# Deep learning for 3D fault detection within virtual reality visualization seismic volumes

Ali Fathalian, Marcelo Guarido, Daniel O. Trad, and Kristopher A. Innanen

## ABSTRACT

An important key for seismic structural interpretation and reservoir characterization is the delineating faults that are considered as seismic reflection discontinuities in conventional methods. Fault detection considers as a binary image segmentation problem of labeling a seismic image with ones on faults and zeros on non-faults using a fully supervised convolutional neural network. The network is trained by using 3D synthetic seismic images and their corresponding binary labels images. The network learns to calculate features that are important for fault detection after training with a synthetic data set. We apply this method to a migrated 3D volume from Australia. The results indicate that the neural network can predict faults from 3D seismic images. Effective Visualization analysis of 3D seismic data volumes is challenging because of their large volumes and highly complex nature. 3D virtual reality (VR) visualization is a useful tool that can benefit seismic data interpretation. In this paper, the seismic information extended reality analytics (SIERA) presents a seismic data visualization in an extended reality environment. Because it is highly customizable, it provides an effective way to interact with seismic data and machine learning results.

## INTRODUCTION

Faults are one type of geological structural surfaces that are typically recognized as lateral reflection discontinuities in a 3D seismic image (Hale, 2013). Hence, to automatically detect faults, some methods have been proposed to calculate signal continuity attributes such as semblance (Marfurt et al., 1998) and coherency (Marfurt et al., 1999; Bakker, 2002; Wu, 2017), or the opposite, reflection discontinuity such as variance (Van Bemmel and Pepper, 2000; Randen et al., 2001) and gradient magnitude (Aqrawi and Boe, 2011). However, reflector discontinuity alone is insufficient to detect faults, because incoherent noise and stratigraphic features can also correspond to reflection discontinuities in a seismic image (Hale, 2013).

Gersztenkorn and Marfurt (1999) suggest measuring continuity or discontinuity using vertically elongated windows for fault detection while using a larger horizontal window for detecting stratigraphic features. In this method, fault features can be enhanced and stratigraphic features can be suppressed because faults are often more vertically aligned in a seismic volume than the stratigraphic features. By assuming that faults are typically normal to reflections, Wu (2017) applies smoothing in directions perpendicular to seismic reflections in computing coherence or semblance. However, faults are rarely vertical or are not necessarily perpendicular to seismic reflections. Based on this observation, Hale (2013) proposes an efficient implementation of such a scanning processing to compute a fault-oriented semblance or fault likelihood volume to highlight fault positions from a seismic

volume. However, this method is expensive because it requires scanning over all possible combinations of fault strikes and dips to find the maximum fault likelihoods.
The convolutional-neural-network (CNN) method is another way that has been introduced to detect faults by pixel-wise fault classification (fault or non-fault) with multiple seismic attributes (Huang et al., 2017; Di et al., 2018; Guitton, 2018; Guo et al., 2018; Zhao and Mukhopadhyay, 2018). Wu et al. (2018) propose a CNN-based pixel-wise classification method to predict the fault probability and estimate the fault orientations at the same time. However, to predict fault at every image pixel, a local window or cube is required, which is computationally expensive.
In this paper, we present a CNN to detect faults from 3D seismic images, in which the fault detection is considered as a binary segmentation problem (Wu et al., 2019). We simplify the original U-Net by reducing the number of convolutional layers and features at each layer to save GPU memory and computational time, but still preserve high performance in the 3D fault detection tasks. The neural network is trained by using only 200 pairs of 3D synthetic seismic and fault volumes. After training with only the synthetic data sets, the neural network can accurately detect faults from 3D field seismic volumes.

The analysis of large amounts of complex data with unknown patterns and multiple interrelated parameters is a challenging task that requires scientific visualization tools with the potential for multidimensional data processing. Virtual reality (VR) is a scientific visualization tool that permits the data to be presented in dynamic images, revealing intrinsic patterns and dependencies. VR is quite demanding computationally but is made possible by modern high-performance computing and advanced computer graphics hardware and software. VR environments are the result of a demand for interactivity in computer visualization. Hence, the purpose of scientific computing visualization is drawing the three-dimensional space data field to make an intuitive image or graphics and making a three-dimensional interpretation for 3d seismic data. Therefore, detailed information about the geological structure is vividly displayed helping on the interpretation and analysis of the geological structure (Kaufman, 1994). Even for large seismic datasets, trained geoscientists can view the data details and make informed decisions concerning whether or not important features are present within the geologic structures. The limitation of conventional visualization tools is using computer monitors to act as the interface between users and data, and thus trust on two-dimensional projections to display these large, three-dimensional datasets. The limited screen space of monitors and representation of 3D data using a 2D screen puts large constraints on how one can interact with seismic data effectively. In this paper, we apply the SIERA tool (Lawton et al., 2020), an immersive analytics application helping geoscientists understand and visualize seismic data and the results of machine learning (ML) applied to seismic datasets (Douglas et al., 2019). By using virtual reality (Kwon et al., 2016; Usher et al., 2017), SIERA can improve the results by providing one with a virtual 3D environment in which to more naturally, efficiently, and effectively analyze 3D seismic data and higher-level ML results.

