International Journal of GEOMATE, Aug. 2023, Vol. 25, Issue 108, pp.241-248
ISSN: 2186-2982 (P), 2186-2990 (O), Japan, DOI: https://doi.org/10.21660/2023.108.s8569
Geotechnique, Construction Materials and Environment

# FRACTURE-FAULT DETECTION USING DEEP LEARNING WITH STEPWISE ELIMINATION FROM SATELLITE IMAGES IN DJIBOUTI

Denis Pastory Rubanga[1], *Sergio Azael May-Cuevas[1], Yessy Arvelyna[2] and Sawahiko Shimada[1]

[1]Faculty of Regional Environment Science, Tokyo University of Agriculture, Japan;
[2] Remote Sensing Technology Center of Japan, Japan

*Corresponding Author, Received: 20 Nov. 2022, Revised: 24 April 2023, Accepted: 28 May 2023

**ABSTRACT:** Accurate estimation of groundwater flow is crucial in arid regions where permanent surface water is absent. In several groundwater simulation models, an important parameter for identifying areas with high potential for groundwater resources is the accurate fracture-fault detection. In the present study we propose a deep learning approach to detect fracture-fault structures in the Ali Faren sub-catchment of Ambouli Wadi in Djibouti. Our deep convolutional neural network (Deep-CNN) model is trained on high-spatial resolution multispectral satellite images using wadi streamline as labels. Fracture-fault structures are extracted using stepwise elimination based on geological characteristics observed in relief images derived from PALSAR-1/2 data. Our results demonstrate that the proposed Deep-CNN model accurately detects fracture-fault lines, achieving a validation accuracy of 0.9684, precision of 0.9124, recall of 0.9701, and F1 of 0.8997. The proposed model has the potential to identify potential areas for groundwater resources across the country, contributing to sustainable water management and improving Djibouti's water security. Our study highlights the potential of deep learning techniques in addressing challenges related to sustainable water management in arid regions.

*Keywords: Fracture fault, High-resolution satellite, Deep learning, Stepwise elimination method*

## 1. INTRODUCTION

Djibouti is a country in the Horn of Africa that covers a land area of 23,200 $km^2$, with a rainfall of 130 mm/year and an arid climate, so it relies mainly on groundwater resources, estimated to be around 300 million $m^3$/year [1].

The country's geology consists of ancient sedimentary lands from the Mesozoic age (secondary era) and is characterized by fracture-fault lines. The Ali Sabieh horst (ca, 27 to 23 Ma) extends from Djibouti city south into the north of the Aysha horst in Ethiopia, and the fracture-fault systems within its borders are composed of rhyolitic lavas in the east (ca, 25 to 19 Ma) and in the west (14 Ma to 10 Ma), mainly covered by basaltic flows such as the Dalha series (ca, 9 to 4 Ma) and Somalis basalt (ca, 7 to 3 Ma) in the east, and Tadjourah basalts (3 Ma to 1 Ma) in the central region [2].

Fracture-fault volcanic aquifers have been analyzed as the main water resources in Djibouti, with pumping test data showing the major basaltic series [3].

Under these conditions, it is important to conduct groundwater modeling to estimate the availability of local water resources. It has been observed that success of groundwater flow simulation using numeric models such as GETFLOWS is greatly impacted by poor fracture-fault delineation, which is an important parameter for identifying areas with high potential for groundwater resources [4] [5].

In 2015 a 1:200 000 geological map was created [6], based on nine maps of 1:100 000 published by ISERT-ORSTOM, from 1983 to 1995 [7]. The fault networks on this map were extracted from satellite images with a vertical resolution of 15 m.


Fig. 1. Unmatching wadi when overlapping Geological map 1:200 000 and map on the WGS 1983 projected coordinated system

When georeferencing this map, it was found that the wadis location was not matching completely with the projected coordinate system WGS 1984 (Fig. 1).

For this reason, to improve fracture-fault delineation, we propose a Deep-CNN approach [8], which is part of Machine Learning, an emergent technology that uses big amounts of data and high-performing computers, and have applications in agriculture, such as crop, livestock, water, and soil management [9]. Deep-CNN is based on Deep Artificial Neural Networks (Deep ANNs), widely referred as deep learning (DL), which is composed of multiple processing layers that allows to learn complex data representations, and it is useful in image recognition.

