

## A CNN ARCHITECTURE FOR DISCONTINUITY DETERMINATION OF ROCK MASSES WITH CLOSE RANGE IMAGES

I. Yalcin <sup>1,2</sup>, R. Can <sup>1</sup>, S. Kocaman <sup>3,4,\*</sup>, C. Gokceoglu <sup>5</sup>

<sup>1</sup> Hacettepe University, Graduate School of Science and Engineering, Beytepe, Ankara, Turkey – [ilyas.yalcin@hacettepe.edu.tr](mailto:ilyas.yalcin@hacettepe.edu.tr)

<sup>2</sup> Hacettepe University, Başkent OSB Technical Sciences Vocational School, 06909 Sincan Ankara, Turkey

<sup>3</sup> Hacettepe University, Dept. of Geomatics Engineering, 06800 Beytepe Ankara, Turkey – [recepcan@hacettepe.edu.tr](mailto:recepcan@hacettepe.edu.tr)

<sup>4</sup> ETH Zurich, Institute of Geodesy and Photogrammetry, 8093 Zurich, Switzerland - [sultankocaman@hacettepe.edu.tr](mailto:sultankocaman@hacettepe.edu.tr)

<sup>5</sup> Hacettepe University, Dept. of Geological Engineering, 06800 Beytepe Ankara, Turkey - [cokke@hacettepe.edu.tr](mailto:cokke@hacettepe.edu.tr)

### Commission II, WG II/10

**KEY WORDS:** Rock Mass Evaluation, Discontinuity Determination, Convolutional Neural Networks, Close-Range Photogrammetry, Structure from Motion

### ABSTRACT:

Determination of discontinuities in rock mass requires scan-line surveys performed in in-situ that can reach up to dangerous and challenging dimensions. With the development of novel technological equipments and algorithms, the studies related to rock mass discontinuity determination remain up-to-date. Depending on the development of the Structure from Motion (SfM) method in the field of close-range photogrammetry, low-cost cameras can be used to produce 3D models of rock masses. However, the determination of rock mass discontinuity parameters must still be carried out manually on these models. Within the scope of this study, a Convolutional Neural Network (CNN) architecture is proposed to identify the discontinuities automatically as the first step for fully automated processing. The Kızılcahamam/Güvem Basalt Columns Geosite near Ankara, Turkey was determined as the study area. The orthophoto of this study area was produced using close-range photogrammetric methods and the training data was produced by manual mensuration. The dataset consists of labeled binary masks and images containing corresponding Red-Green-Blue (RGB) bands. Furthermore, the amount of data was increased by applying augmentation methods to the dataset. The U-Net architecture was used to detect rock discontinuities based on the produced orthophoto. The preliminary results presented here reveal that the discontinuity determination capability of the proposed method is high based on the visual assessments, while problems exist with image quality and discontinuity identification. In addition, considering the small size of the training dataset, the accuracy of the model would increase when a larger dataset could be employed.

### 1. INTRODUCTION

The rock mass classification systems are mainly used for the determination of the strength and deformability characteristics of rock masses. The studies on the rock mass classification include Rock Mass Rating (RMR) system (Bieniawski, 1989), Q system (Barton, 2002), and Geological Strength Index (GSI) (Hoek and Brown, 1997). The parameters defined in these studies are mainly related to discontinuities in rock masses. In this respect, accurate determination of the discontinuities and calculating their orientations are essential for engineering structures constructed in and on rock masses and monitoring their stability.

The discontinuities constitute the basis of works such as rockfall, tunnelling, and slope stability, etc. Compasses have conventionally been used for the measurement of discontinuity orientation during fieldwork. The field-based approach has several difficulties and limitations caused by environmental conditions and forms possible threats to human lives resulting from manual operation in inaccessible, hazardous, or dangerous regions. In the light of technological developments, geometric and semantic information about rock masses can be obtained without the need for physical intervention or access. Remote sensing instruments and methods have the advantage of producing information safely when compared with in-situ studies.

Over the last decade, Light Detection and Ranging (LiDAR) (e.g., Riquelme et al., 2014; Singh et al., 2021; Chen et al., 2017) and optical photogrammetric methods (e.g., Bogdanowitsch et al., 2022, Winkelmaier et al., 2020) have been widely used in the detection of rock mass discontinuities. Although LiDAR technology has the capability to produce accurate models, it has high costs in terms of hardware, software and computation, which negatively affect its widespread use. The development of the Structure from Motion (SfM) method in the field of close-range photogrammetry enabled the use of images obtained from low-cost cameras, which eliminate such problems to a broad extent. Compared to LiDAR equipment, photogrammetric systems have a lower cost (Cawood et al., 2017). Advances in Unmanned Aerial Vehicle (UAV) and camera equipment have helped to reduce costs and ensured the widespread use of these systems. Moreover, the orientation of rock discontinuities can be determined by using overlapping images obtained with mobile phone cameras (Ozturk et al., 2019). While precise camera parameters and image orientation data are required in the workflow of the traditional photogrammetric approaches, the SfM method can solve these parameters by using a large number of overlapping images.