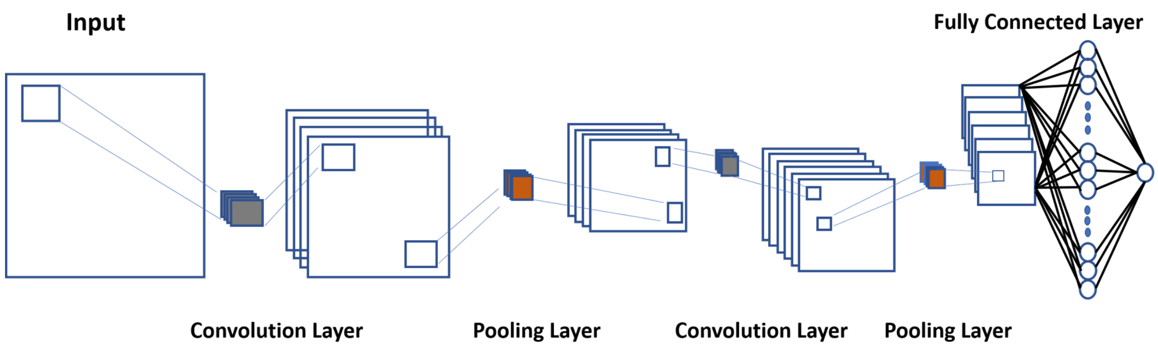


FIG. 1. Convolutional neural network architecture.

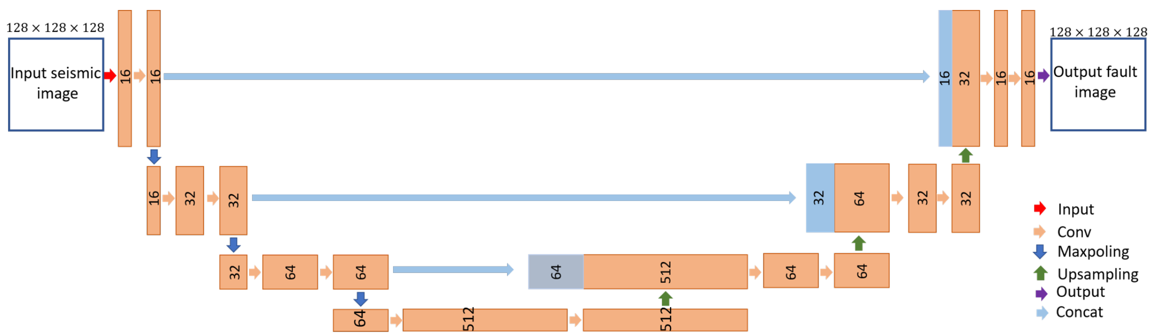


FIG. 2. A simplified end-to-end convolutional neural network (U-Net).

## DEEP LEARNING METHODOLOGY

Deep learning is a branch of machine learning the tries to model data using a a high-level representation by multiple layers of neurons with nonlinear transformations (LeCun et al., 2015). By increasing the number of layers in the deep learning models, this method becomes very powerful and is one of the most popular techniques to process big data. Using deep learning allows the computer to build complex concepts out of simpler concepts by constructing deeper neural networks. There are different deep learning architectures, such as deep neural networks, convolutional deep neural networks, deep belief networks, and recurrent neural networks with deep feature layers.

Convolutional neural networks (CNN) are a kind of feed-forward artificial neural network that can be formed among artificial neurons inspired by the organization of human neurons (LeCun et al., 2015). The CNN is applicable to solve problems such as image and video recognition (Krizhevsky et al., 2012; Karpathy et al., 2014) and recommender systems (Van den Oord et al., 2013). Figure 1 shows the CNN architecture formed by a stack of distinct hidden layers that transform the input data into an output volume. In this architecture, each convolution layer consists of a set of trainable filters. It is common to periodically insert a pooling layer in between convolution layers to reduce the spatial size of the representation which helps in decreasing the amount of parameters and therefore computation load.

