HAL
open science

Generative adversarial networks as a novel approach for tectonic fault and fracture extraction in high resolution satellite and airborne optical images

Bahram Jafrasteh, Isabelle Manighetti, Josiane Zerubia

To cite this version:

Bahram Jafrasteh, Isabelle Manighetti, Josiane Zerubia. Generative adversarial networks as a novel approach for tectonic fault and fracture extraction in high resolution satellite and airborne optical images. ISPRS 2020 - XXIV International Society of Photogrammetry and Remote Sensing Congress, Aug 2020, Nice / Virtual, France. pp.1219-1227, 10.5194/isprs-archives-XLIII-B3-2020-1219-2020 . hal-02982510

HAL Id: hal-02982510

https://inria.hal.science/hal-02982510v1

Submitted on 28 Oct 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

GENERATIVE ADVERSARIAL NETWORKS AS A NOVEL APPROACH FOR TECTONIC FAULT AND FRACTURE EXTRACTION IN HIGH-RESOLUTION SATELLITE AND AIRBORNE OPTICAL IMAGES

Bahram Jafrasteh[1,2] *, Isabelle Manighetti[1], Josiane Zerubia[2]

[1] Géoazur, Université Côte d'Azur, Observatoire de la Côte d'Azur, IRD, CNRS, Sophia Antipolis, France - (manighetti, jafra)@geoazur.unice.fr

[2] INRIA, Université Côte d'Azur, Sophia Antipolis, France - (Josiane.Zerubia, Bahram.Jafrasteh)@inria.fr

KEY WORDS: Remote sensing, Deep learning, Curvilinear feature extraction, Image processing, Generative adversarial networks, High resolution, Tectonic fault and fractures, Fault mapping

ABSTRACT:

We develop a novel method based on Deep Convolutional Networks (DCN) to automate the identification and mapping of fracture and fault traces in optical images. The method employs two DCNs in a two players game: a first network, called Generator, learns to segment images to make them resembling the ground truth; a second network, called Discriminator, measures the differences between the ground truth image and each segmented image and sends its score feedback to the Generator; based on these scores, the Generator improves its segmentation progressively. As we condition both networks to the ground truth images, the method is called Conditional Generative Adversarial Network (CGAN). We propose a new loss function for both the Generator and the Discriminator networks, to improve their accuracy. Using two criteria and a manually annotated optical image, we compare the generalization performance of the proposed method to that of a classical DCN architecture, U-net. The comparison demonstrates the suitability of the proposed CGAN architecture. Further work is however needed to improve its efficiency.

1. INTRODUCTION

Fractures and faults are ubiquitous on Earth. Fractures are generally small-scale (lengths $< 10^{-3}$-10^{3} m), shallow (a few km depth) cracks in the rocks, whereas faults are larger features (lengths 10^{-3}-10^{3} km, widths $1 - 10^{2}$ km from ground surface to depth) accommodating significant cumulative displacements (10^{-3}-10^{3} km) over millions of years (e.g., Manighetti et al., 2001a). Faults have a 3D architecture, including fault elements of different lengths and a myriad of fractures off the principal fault plane (e.g., Perrin et al., 2016). When submitted to large stresses, as those imposed by Plate Tectonics at the global scale, faults may produce earthquakes. Because these earthquakes can be devastating, decades of scientific works have been done to understand them and provide clues to anticipate them. Among key information, the way fractures and faults are distributed in a region partly controls the location, distribution and magnitudes of the earthquakes in that region. Locating faults and fractures accurately is thus of upmost importance to assess seismic hazard (e.g., Benedetti et al., 2013; Walpersdorf et al., 2014). This can be done by identifying the traces that faults form as they intersect the ground surface. These surface traces have specific properties that can be recognized visually. Therefore, many studies have been conducted to identify fault and fracture surface traces in many regions worldwide. Trace identification was done visually, from direct observation in the field or from analysis of remote sensing image data, while mapping of the fault and fracture traces, commonly called "annotation", was done manually (e.g., Manighetti et al., 1998, 2001b, 2015, 2020).

However, manual annotation on image data requests a tremendous time, and relies on expert knowledge, which might not be always available. This calls for the need to develop a fast, reliable and accurate automatic fault identification and mapping method in image data, that is derived from expert knowledge yet can be operated without further inputs of the experts. Such methods have been developed in earlier works (Cartabia et al., 1994; Mavrantza and Argialas, 2003; Aghaee Rad, 2019; Farahbakhsh et al., 2020), and are generally divided in two categories: unsupervised and supervised. The latter takes advantage of the expert knowledge and is thus expected to generate reliable results; however, it requests annotations produced by the experts, what may be a long and difficult task. On the contrary, the unsupervised methods do not involve any annotation inputs from the experts, and are thus of more simple use; however they are more prone to mistakes such as detecting irrelevant features. In the domain of fault detection, the unsupervised methods might for instance mistakenly consider any linear feature, such as image edges or roads, as fault or fracture traces. Correcting these errors would then request significant post-processing by the experts.