The Deep Learning (DL) (LeCun et al., 1998) and in particular the Convolutional Neural Networks (CNN) have offered new possibilities in the field of image processing. They become

\* Corresponding author

available for researchers in a wide range of applications thanks to the databases such as ImageNet (Deng et al., 2009). Additionally, with popular CNN architectures such as U-Net (Ronneberger et al., 2015), AlexNet (Krizhevsky et al., 2012), VGG (Simonyan and Zisserman, 2014), ResNet (He et al., 2016), etc.; the semantic segmentation (Mei et al., 2020), object recognition (Girshick et al., 2014) and image classification studies can be carried out. CNNs also have the potential to detect rock discontinuities based on images obtained with aerial or close-range photogrammetry techniques.

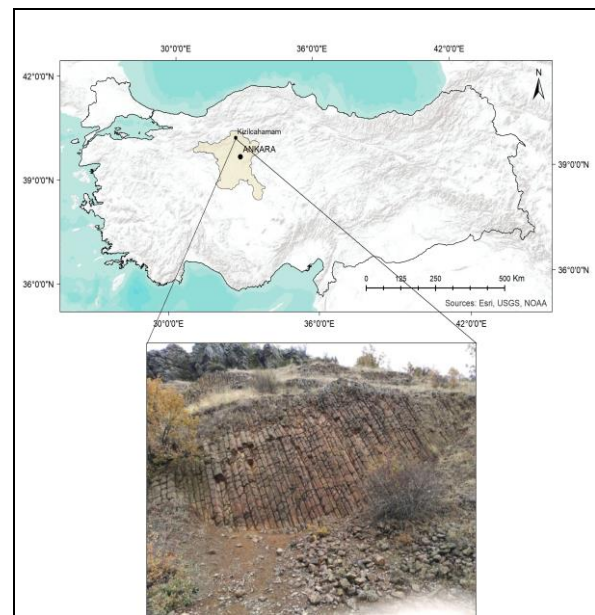
With its known success in biomedical applications and in the field of image segmentation, U-Net was used also in crack detection studies (Hamishebahar et al., 2022). Liu et al. (2019) completed the crack detection work with high accuracy via U-Net. In their studies, it has been seen that U-Net has achieved a successful result despite using small datasets compared to other architectures. Furthermore, it has been observed that CNN architectures unearth more successful results than edge detection operators such as Canny and Sobel in crack detection studies (Jenkins et al., 2018; Mei et al., 2020). Jogin et al. (2018) revealed that CNNs provided higher success compared to image classification methods such as logistic regression, K-Nearest Neighborhood (KNN), Support Vector Machine (SVM), etc.

The aim of this study is to detect rock discontinuities by close-range photogrammetric images and a CNN architecture. In this context, overlapping images were taken from a study area near Ankara, Turkey. The digital surface model (DSM) and the orthophoto of the study area were produced using the SfM method from terrestrial images. The training dataset was prepared manually from the orthophoto, and the discontinuities were detected with CNN architecture as explained in this paper.

## 2. STUDY AREA

The Kızılcahamam/Güvem Basalt Columns Geosite near Ankara, Turkey was chosen as the study area. The site was selected due to the basalts, which have durability, generally smooth shapes, and clean faces. Furthermore, the study area is located 110 km from the center of Ankara and has good accessibility. There are also different basalt outcrops in the

region. Figure 1 shows the location of the study area and an outcrop of the basalts to be evaluated.

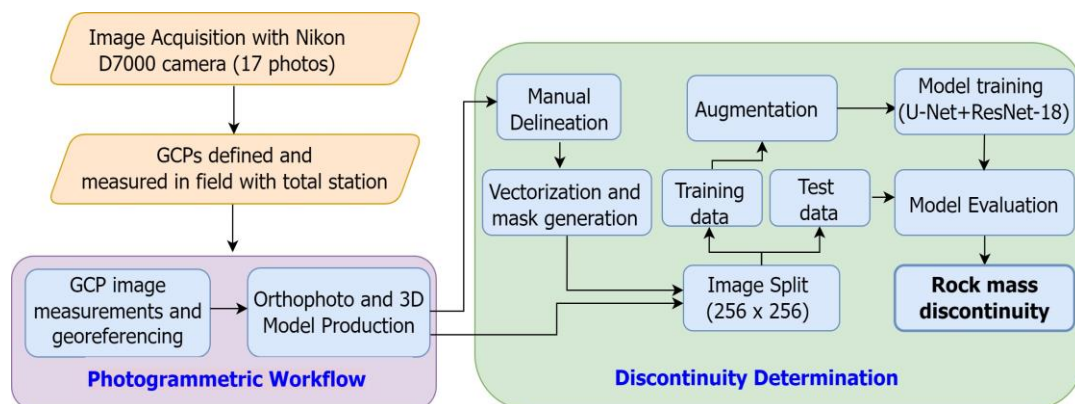


**Figure 1.** The location map of the study site and a typical outcrop of the columnar basalts employed.

The size of the working area is approximately 7 m x 17 m. The lengths of each rock block in the basalts range between 5 cm to 85 cm. In addition, as can be seen in Figure 1, some rockfalls were observed in the toe of the slope, which indicates that carrying out traditional scan-line surveys are dangerous for engineers and researchers.