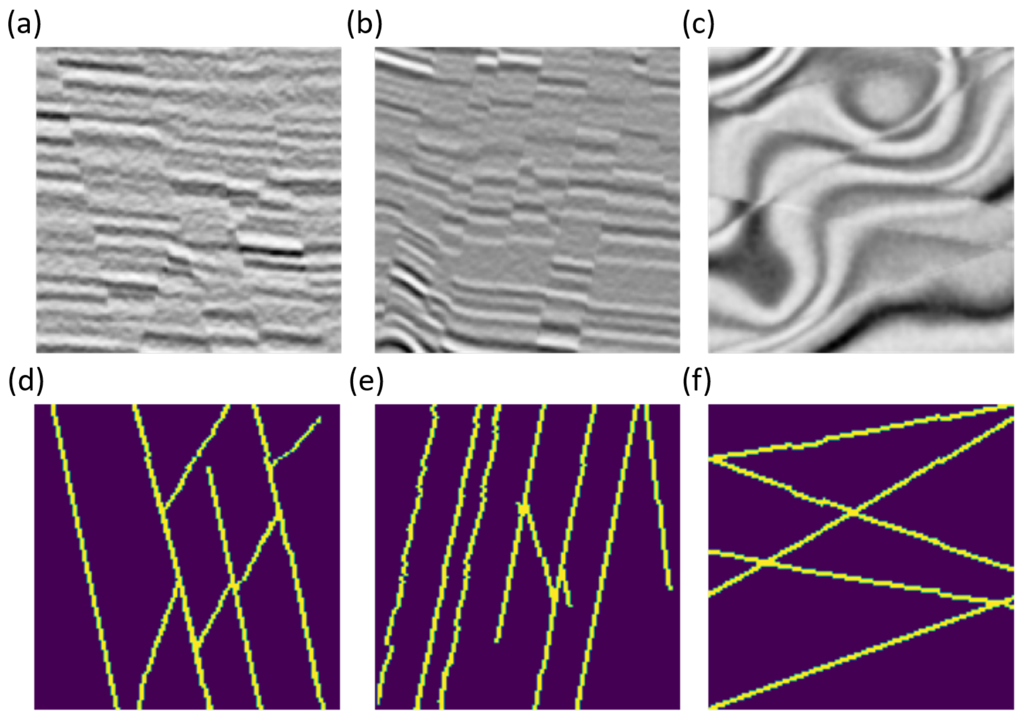


FIG. 3. The top row correspond to (a)inline, (b) crossline and (c)time slice synthetic seismic images that are cropped from the 3D synthetic training data sets. The bottom row shows corresponding true fault images (with labeling ones on faults and zeros elsewhere).

The main purpose of a CNN is to learn the feature mapping of an image and exploit it to make more nuanced feature mapping. In image segmentation using CNN, the feature map needs to convert the input into a vector and also reconstruct an image from this vector. The whole idea of U-Net (Ronneberger et al., 2015) is revolved around this problem. When converting the image into a vector, we learn the feature mapping of the image. In U-Net we use the same feature mapping to convert a vector again to the image. The workflow of U-Net preserves the structural integrity of the image which reduces distortions enormously.

In this paper, we use the U-Net architecture proposed by Wu et al. (2019). U-Net architecture is shown in Figure 2. This network contains a contracting path and an expansive path. In the contracting path (left side), each step contains two $3 \times 3 \times 3$ convolutional layers and a $2 \times 2 \times 2$ max pooling operation with stride 2 for downsampling and the number of features is doubled after each step. In the expansion path (right side), every step contains a $2 \times 2 \times 2$ upsampling operation, a concatenation with features from the left contracting path, and two $3 \times 3 \times 3$ convolutional layers. The final output layer is a $1 \times 1 \times 1$ convolutional layer, which has the same size as the input seismic image.