Tectonic faults have curvilinear traces at the ground surface. We claim here that these curvilinear traces can be extracted from image data using a combination of computer vision and machine learning methods, provided that we have access to expert annotations. We are thus conducting supervised learning. Supervised learning is a machine learning task that learns how to model a series of input variables to one or more output variables. One of the most common non-linear methods of supervised learning is called Artificial Neural networks (ANN). ANN, inspired by biological neurons, is an information processing unit having artificial neurons and "hidden layers". Each hidden layer can contain one or more neurons. The neurons are dedicated to collapse and map the received input variables into one single variable. For that, they are conducting different operations to the input data such as attributing them some weights, modifying them through bias, or enhancing them through linear or non-linear functions called activation functions. The neuron

* Corresponding author

output is then sent to one or more neurons of the next hidden layer. These operations are repeated until the total information reaches the last output layer. Weights and biases are the main variables that need to be adjusted during the calculations, and these adjustments are called the ANN training. This training is usually done with a back-propagation algorithm (details can be found in Leung and Haykin, 1991). An ANN with more than one hidden layer is known as a Deep Neural Network (DNN).
Image processing techniques can be combined with DNN to improve the performance of both methods. One of the most commonly used image processing techniques is image convolution. Image convolution is a process through which one pixel is added to its local neighboring pixels, using a weighted matrix known as kernel or filter. If we replace the DNN neurons with an image convolution operation, then the weights of the network are replaced by convolution kernels. This architecture, which benefits from the high computational capacity of neural networks and the local feature extraction capability of image convolution, is called Deep Convolutional Network or DCN. Here we propose to use a particular architecture of DCN: a Generative Adversarial Network (GAN) (Goodfellow et al., 2014). GAN, inspired from game theory, combines two DCNs in a min-max alternating fashion: A first network, called Generator (G), is trained to generate synthetic images, also called fake images in GAN, that resemble the annotated ones; then a second network, called Discriminator (D), is trained to discriminate, through comparisons between synthetic and annotated images, which synthetic images produce the most reliable results. Because we need to present an annotated ground truth image to the GAN system, the architecture is called a "Conditional generative adversarial network" or CGAN. The CGAN approach has been used in earlier works to successfully extract roads and buildings from aerial and spatial images (Pan et al., 2019; Zhang et al., 2019), but, to our knowledge, it has never been applied to fault and fracture detection.

Therefore, in this study, we develop a procedure to adapt CGAN for fault and fracture extraction in optical image data. We also investigate the relative efficiency of CGAN and DCN for fault and fracture extraction.

2. AVAILABLE APPROACHES

As said earlier, the available approaches are classified in two main categories: unsupervised (Lacoste et al., 2005; Jeong et al., 2015, 2016) and supervised methods (Parikh et al., 1990; Cartabia et al., 1994; Aghaee Rad, 2019).

Most unsupervised methods rely on edge detection. Edge detection aims to identify pixels in an image where intensity and brightness change sharply. Algorithms based on edge detection have been widely used to extract lineaments, including faults, from satellite images (Parikh et al., 1991; Mavrantza and Argialas, 2003; Marghany and Hashim, 2010; Corgne et al., 2010; Rahnama and Gloaguen, 2014). Introducing pre-processing steps such as histogram equalization (Marghany and Hashim, 2010) or directional filters (Mavrantza and Argialas, 2003; Marghany and Hashim, 2010; Soto-Pinto et al., 2013; Borisova et al., 2014; Farahbakhsh et al., 2020) has helped to improve the lineament extraction. However, edge detection generally fails detecting faults accurately because it identifies as faults other irrelevant linear features.
ANN has also been applied to lineament detection in optical images (Parikh et al., 1990; Cartabia et al., 1994; Aghaee Rad, 2019). The approach can handle multiple inputs and outputs and clearly provides more satisfying results. Among these studies, Aghaee Rad, 2019 used a DCN to compute the pixel probability in an image to be a fault. However, they did not use optical images.

In this study, we go beyond these prior works and propose a novel approach to accurately extract lineaments, here faults and fractures, from optical images, based on a generative adversarial neural network.

3. NEW METHODOLOGY

We have developed a conditional GAN structure (Figure 1). The network consists of two DCN sub-networks: a Generator network G and a Discriminator network D. The Generator, also called Segmentor, annotates each pixel of the input image with a generated label probability. The Discriminator, also called Classifier, compares the generated labels for each pixel in the image with the ground truth annotated image and sends its feedback to the Generator.

3.1 Deep Convolutional Neural networks

Generally, three datasets are introduced in a neural network architecture: a training dataset, made to learn the DCN with expert annotations of the elements of concern in a part of the input image; a test dataset made to quantify the ability of the trained DCN (generalization ability) to predict the elements of concern in another unseen part of the image; and a previously unseen validation dataset made to both check the generalization performance as training is progressing, and to stop the training before we over-train the DCN ("overfitting"); overfitting indeed happens when the network is trained so much on a set of images that it can only reproduces that images, with high accuracy. As the computational cost to feed a large size image is very high, the training is done on smaller sub-images extracted from the source image. Each sub-image is a matrix of pixel values that form the network input. We use here a DCN, called U-net (upper right part of Figure 1) (Ronneberger et al., 2015). U-net has two major parts, a first one which is a contraction path called Encoder, dedicated to identify the elements of concern, and a second expansion part, called Decoder, dedicated to locate accurately these identified elements within the image. The first layer of the U-net includes the input image as a multitude of matrices. The next layer, which is the first hidden layer, convolves (i.e., applies a filter) each input matrix to highlight some basic features in the image (such as edges, sharp brightness changes, etc.). The convolution is done using a sliding window (called stride) of 1 pixel that rolls over the sub-image, and a filter weight, as shown in Figure (2). This process reduces the final size of the image (downsampling). In the next step, the convolved image (also called "feature") is presented to the next layer, and the same operation is performed with more complex filters. The process is repeated until the multi-convolved and downsampled features reach the end of the contraction path. From this point forward, the image sustains a transposed convolution, that is, similar to the convolution but in a backward direction (Figure (2)). The objective here is to find the accurate position of the identified elements in the image. It results that the image size is increased across the successive layers (upsampling), until it is back to its original size. The main objective of the training phase is to optimize the weights of the convolutions (Figure (2)). As these weights are optimized, the network generates images that resemble more closely to the annotated ground truth. This optimization is done to minimize,