Deep-CNN relies on easily accessible satellite-derived datasets. Our approach uses WorldView-3 (WV-3) [10], PALSAR-1/2 [11], 5 m resolution ALOS World 3D level 1 DSM [12], and slope derived from DEM to train a multi-input Deep-CNN model using wadi streamline derived from AW3D-DEM as fracture-fault labels of the target area [13].

Deep learning techniques in geological feature settings have been widely used, such as automatic extraction of seismic landslides in large earthquake areas with complex environments [14] [15]. Other works have used object-oriented classification techniques to segment homogeneous images using high-resolution multi-spectral data [16].

Furthermore, studies have used deep learning to establish a reliable river environment by observing the complicated relationships between water environment factors [17], to detect defects on asphalt [18] and used remote sensing for mapping fault distribution based on DEM data [19]. However, the application of deep learning to surface geological structures such as fracture-faults is limited.

This study differs from previous works in that it proposes a Deep-CNN approach for fracture-fault detection that is trained using multispectral satellite images, which provides a more accurate method for identifying fracture-fault lines. The proposed approach has the potential to improve groundwater flow model simulations and aid in identifying potential areas for groundwater resources across Djibouti, contributing to sustainable water management in arid regions, particularly in the context of Djibouti's water security.

## 2. RESEARCH SIGNIFICANCE

This study's significance lies in its contribution to the application of deep learning and stepwise elimination methods in underground water simulation. It is the first of its kind in sub-Saharan Africa's dryland, where limited data poses a challenge to groundwater resource management.

The use of advanced technologies like deep learning can support sustainable water management in arid regions like Djibouti. The novelty of this study rests on the combination of deep learning and multispectral satellite imagery to accurately detect fracture-fault structures in the Ali Faren sub-catchment of Ambouli Wadi.

The proposed approach can improve groundwater flow model simulations and aid in identifying potential areas for groundwater resources, contributing to sustainable water management.

In the following sections, methodology of fault detection is explained for the study site, which is an arid zone that heavily relies on underground water. Processes applied to data are detailed, and settings of the Deep-CNN training model are enumerated. Then results of the model training and evaluation are discussed. Lastly, it is concluded that this model can detect fault structures.

## 3. STUDY AREA

### 3.1 Study Area Data

The target area of this study is the Ali Faren catchment, which is located 40 km to the west of Djibouti city, and it is part of the Ambouli watershed (Fig. 2).


Fig. 2 Study area: Ali Faren Catchment, part of the Ambouli Watershed, Djibouti.

For this study, we used WorldView-3 (WV-3), the latest in a constellation of commercial high-spatial resolution Earth imaging satellites developed by Digital Globe Inc (Longmont, Colorado, USA) [12]. High-spatial images collected between 12th May 2018 and 5th April 2021 with 0% cloud cover between N.W Latitude 11.55372100 and Longitude 42.77572500, 8-band multispectral image (red, red edge, coastal, blue, green, yellow, near-IR1, and near-IR2) 400 nm – 1040 nm of 1.2-m spatial resolution, and 5-m spatial resolution AW3D DEM (ALOS World 3D level 1 DSM).

The original DEM resolution of 5-m was downscaled to same resolution as the PALSAR-1/2 images. Using ESRI's ArcGIS Pro v 2.7 (ESRI 2021), we created the slope images derived from the DEM and PALSAR-1/2, that were stacked together to create a composite of 11-band raster image. The PALSAR-1/2 Global Mosaic data is a seamless global SAR image created by mosaicking the SAR image in backscattering coefficients measured by PALSAR-1/2 with 25-m resolution.

### 3.2 Relief Image for delineation of fault system

Since PALSAR-1/2 are filled with speckle noise due to the interference during signal transmission, the adaptive filter Enhanced Lee Filter was applied to reduce speckle noise on the image [20], [21]. SAR images can penetrate clouds and hence observe the ground surface during day and night because they are composed of radar signals. For PALSAR-1/2 images, geometric distortion correction (ortho-rectification) and topographic effect on image intensity (slope correction) were applied.