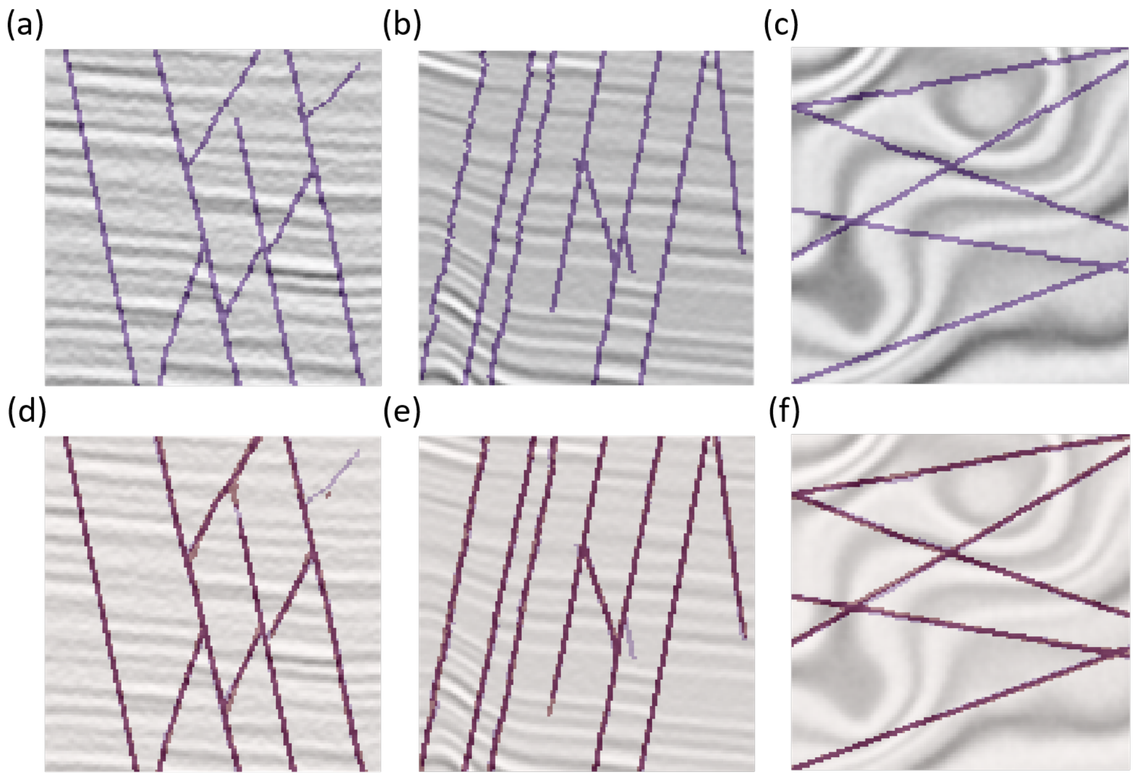


FIG. 4. Illustrate (a) the inline, b) crossline, and c) time slice of the validation seismic and label fault images, and d), e) and f) the corresponding predictions by using the CNN method. The fault locations and their predictions are indicated by the purple lines and red lines, respectively. The CNN method achieves the best performance in obtaining an accurate, clean, and complete fault detection.

## TRAINING AND VALIDATION

In a CNN, training a model often needs a large number of images and corresponding labels. To avoid the manual labor involved in labeling or interpreting faults in a 3D seismic image and inaccurate manual interpretation, Wu and Hale (2016) and Wu et al. (2019) propose an effective and efficient way to create synthetic seismic images and corresponding fault labels for training and validating a CNN model. With this approach, they created a set of synthetic seismic and fault images which we used to train our model. A final training set of seismic images (Figure 3) with size of $128 \times 128 \times 128$ can be obtained by cropping windows from the 3D Synthetic seismic image. In the second row of Figure 3, the corresponding binary fault labeling images are shown.

We used 200 pairs of unique seismic images and corresponding fault labeling images for training, and 20 pairs of seismic and fault labeling images for validation. Also, to increase the variety of the data sets and to prevent the neural network from learning incoherent patterns, we use data augmentation during the training. Simple data augmentations including vertical flip and rotation around the vertical time or depth axis during training are applied without rotating the seismic and fault volumes around the inline or crossline axis. The rotation around the inline or crossline axis causes vertical seismic structures and flat faults, which are geologically unrealistic.

The top row of Figure 4 (Figure 4a, 4b, and 4c) shows the inline, crossline, and time slice of the validation seismic and fault images. The corresponding true fault image (purple lines)is overlaid with the cropped seismic image. We choose a size $128 \times 128 \times 128$ for each 3D seismic or fault image because of limitations in memory and GPU. The network will train with 25 epochs, and all training images process at each epoch. We apply the trained model to the synthetic seismic volume (Figure 3a, 3b, and 3c) to verify the CNN model trained with 25 epochs. The second row of Figure 4 (4d, 4e, and 4f) Shows fault detection results that are computed by using the CNN-based segmentation. In these figures, the red lines refer to the prediction of the fault. By comparison between the fault detection results and the validation fault images, we found that the CNN method achieves the best performance in computing an accurate, clean, and complete fault detection, which is most consistent with the true fault labeling shown in Figure 3 (Figure 3d, 3e, and 3f).

