point (GCP) establishment and image acquisition with an RPAS. A Global Navigation Satellite System (GNSS) receiver that operates with the Real-Time Kinematic (RTK) Network method was used to measure the GCP ground coordinates with high accuracy. Due to access limitations, 2 GCPs were measured with a GNSS device, and an additional 18 GCPs were measured by means of a total station. Image acquisition was performed with a DJI Phantom 4, which was also aided with an RTK module [34]. Although RPASs equipped with RTK systems are known to provide high accuracy in the order of a few centimeters [35], the GCPs were added in order to observe the full accuracy potential of the 3D model. The motion blur effect in image acquisition was minimized through the 3-axis gimbal on the DJI. In Table 2, the technical specifications of the DJI RTK are summarized. A total of 34 overlapping images were captured from the study area, 26 of which were obtained with a camera–object distance of approximately 30 m, while the remaining images were taken at 45 m. The images have a size of 5472 × 3648 pixels. The camera focal length was 24 mm. A number of GCPs were used as check points (CPs) in the bundle block adjustment. Please see Figures 3 and 4 for sample images and the GCP distribution.

Table 2. Technical specifications of the DJI Phantom 4 RTK [34].

Specification	Value
Weight	1391 g
Crosswise distance	350 mm
Ascent/Descent speed	6/3 (m/s)
RTK module	Enabled
GNSS module	GPS, GLONASS, Galileo, BeiDou
Fixed time	<50 s
Sensor size/type	1″/CMOS
Effective image pixel resolution	20 MP
Image Frame Size (pixel)	5472 × 3648
Focal length	8.8–24 mm
Image bands included in the camera	Red, Green, Blue (RGB)

The 3D model and the orthophoto were generated with Agisoft Metashape Professional version 1.8.4 [36]. Six out of twenty GCPs were used as CPs in the bundle block adjustment. The blue rectangle in Figure 4 shows the actual model area used as no GCP could be marked on the rock surface on the left side due to limited accessibility. The GCPs and the photogrammetric products were defined in the Transverse Mercator (TM) central meridian, which is 33° east from the Greenwich Meridian projection system. The datum was defined as the World Geodetic System 1984 (WGS 84) reference framework. Figure 5 illustrates the camera locations of the study area and the number of overlapping images.

2.4. Data Augmentation Techniques and Training Data Preparation

CNN models have recently been widely used in image segmentation [37] and classification [38] research. The main success criteria of a CNN are high prediction scores and the prevention of overfitting. The latter can be avoided with adjustments to the model architecture and dataset. While the possibility of overfitting can be reduced by using the dropout layer in the model design, the model performance can also be improved [39]. Overfitting can also be prevented through the utilization of transfer learning, as proposed by Weiss et al. [40] and discussed in depth by Shorten and Khoshgoftaar [41]. Moreover, data augmentation techniques, such as rotation, scaling, color jittering, cropping, flipping, translation, noise injection, and contrast change, can be helpful for this purpose. The augmentation techniques can be classified as image manipulation-based and deep learning-based. Generative adversarial networks (GANs) such as CycleGAN and Pix2Pix can be

given as examples of deep learning-based models [41,42]. Image-manipulation-based augmentations mainly introduce radiometric and geometric variations to images.

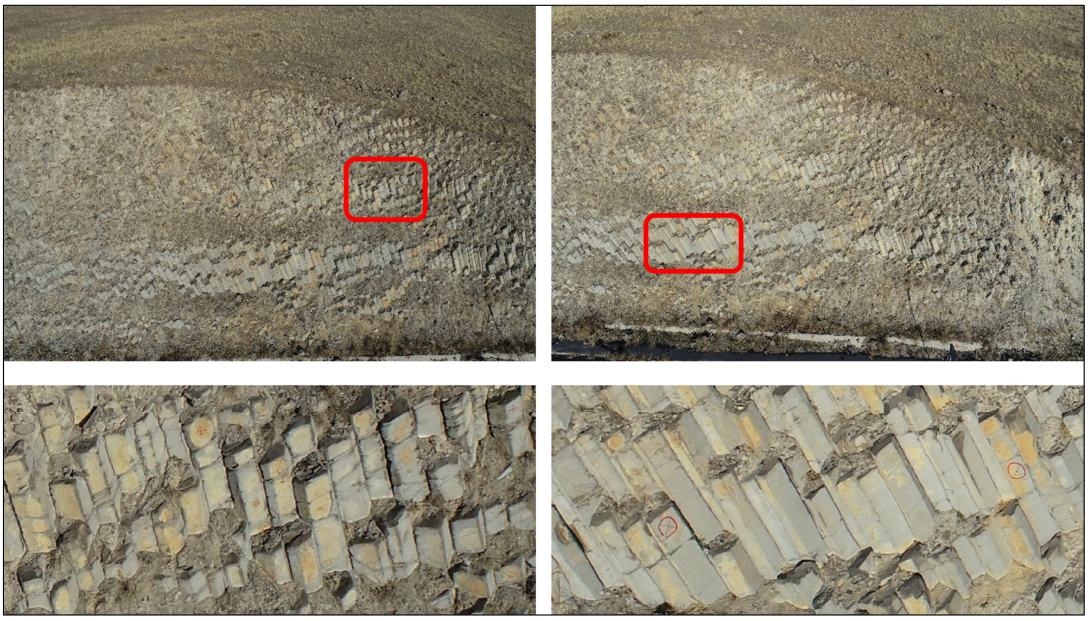


Figure 3. Two images collected in the study area (**above**), and parts of them marked with red polygons (**below**).

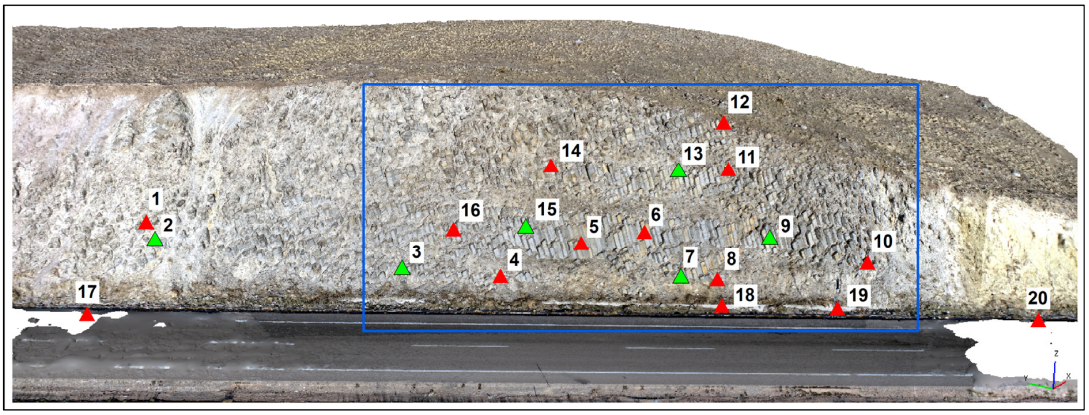
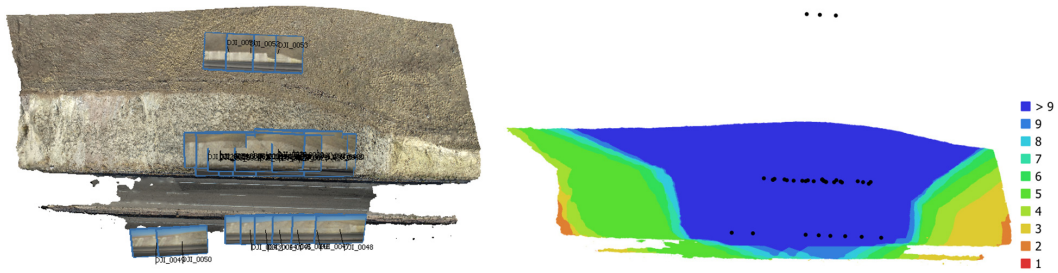


Figure 4. The GCP (red triangles) and CP (green triangles) distribution in the study area. The blue rectangle depicts the model area. The identification number of each point is labelled in the image.



(**a**) (**b**)

Figure 5. Images obtained from the working area: (**a**) image perspective center locations in the study area and (**b**) the number of overlapping images. The black dots in (**b**) also represent the perspective center locations of the images.

In this study, the Albumentations library [43], which was developed in the Python programming language, was used to implement the Radiometric/Geometric (Rad/Geo) augmentation methods. Geometric methods such as flipping, perspective transformation, transposing, shifting, scaling, rotation, affine transformation, resizing, and optical and grid distortion, and radiometric methods such as Gaussian noise injection, hue–saturation–value change, motion blur, sharpening, and histogram equalization were applied to the images and their masks. Since the rock discontinuities were manually delineated on the 3D model, back projection was applied to the vector data to generate the masks (Figure 6). Furthermore, 3D lines with known X, Y, and Z coordinate values of the start and end points were produced after manual delineation. Thus, instead of training a CNN model from a single orthoimage, all 34 raw images were used in the model training, which is called multi-view or alternatively 3D augmentation here. Another advantage of working with the original (raw) images is eliminating the radiometric and geometric issues in orthophotos caused by DSM quality, which is usually poor at the discontinuities due to shadows and illumination problems, as explained by Yalcin et al. [27].