From this, a relief image was obtained, and it was possible to delineate the fault system in the study area. The relief image shows that fault lines and dry rivers are associated with rugged features on image, where fault lines appears as straight lines and rivers as curved lines (Fig. 3).

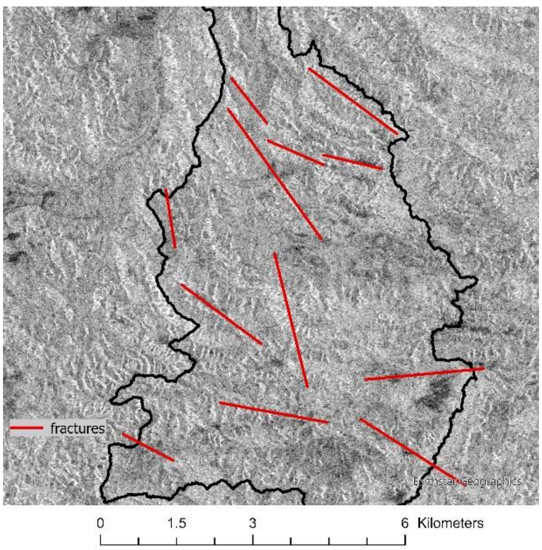


Fig. 3 Relief image (PALSAR 1/2) showing fault-fractures in Ali Faren

Fracture-fault labels as wadi streamline were derived from AW3D-DEM of the target area. The coverage of the created composite layer included the target Ali Faren catchment (Fig. 4).

### 3.3 Labeling the Training Data

To train the Deep-CNN model on a generated multi-band raster image, a group of labeled raster cells of the 11-multi-band, which can indicate the characteristics of fracture-fault, must be used to train the model. We obtained the fracture-fault label by using ArcGIS Pro software, applying the buffer method of 50 m resulting into a polygon shape as shown in Fig. 4. The derived polygon shape files were used as Region of Interest (ROI) in ENVI Deep Learning in creating training labels [22], [23].

The data used in this research, and the processes conducted are expressed in Fig. 5.

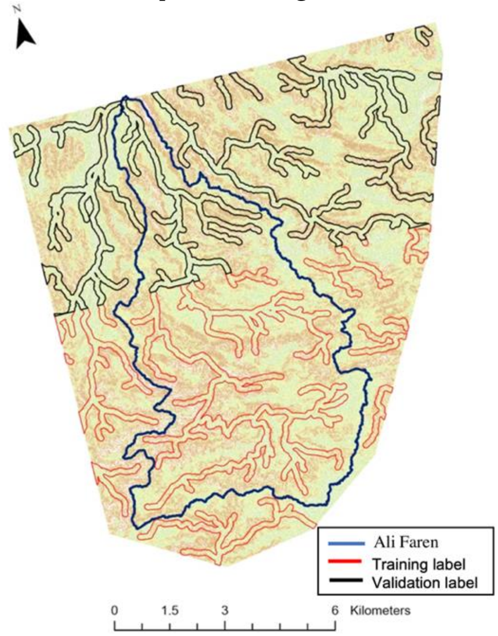


Fig. 4 Stacked 11-band raster image training (red) and validation (black) data for training Deep-CNN model

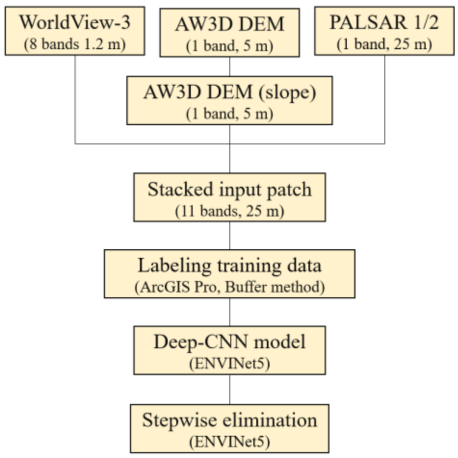


Fig. 5 Data used in this study

### 3.4 Stepwise elimination

Stepwise elimination is a technique commonly used for identifying a subset of features that are related to a set of input observations to output measurements [24], which consists of finding a small subset of features, so the resulting model provides and accurate representation of the measurements. This is achieved by eliminating variables, in a stepwise manner.