using a loss function, the difference between the ground truth and the network output labels. The most common optimization is done by back-propagation using a Stochastic Gradient Descend algorithm (Bottou, 2010). Over the entire calculation, different procedures are applied to accelerate the training (such as mini-batch and batch normalization (Ioffe and Szegedy, 2015)). Skip connections are used to transfer the information between the encoder and the decoder parts. They thus connect the spatial feature maps in the Encoder with the reconstructed feature maps in the Decoder, and improve the generalization ability of the U-net (Ye and Sung, 2019).

One of the commonly used activation functions in the U-net architecture is Leaky ReLu (Xu et al., 2015). Leaky ReLU maps negative value of x to a linear function with slope λ and positive values of x to $f(x) = x$. It is defined as :

$$f(x) = \begin{cases} \lambda x & x < 0 \\ x & x \geq 0 \end{cases} \tag{1}$$

Here, instead of using one DCN architecture such as U-net, we use two DCNs organized as a generative adversarial network, similarly to a two players game. In this architecture, one of the networks, called Generator, learns to produce images as close as possible to the ground truth ones. The other network, called Discriminator, has access to both the ground truth images and the images produced by the Generator, and learns to discriminate them. The Generator changes its parameters according to the feedbacks it receives from the Discriminator, making it to progressively learn to produce images that resemble more and more to the ground truth ones.

3.2 GAN versus conditional GAN

Let's assume we have a dataset with N training images x_n and the corresponding ground truth label y_n, with probability distributions P_{data_x} and P_{data_y}, respectively. Then, the conventional loss function (Goodfellow et al., 2014) of the GAN network is defined as follows:

$$\min_D \max_G L_{GAN}(D,G) = \mathbf{E}_{\mathbf{x}\sim P_{data_\mathbf{x}}}[logD(\mathbf{x})] + \mathbf{E}_{\mathbf{z}\sim P_\mathbf{z}}[log(1 - D\big(G(\mathbf{z})\big)] \tag{2}$$

where $\mathbf{z}$ is an input noise variable drawn from a probability distribution $P_\mathbf{z}$. Conditional version of GAN can be constructed by feeding extra information y_n to both the Generator and Discriminator networks, instead of noise variable. One can define the following loss function for the Conditional GAN (CGAN):

$$\min_D \max_G L_{cGAN}(D,G) = \mathbf{E}_{\mathbf{x}\sim P_{data_\mathbf{x}}, P_{data_\mathbf{y}}}[logD\big((\mathbf{x},\mathbf{y})\big)] + \mathbf{E}_{\mathbf{x}\sim P_{data_\mathbf{x}}}[log(1 - D\big((\mathbf{x}, G(\mathbf{x}))\big)] \tag{3}$$

The first term in the above equation maximizes the probability to correctly predict the output variable $\mathbf{y}$. The second term is used to train the Generator network G to accurately segment the image and fool the Discriminator network D.

The two networks G and D are adversarially trained, that is, the Generator network minimizes the difference between the network-annotated images and the ground truth, while the Discriminator network separates the network-annotated images from the ground truth ones.

3.3 Conditional GAN for optical image segmentation

Let's H and W be the height and the width of the n_{th} image x_n, respectively. The dimension of the input image is $H \times W \times 3$, as we use three bands of the image, and the dimension of the output segmented image is $H \times W \times 1$.

The proposed loss function with N training samples is defined as:

$$\min_D \max_G L_{cGAN}(D,G) = \frac{1}{N}\sum_{n=1}^{N}\sum_{l=1}^{L}\Big\{||f_{l_D}(x_n, G(x_n)) - f_{l_D}(x_n, y_n)||\Big\} + \frac{1}{N}\sum_{n=1}^{N}||G(x_n) - y_n||^2 \tag{4}$$

Where f_{l_D} represents the features extracted from the l_{th} layer of the Discriminator network. The first part of the equation refers to the mean absolute error (MAE), called the L_1 loss function and the second part represents the mean squared error (MSE), called the L_2 loss function. Xue et al., 2018 mentioned that the first term of the above loss equation captures image features at different scales, thus hierarchical features. We defined the last term to empower the Generator network in extracting faults and fractures.

3.4 Network architecture and training specifications

As said earlier, the Generator network is a fully connected convolutional neural network with an encoder followed by a decoder. Transposed convolution layers are selected to do upsampling. The parameters of convolutional and transposed convolutional layers are as follows: kernel size: 4×4, stride: 2, activation function: leaky ReLU. The network takes advantage of the skip connection layers from U-net to recover the spatial information that can be lost otherwise from down-sampling at various resolutions (Drozdzal et al., 2016). The Discriminator network uses the encoder part of the Generator network and merges hierarchical features of the image to capture both long-range and short-range spatial relations among pixels (Figure 1). The parameters of this network are as follows: kernel size: 4×4, stride: 2, activation function: leaky ReLU. The back-propagation method is used to optimize the network parameters. There are two steps for the training: in the first step, we fix the Discriminator parameters and train the Generator. In the next step, we fix the parameters of the Generator and train the Discriminator using the same gradient as that obtained in the previous step. The two networks are competing to alternatively maximize and minimize the same loss function. As the epochs (i.e., number of iterations) of training increase, the two networks learn to better predict the output. In the final steps, the Generator produces probability maps that resemble the ground truth images manually annotated by the experts.