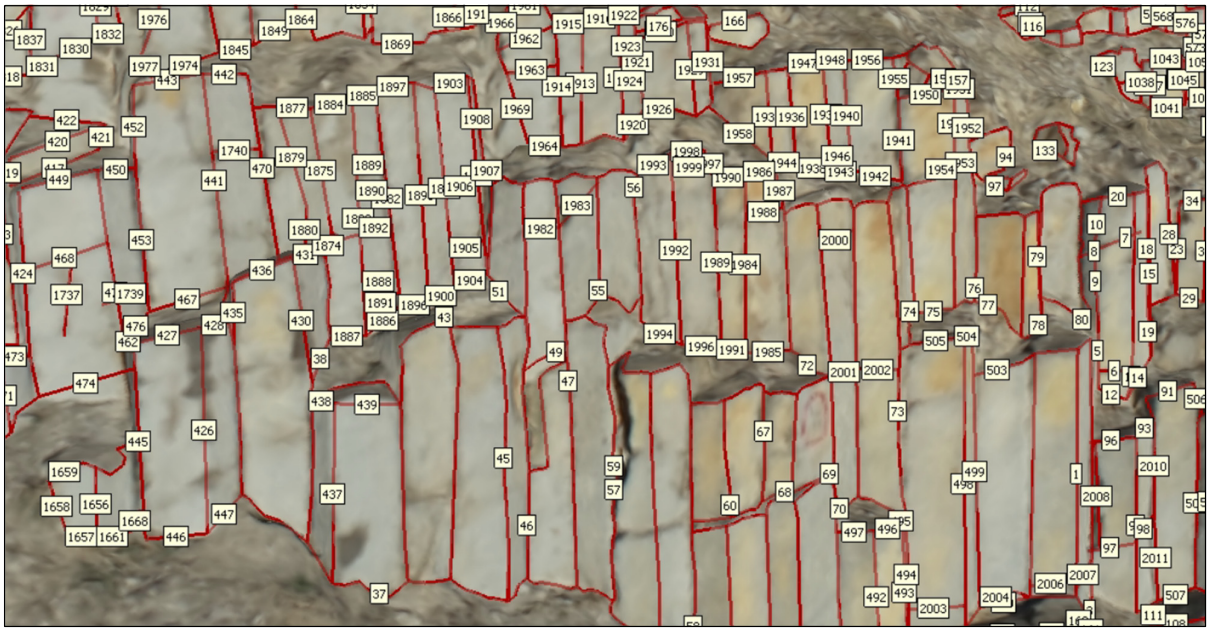


Figure 6. The manually delineated rock discontinuities shown on the 3D model.

The photogrammetric back projection used for the 3D augmentation method is based on collinearity equations (Equations (1) and (2)) [44]. In the frame camera, the principal feature of each ray responsible for generating an image was the perspective center, image point, and corresponding object point aligning along a single line in space. The relationship between the image and object coordinate systems was established with the collinearity equations that were obtained with the help of these lines. The object space coordinates (X_L, Y_L, and Z_L) and orientations (ω, φ, and κ) of the camera perspective center define the exterior orientation, while the focal length (f) and the location of the principal point (x_0, y_0) define the interior orientation.

$$x - x_0 = -f \frac{m_{11}(X - X_L) + m_{12}(Y - Y_L) + m_{13}(Z - Z_L)}{m_{31}(X - X_L) + m_{32}(Y - Y_L) + m_{33}(Z - Z_L)}, \quad (1)$$

$$y - y_0 = -f \frac{m_{21}(X - X_L) + m_{22}(Y - Y_L) + m_{23}(Z - Z_L)}{m_{31}(X - X_L) + m_{32}(Y - Y_L) + m_{33}(Z - Z_L)}, \quad (2)$$

where m denotes the rotation matrix elements created from the omega, phi, and kappa angles of the camera external orientation elements [45].

With the SfM results, the transformation between object space coordinates and image space coordinates can be performed by providing the exterior orientation along with the camera interior orientation parameters into collinearity equations. In this study, back projection was performed using the Agisoft Metashape Application Programming Interface (API) [36], and all lines were transformed into the pixel coordinates of all 34 images from the image space coordinates. To evaluate the contribution of the 3D augmentation, two models were trained using a single image (mono-view model) and all 34 images (multi-view model). The masks with a size of 5472 × 3648 pixels were gridded into 256 × 256-pixel tiles. The test tiles were selected from the southwest part of the 3D model (see Figure 7) to reduce the complexity caused by overlapping images. After the resizing and ensuring no data elimination at the site edges, 146 training images were obtained for a mono image, while 4229 training images were available from the multi-view images (see Table 3). The same test tiles (33 images) were used for both models for the purpose of a proper comparison. In both mono image and multi-view image pairs, 80–20% of the training data were split for training and validation. Furthermore, we evaluated the contribution of Rad/Geo augmentation for both models. Subsequently, the mono (dataset-4) and multi-view (dataset-3) images were used as inputs to the CNN model (see next section). Examples of Rad/Geo augmentation results are shown in Figure 8.

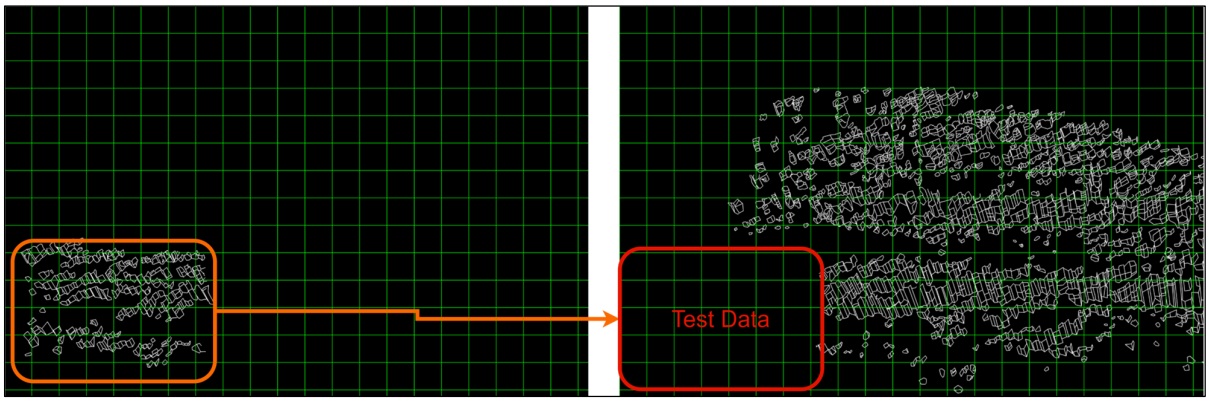


Figure 7. The training (**right**) and test (**left**) data masks, and the gridding scheme (green squares) for one image.

Table 3. The number of image tiles used for training and testing for different augmentation configurations.

Augmentation Configuration	Train	Test
Dataset-1 (3D + Rad/Geo augmentation)	8458	711
Dataset-2 (Mono + Rad/Geo augmentation)	292	33
Dataset-3 (Multi-view augmentation)	4229	711
Dataset-4 (Mono only)	146	33

2.5. The CNN Models

A CNN basically consists of convolutional, activation, pooling, and fully connected layers, which may change with the addition of batch normalization and dropout layers to the network [46]. CNN architecture is similar to ANNs in that it uses feed-forward ANNs in the fully connected layer. The main difference from an ANN is that instead of using perceptrons, CNNs produce feature maps by applying kernels or image filters to the input data. Among the CNN models, U-Net and LinkNet are the architectures used in image segmentation. U-Net was first used in biomedical image segmentation by Ronneberger et al. [47]. The U-Net architecture, which consists of two main stages, the encoder–decoder part, received its name from its similarity to a U shape. U-Net, with its 8 million parameters, emerged as a modification of the fully convolutional network architecture [48]. In this architecture, the fully connected layer is replaced with convolutional layers, providing flexibility

in the input and output dimensions. Thus, the segmentation map, which is the output of the model, is produced in place of the classification score [49]. The LinkNet architecture has 11.5 million parameters. Unlike U-Net, LinkNet directly connects the decoder blocks to the corresponding encoder blocks without making connections to different layers. Thus, it aims to shorten the computation time of the architecture [50]. The encoder and decoder parts of different CNN architectures can be combined with each other. Badrinarayanan et al. [51] developed SegNet using 13 convolutional layers of the VGG16 [52] architecture as the encoder. Furthermore, SegNet, which was developed for image segmentation, achieves better scores compared to other methods. A new CNN model was trained by Ramasamy et al. [53] using the Squeeze and Excitation (SE) ResNet152 architecture as the backbone of the LinkNet architecture. This model achieves a higher score in semantic segmentation than other models. Combinations of the mentioned methods were utilized in this study for training from scratch with local features and transfer learning with domain adaptation, as detailed below.