## 4. DEEP-CNN TRAINING

ENVINet5 (built in ENVI version 5.6, and ENVI Deep Learning version 1.2) is used to train the Deep-CNN model on the multi-band raster image. The ENVINet5 architecture shown in Fig. 6 is a mask-based encoder-decoder fully convolutional network with 5 levels and 23 convolutional layers. Its architecture is based on U-Net [25] with some modifications on layers of convolution and the size of input and output.

The input of ENVINet5 is a patch with fracture-fault polygon sampled from the 11-band multi-input raster images. Training raster is used to train the initialization model. Model training is to expose training raster to the model. The model learns to convert the spectral and spatial information in the training raster into a class activation map, a probability map, highlighting the target to be extracted as shown in Fig. 6.

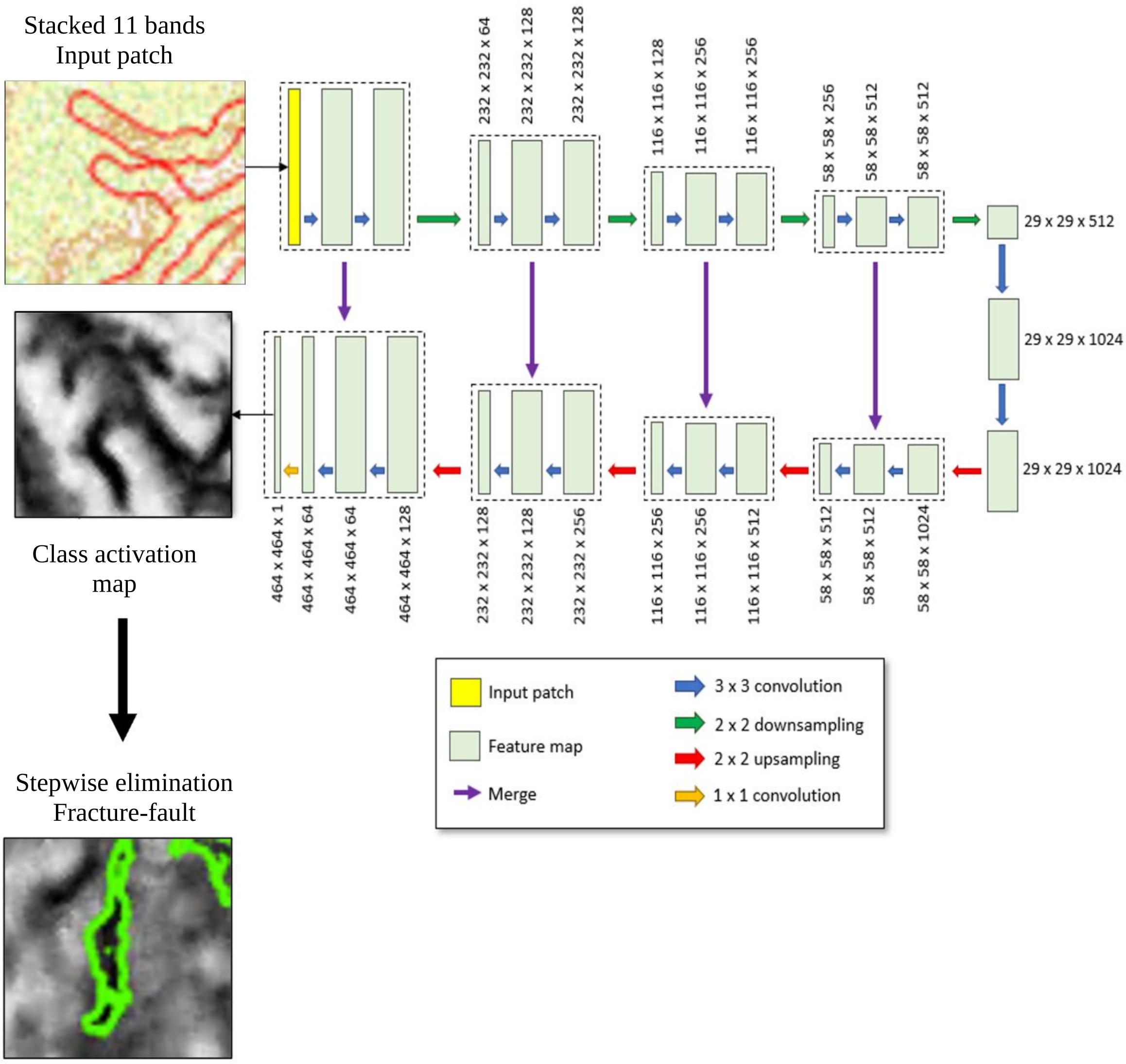


Fig. 6 ENVINet5 architecture. Source; Harris Geospatial Solution (2021) [23]