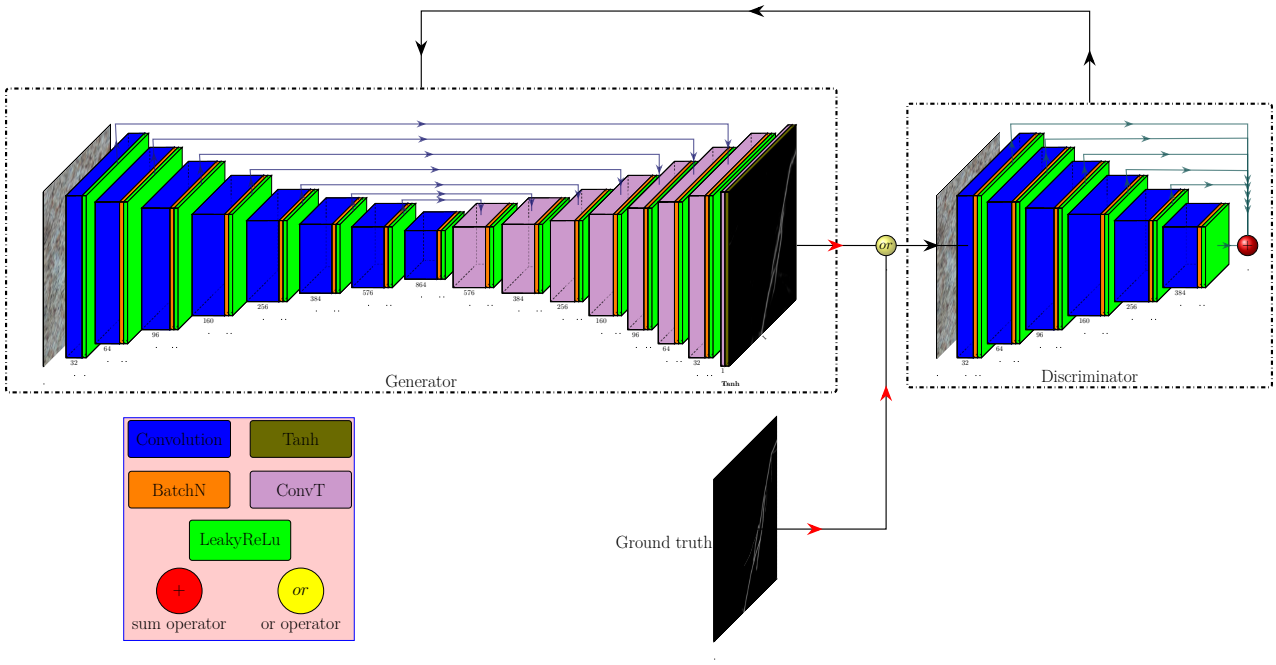


Figure 1. Proposed conditional GAN architecture and methodology. The network has one Generator and one Discriminator sub-networks. The input of the Discriminator is the output of Generator combined with the input image. As training progresses, the Generator produces segmented images resembling more and more to the ground truth. The Discriminator then learns to discriminate between the ground truth image and the images produced by the Generator.

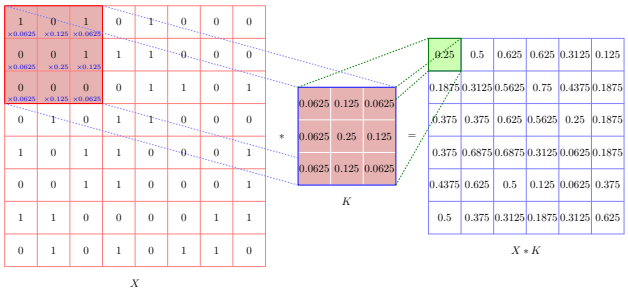


Figure 2. Example of a convolution of an image $\mathbf{X}$ with a Gaussian blur kernel (or filter) $\mathbf{K}$. $\mathbf{X}$ with size 8×8 passes through a convolution kernel function $\mathbf{K}$ with size 3×3. The filter moves toward the right part of the image and new pixel values are computed. The sliding window (stride) is of 1 pixel.

Training specifications are as follows: batch size number: 64; $\lambda = 0.2$; optimization algorithm: Adam (Kingma and Ba, 2014) with initial learning rate: 0.002, $\beta_1 = 0.5$ and $\beta_2 = 0.99$; number of epochs: 200. However, we allow the network to continue the training until the performance stops increasing over twenty consequent epochs.