## 3. METHODOLOGY

This study has two stages as close-range photogrammetric work and the detection of discontinuities with the CNN architecture. The overall workflow of the study is presented in Figure 2 and explained in detail in the next sub-headings.



**Figure 2.** The overall workflow of the study.

### 3.1 Close-Range Photogrammetric Workflow

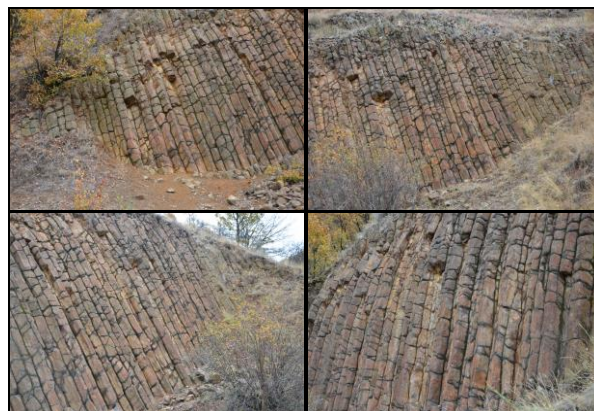
The photogrammetry technique with a distance up to 300 m between the object and the camera is called close-range photogrammetric technique (Wolf et al., 2014). As in aerial photogrammetry, this method requires exterior orientation parameters including image rotation (Yaw, Pitch, and Roll) and position (X, Y, Z) at the time of exposure. These parameters can be determined with the SfM or self-calibration methods together with the camera interior orientation parameters in a bundle block adjustment process. The main requirement for 3D scene reconstruction is acquisition of overlapping images of the object of interest from different positions and angles (Westoby et al., 2012). The images of the basalts in the study were captured with a Nikon D7000 professional camera. The technical specifications of the camera are given in Table 1, and examples to the images taken in the site are given in Figure 3.

<b>Camera type</b>	D-SLR
<b>Sensor</b>	CMOS
<b>Effective image pixel</b>	16.2 megapixel
<b>Image size (pixel)</b>	4928 x 3264
<b>Lens</b>	18-105 mm VR lens

**Table 1.** Camera specifications (Nikon, 2022).

Within the study, in-situ measurements were carried out by an engineering geology expert (last author) to collect the ground

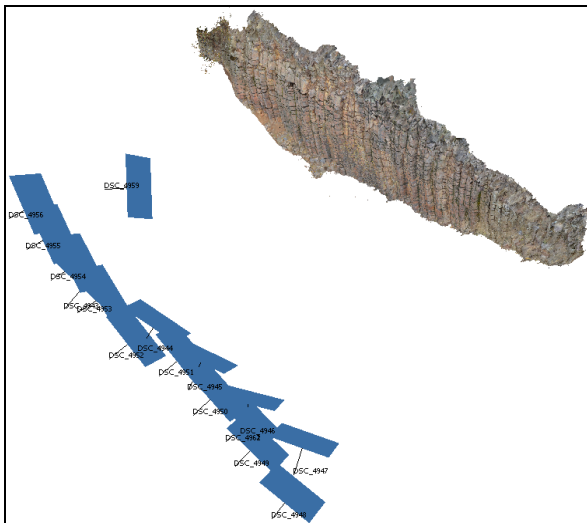
truth. A GNSS (Global Navigation Satellite System) and a total station instrument were used to measure a number of Ground Control Points (GCPs) in the field. A total of 9 GCPs were defined and measured in the study area (Figure 4). In addition, a total of 17 images were captured from the basalt rocks area (Figure 5).



**Figure 3.** Sample scene samples taken from the study area for 3D modelling.



**Figure 4.** GCPs measured in the field and in the images.



**Figure 5.** Camera positions with respect to the site.

The 3D model and the orthophoto were produced by using Agisoft Metashape Professional Software version 1.8.1 (Agisoft LLC, 2022). In order to produce orthophoto in the software, a total of five steps can be followed such as; align photos and sparse point cloud generation with tie points, measuring GCPs and Check Points (CPs), dense cloud generation, the generation of digital elevation model, and orthophoto generation (Agisoft, 2022). Out of nine GCPs, three of them were utilized as CPs in the bundle block adjustment process. The GCPs and the final products were defined in the Universal Transverse Mercator (UTM) zone 36N projection system referenced on the World Geodetic System 1984 (WGS 84).