ENVINet5 refers to the binary cross-entropy loss function with the weighted map used by U-Net.

$$E = \sum_{x \in \Omega} \omega(x) \log\left(p_{l(x)}(x)\right), \tag{1}$$

where $p_{l(x)}(x)$ is the SoftMax loss function; $l : \Omega \rightarrow \{1,\cdots,K\}$ is the label value of the pixel and ω:ΩεR is the weight of the pixel that gives a higher weight to the pixel close to the boundary point in the raster image [26].

The number of patches per epoch determines the amount of training. The settings are lower for small datasets and higher for large datasets. The number of patches per epoch is set to 414 and number of patches per image set to 32.

The number of patches per batch system is automatically set to 1. For the rest of the parameters, ENVI automatically determines the appropriate values. To prevent overfitting, data augmentation is used. Data augmentation is a technique commonly used with deep learning to supplement the original training data. By having more information to extract from the training data, the trainer and classifier can more effectively learn the appearance of features of interest. During each epoch, ENVI creates a new training dataset with a randomly assigned angle per training example. Likewise, it creates a new training dataset with a randomly assigned scale factor per training example.

In traditional Deep-CNN training, an epoch represents the period in which all data sets are passed into the training model. However, it is different in the ENVI Deep Learning Module, which intelligently extracts patches from the training raster, so high-brightness characteristic pixel areas are more often encountered than low-brightness regions at the beginning of training. At the end of the training, all areas look more uniform. Because there is bias determination in patch extraction, an epoch in ENVI deep learning refers to the number of patches trained before bias decision adjustment.

To get a better model, multiple epochs are needed to fully train the model. The number of epochs and the number of patches per epoch depend on the diversity of feature sets to be learned, which has no exact value. In general, enough epochs are needed to adjust the weight to ensure smooth progress. We set number of epochs to 50.

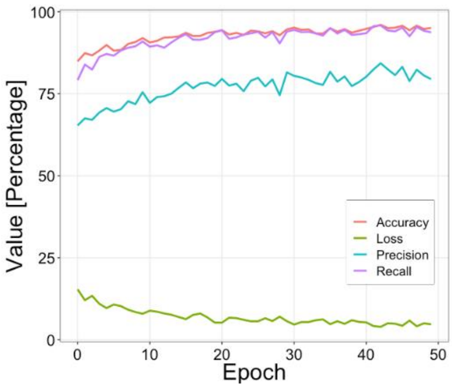


Fig. 7 Showing Training curve for accuracy, loss, precision, recall of the experiment

Model training were run on a Windows 10 Pro Operating System Intel (R) Core 9TM i9-990K CPU processor clocking 3.6GHz and with a NVIDIA GeForce RTX 2080i GPU. The training time of the model is 50 min. The curves of training accuracy, training loss, training precision and training recall are shown in Fig. 7.

## 5. RESULTS AND DISCUSSION

### 5.1 Training and Evaluation Result

The training curve showing accuracy, loss, precision, and recall is shown in Fig 7. Similarly, results of validation of the raster images showing validation accuracy, validation loss, validation precision and validation recall curves are shown in Fig. 8.