4. EXPERIMENTAL STUDIES

Dataset: We use a georeferenced optical four band image with spatial resolution 5×10^{-4} m and size 20164×6807, constructed with photogrammetry from hundreds of hand-held camera pictures taken on the field (Figure 3(a)). Here, we use only the red (R), green (G) and blue (B) bands of the image. The image encompasses faults and fractures cutting a granite rock. The distorted or incomplete parts of the image at the site corners were not used. The entire image is divided in three regions, based on the geographical location, for training, validation and test, including 384, 94, 222 sub-images, respectively. This type of division which generates test sub-images that are far from the training ones is also recommended by Jafrasteh et al., 2018; Bastani et al., 2018. Figure 3(c)-3(d) shows one of the patches in the image. Figure 3(b) shows the available expert annotations. As these annotations are not homogeneous (parts of the image best annotated than others), we imposed a threshold to the input images: an image is presented to the network only if the number of its annotated pixels is higher than one hundred. Furthermore, there are different uncertainties in manual fault mapping: one is that actual fault traces cannot be located with a greater precision than a few pixels, even by the best experts; a second one is that some fault traces are uncertain, even for the best experts (see different colors in Figure 3(d), the red lines represent major certain faults, yellow lines are for minor certain faults, and green color with dashed lines represent uncertain faults). To address the former issue, we applied a Gaussian filter with $\sigma = 1.0$ to the annotated fault traces. The value of σ was selected based on trial and error in range 0.5-2.5, to find the best compromise between actual line thickness and efficient training of the model. To address the second issue, we basically considered two pixel classes: certain fault, and not-a-fault

Data Augmentation: Since the training samples are few, 384, we used different augmentation techniques to artificially increase the number of training samples, while making the algorithm invariant to changes in image brightness, contrast, fault orientations and positions, etc. (Perez and Wang, 2017). More specifically, we applied to the input images some horizontal and vertical reversals, crops and rotations, along with brightness, contrast, hue, and saturation changes (Krizhevsky et al., 2012; Taylor and Nitschke, 2017). An input image sustains each of these changes with a probability of 0.5.

Baselines: We compare the performance of our proposed method with that obtained with classical U-net. The architec-

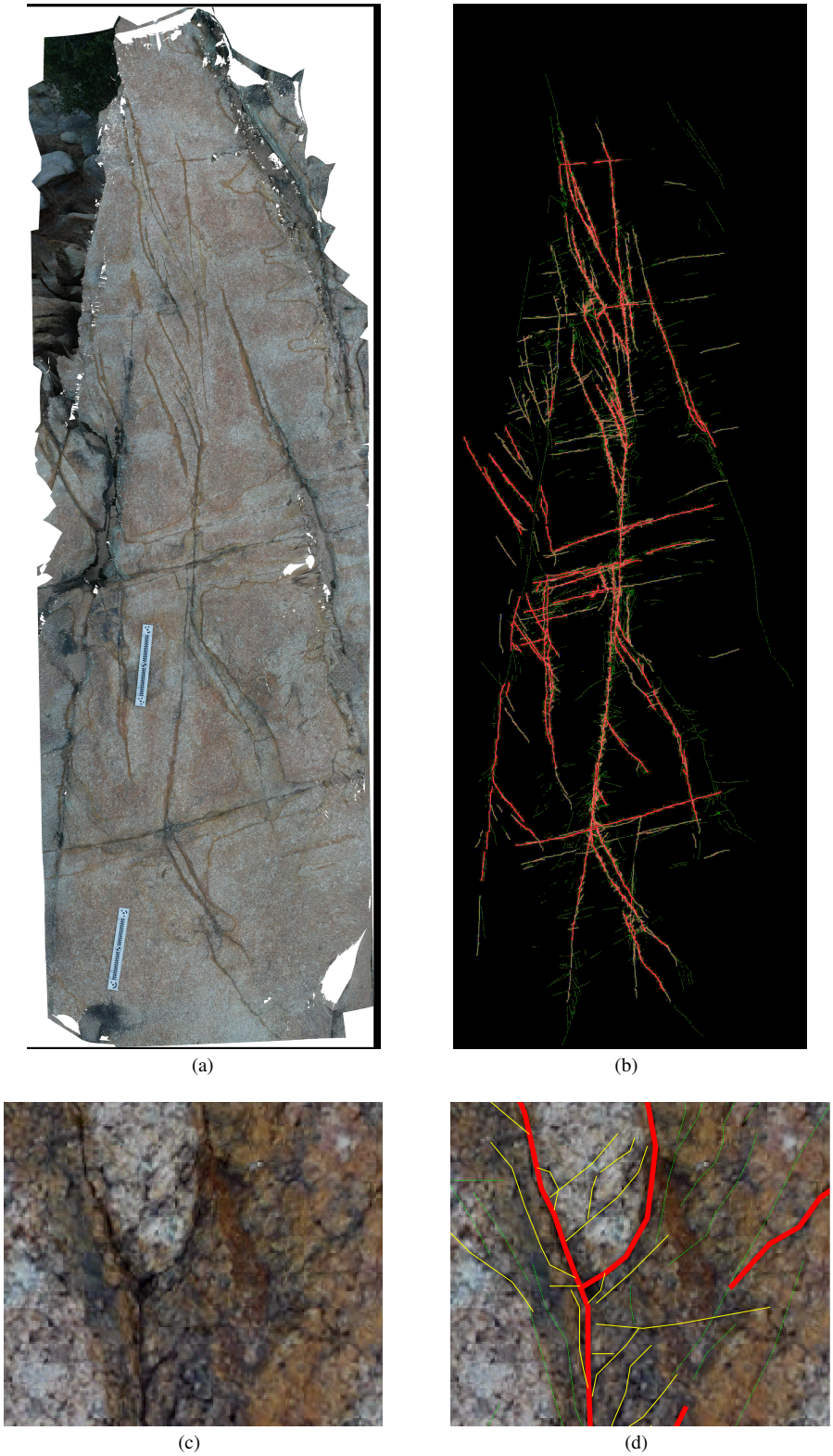


Figure 3. Optical image and manual expert annotations. (a) image of entire site; (b) annotations with major faults in red/pink and more minor faults in green; (c) patch of size 256×256 within input image; (d) manual annotations of the patch, with in red the major faults, in yellow the more minor faults, and in green the uncertain fault traces.

ture of U-net and the network parameters are the same as in our Generator network (Figure 1).