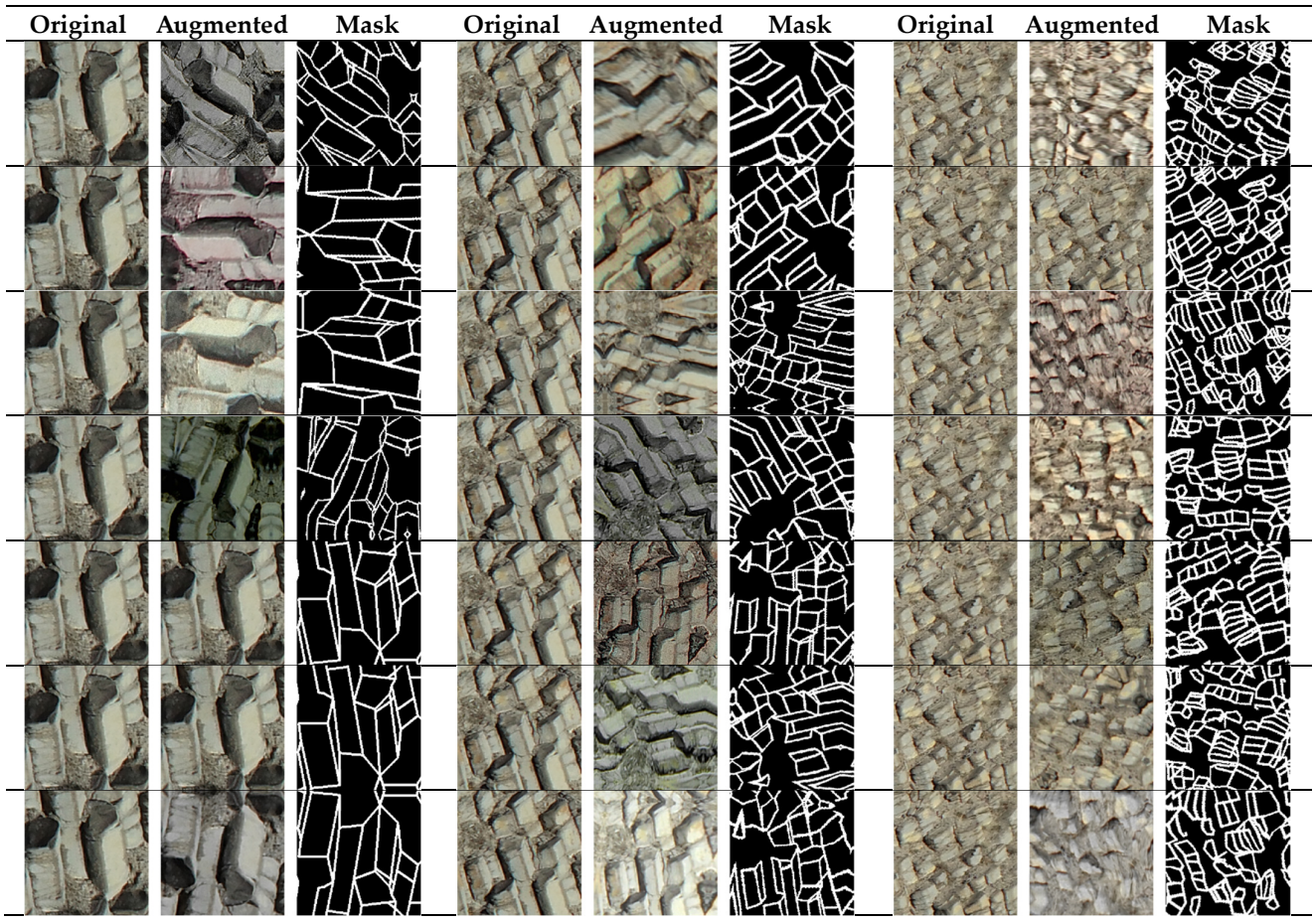


Figure 8. *Cont.*

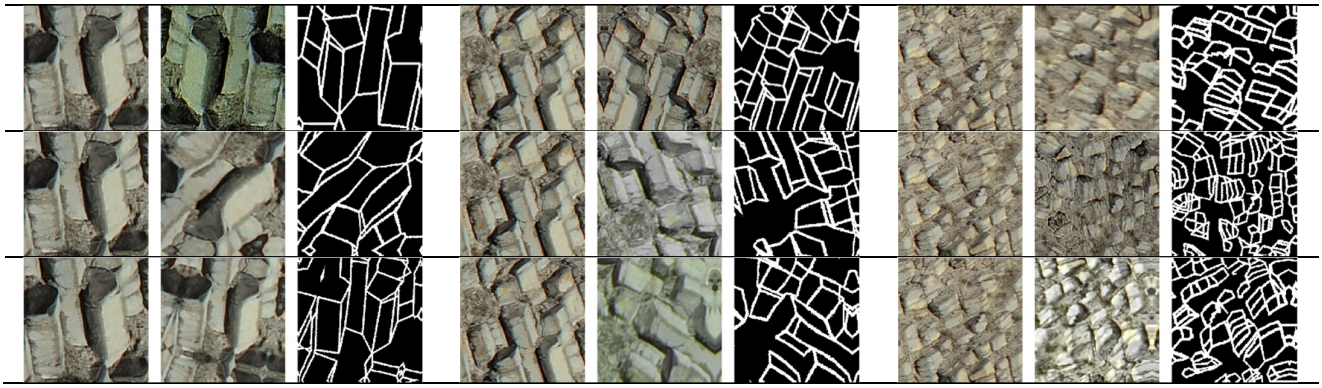


Figure 8. The original (**left**), augmented image (**middle**), and manually delineated mask (**right**) of each column.

2.5.1. Training from Scratch with Local Features

In the first stage of the CNN part of this study, Dataset-1, Dataset-2, Dataset-3, and Dataset-4 (see Table 3) were trained with the modified U-Net architecture that was initialized with random weights by using ResNet-18 as a backbone instead of the encoder part of the U-Net architecture. ResNet [54] achieved first place in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [55] with an error rate of 3.57%. The model was trained on more than three million labeled images. The ResNet architecture is named according to the differences in the number of layers, such as the ResNet-152 model with its 152 layers (which was used in the ImageNet dataset). It is approximately 8 times deeper than VGG models. In addition, the shortcut connections used in the model can prevent gradient reduction problems [56]. The model training parameters are given in Table 4. Adaptive Moment Estimation (Adam) was preferred as the learning algorithm in this study. Adam consists of two different optimizer algorithms, requires less memory, and has a strong learning performance [57]. Rectified Linear Unit (ReLU), which is used as an activation function in CNN architectures, produces results between zero (inclusive) and infinity [58]. Compared to other activation functions on large datasets, ReLU has been observed as being faster and performing better [59]. In the last layer of the CNN model in this study, the sigmoid activation function was used. Binary cross-entropy and dice loss are combined in the model [60]. Thus, the proposed model has a more balanced loss function with higher training stability.

Table 4. Hyperparameters of the CNN trained from scratch with local features.

Model Parameters	
Model	U-Net
Backbone	ResNet-18
Epochs	100
Batch Size	8
Optimizer	Adam
Activation Layer	ReLU and Sigmoid
Loss Function	Binary Cross Entropy and Dice Loss

2.5.2. Transfer Learning with Domain Adaptation

CNNs require large datasets with thousands of images to work effectively, such as ILSVRC [55] and the Canadian Institute for Advanced Research (CIFAR10-100) [61], which poses a challenge for their use. A CNN model trained on large datasets can make predictions on another dataset, or the model can be re-trained using the biases and weights it has learned. This method, which is used to increase model performance and avoid overfitting in cases where there are not enough data, is called transfer learning [40]. The similarity between the source and target domains enables this method to produce strong

results [62]. As a further investigation, a model trained with the crack dataset on the Kaggle platform [63] was re-trained with Dataset-3 (3D augmented) using fine tuning based on transfer learning. Figure 9 shows sample images and corresponding masks from the crack dataset, which contains images of road and pavement surfaces with cracks along with their corresponding masks. This dataset, consisting of 11,298 images with a total size of 448 × 448, was primarily divided into 9603 images for training and 1695 images for testing (85%/15%). This train split was further divided as 90%–10% for training and validation, respectively.

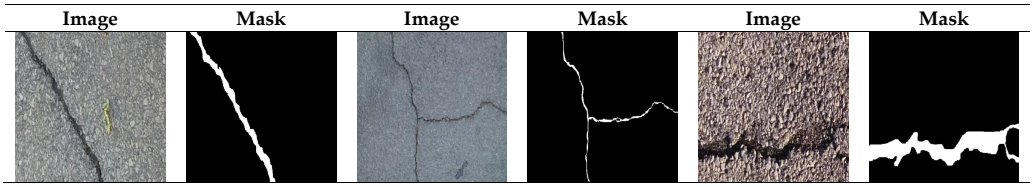


Figure 9. Crack dataset images and their corresponding masks from Kaggle [61].

The 448 × 448-sized images and masks were resized to 256 × 256 using the Albumentations library so that they could be processed with Dataset-3. The reason for using Dataset-3 in this study was that the data size in Dataset-1 with Rad/Geo augmentation was larger than the crack dataset. The crack dataset was trained with two different architectures, U-Net and LinkNet, using three different backbones. The model weights were obtained by fine tuning the pre-trained model for 30 epochs with Dataset-3, which contains rock discontinuities, using the same model configuration. A total of 711 test tiles in Dataset-3 were used to assess the prediction results. The backbones used in this study were SE-ResNet-18, SE-ResNext-50, and VGG16. The SE block reduces the significance of less important feature maps while assigning greater importance to the feature maps that specify the class [46]. The SE block, which was implemented by Hu et al. [64] on the ILSVRC dataset, provided a higher score with small differences at different depths compared to ResNet. Furthermore, high scores were obtained when the SE network was used as a backbone in the brain tumor segmentation area [53]. Table 5 shows the parameters of the model architectures used in this study.

Table 5. Training parameters of the CNN model.

Model Parameters	Crack Dataset for Random Initialization	Dataset-3 for Domain Adaptation
Model	U-Net and LinkNet	U-Net and LinkNet
Backbones	SE-ResNet-18, SE-ResNext-50, VGG16	SE-ResNet-18, SE-ResNext-50, VGG16
Epochs	100	30
Batch Size	8	8
Optimizer	Adam	Adam
Activation Layer	ReLU and Sigmoid	ReLU and Sigmoid
Loss Function	Binary Cross Entropy and Dice Loss	Binary Cross Entropy and Dice Loss

2.6. Validation

The CNN models were assessed quantitatively with the F1-score and Jaccard index (Intersection over Union—IoU), while the root mean square error (RMSE) obtained from the CPs in the bundle block adjustment was employed to evaluate the photogrammetric point positioning accuracy. The F1-score, which is calculated from the harmonic mean of recall and precision ratios, is widely used for evaluating CNN models. The F1-score is also referred to as the Dice similarity coefficient [65]. In the evaluation of CNN models conducted on image segmentation, the Jaccard index was found to be more reliable [66]. Equations (3) and (4) were used to compute the F1-score and the Jaccard index.

$$\mathrm{F1-score} = \frac{2\mathrm{TP}}{2\mathrm{TP} + \mathrm{FP} + \mathrm{FN}}, \quad (3)$$

$$\text{Jaccard index} = \frac{TP}{TP + FP + FN}, \tag{4}$$

where true negatives (TN) and true positives (TP) represent the number of correctly classified negative and positive samples, respectively. Also, the false negatives (FN) and false positives (FP) represent the number of misclassified positive and negative samples [56]. In Equations (5) and (6), the RMSE formulas are shown for the X, Y, and Z coordinates. For each axis, the estimated coordinate value resulting from the bundle block adjustment was subtracted from the manually measured (accepted correct) value for the CPs.

$$RMSE_X = \pm\sqrt{\frac{\sum_{i=1}^{n}(X_E - X_M)^2}{n}},\ RMSE_Y = \pm\sqrt{\frac{\sum_{i=1}^{n}(Y_E - Y_M)^2}{n}},\ RMSE_Z = \pm\sqrt{\frac{\sum_{i=1}^{n}(Z_E - Z_M)^2}{n}}, \tag{5}$$

$$RMSE_{XYZ} = \pm\sqrt{RMSE_X^2 + RMSE_Y^2 + RMSE_Z^2}, \tag{6}$$

where X_E, Y_E, and Z_E represent the estimated coordinates, while X_M, Y_M, and Z_M represent the values from the manual measurements. The "n" in the equations defines the number of points.

3. Results

Here, we present our results both quantitatively and qualitatively. The qualitative assessments are based on a visual inspection of the predicted discontinuities with respect to the ground truth. The quantitative assessments involve performance measurements (F1-score and Jaccard index) of the CNN models based on different data augmentation scenarios, as well as comparisons of the training from scratch and transfer learning with domain adaptations.