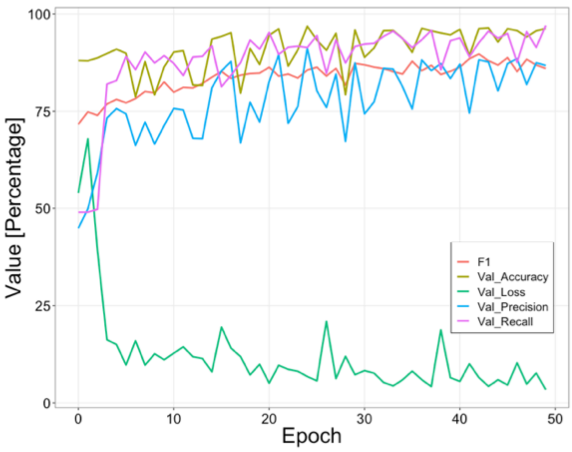


Fig. 8 Validation curves for accuracy, loss, precision, recall and F1

ENVI generates a model based on the lowest point of the validation loss value, that is, the epoch with the best match between the classifier and the validation data. The lowest loss value of the trained model is 0.0393, accuracy is 0.9604, precision is 0.8438, recall is 0.9582 and F1 is 0.8997. Validation of data yielded a loss 0.03391556, accuracy of 0.9684, precision of 0.9124, recall of 0.9701 and F1 of 0.8997

### 5.2 Image Classification

To classify other raster images for fracture-faults using the trained model, class activation map raster represented in Fig. 9 was generated.

The class activation map shows each pixel in the grayscale image roughly represented as probability of belonging to the fracture-fault, and the threshold ranges from 0 to 1. The black area in the class activated grayscale image represents the area with high probability, which means it is identified as a fracture-fault. The result identified by the trained model is a fracture-fault probability map; the larger the values in this probability map, the more confident the identified fracture-fault.

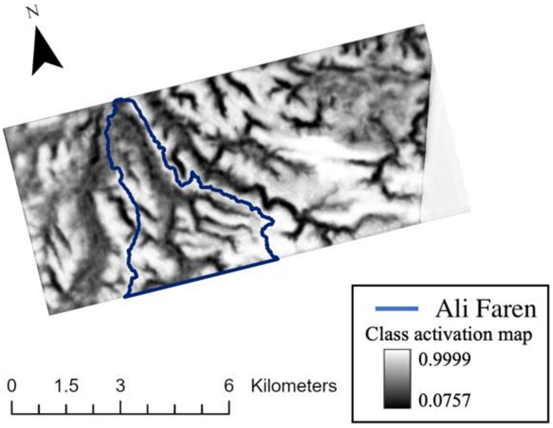


Fig. 9 Image class activation raster of selected validation area. In blue: Ali Faren watershed

### 5.3 Stepwise Elimination in Fracture-fault Classification

Fracture-fault classification from the classified image based on activation map is not a straightforward approach. To identify the real fracture-fault a stepwise elimination approach was used to identify the fracture-fault from the wadi streamlines class activation map since the result of the activation map are a mixture of wadi streamline and fault fracture.

Using the stepwise elimination, we set conditions based on geological characteristics of fracture-fault. Fracture-fault will not winded as much as wadi. Wadis will not go beyond their watershed regions and some parts of the fracture-faults might be used as wadi. Fracture-fault shapes as defined in geological maps also tend to be straight.

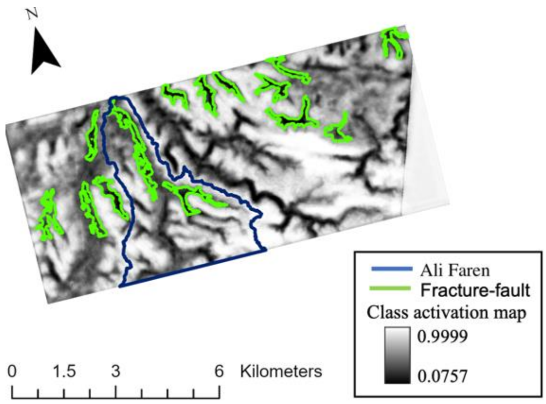


Fig. 10 Image class activation raster showing fracture-fault (green) after stepwise elimination

The predefined assumption of the stepwise elimination considered setting threshold to the class activation maps. With the set threshold of class value of given pixel values of greater than or equal to the threshold value, the polygon will be designated as the fracture-fault feature class within the polygon.

Through trial and error, the appropriate threshold used was between 0.0588 and 0.8409 for a mixture of wadi and fracture-fault polygons. The best threshold for fracture-fault activation were between 0.186343 to 0.499941.