## APPLICATION TO THE FIELD SEISMIC DATA

To verify the CNN we apply the same CNN model on the seismic images that are acquired from the Australian seismic data. The 3D seismic volume is a subset (128 [vertical] × 128 [inline] × 128 [crossline] samples) extracted from the Australian's Offshore seismic data, which is graciously provided by the geoscience Australia's data repository using the NOPIMS. Figure 5 shows the fault probability image predicted for crossline (Figure 5b), inline (Figure 5d), and time slice (Figure 5f) by using the trained CNN model. Our results show that the CCN model, although trained by only synthetic data sets, works very well to provide a clean and accurate prediction of faults in this field seismic image. In CNN

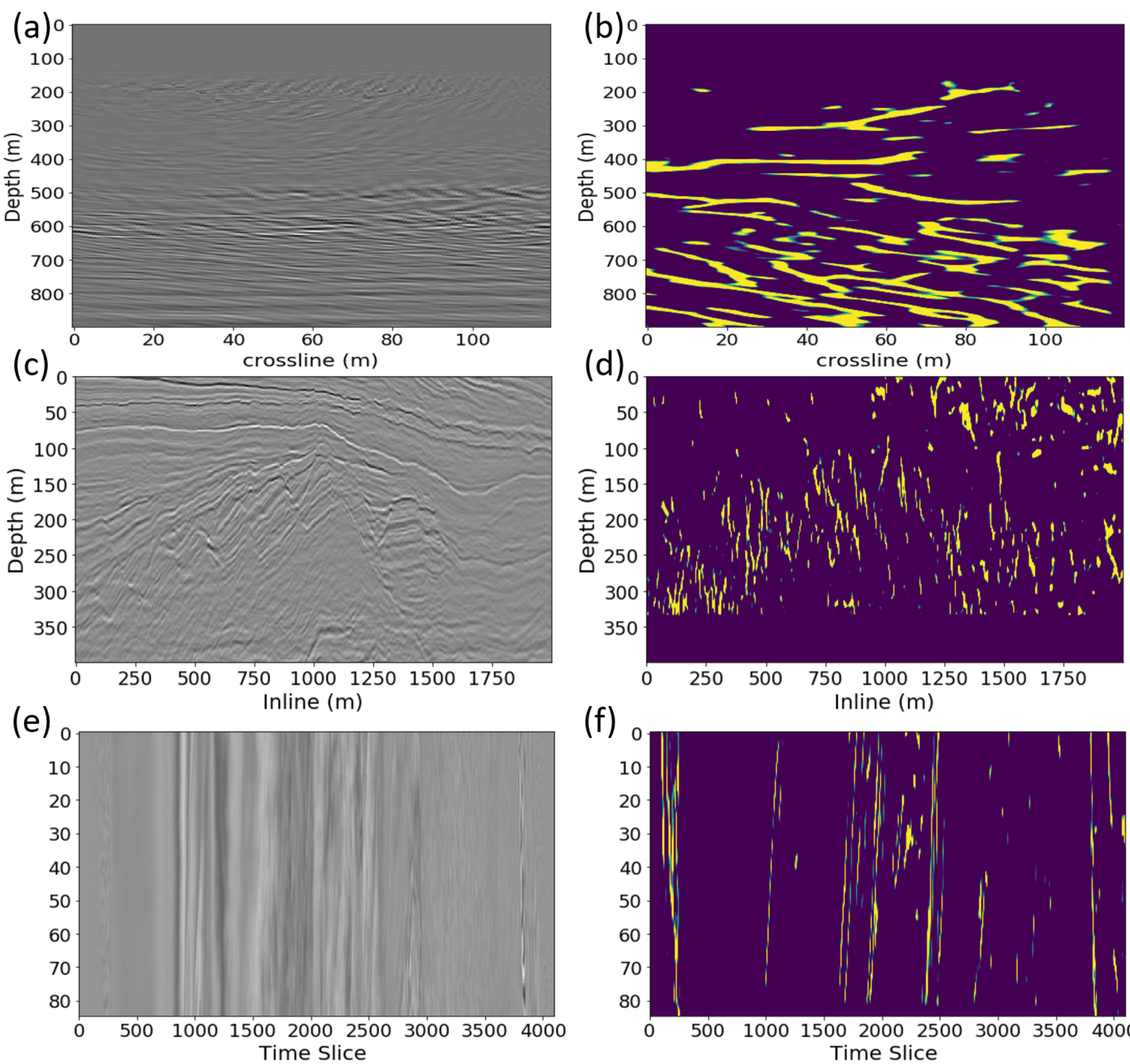


FIG. 5. The crossline, inline, and time slice of 3D seismic image are displayed with faults that are detected by using the trained CNN model.