### 3.2 CNN Architecture

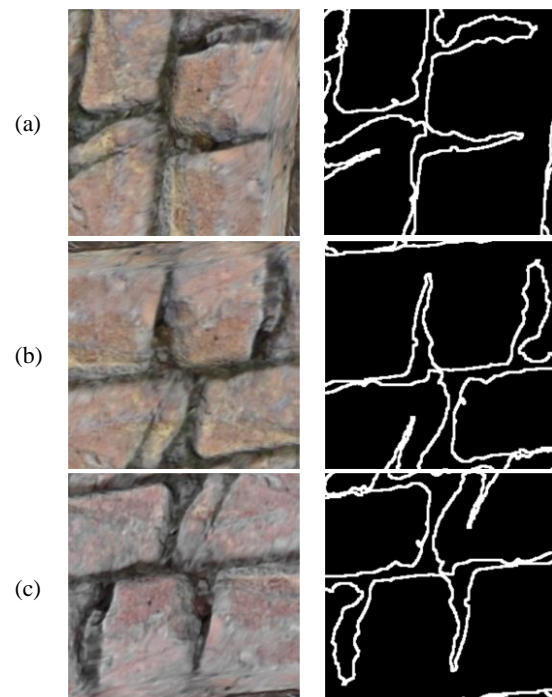
With the technological advancements, the image data collection platforms and devices in close-range photogrammetry, such as well-equipped smart mobile phones and UAVs, have increased. In order to process the massive amount of data, efficient methods and computational environments are also required. The CNNs belong to the family of Artificial Neural Networks (ANNs) and DL methods. A CNN has a convolutional multi-layer structure, consists of input, convolutional, pooling, activation and fully connected layers. When compared to ANN, in CNN the multi-layered image data is used as input in the architecture and the information in the dataset is extracted by applying image filters. The method is similar to ANN in that it performs the classification process in the form of forward-feed in the fully connected layer (Alzubaidi et al., 2021).

Not having a fully connected layer, the U-Net CNN architecture was proposed by Ronneberger et al. (2015) initially for the segmentation of biomedical images. The Rectified Linear Unit (ReLU) function is used as the activation layer in U-Net. The ReLU (Nair and Hinton, 2010) is frequently used in CNN studies compared as activation function since it works faster and provides higher performance (Krizhevsky et al., 2012). The ReLU function changes the values as a result of the learning process between zero (included) and infinity.

The fact that CNN architectures require large amount of training data compared to other image segmentation methods is an obstacle for the wider use of CNN. In order to minimize this disadvantage, image databases have been created. The databases

can be categorized according to their subjects, promoting the work of researchers. An example of Crack detection is EdmCrack600 (EdmCrack600, 2022). Moreover, the success of CNN can be enhanced by increasing the number of train data with data augmentation methods. Techniques such as flipping, rotating the image, changing the color space, adding noise and cutting can be applied within the scope of data augmentation methods (Alzubaidi et al., 2021).

Here, the U-Net architecture was employed with a training dataset generated by manual delineation of the discontinuity lines on the orthophotos. A raster ground-truth mask was generated and split as 256 x 256 images to be used as input. The input features include RGB images of the same size. In addition, data augmentation techniques were also applied to obtain better predictions with a relatively small training data size. In total, 25 out of 259 images were reserved for testing, 24 of the remaining 234 images were used for validation and 210 for training. Images augmentation was applied to the training set using Albumentations library (Albumentations, 2022), and a total of 2100 images were obtained. An example to the training images, the augmented images and the produced masks are shown in Figure 6.



**Figure 6.** (a) A sample training image, (b,c) the augmented images and their corresponding masks.

The CNN architecture developed here is based on U-Net, similar to Lee et al. (2022). The CNN architecture is a modified version of U-Net architecture. Instead of the encoder part that is proposed in the original U-Net paper, ResNet-18 model was used as the encoder (He et al., 2016). Also, in decoder part, transposed convolutional layers were used instead of up sampling layers. In the CNN model, weights were produced with the data set in the study instead of pre-trained weights. Furthermore, the use of the ResNet-18 model in the encoder part of the model made learning more effective. The ResNet model won first place in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2015. Additionally,

ResNet has a deeper structure than the VGG model (He et al., 2016). As in the study, various CNN models can be combined. This was also observed in the SegNet study, in which VGG-16 was used in the encoder part (Badrinarayanan et al., 2017).

Within the scope of the study, model hyper-parameters; 50 epochs, batch size 8 and Adaptive Moment Estimation (Adam) were used and Adam was determined as the optimizer. Defined as a learning algorithm, Adam uses the gradient-descent method (Liu et al., 2019). It also reduces memory usage while ensuring that learning is fast (Alzubaidi et al., 2021). ReLU was used as the activation layer of the study and the sigmoid function was used in the last layer. In the model, the combined form of binary cross entropy and Dice loss was used as the loss function. The loss function was used to measure the closeness of the model to the true value and the combination of the two loss function types is widely used (Yeung et al., 2022).