Metrics: We use a confusion matrix (Table 1) to classify the pixels as True Positive (TP), True Negative (TN), False Negative (FN) and False Positive (FP). TP and TN are predicted pixels in agreement with the same manual annotations from an expert (being a fault and not being a fault, respectively). FP and FN are predicted pixels in disagreement with the manual annotations (prediction of a fault and of absence of a fault where the manual annotations say the opposite). We investigate the generalization performance of our method using Dice similarity coefficient and Intersection Over Union (IOU). The higher the values of these two criteria the better is the accuracy of the proposed method.

Dice similarity coefficient (Sudre et al., 2017) is:

$$Dice = \frac{2TP}{2TP + FN + FP}. \tag{5}$$

IOU (Jaccard, 1912) is :

$$IOU = \frac{TP}{TP + FN + FP}. \tag{6}$$

A perfect match between ground truth and predicted maps would make these two criteria equal to one.

		Predicted pixel value	
		Fault	no Fault
Ground truth	Fault	**True Positive**	**False Negative**
	no Fault	**False Positive**	**True Negative**

Table 1. The confusion matrix to classify pixel values.

5. RESULTS AND DISCUSSION

Table 2 shows IOU and Dice values for our proposed CGAN and the most classical U-net architecture, from the test images. The two performance indicators are better for the novel method we proposed than for U-net.

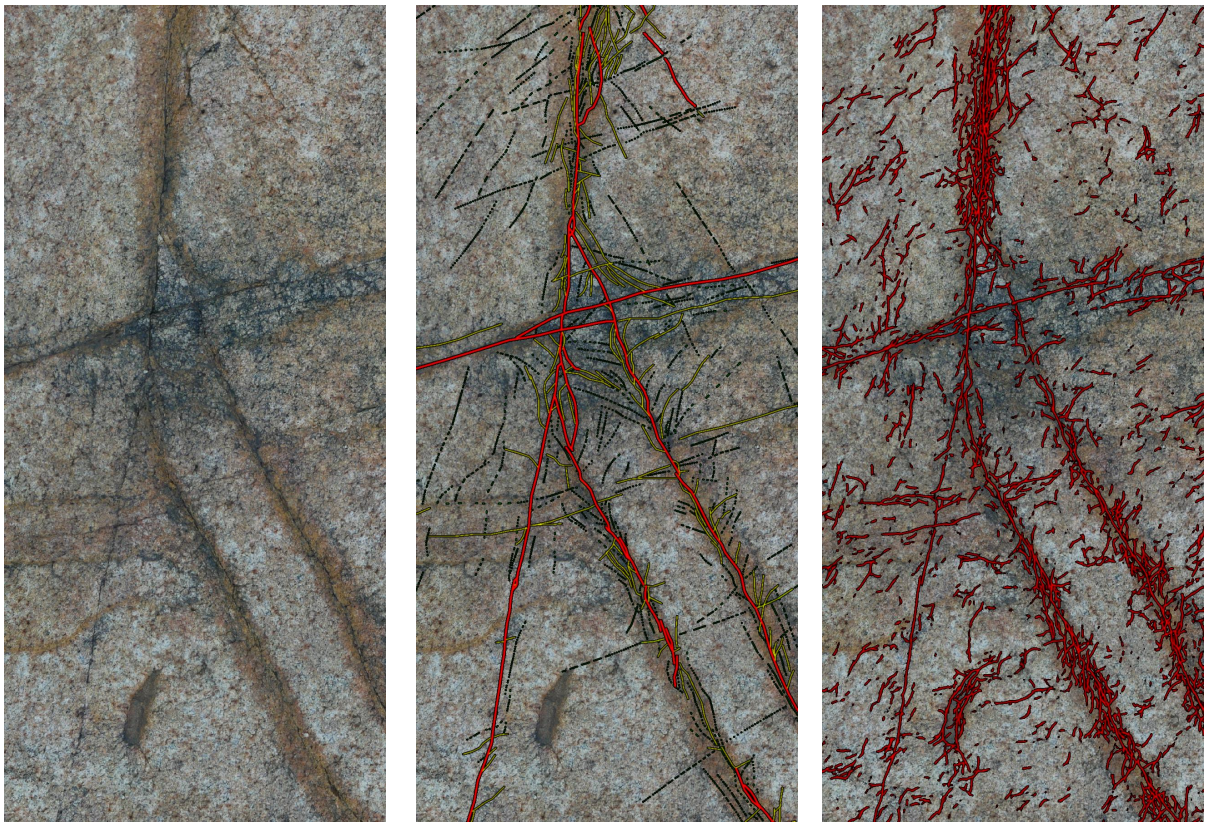
	IOU	Dice
U-net	0.057	0.108
Proposed CGAN	**0.063**	**0.119**

Table 2. Performance of the proposed conditional GAN and U-net methods.