3.1. Qualitative Assessments

The datasets described in Table 3 were trained using U-Net architecture with ResNet-18 as the backbone, and it was tested on 33 tiles using the learned model weights. The model prediction results for Dataset-1 (multi-view images with 3D + Rad/Geo augmentation), Dataset-2 (mono image with Rad/Geo augmentation), Dataset-3 (multi-view images with 3D augmentation), and Dataset-4 (mono image only) are depicted in Figure 10. Based on visual assessments, it is evident that 3D augmentation is effective in detecting rock discontinuities. Furthermore, the Rad/Geo augmentation techniques (Dataset-1) significantly contributed to enhancing the results by augmenting the dataset with various radiometric and geometric modifications.

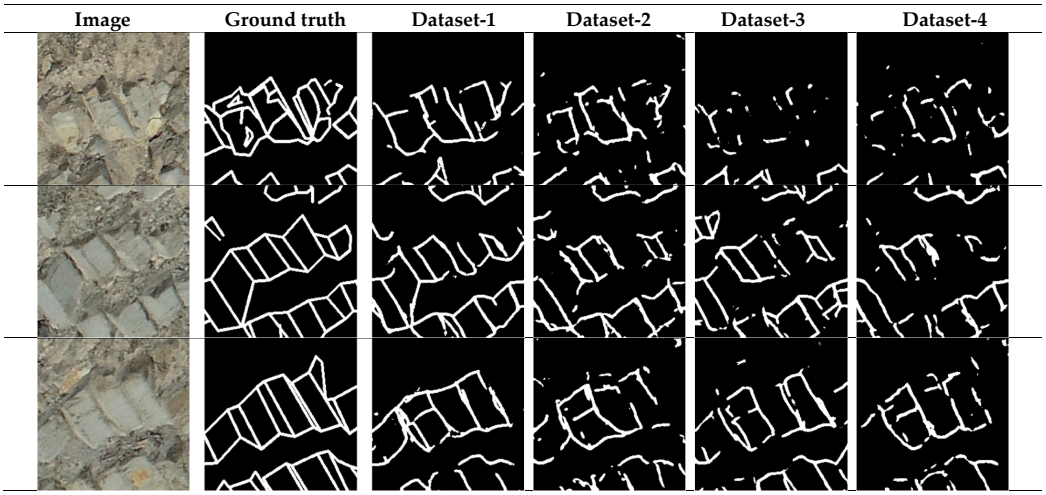


Figure 10. *Cont.*

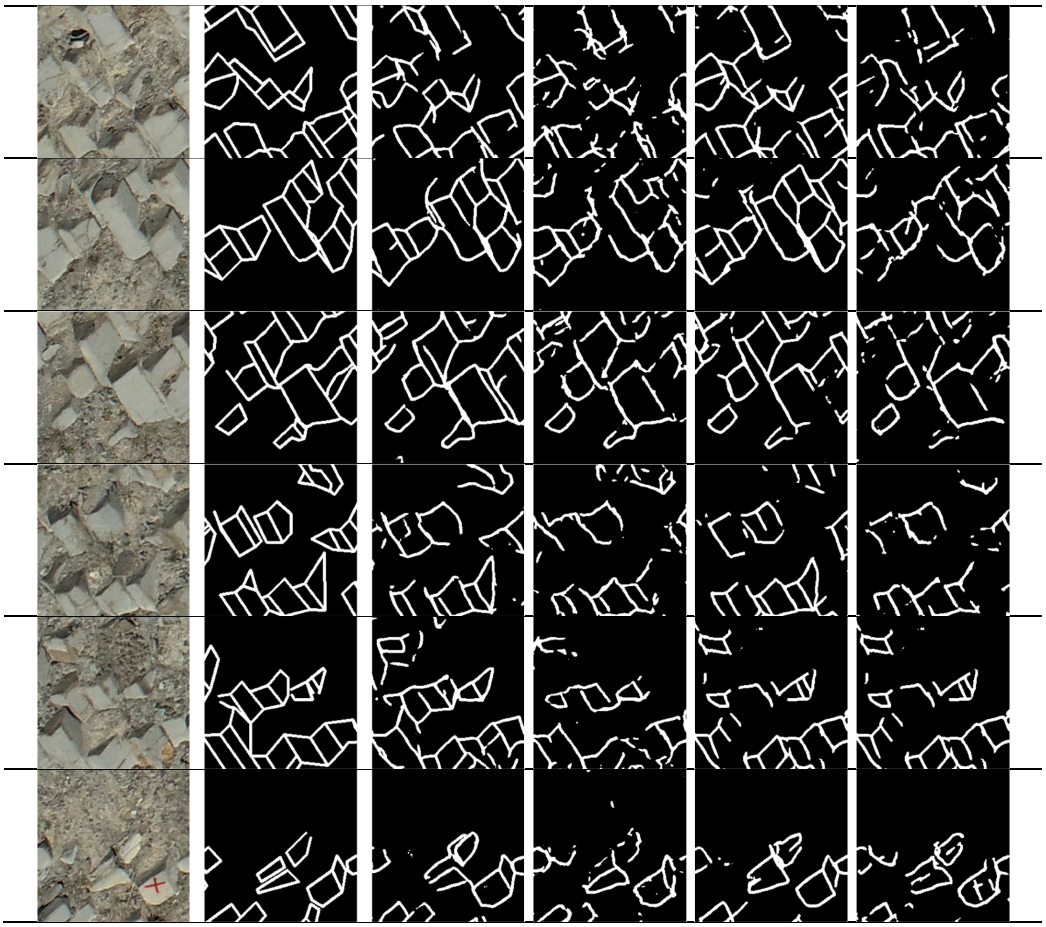


Figure 10. The effect of 3D augmentation on the model prediction results.

The transfer learning approach employed for rock discontinuity detection involved using the crack dataset as the source for random initialization and Dataset-3 as the target for domain adaptation. In this study, U-Net and LinkNet architectures were modified with different backbone types. By using the model weights obtained through the training process, predictions were made on 711 test data points. The model estimation results of the six different CNN models with transfer learning and domain adaptation are provided in Figure 11. The U-Net + SE-ResNeXt-50 model was found to yield better predictions. For further results, please refer to Figure A2 in Appendix A.

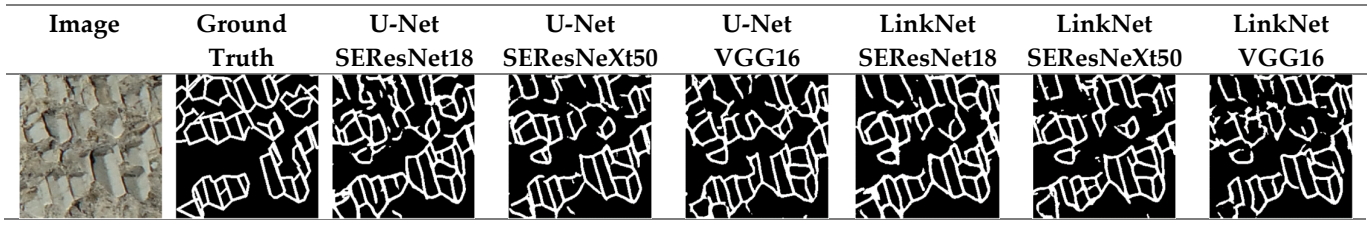


Figure 11. *Cont.*

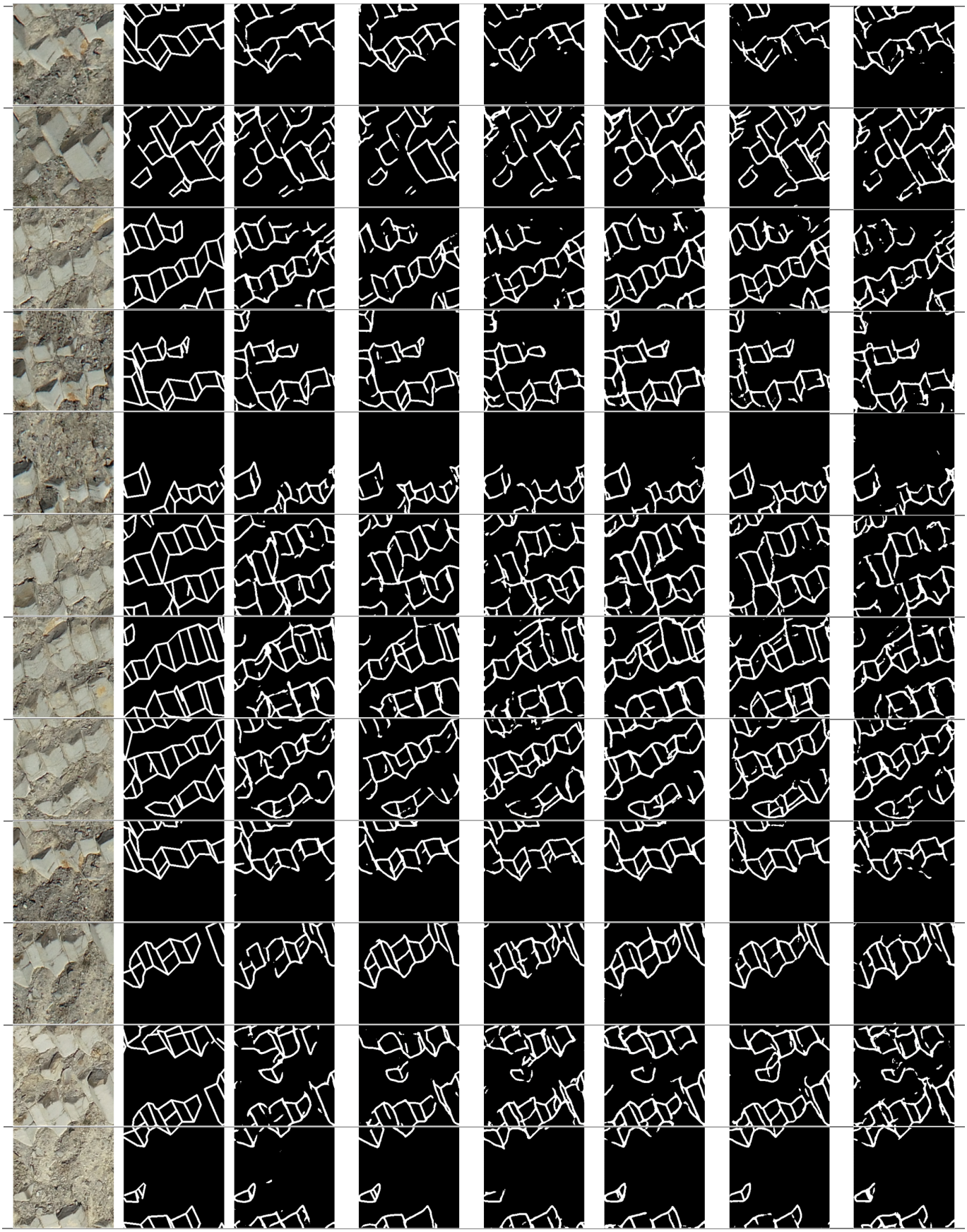


Figure 11. Prediction results of the six different CNN models used for transfer learning and domain adaptation.