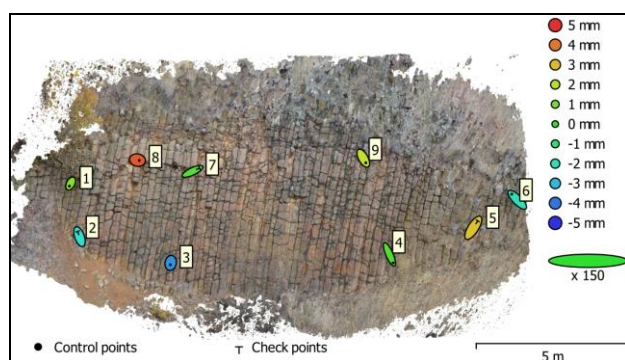
#### 4. RESULTS AND DISCUSSIONS

##### 4.1 Photogrammetric Processing Results

The photogrammetric processing accuracy is evaluated based on the CP coordinate differences between the ground-surveyed and those obtained from the bundle block adjustment method. The differences for the three CPs are given in Table 2. The 3D root mean square error (RMSE) obtained from the CPs was 3.88 mm, which indicates high positioning accuracy of the model. The average image GSD was 1.74 mm. Additionally, the error ellipses of all GCPs are shown in Figure 7. Points 2, 5 and 7 shown in Figure 7 were designated as CPs. In the Figure, the color tones of the ellipses represent the height error and the dimensions represent the point positioning error.

Point No	X (mm)	Y (mm)	Z (mm)	Total (mm)	Total (Pixel)
2	-0.83	2.31	-2.33	3.38	0.07
5	1.96	3.06	2.91	4.66	0.07
7	3.07	1.47	0.55	3.45	0.14
<b>RMSE</b>	<b>2.16</b>	<b>2.37</b>	<b>2.18</b>	<b>3.88</b>	<b>0.10</b>

**Table 2.** CP ground coordinate differences obtained from the bundle block adjustment.



**Figure 7.** GCPs error values and ellipses.

The orthophoto produced in the study is shown in Figure 8. The dimensions of the orthophoto consisting of RGB bands are 3707 x 9648 with a spatial resolution of 1.7 mm. The image

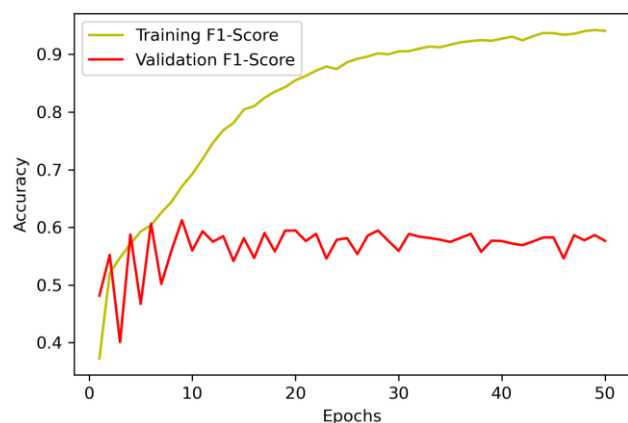
radiometric resolution is 8 bits. The orthophoto was cropped in the edges to avoid the image quality issues in those parts.