We further eliminated the classified raster based on the prior defined activation map, result of the final fracture-fault are shown in Fig. 10 and Fig. 11

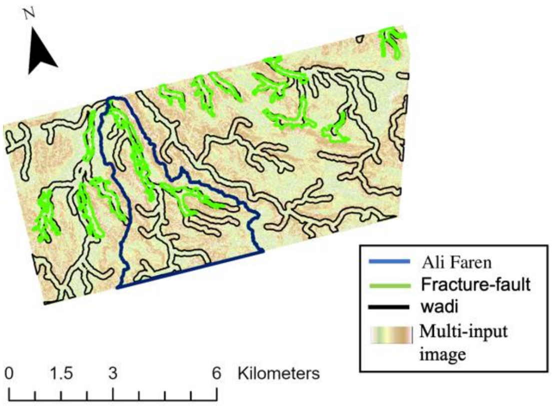


Fig. 11 Fracture-fault as seen on a wadi after stepwise elimination

## 6. CONCLUSION

This study demonstrates the successful application of the Deep-CNN technique in detecting fracture-fault structures in the Ali Faren catchment of Djibouti.

The stepwise technique applied to the Deep-CNN generated class activation map provided accurate fracture-fault structure detection, and with the Geological maps available, can improve groundwater flow model simulations and aid in identifying potential areas for groundwater resources.

Future research will evaluate the developed deep-CNN models in wider areas within Djibouti and explore the implementation of other deep learning architecture, such as feature engineering of appropriate parameters.

The ultimate goal is to develop a comprehensive deep-CNN model to be used in other water catchments in Djibouti and integrate the fracture-faults detected results in groundwater flow model simulation.

This study demonstrates the potential of deep learning techniques for groundwater resource management in arid regions like Djibouti, which can support sustainable water management practices.

## 7. ACKNOWLEDGMENTS

This study was supported by Japan Science and Technology Agency (JST) and Japan International Cooperation Agency (JICA) under the project of Science and Technology Research Partnership for Sustainable Development (SATREPS GRANT NUMBER: JPMJSA1802) SATREPS (Science and Technology Research Partnership for Sustainable Development).

This study acknowledges the copyright Notice for ALOS PALSAR RTC: © NASA 2006-2007, Includes Material © JAXA, METI [2006-2007]. The PALSAR-2 Global Mosaic data have been provided by the Japan Aerospace Exploration Agency

## 8. REFERENCES

[1] Malow F. A., Development of a 3D Water Flow Modelling based on Scarce Data for Arid Land Water Resources Management: Case Study of Ambouli and Kourtimalei Watersheds in Djibouti, Ph.D. Thesis, Tokyo University of Agriculture, Tokyo, Japan, pp. 22 (2018).

[2] Razack M., Jalludin M., Houmed-Gaba A., Simulation of climate change impact on a coastal aquifer under arid climate. The Tadjourah Aquifer (Republic of Djibouti, Horn of Africa). Water, Vol. 11 (11), pp. 2347 (2019).

[3] Awaleh M.O., Baudron P., Soubaneh Y.D., Boschetti T., Hoch F.B., Egueh N.M., Mohamed J., Dabar O.A., Dufresne, J.M., Cassani, J., Recharge, groundwater flow pattern and contamination processes in an arid volcanic area: Insights from isotopic and geochemical tracers (Bara aquifer system, Republic of Djibouti), Journal of Geochemical Exploration 175, pp. 82-98 (2017).

[4] Malow F.A., Shimada S., Hazart A., Event-based Rainfall-runoff Simulations using GETFLOWS for Kourtimalei Catchment in Djibouti, International Journal of Environmental and Rural Development, Volume 8, Issue 1, pp. 169-176, (2017).

[5] Tosaka H., Mori K., Tada K., Tawara Y., Yamashita K., A general-purpose terrestrial fluids/heat flow simulator for catchment system management. In IAHR International Groundwater Symposium (2010).

[6] CERD, Geological map of the Republic of Djibouti, 1:200,000, Centre d'Etude et de Recherche de Djibouti, Djibouti, (2015).

[7] ISERST, Carte Geologique de la Republique de Djibouti a 1:100,000, ORSTOM, Office de la Recherche Scientifique et Technique Outre Mer, Paris, France, (1983, 1985, 1986, 1987, 1995).

[8] LeCun Y., Yoshua B., Geoffrey H., Deep learning. Nature 521.7553, pp. 436-444, (2015).