3.2. Quantitative Results

The accuracy of the 3D model generated through photogrammetric processing was assessed by comparing the coordinate differences between the CPs measured in the field and the points obtained from the bundle block adjustment method. The RMSE values calculated using CPs are provided in Table 6. An RMSE value of 8.95 mm indicates that the 3D model has a high point coordinate accuracy. The error ellipses that were obtained from a second run of bundle block adjustment by utilizing all GCPs as the control are shown in Figure 12.

Table 6. The RMSE values calculated from the independent CPs.

Check Points	RMSEx (mm)	RMSEy (mm)	RMSEz (mm)	RMSExyz (mm)	Image Pixel
2	2.42	0.40	3.40	4.19	0.32
3	−6.31	−2.29	−2.52	7.18	0.32
7	12.72	4.41	−7.25	15.29	0.21
9	6.69	−0.77	2.18	7.08	0.28
13	7.42	−1.23	−6.89	10.20	0.17
15	−0.44	−4.28	−2.26	4.86	0.28
Total RMSE	7.16	2.75	4.62	8.95	0.26

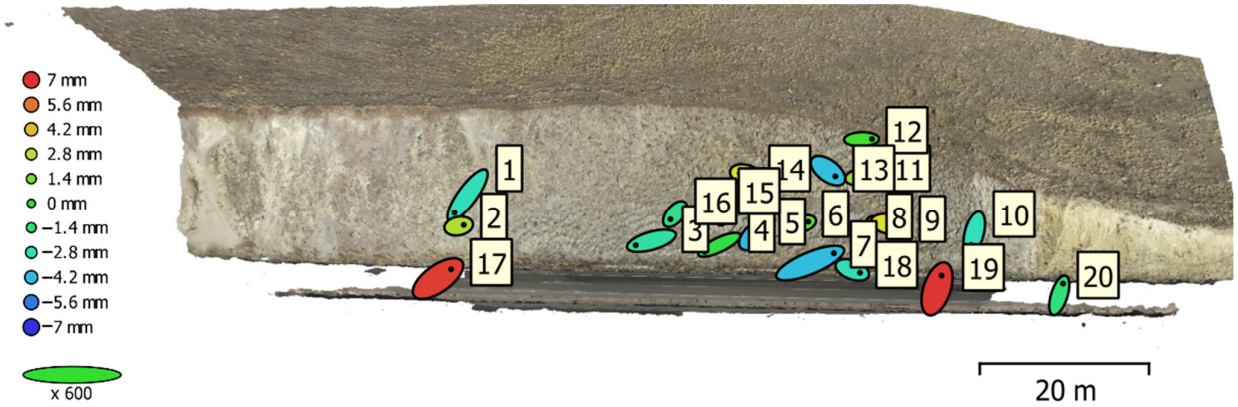


Figure 12. The errors and ellipses of the calculated GCP values.

The results obtained from the CNN model trained from scratch and the transfer-learned and domain-adapted CNN models are presented in Tables 6 and 7, respectively. The 3D augmentation method demonstrated a high score, as depicted in Table 7. The score differences between the results were small since the comparison is based on pixels, and the number of black (negative) values was greater (imbalanced). Thus, the Jaccard index provided a more realistic interpretation of the results. The transfer learning and domain adaptation results given in Table 8 showed that the U-Net + SE-ResNeXt-50 model yielded higher performance scores than the others.

Table 7. Evaluation of the 3D augmentation effect with measurements.

	Dataset-1	Dataset-2	Dataset-3	Dataset-4
F1-Score	0.914	0.910	0.917	0.916
Jaccard Index	0.842	0.834	0.847	0.845

Table 8. Comparison of the transfer-learned CNNs with domain adaptation.

	U-Net SEResNet18	U-Net SEResNeXt50	U-Net VGG16	LinkNet SEResNet18	LinkNet SEResNeXt50	LinkNet VGG16
F1-Score	0.891	0.906	0.897	0.888	0.898	0.894
Jaccard Index	0.803	0.829	0.813	0.799	0.815	0.808

In this study, the predictions were used to calculate the manually defined scan-line distances (Figure 13). The distances were calculated from the 3D coordinates that were obtained from the transformation of image coordinates to the object space coordinates based on Equations (1) and (2). Afterward, they were validated from the manually produced ground truth, which was obtained from the point cloud produced in the study area. In Table 9, the discrepancies ranging from 0.5 cm to 1.5 cm are presented. The differences mainly stemmed from the image orientation accuracy and the image and point cloud measurement accuracy sourced from a manual identification of the discontinuity lines (and intersection points).

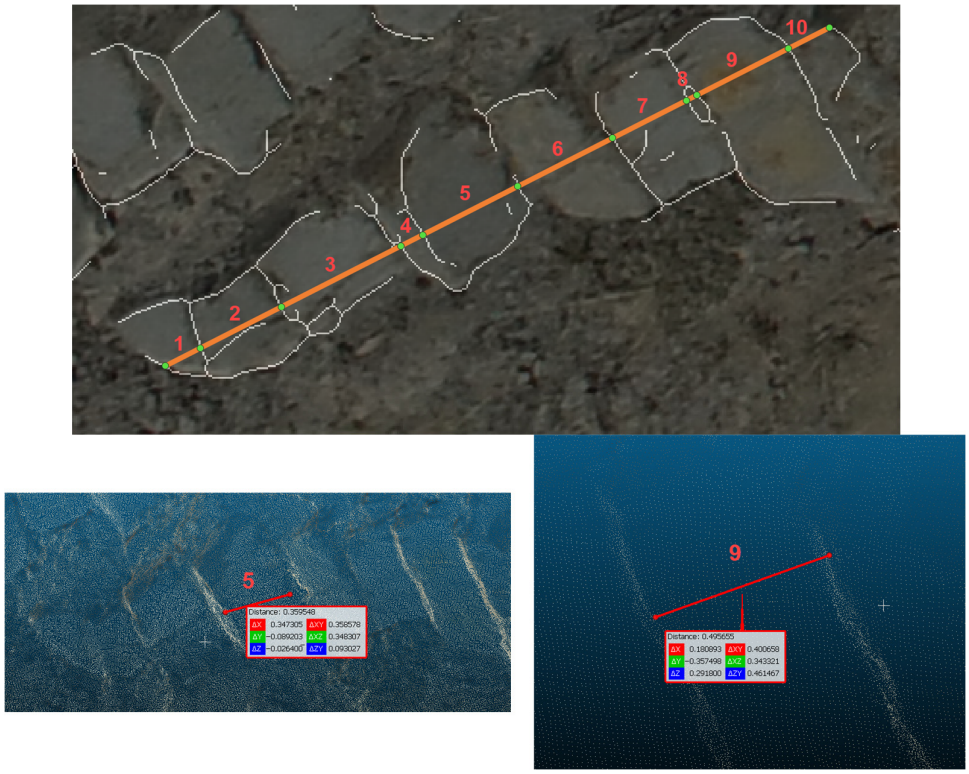


Figure 13. Line and intersection points whose length was calculated on the image taken from the study area (**above**) and from samples of the control measurements (**below**).

Table 9. Comparison of the processing results and control measurements.

Line Segments	Processing Results (cm)	Control Measurements (cm)	Differences (cm)
1	15.36	15.15	0.21
2	33.87	33.30	0.57
3	48.70	48.49	0.21
4	21.46	22.25	−0.79

Table 9. *Cont.*

5	35.42	35.95	−0.53
6	53.55	52.81	0.74
7	34.93	33.45	1.48
8	11.17	10.87	0.30
9	50.84	49.56	1.28
10	16.68	17.63	−0.95
		RMSE	0.82

4. Discussions

Characterizing rock masses in geotechnical and engineering geological studies through conventional fieldwork is time-consuming and can pose life-threatening risks. Accessibility to the site presents another major obstacle to the successful realization of such projects. LiDAR and optical data can reduce the duration of fieldwork and mitigate access problems. Ozturk et al. [22] demonstrated the usability of smartphone images in reducing costs and eliminating the need for complex equipment. However, depending on the site characteristics, terrestrial imaging may be unsuitable, and RPAS platforms ensure remote data collection, typically with optical cameras. Yet, manual interpretation of the data from scan-line surveys to detect and measure discontinuities in rock masses has the major drawbacks of requiring expertise and being time-consuming. On the other hand, while deep learning methods, particularly CNNs for image segmentation and classification, provide promising results for discontinuity detection (see Yalcin et al. [27,28]), they are also limited by the manual labeling required to obtain the necessary amount of data for learning the model parameters. Data augmentation techniques and transfer learning approaches can help overcome this obstacle.

In this study, we proposed a multi-view image augmentation approach for detecting discontinuities in rock masses with a CNN that was trained from scratch with local features and was also trained with transfer learning from the crack dataset after domain adaptation. The images were taken in stereo fashion with an RPAS in a part of Bala, Ankara. Based on the 3D model of the rock blocks through photogrammetric processing, manual delineation can be utilized to identify rock discontinuities. Most rock mass discontinuity determination studies are based on 3D point clouds, either from LiDAR sensors or photogrammetric DSMs (albeit to a lesser extent). The main disadvantage of point cloud-based approaches is the increase in data size in the algorithms, which leads to higher computational cost.

On the other hand, using conventional edge operators such as Canny and Sobel for image-based discontinuity detection suffers from a high level of noise in identified edges due to the color variations and textural characteristics of rock surfaces (see Lee et al. [25]). Deep learning models, especially CNNs, were used for this purpose by Yalcin et al. [27,28]. The preliminary results presented by Yalcin et al. [27] in the Kızılcahamam/Güvem Basalt Columns Geosite show that using orthophotos with data augmentation as the input to a U-Net architecture yielded an F1-score of 58%. The images were taken with an off-the-shelf camera. In the aforementioned study, the training dataset was produced manually from orthophotos. However, image artifacts were visible, especially at the discontinuities, which decreased the accuracy significantly. Yalcin et al. [27] emphasized the importance of data preparation and the potential of using original (raw) images instead of orthoimages. Yalcin et al. [28] evaluated the transfer learning capability of a modified version of the U-Net architecture with ResNet-18, which was also used on the Kızılcahamam/Güvem site. The proposed model also learned from the crack dataset, and it yielded a Jaccard index of 88% and was able to overcome the requirement of a high amount of data. However, it also indicated the need for investigating different rock mass types, as a single type of basalt is not sufficient for drawing a firm conclusion. In addition, although the efficiency of transfer learning without domain adaptation was also investigated by predicting the discontinuities

after training the model weights with the crack dataset only, the results were very poor and therefore these preliminary investigation results were not presented in this paper.