**Figure 8.** Orthophoto produced from the close range images

##### 4.2 The CNN Results

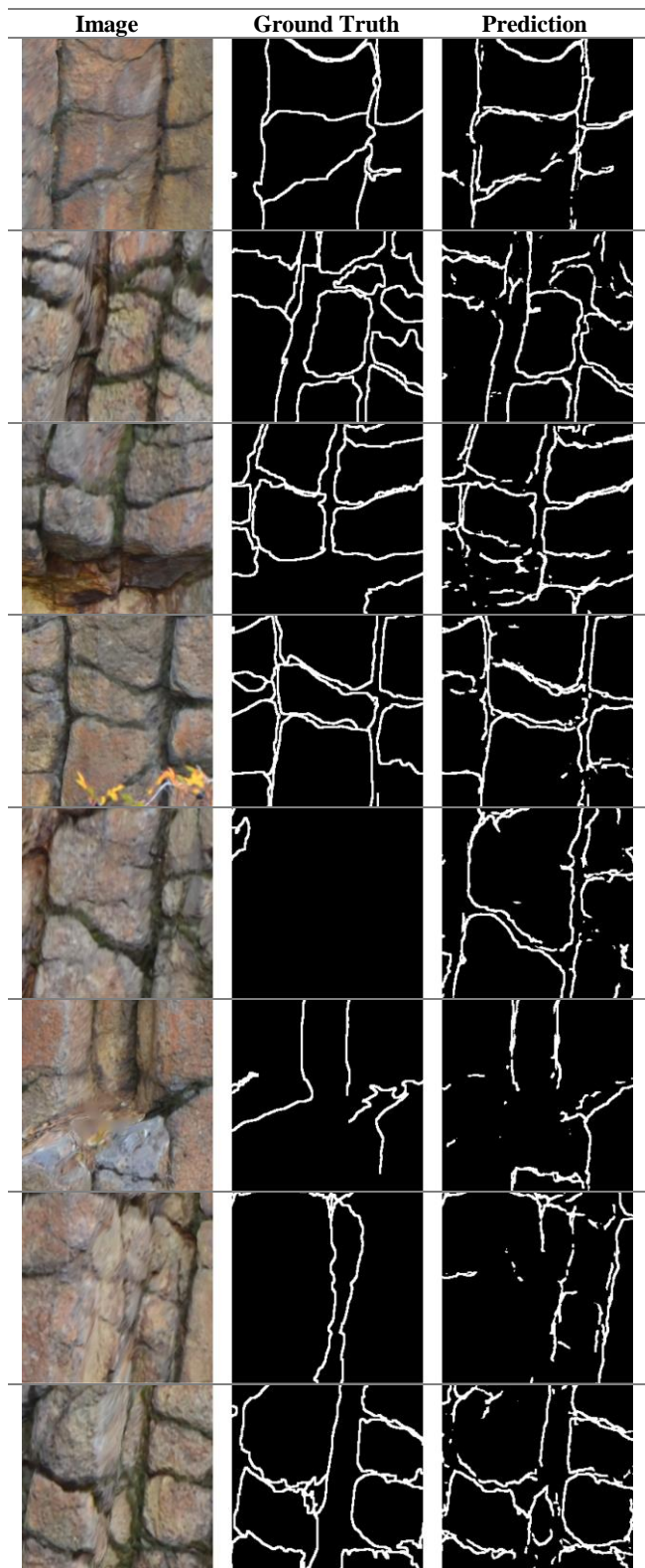
The model training - validation accuracy obtained from the model are given in Figure 9. In the study, the accuracy of the CNN model was calculated as 58% according to the F1-Score. According to the Figure, it was observed that the model was overfitting after the 10<sup>th</sup> epoch. However, no image enhancement was applied on the input data and the size of the training dataset is relatively small as it was obtained from one site only. Examples to the test images, ground truth and prediction results are presented in Figure 10.



**Figure 9.** Training and validation accuracy

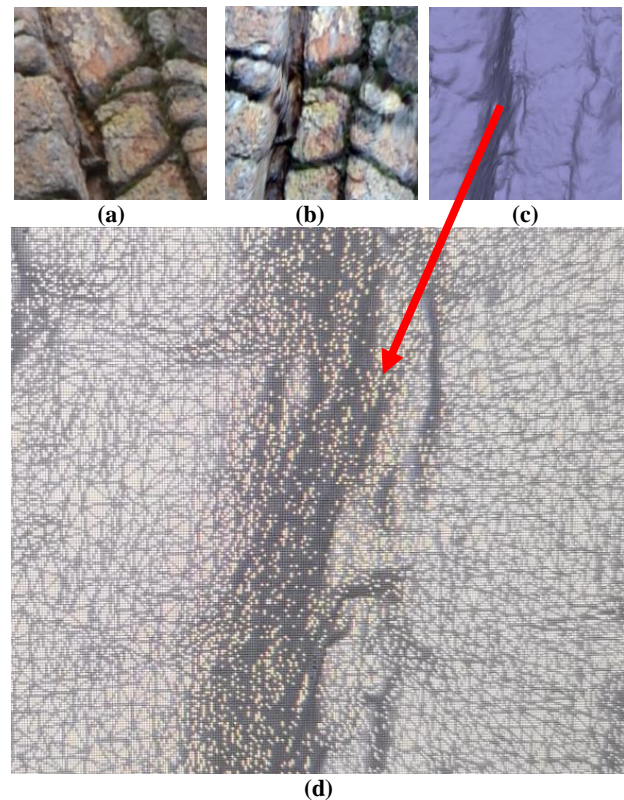
When these results are compared with those of crack detection studies; the latter ones revealed the high predictive performance of CNNs for crack detection on smooth surfaces such as buildings and roads (e.g., see Liu et al., 2019; Chen et al., 2020; Mei et al., 2020). The performance of CNN on complex surfaces such as rocks has not been sufficiently investigated yet. Lee et al. (2022) applied the method with 57,024 images and the validation results of the model according to the Intersection over Union (IoU) metric was 0.611. The main reasons are the small size (thinness) of rock discontinuity delineations, higher dimensionality (3D) of the rock surface when compared with building façades and road surface, and the variations on the rock surface characteristics caused by colour reflectance and roughness, such as shadows. In addition, since the IoU and F1-score metrics are calculated on a pixel-by-pixel comparison basis, the scores are low even though a visual inspection on the results reveals a better prediction performance. As an example, in Figure 10, it can be observed that the main discontinuity structures could largely be detected by the model. As future work, issues such as line completeness and noisy observations

could be largely eliminated by applying post-processing methods, e.g., morphological filters. Thus, the accuracy metrics could also be improved. Furthermore, instead of using pixel-wise metrics, problem-specific accuracy parameters such as line completeness in terms of percentage can be developed.