[9] Liakos K.G., Busato P., Moshou D., Pearson S., Bochtis D., Machine Learning in Agriculture: A Review. Sensors, 18 (8), pp. 2674, (2018). https://doi.org/10.3390/s18082674

[10] WordlView-3 Specifications. Available online: https://earth.esa.int/eogateway/missions/worldview-3 (accessed on 24 January 2022).

[11] JAXA, Global PALSAR-2/PALSAR/JERS-1 Mosaic and Forest/Non-Forest Map (FNF) Dataset Description, pp. 18, (2022). https://www.eorc.jaxa.jp/ALOS/en/dataset/pdf/DatasetDescription_PALSAR2_FNF_v200a.pdf

[12] Digital Globe WorldView-3 datasheet, pp. 2, (2021). https://www.l3harrisgeospatial.com/Portals/0/pdfs/DG2017_WorldView-3_DS.pdf

[13] Prakash N., Manconi A., Loew S., Mapping landslides on EO data: Performance of deep learning models vs. traditional machine learning models. Remote Sens., 12, 346, pp. 12, (2020)

[14] Liu P., Wei Y., Wang Q., Chen Y., Xie J., Research on post-earthquake landslide extraction algorithm based on improved U-Net model. Remote Sens. 12, pp. 894, (2020).

[15] Wu J., Shi Y., Wang W., Fault Imaging of Seismic Data Based on a Modified U-Net with Dilated Convolution. Appl. Sci. 12, pp. 2451, (2022). https://doi.org/10.3390/ app12052451

[16] Lin L., Zhong Z., Cai Z., Sun A.Y., Li C., Automatic Geological Fault Identification from Seismic Data Using 2.5 D Channel Attention U-net. Geophysics, 87 (4), pp. 1-58, (2022).

[17] Zhang S., Qi J., A Sensitivity Analysis of River Environment Factors Through Deep Learning. International Journal of GEOMATE, Vol.23, Issue 97, pp. 146-153. (2022).

[18] Opara J.N., Thein A.B.B., Izumi S., Yasuhara H., Chun P., Defect detection on asphalt pavement by deep learning, International Journal of GEOMATE, Vol.21, Issue 83, pp.-87-94, (2021).

[19] Nanda M., Rizal S., Abdullah F., Idroes R., Ismail N., Mapping faults distribution based on DEM data for regional spatial plan assessment of Sabang municipality, Indonesia., International Journal of GEOMATE, Vol.19, Issue 76, pp. 197–204, (2020). DOI: https://doi.org/10.21660/2022.97.3357

[20] Kulkarni S., Kedar M., Rege P.P., Comparison of Different Speckle Noise Reduction Filters for RISAT-1 SAR Imagery. International

conference on communication and signal processing (ICCSP), pp. 537-541, IEEE, (2018).
[21] Jiang Z., Pan W.D., Shen H., Spatially and spectrally concatenated neural networks for efficient lossless compression of hyperspectral imagery. Journal of Imaging, 6 (6), pp. 38, (2020).
[22] Fetai B., Račič M., Lisec A., Deep Learning for Detection of Visible Land Boundaries from UAV Imagery. Remote Sensing, 13 (11), pp. 2077, (2021).
[23] Exelis Visual Information Solutions. ENVI Deep Learning—Pixel Segmentation Training Background. Available online: https://www.l3harrisgeospatial.com/docs/PixelSegmentationTrainingBackground.html (accessed on 11 May 2023).
[24] Sauk B., Sahinidis N.V., Backward Stepwise Elimination: Approximation Guarantee, a Batched GPU Algorithm, and Empirical Investigation. SN COMPUT. SCI. 2, pp. 396 (2021). https://doi.org/10.1007/s42979-021-00788-1
[25] Subraja N., Venkatasekhar D., Satellite Image Segmentation using Modified U-Net Convolutional Networks. In 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), pp. 1706 – 1713, IEEE, (2022).
[26] Zhang P., Xu C., Ma S., Shao X., Tian Y., Wen B., Automatic Extraction of Seismic Landslides in Large Areas with Complex Environments Based on Deep Learning: An Example of the 2018 Iburi Earthquake, Japan. Remote Sens., 12, pp. 3992, (2020). https://doi.org/10.3390/rs12233992