Here, we present a new type of data augmentation (called multi-view or 3D augmentation) for multi-view imaging, which is designed to learn from a small amount of data locally using perspective geometric principles (i.e., photogrammetric back projection). The proposed data augmentation method overcomes the need for working with point clouds, as well as the poor image quality of orthophotos of discontinuities that is caused by illumination and shadows. The proposed method is expected to contribute to the detection of rock discontinuities with deep learning by overcoming the small data size barrier, which was pursued as manual delineation is time-consuming and requires a high level of expertise. We also employed transfer learning with domain adaptation to obtain pre-computed weights to avoid overfitting due to factors such as texture differences, brightness, complexity, and topographic variations on rock surfaces. These factors also cause difficulties in the manual delineation of discontinuities. In a study by Zhang et al. [67] on rockfall areas, rock discontinuities were manually delineated under expert supervision. In that study, it was emphasized that point cloud-based automatic algorithms are insufficient for detecting rock discontinuities. Thus, the use of CNN models with pre-computed weights can also support the delineation of local data to improve the process and reduce the requirement of data for learning on complex rock mass types.

In our study, besides the efficiency of multi-view augmentation, we evaluated the transfer learning capability of two different CNN architectures on three different backbones with domain adaptation and using a small amount of data. Although the model trained with the crack dataset was not successful, fine tuning by using local data with as small a data size as 34 multi-view images and the 3D augmentation technique ensured a high level of accuracy, similar to when U-Net with ResNet-18 yielded an F1-score of 91.7%. Further, the U-Net and LinkNet architectures tested with SE-ResNet-18, SE-ResNeXt-50, and VGG16 backbones with a crack dataset and 3D augmentation yielded F1-scores between 88.8% and 90.6%. These are the highest performance values observed in the literature.

The performance evaluation of the CNN models for rock discontinuities was also a difficult task, and it must be handled differently from aerial/satellite image classification or segmentation studies. The pixel-based F1-score and Jaccard index can be misleading in detecting discontinuities. According to Yalcin et al. [27], although the CNN model yielded a rather low F1-score, it was seen that some discontinuities were predicted correctly. However, the model was overfit due to an insufficient amount of data. The proposed augmentation techniques are more suitable when compared to existing studies [25,27,28,68,69], and they have also resulted in significantly higher performance measures. Byun et al. [68] obtained a Jaccard index of 38% for extracting discontinuity lines with a CNN. Asadi et al. [69] obtained an F1-score of 84%. Lee et al. [25] obtained a Jaccard index of 62% for a similar purpose. With the development of geometry-based evaluation criteria, the model scores for rock discontinuities can gain more reliability [70]. When visually inspecting the model predictions, it was observed that discontinuities were detected as partial lines. Future studies focusing on line completeness should be conducted for the prediction of discontinuities. It is expected that image-based CNNs can be even more successful in line completeness when layers containing different features such as height are added in addition to RGB image layers.

In this study, we included only the RGB bands in the model training. Recent geospatial machine learning applications have integrated further geometric properties such as elevation [29], which can also be proposed for discontinuity detection with CNNs. Examples of such features are hill shade representations of surfaces, depth maps, and different spectral bands.

On the other hand, the properties of discontinuities show a wide variety. In general, systematic joints are formed by tectonic activity in all types of rocks, metamorphism in metamorphic rocks, and the cooling process in igneous rocks. In addition, the mineral content and texture of rocks also affect some of the properties of all types of discontinuities,

such as joints. Consequently, it is almost impossible to generalize discontinuity properties, and training or re-training with local features is recommended.

5. Conclusions and Future Work

In this study, we proposed a novel data augmentation technique based on multi-view imaging and perspective geometry for detecting rock mass discontinuities to reduce the amount of data required for training a CNN. We also applied transfer learning with domain adaptation to avoid overfitting and evaluated six different CNN architectures for this purpose. We demonstrated the results using aerial images of the studied basalts, which were taken from an RPAS over Bala, Ankara (Türkiye). A total of 34 multi-view images were collected, and the discontinuity lines were manually delineated on a 3D point cloud. The lines were back-projected onto the raw images to increase the amount of data. Further, radiometric and geometric augmentation methods were also experimented with, and the use of 3D augmentation was found to be sufficient for the studied case. The CNN trained from scratch with local features based on U-Net with ResNet-18 yielded an F1-score of 91.7%. The U-Net and LinkNet architectures were tested with different backbones, such as SE-ResNet-18, SE-ResNeXt-50, and VGG16, with a crack dataset populated with road and building fractures; in addition, after domain adaptation, multi-view images were used (with 3D augmentation). The highest performance was achieved with U-Net + SE-ResNeXt-50 with an F1-score of 90.6%. Although the results were found to be comparable, it is recommended to use transfer learning over training from scratch with local features in different sites with a small amount of labeled data. This is because this type of approach can be expected to prevent overfitting depending on the data size and site characteristics. Yet, 3D augmentation was proven to be successful and yielded the highest performance scores for rock mass discontinuity determination.

As the basis for future research, other rock mass types should be classified to comprehend the limitations of the proposed method. In addition, as the evaluations were based on binary pixel information, line-based measures could be utilized to improve the accuracy and reliability of the assessments. Further research is also needed to ensure the line completeness of the detected discontinuities. Yet, the proposed study revealed the potential of photogrammetric image analysis for rock mass characterization, and the discontinuities that were detected in raw images could be transformed to object space for obtaining further rock mass parameters such as dip and dip orientation. Additionally, different rock types should be employed to develop the methodology proposed herein.

As a final remark, use of photogrammetric methods instead of conventional scan-line surveys may decrease the possible errors and biases.

Author Contributions: Conceptualization and validation: Candan Gokceoglu and Sultan Kocaman; methodology, Ilyas Yalcin, Recep Can, and Sultan Kocaman; data curation, Ilyas Yalcin; software, Ilyas Yalcin and Recep Can; formal analysis and supervision, Sultan Kocaman; writing—original draft preparation, Ilyas Yalcin; writing—review and editing, Sultan Kocaman and Candan Gokceoglu. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data of this study are available upon reasonable request.

Acknowledgments: This study is part of the PhD thesis research of Ilyas Yalcin. The authors thank Mehmet Dogruluk for his support in performing field measurements.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

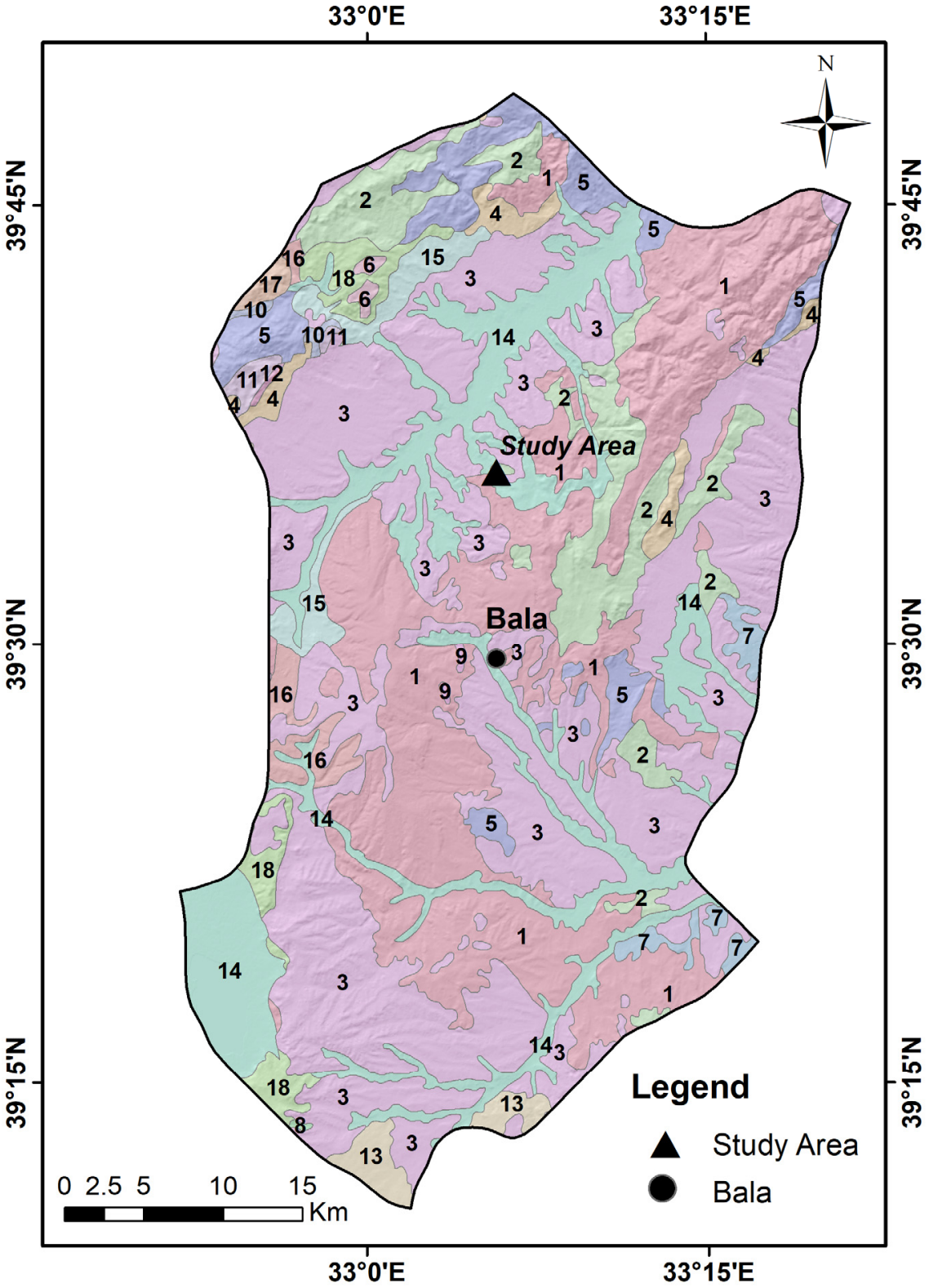


Figure A1. Lithological unit map of Bala town (see Table A1 for explanations) [33].