**Figure 10.** Sample images from the test data (left), ground truth (middle) and the CNN model results (right).

On the other hand, although orthophotos were used for training data preparation in this study for its practicality, image blurring was observed on the discontinuities as gaps occur in these areas in the DSM (Figure 11). The lower image quality also affects the model prediction quality. Although it is possible to delineate discontinuities on raw images as an alternative, the selection of the same discontinuity lines/surfaces on multiple images would cause problems that may introduce further uncertainty on the detected discontinuities.



**Figure 11.** (a) original (raw) image; (b) orthophoto; (c) shaded DSM; (d) mesh model.

## 5. CONCLUSIONS

The backbone of rock mass classification systems and analysis of discontinuity-controlled failures is the accurate determination of discontinuity orientation. However, sometimes, due to the inaccessibility of high and steep slopes, the measured number of discontinuities is limited. For this reason, in this study, it was aimed to detect discontinuities in rock mass by semantic segmentation of orthophoto with a CNN architecture, i.e., U-Net + ResNet-18. The orthophoto was produced by close-range photogrammetric method. According to the results, the F1-score was 58%. Image artefacts were observed in orthophotos at the discontinuities, which is a major reason for accuracy deterioration. In addition, no image enhancement or post-processing to prediction results such as morphological filters or smoothing was applied. Furthermore, the accuracy score was obtained from a pixel-wise comparison.

Yet, the visual inspection results indicate a higher performance and attention needs to be paid to training data preparation, use of raw images as an alternative to orthophotos, application of morphological and image filtering methods to the results to ensure line completeness and reduction of the noise, and the selection of appropriate spatial metrics for the evaluation of the

results. In addition, the training data size needs to be increased for obtaining higher accuracy. The fact that the visuals are better than the score supports this idea. The future work of the study also includes investigations on different rock mass types and the determination of the rock mass boundaries in 3D. The results of the present study may help engineering geologists when applying scan-line surveys.

#### ACKNOWLEDGEMENTS

Ilyas Yalcin acknowledges the support of the Scientific and Technological Research Council of Turkey 2224-A Grant Program for Participation in Scientific Meetings. The paper is part of the Ph.D. thesis research of the Ilyas Yalcin.

#### REFERENCES

- Agisoft LLC, 2022. Agisoft Metashape Professional Software, Version 1.8.1, <https://www.agisoft.com/> (last accessed on 01 April 2022).
- Agisoft, 2022. <https://agisoft.freshdesk.com/support/solutions/articles/31000153696> (last accessed on 01 April 2022).
- Albumentations, 2022. <https://albumentations.ai/>. (last accessed on 01 April 2022).
- Alzubaidi, L., Zhang, J., Humaidi, A.J. et al., 2021. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *J Big Data*, 8, 53. DOI:10.1186/s40537-021-00444-8.
- Badrinarayanan, V., Kendall A., Cipolla, R., 2017. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(12), 2481-2495. DOI:10.1109/TPAMI.2016.2644615.
- Barton, N., 2002. Some new Q-Value correlations to assist in site characterisation and tunnel design. *International Journal of Rock Mechanics and Mining Sciences*, 39 (2), 185–216. DOI:10.1016/S1365-1609(02)00011-4.
- Bieniawski, Z.T., 1989. *Engineering rock mass classifications*. Wiley, New York.
- Bogdanowitsch, M., Sousa, L., Siegesmund, S., 2022. Building stone quarries: resource evaluation by block modelling and unmanned aerial photogrammetric surveys. *Environ Earth Sci*, 81, 16. DOI:10.1007/s12665-021-10031-y.
- Cawood, A. J., Bond, C. E., Howell, J. A., Butler, R. W. H., Totake, Y., 2017. LiDAR, UAV or compass-clinometer? Accuracy, coverage and the effects on structural models. *Journal of Structural Geology*, 98, 67–82. DOI:10.1016/j.jsg.2017.04.004.
- Chen, N., Kemeny, J., Jiang, Q., Pan, Z., 2017. Automatic extraction of blocks from 3D point clouds of fractured rock. *Computers & Geosciences*, 109, 149-161. DOI:10.1016/j.cageo.2017.08.013.
- Chen, T., Cai, Z., Zhao, X., Chen, C., Liang, X., Zou, T., Wang, P., 2020. Pavement crack detection and recognition using the architecture of segNet. *Journal of Industrial Information Integration*, 18, 100144. DOI: 10.1016/j.jii.2020.100144.
- Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L., 2009. ImageNet: A large-scale hierarchical image database. *IEEE Conference on Computer Vision and Pattern Recognition*, 248-255, DOI:10.1109/CVPR.2009.5206848.
- EdmCrack600, 2022. <https://github.com/mqp2259/EdmCrack600> (last accessed on 01 April 2022).
- Girshick, R., Donahue, J., Darrell, T., Malik, J., 2014. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. The IEEE Computer Society Conference on Computer Vision and Pattern Recognition. DOI:10.1109/CVPR.2014.81.
- Hamishebahar, Y., Guan, H., So, S., Jo, J., 2022. A Comprehensive Review of Deep Learning-Based Crack Detection Approaches. *Appl. Sci.*, 12, 1374. DOI:10.3390/app12031374.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. The IEEE Conference on Computer Vision and Pattern Recognition, 770-778. DOI:10.1109/CVPR.2016.90.
- Hoek, E., Brown, E.T., 1997. Practical estimates of rock mass strength. *International Journal of Rock Mechanics and Mining Sciences*, 34 (8), 1165–1186. DOI:10.1016/S1365-1609(97)80069-X.
- Jenkins, M.D., Carr, T.A., Iglesias, M.I., Buggy, T., Morison, G., 2018. A Deep Convolutional Neural Network for Semantic Pixel-Wise Segmentation of Road and Pavement Surface Cracks. 26th European Signal Processing Conference (EUSIPCO), 2120-2124, DOI: 10.23919/EUSIPCO.2018.8553280.
- Jogin, M., Mohana, Madhulika, M.S., Divya, G.D., Meghana, R.K., Apoorva, S., 2018. Feature Extraction using Convolution Neural Networks (CNN) and Deep Learning. 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), 2319-2323, DOI: 10.1109/RTEICT42901.2018.9012507.
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012 ImageNet Classification with Deep Convolutional Neural Networks. *Neural Information Processing Systems*, 25. DOI: 10.1145/3065386.
- LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86 (11), 2278-2324. DOI: 10.1109/5.726791.
- Lee, Y. K., Kim, J., Choi, C. S., Song, J. J., 2022. Semi-automatic calculation of joint trace length from digital images based on deep learning and data structuring techniques. *International Journal of Rock Mechanics and Mining Sciences*, 149, 104981. DOI:10.1016/j.ijrmms.2021.104981.
- Liu, Z., Cao, Y., Wang, Y., Wang, W., 2019. Computer vision-based concrete crack detection using U-net fully convolutional