Figure 4 shows a patch of the optical image in the test region, and compares the expert annotations with our predictions. The latter are presented as probabilities. We see that our predictions well produce the major faults. They also detect many of the more minor faults, even though their continuity is not totally well predicted. Many short fractures are predicted at a scale that goes beyond the annotated map. They might identify real, unmapped features, but this needs to be examined in subsequent work by the experts. We also note that the network mistakenly identified the edges of a rock inclusion as a fault set. This problem arises from the training set not including such inclusion. Therefore this problem will be solved easily in subsequent work by introducing images containing rock inclusion into the training set. Altogether our results demonstrate that Conditional GAN architecture with careful specifications can be employed to analyze optical images and extract curvilinear fault traces accurately. To the best of our knowledge, this is the first time that GAN architecture is adapted to extract geological faults and fractures from optical images. The main advantage of this approach compared to unsupervised methods is its ability to capture both linear and non-linear structures in a meaningful manner.

6. CONCLUSIONS

In this study, we proposed a novel methodology to extract fault and fracture traces in optical images. The methodology is based on conditional GAN architecture. We conditioned the network from a manually annotated image. As the actual fault traces have a certain "thickness", we represented them with a Gaussian filter. Then, we introduced these classified annotated images to CGAN. The CGAN takes advantage of a novel loss function to improve the performance and extract hierarchical features from the input image. Using the Dice and IOU criteria, we compared the performance of our CGAN approach with that of the most classical U-net architecture, both applied to a large size and high resolution optical image. The higher values of these criteria for the CGAN approach demonstrate its higher performance. More work still needs to be done however, especially to include more fault classes reproducing their natural hierarchy (major faults, more minor faults, faults uncertain at different degrees, etc.), and to take into account that manual fault maps are inherently incomplete; this might bias the performance criteria analysis. Furthermore, most faults form a topographic imprint, and therefore, integrating topography in the input data would be useful. Eventually, we will need to apply our new method to different datasets (different image types and resolutions, including satellite and aerial images), so as to verify that our novel approach can be generalized.



(a) (b) (c)

Figure 4. An image patch, 1800×840, taken from the test region (a), the corresponding annotation made by human expert (major faults in red, more minor faults in yellow, and uncertain faults in black) (b) and our CGAN (c) prediction.

References

Aghaee Rad, M. A., 2019. Machine learning of lineaments from magnetic, gravity and elevation maps. PhD thesis, University of British Columbia.

Bastani, F., He, S., Abbar, S., Alizadeh, M., Balakrishnan, H., Chawla, S., Madden, S., DeWitt, D., 2018. Roadtracer: Automatic extraction of road networks from aerial images. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4720–4728.

Benedetti, L., Manighetti, I., Gaudemer, Y., Finkel, R., Malavieille, J., Pou, K., Arnold, M., Aumaître, G., Bourlès, D., Keddadouche, K., 2013. Earthquake synchrony and clustering on Fucino faults (Central Italy) as revealed from in situ 36Cl exposure dating. *Journal of Geophysical Research: Solid Earth*, 118(9), 4948–4974.

Borisova, D., Jelev, G., Atanassov, V., Koprinkova-Hristova, P., Alexiev, K., 2014. Algorithms for lineaments detection in processing of multispectral images. *Earth Resources and Environmental Remote Sensing/GIS Applications V*, 9245, International Society for Optics and Photonics, 92451L.

Bottou, L., 2010. Large-scale machine learning with stochastic gradient descent. *Proceedings of COMPSTAT'2010*, Springer, 177–186.

Cartabia, G., Zerilli, A., Apolloni, B., 1994. Lineaments recognition for potential fields images using a learning algorithm for Boltzmann machines. *SEG Technical Program Expanded Abstracts 1994*, Society of Exploration Geophysicists, 432–435.

Corgne, S., Magagi, R., Yergeau, M., Sylla, D., 2010. An integrated approach to hydro-geological lineament mapping of a semi-arid region of West Africa using Radarsat-1 and GIS. *Remote Sensing of Environment*, 114(9), 1863–1875.

Drozdzal, M., Vorontsov, E., Chartrand, G., Kadoury, S., Pal, C., 2016. The importance of skip connections in biomedical image segmentation. *Deep Learning and Data Labeling for Medical Applications*, Springer, 179–187.

Farahbakhsh, E., Chandra, R., Olierook, H. K. H., Scalzo, R., Clark, C., Reddy, S. M., Müller, R. D., 2020. Computer vision-based framework for extracting tectonic lineaments from optical remote sensing data. *International Journal of Remote Sensing*, 41(5), 1760-1787.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial nets. *Advances in Neural Information Processing Systems*, 2672–2680.

Ioffe, S., Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37*, ICML'15, JMLR.org, 448–456.

Jaccard, P., 1912. The distribution of the flora in the alpine zone. 1. *New Phytologist*, 11(2), 37–50.

Jafrasteh, B., Fathianpour, N., Suárez, A., 2018. Comparison of machine learning methods for copper ore grade estimation. *Computational Geosciences*, 22(5), 1371–1388.

Jeong, S.-G., Tarabalka, Y., Nisse, N., Zerubia, J., 2016. Progressive Tree-like Curvilinear Structure Reconstruction with Structured Ranking Learning and Graph Algorithm. *arXiv preprint arXiv:1612.02631*.

Jeong, S.-G., Tarabalka, Y., Zerubia, J., 2015. Stochastic model for curvilinear structure reconstruction using morphological profiles. *2015 IEEE International Conference on Image Processing (ICIP)*, IEEE, 3575–3579.

Kingma, D. P., Ba, J., 2014. Adam: A Method for Stochastic Optimization. *arXiv preprint arXiv:1412.6980*.

Krizhevsky, A., Sutskever, I., Hinton, G. E., 2012. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 1097–1105.

Lacoste, C., Descombes, X., Zerubia, J., 2005. Point processes for unsupervised line network extraction in remote sensing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10), 1568–1579.