Table A1. Lithological units in Bala town.

ID	Lithological Unit	Area (km^2)
1	Dizilitaslar formation: Conglomerate, sandstone, and claystone	504.66
2	Evciler Volcanics: Basaltic lava and pyroclastics	138.29
3	Ic anadolu group: Undifferentiated middle miocene–pliocene continental sediments	740.43
4	Cavuslu volcanics: Basalt and spilitic basalt	28.67
5	Incik formation: Conglomerate, sandstone, mudstone, and gypsum	89.52
6	Limestone blocks of Triassic age	3.96
7	Orta Anadolu granitoids: Granite, granodiorite, and monzonite	18.83
8	Golbasi formation: Conglomerate, sandstone, and mudstone	0.61
9	Basalt	1.43
10	Sekili evaporite member: Gypsum, anhydrite, mudstone, and sandstone	3.40
11	Barakli formation: Continental conglomerate, sandstone, and mudstone	7.25
12	Artova ophiolitic melange: Serpantinite, harzburgite, dunite, gabbro, diabase, radiolarite, chert, and limestone	1.25
13	Cayraz formation: Conglomerate, sandstone, claystone, limestone, and mudstone	22.90
14	Alluvium	287.47
15	Kumartas formation: Conglomerate, sandstone, and siltstone	35.89
16	Hancili formation: Sandstone, siltstone, marl, clayey limestone, tuff, and gypsum	23.96
17	Tohumlar volcanics: Dacite, rhyolite, and pyroclastics	5.69
18	Ophiolitic melange	41.10

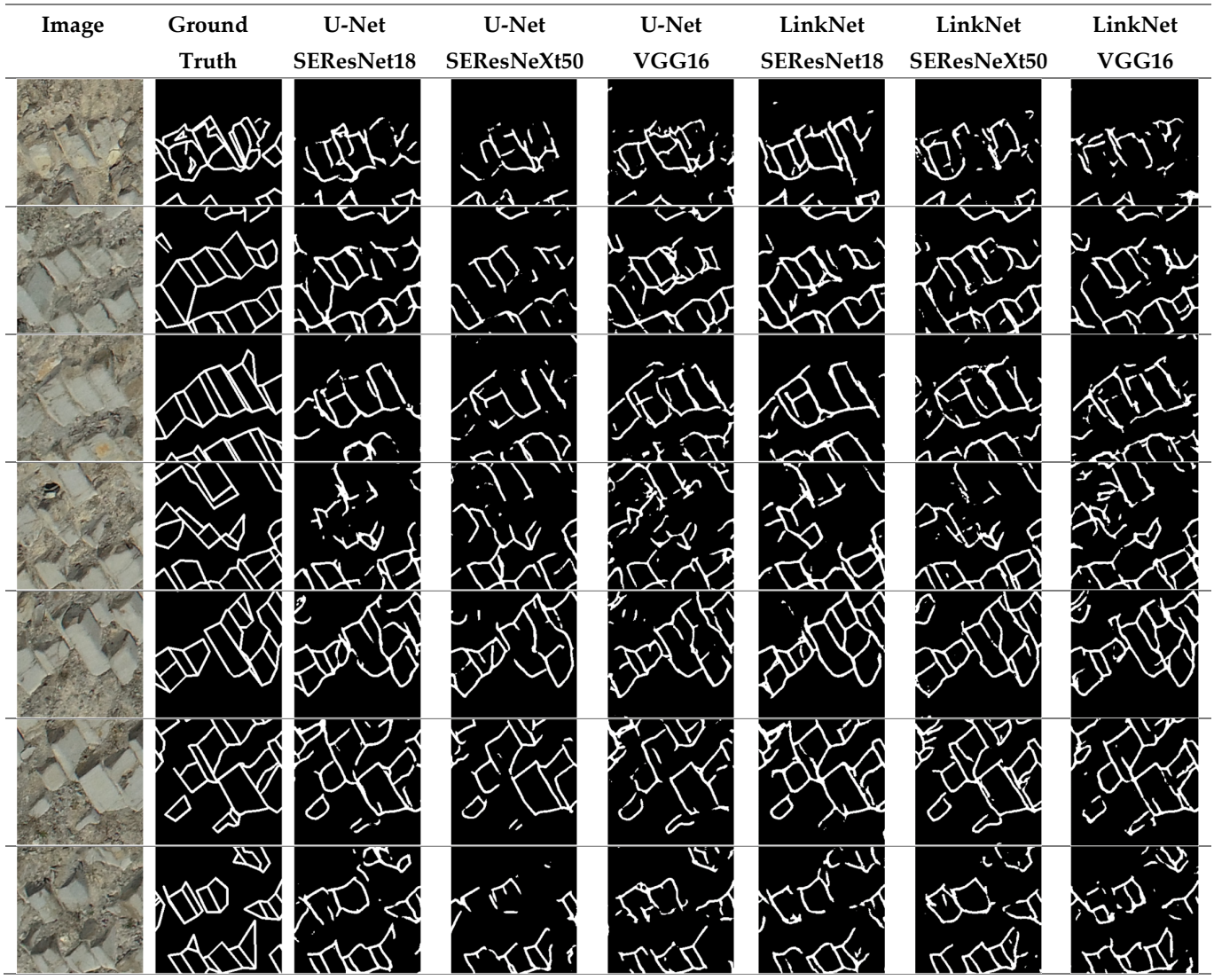


Figure A2. *Cont.*

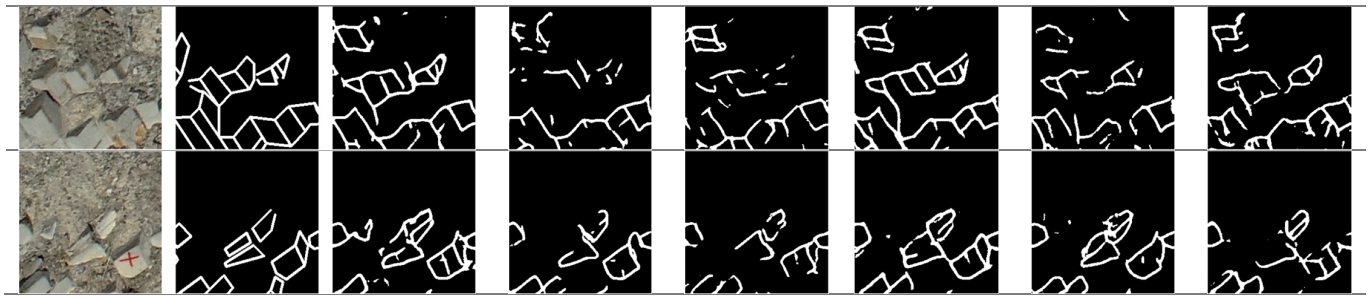


Figure A2. Further results of the transfer learning for qualitative assessment.

References

1. Bostanci, H.T.; Alemdag, S.; Gurocak, Z.; Gokceoglu, C. Combination of discontinuity characteristics and GIS for regional assessment of natural rock slopes in a mountainous area (NE Turkey). *Catena* **2018**, *165*, 487–502. [CrossRef]
2. AFAD. AFAD (Disaster and Emergency Management Pesidency). Available online: https://www.afad.gov.tr/kurumlar/afad.gov.tr/e_Kutuphane/Istatistikler/2022-Yili-Doga-Kaynakli-Olay-Istatistikleri.pdf (accessed on 31 July 2023).
3. Gokceoglu, C. 6 February 2023 Kahramanmaraş—Türkiye Earthquakes: A General Overview. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2023**, *XLVIII-M-1-2023*, 417–424. [CrossRef]
4. Bieniawski, Z.T. *Engineering Rock Mass Classifications: A Complete Manual for Engineers and Geologists in Mining, Civil, and Petroleum Engineering*; John Wiley & Sons: New York, NY, USA, 1989.
5. Hoek, E.; Brown, E.T. Practical estimates of rock mass strength. *Int. J. Rock Mech. Min. Sci.* **1997**, *34*, 1165–1186. [CrossRef]
6. Barton, N. Some new Q-value correlations to assist in site characterisation and tunnel design. *Int. J. Rock Mech. Min. Sci.* **2002**, *39*, 185–216. [CrossRef]
7. ISRM. *Rock Characterization, Testing and Monitoring: ISRM Suggested Methods*; Pergamon Press: Oxford, UK, 1981.
8. Zhang, Q.; Huang, X.; Zhu, H.; Li, J. Quantitative assessments of the correlations between rock mass rating (RMR) and geological strength index (GSI). *Tunn. Undergr. Space Technol.* **2019**, *83*, 73–81. [CrossRef]
9. Priest, S.D.; Hudson, J.A. Estimation of discontinuity spacing and trace length using scanline surveys. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* **1981**, *18*, 183–197. [CrossRef]
10. Pahl, P.J. Estimating the mean length of discontinuity traces. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* **1981**, *18*, 221–228. [CrossRef]
11. Watkins, H.; Bond, C.E.; Healy, D.; Butler, R.W.H. Appraisal of fracture sampling methods and a new workflow to characterise heterogeneous fracture networks at outcrop. *J. Struct. Geol.* **2015**, *72*, 67–82. [CrossRef]
12. ISRM. International society for rock mechanics commission on standardization of laboratory and field tests: Suggested methods for the quantitative description of discontinuities in rock masses. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* **1978**, *15*, 319–368. [CrossRef]
13. Battulwar, R.; Zare-Naghadehi, M.; Emami, E.; Sattarvand, J. A state-of-the-art review of automated extraction of rock mass discontinuity characteristics using three-dimensional surface models. *J. Rock Mech. Geotech. Eng.* **2021**, *13*, 920–936. [CrossRef]
14. Riquelme, A.; Cano, M.; Tomás, R.; Abellán, A. Identification of Rock Slope Discontinuity Sets from Laser Scanner and Photogrammetric Point Clouds: A Comparative Analysis. *Procedia Eng.* **2017**, *191*, 838–845. [CrossRef]
15. Snavely, N.; Seitz, S.M.; Szeliski, R. Modeling the World from Internet Photo Collections. *Int. J. Comput. Vis.* **2008**, *80*, 189–210. [CrossRef]
16. Salvini, R.; Vanneschi, C.; Coggan, J.S.; Mastrorocco, G. Evaluation of the Use of UAV Photogrammetry for Rock Discontinuity Roughness Characterization. *Rock Mech. Rock Eng.* **2020**, *53*, 3699–3720. [CrossRef]
17. Kong, D.; Saroglou, C.; Wu, F.; Sha, P.; Li, B. Development and application of UAV-SfM photogrammetry for quantitative characterization of rock mass discontinuities. *Int. J. Rock Mech. Min. Sci.* **2021**, *141*, 104729. [CrossRef]
18. Wang, W.; Zhao, W.; Chai, B.; Du, J.; Tang, L.; Yi, X. Discontinuity interpretation and identification of potential rockfalls for high-steep slopes based on UAV nap-of-the-object photogrammetry. *Comput. Geosci.* **2022**, *166*, 105191. [CrossRef]
19. Song, J.; Du, S.; Yong, R.; Wang, C.; An, P. Drone Photogrammetry for Accurate and Efficient Rock Joint Roughness Assessment on Steep and Inaccessible Slopes. *Remote Sens.* **2023**, *15*, 4880. [CrossRef]
20. Riquelme, A.J.; Abellán, A.; Tomás, R.; Jaboyedoff, M. A new approach for semi-automatic rock mass joints recognition from 3D point clouds. *Comput. Geosci.* **2014**, *68*, 38–52. [CrossRef]
21. Liu, H.; Yao, M.; Xiao, X.; Xiong, Y. RockFormer: A U-shaped transformer network for Martian rock segmentation. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 1–16. [CrossRef]
22. Ozturk, H.S.; Kocaman, S.; Gokceoglu, C. A low-cost approach for determination of discontinuity orientation using smartphone images and application to a part of Ihlara Valley (Central Turkey). *Eng. Geol.* **2019**, *254*, 63–75. [CrossRef]
23. Chen, K.; Jiang, Q. A non-contact measurement method for rock mass discontinuity orientations by smartphone. *J. Rock Mech. Geotech. Eng.* **2022**, *15*, 2892–2900. [CrossRef]