networks. *Automation in Construction*, 104, 129-139. DOI:10.1016/j.autcon.2019.04.005.

Mei, Q., Gül, M., Azim, M.R., 2020. Densely connected deep neural network considering connectivity of pixels for automatic crack detection. *Automation in Construction*, 110, 103018. DOI: 10.1016/j.autcon.2019.103018.

Nair, V., Hinton, G.E., 2010. Rectified Linear Units Improve Restricted Boltzmann Machines. *The 27th International Conference on Machine Learning (ICML-10)*, June 21-24, 2010, Haifa, Israel, 27, 807-814.

Nikon, 2022.

[https://www.nikon.co.uk/en\\_GB/product/discontinued/digital-cameras/2015/d7000#overview](https://www.nikon.co.uk/en_GB/product/discontinued/digital-cameras/2015/d7000#overview) (last accessed on 01 April 2022).

Ozturk, H.S., Kocaman, S., Gokceoglu, C., 2019. A low-cost approach for determination of discontinuity orientation using smartphone images and application to a part of Ihlara Valley (Central Turkey). *Engineering Geology*, 254. DOI:10.1016/j.enggeo.2019.04.011.

Riquelme, A.J., Abellán, A., Tomás, R., Jaboyedoff, M., 2014. A new approach for semi-automatic rock mass joints recognition from 3D point clouds. *Computers & Geosciences*, 68, 38-52. DOI:10.1016/j.cageo.2014.03.014.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 9351, 234-241. DOI:10.1007/978-3-319-24574-4\_28.

Simonyan, K., Zisserman, A., 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv preprint arXiv e-print*, 1409.1556. DOI: 10.48550/arXiv.1409.1556.

Singh, S.K., Raval, S., Banerjee B.P., 2021. Automated structural discontinuity mapping in a rock face occluded by vegetation using mobile laser scanning. *Engineering Geology*, 285, 106040. DOI: 10.1016/j.enggeo.2021.106040.

Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., Reynolds, J.M., 2012. 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, 179, 300-314. DOI: 10.1016/j.geomorph.2012.08.021.

Winkelmaier, G., Battulwar, R., Khoshdeli, M., Valencia, J., Sattarvand J., Parvin, B., 2020. Topographically Guided UAV for Identifying Tension Cracks Using Image-Based Analytics in Open-Pit Mines. *IEEE Transactions on Industrial Electronics*, 68 (6), 5415-5424. DOI: 10.1109/TIE.2020.2992011.

Wolf, P.R., Dewitt, B.A., Wilkinson, B.E., 2014: *Elements of Photogrammetry with Applications in GIS*. McGraw Hill Education, USA.

Yeung, M., Sala, E., Schönlieb, C.B., Rundo, L., 2022. Unified Focal loss: Generalising Dice and cross entropy-based losses to handle class imbalanced medical image segmentation. *Computerized Medical Imaging and Graphics*, 95 DOI: 10.1016/j.compmedimag.2021.102026.