Leung, H., Haykin, S., 1991. The complex backpropagation algorithm. *IEEE Transactions on Signal Processing*, 39(9), 2101–2104.

Manighetti, I., King, G., Gaudemer, Y., Scholz, C., Doubre, C., 2001a. Slip accumulation and lateral propagation of active normal faults in Afar. *Journal of Geophysical Research: Solid Earth*, 106(B7), 13667–13696.

Manighetti, I., Perrin, C., Dominguez, S., Garambois, S., Gaudemer, Y., Malavieille, J., Matteo, L., Delor, E., Vitard, C., Beauprêtre, S., 2015. Recovering paleoearthquake slip record in a highly dynamic alluvial and tectonic region (Hope Fault, New Zealand) from airborne lidar. *Journal of Geophysical Research: Solid Earth*, 120(6), 4484–4509.

Manighetti, I., Perrin, C., Gaudemer, Y., Dominguez, S., Stewart, N., Malavieille, J., Garambois, S., 2020. Repeated giant earthquakes on the Wairarapa fault, new Zealand, revealed by Lidar-based paleoseismology. *Scientific Reports*, 10(1), 1–11.

Manighetti, I., Tapponnier, P., Courtillot, V., Gallet, Y., Jacques, E., Gillot, P.-Y., 2001b. Strain transfer between disconnected, propagating rifts in Afar. *Journal of Geophysical Research: Solid Earth*, 106(B7), 13613–13665.

Manighetti, I., Tapponnier, P., Gillot, P., Jacques, E., Courtillot, V., Armijo, R., Ruegg, J., King, G., 1998. Propagation of rifting along the Arabia-Somalia plate boundary: Into Afar. *Journal of Geophysical Research: Solid Earth*, 103(B3), 4947–4974.

Marghany, M., Hashim, M., 2010. Lineament mapping using multispectral remote sensing satellite data. *International Journal of the Physical Sciences*, 5(10), 1501–1507.

Mavrantza, O., Argialas, D. P., 2003. Implementation and evaluation of spatial filtering and edge detection techniques for lineament mapping: case study - Alevrada, Central Greece. M. Ehlers (ed.), *Remote Sensing for Environmental Monitoring, GIS Applications, and Geology II*, 4886, International Society for Optics and Photonics, SPIE, 417 – 428.

Pan, X., Yang, F., Gao, L., Chen, Z., Zhang, B., Fan, H., Ren, J., 2019. Building extraction from high-resolution aerial imagery using a generative adversarial network with spatial and channel attention mechanisms. *Remote Sensing*, 11(8), 917.

Parikh, J. A., DaPonte, J. S., Damodaran, M., Karageorgiou, A., Podaras, P., 1991. Comparison of backpropagation neural networks and statistical techniques for analysis of geological features in landsat imagery. *Applications of Artificial Neural Networks II*, 1469, 526–538.

Parikh, J. A., DaPonte, J. S., Damodaran, M., Sherman, P., 1990. Application of neural networks to pattern recognition problems in remote sensing and medical imagery. *Applications of Artificial Neural Networks*, 1294, International Society for Optics and Photonics, 146–160.

Perez, L., Wang, J., 2017. The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv:1712.04621*.

Perrin, C., Manighetti, I., Ampuero, J.-P., Cappa, F., Gaudemer, Y., 2016. Location of largest earthquake slip and fast rupture controlled by along-strike change in fault structural maturity due to fault growth. *Journal of Geophysical Research: Solid Earth*, 121(5), 3666–3685.

Rahnama, M., Gloaguen, R., 2014. TecLines: A MATLAB-based toolbox for tectonic lineament analysis from satellite images and DEMs, part 1: Line segment detection and extraction. *Remote Sensing*, 6(7), 5938–5958.

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*, Springer, 234–241.

Soto-Pinto, C., Arellano-Baeza, A., Sánchez, G., 2013. A new code for automatic detection and analysis of the lineament patterns for geophysical and geological purposes (ADALGEO). *Computers & Geosciences*, 57, 93–103.

Sudre, C. H., Li, W., Vercauteren, T., Ourselin, S., Cardoso, M. J., 2017. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer, 240–248.

Taylor, L., Nitschke, G., 2017. Improving deep learning using generic data augmentation. *arXiv preprint arXiv:1708.06020*.

Walpersdorf, A., Manighetti, I., Mousavi, Z., Tavakoli, F., Vergnolle, M., Jadidi, A., Hatzfeld, D., Aghamohammadi, A., Bigot, A., Djamour, Y. et al., 2014. Present-day kinematics and fault slip rates in eastern Iran, derived from 11 years of GPS data. *Journal of Geophysical Research: Solid Earth*, 119(2), 1359–1383.

Xu, B., Wang, N., Chen, T., Li, M., 2015. Empirical evaluation of rectified activations in convolutional network. *arXiv preprint arXiv:1505.00853*.

Xue, Y., Xu, T., Zhang, H., Long, L. R., Huang, X., 2018. Segan: Adversarial network with multi-scale l 1 loss for medical image segmentation. *Neuroinformatics*, 16(3-4), 383–392.

Ye, J. C., Sung, W. K., 2019. Understanding geometry of encoder-decoder CNNs. *arXiv preprint arXiv:1901.07647*.

Zhang, X., Han, X., Li, C., Tang, X., Zhou, H., Jiao, L., 2019. Aerial image road extraction based on an improved generative adversarial network. *Remote Sensing*, 11(8), 930.