24. An, P.; Fang, K.; Zhang, Y.; Jiang, Y.; Yang, Y. Assessment of the trueness and precision of smartphone photogrammetry for rock joint roughness measurement. *Measurement* **2022**, *188*, 110598. [CrossRef]
25. Lee, Y.K.; Kim, J.; Choi, C.S.; Song, J.J. Semi-automatic calculation of joint trace length from digital images based on deep learning and data structuring techniques. *Int. J. Rock Mech. Min. Sci.* **2022**, *149*, 104981. [CrossRef]
26. Qiu, L.; Tang, L.; Zhong, R. Toward the recognition of spacecraft feature components: A new benchmark and a new model. *Astrodynamics* **2022**, *6*, 237–248. [CrossRef]
27. Yalcin, I.; Can, R.; Kocaman, S.; Gokceoglu, C. A CNN Architecture for Discontinuity Determination of Rock Masses with Close Range Images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2022**, *XLIII-B2-2022*, 1101–1108. [CrossRef]
28. Yalcin, I.; Can, R.; Kocaman, S.; Gokceoglu, C. Rock Mass Discontinuity Determination With Transfer Learning. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2023**, *XLVIII-M-1-2023*, 609–614. [CrossRef]
29. Buyukdemircioglu, M.; Can, R.; Kocaman, S.; Kada, M. Deep learning based building footprint extraction from very high resolution true orthophotos and nDSM. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2022**, *2*, 211–218. [CrossRef]
30. Sawa, K.; Yalcin, I.; Kocaman, S. Building Detection from SkySat Images with Transfer Learning: A Case Study over Ankara. *PFG–J. Photogramm. Remote Sens. Geoinf. Sci.* **2024**, *92*, 163–175. [CrossRef]
31. Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [CrossRef]
32. Farahani, A.; Voghoei, S.; Rasheed, K.; Arabnia, H.R. A brief review of domain adaptation. *Adv. Data Sci. Inf. Eng.* **2021**, 877–894. [CrossRef]
33. Donmez, M.; Akcay, A.E.; Kara, H.; Yergok, A.F.; Esenturk, K. *Geological maps of Turkey*; General Directorate Mineral Research and Exploration: Ankara, Türkiye, 2008; p. 26.
34. DJI. Available online: https://ag.dji.com/phantom-4-rtk/specs (accessed on 22 June 2023).
35. Taddia, Y.; Stecchi, F.; Pellegrinelli, A. Coastal Mapping Using DJI Phantom 4 RTK in Post-Processing Kinematic Mode. *Drones* **2020**, *4*, 9. [CrossRef]
36. Agisoft. Available online: https://www.agisoft.com/ (accessed on 8 July 2023).
37. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2980–2988.
38. He, T.; Zhang, Z.; Zhang, H.; Zhang, Z.; Xie, J.; Li, M. Bag of Tricks for Image Classification with Convolutional Neural Networks. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 558–567.
39. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958. Available online: https://dl.acm.org/doi/10.5555/2627435.2670313 (accessed on 20 May 2024).
40. Weiss, K.; Khoshgoftaar, T.M.; Wang, D. A survey of transfer learning. *J. Big Data* **2016**, *3*, 9. [CrossRef]
41. Shorten, C.; Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. *J. Big Data* **2019**, *6*, 60. [CrossRef]
42. Yang, S.; Xiao, W.; Zhang, M.; Guo, S.; Zhao, J.; Shen, F. Image Data Augmentation for Deep Learning: A Survey. *arXiv* **2022**, arXiv:2204.08610. [CrossRef]
43. Albumentations. Available online: https://albumentations.ai/ (accessed on 8 July 2023).
44. Mikhail, E.M.; Bethel, J.S.; McGlone, J.C. *Introduction to Modern Photogrammetry*; John Wiley & Sons: New York, NY, USA, 2001.
45. Wolf, P.R.; Dewitt, B.A.; Wilkinson, B.E. *Elements of Photogrammetry with Applications in GIS*, 4th ed.; McGraw-Hill Education: New York, NY, USA, 2014.
46. Khan, A.; Sohail, A.; Zahoora, U.; Qureshi, A.S. A survey of the recent architectures of deep convolutional neural networks. *Artif. Intell. Rev.* **2020**, *53*, 5455–5516. [CrossRef]
47. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI, Munich, Germany, 5–9 October 2015; pp. 234–241.
48. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
49. Minaee, S.; Boykov, Y.; Porikli, F.; Plaza, A.; Kehtarnavaz, N.; Terzopoulos, D. Image Segmentation Using Deep Learning: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 3523–3542. [CrossRef] [PubMed]
50. Chaurasia, A.; Culurciello, E. LinkNet: Exploiting encoder representations for efficient semantic segmentation. In Proceedings of the 2017 IEEE Visual Communications and Image Processing (VCIP), St. Petersburg, FL, USA, 10–13 December 2017; pp. 1–4.
51. Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef] [PubMed]
52. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556. [CrossRef]
53. Ramasamy, G.; Singh, T.; Yuan, X. Multi-Modal Semantic Segmentation Model using Encoder Based Link-Net Architecture for BraTS 2020 Challenge. *Procedia Comput. Sci.* **2023**, *218*, 732–740. [CrossRef]
54. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
55. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Kai, L.; Li, F.-F. ImageNet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.

56. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* **2021**, *8*, 53. [CrossRef]
57. Kingma, D.P.; Ba, J.L. ADAM: A Method for Stochastic Optimization. In Proceedings of the 3rd International Conference for Learning Representations, San Diego, CA, USA, 7–9 May 2015.
58. Nair, V.; Hinton, G.E. Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th International Conference on International Conference on Machine Learning, Haifa, Israel, 21–24 June 2010; pp. 807–814.
59. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097–1105. [CrossRef]
60. Li, X.; He, M.; Li, H.; Shen, H. A Combined Loss-Based Multiscale Fully Convolutional Network for High-Resolution Remote Sensing Image Change Detection. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [CrossRef]
61. Krizhevsky, A. *Learning Multiple Layers of Features from Tiny Images*; University of Toronto: Toronto, ON, Canada, 2009; p. 60.
62. Alzubaidi, L.; Al-Shamma, O.; Fadhel, M.A.; Farhan, L.; Zhang, J.; Duan, Y. Optimizing the Performance of Breast Cancer Classification by Employing the Same Domain Transfer Learning from Hybrid Deep Convolutional Neural Network Model. *Electronics* **2020**, *9*, 445. [CrossRef]
63. Kaggle. Available online: https://www.kaggle.com/datasets/lakshaymiddha/crack-segmentation-dataset (accessed on 27 July 2023).
64. Hu, J.; Shen, L.; Sun, G. Squeeze-and-Excitation Networks. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
65. Al-antari, M.A.; Al-masni, M.A.; Choi, M.T.; Han, S.M.; Kim, T.S. A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification. *Int. J. Med. Inform.* **2018**, *117*, 44–54. [CrossRef] [PubMed]
66. Yu, L.; Chen, H.; Dou, Q.; Qin, J.; Heng, P.A. Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE Trans. Med. Imaging* **2016**, *36*, 994–1004. [CrossRef] [PubMed]
67. Zhang, W.; Zhao, X.; Pan, X.; Wei, M.; Yan, J.; Chen, J. Characterization of high and steep slopes and 3D rockfall statistical kinematic analysis for Kangyuqu area, China. *Eng. Geol.* **2022**, *308*, 106807. [CrossRef]
68. Byun, H.; Kim, J.; Yoon, D.; Kang, I.S.; Song, J.J. A deep convolutional neural network for rock fracture image segmentation. *Earth Sci. Inform.* **2021**, *14*, 1937–1951. [CrossRef]
69. Asadi, M.; Sadeghi, M.T.; Bafghi, A.Y. A Multi-Classifier System for Rock Mass Crack Segmentation Based on Convolutional Neural Networks. In Proceedings of the 2021 26th International Computer Conference, Computer Society of Iran (CSICC), Tehran, Iran, 3–4 March 2021; pp. 1–6.
70. Kim, J.; Lee, Y.K.; Choi, C.S.; Fereshtenejad, S.; Song, J.J. Scanline intersection similarity: A similarity metric for joint trace maps. *Comput. Geosci.* **2023**, *175*, 105358. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.