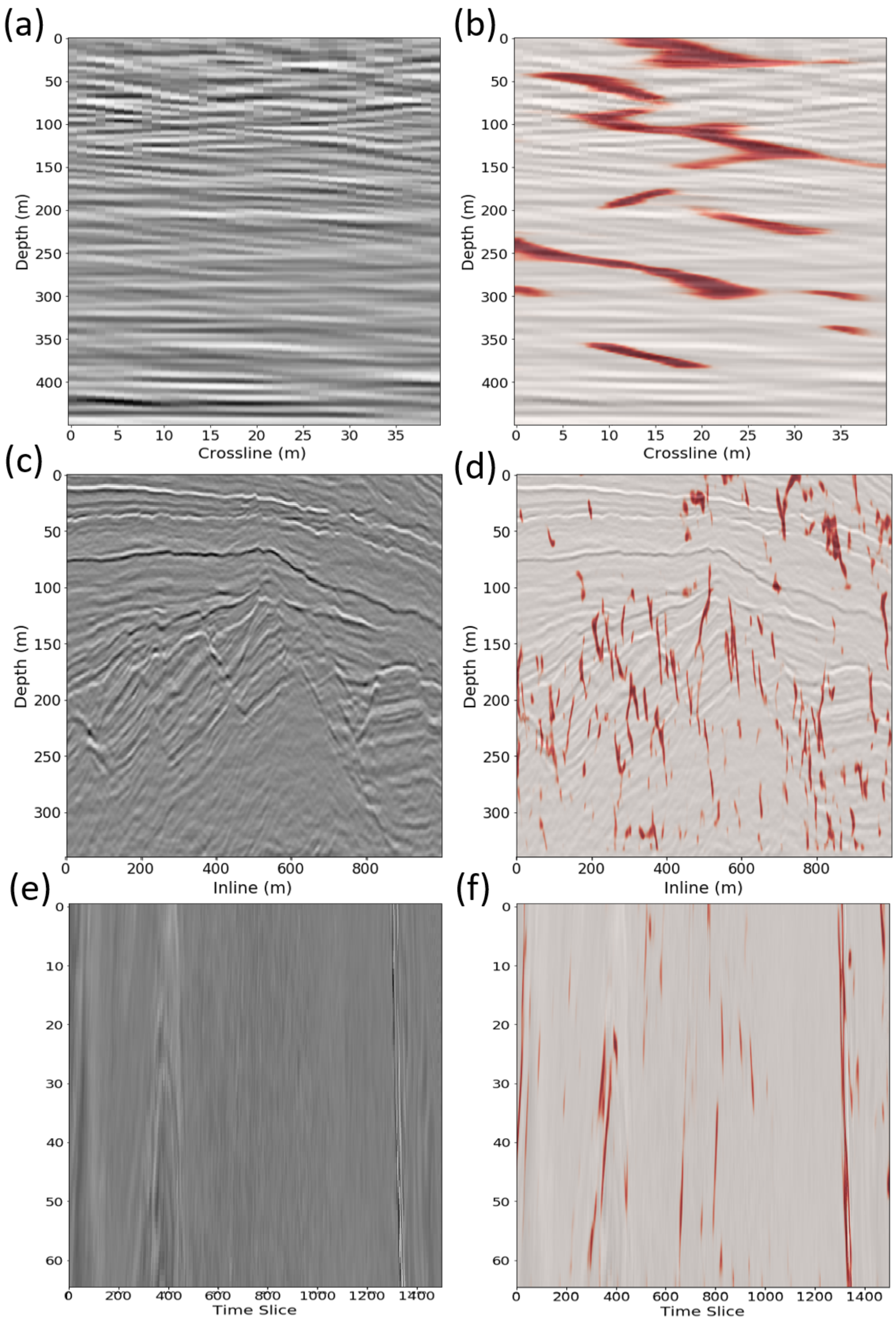


FIG. 6. Magnified view of the seismic images (crossline, inline, and time slice) and correspond fault detection using the CNN method.

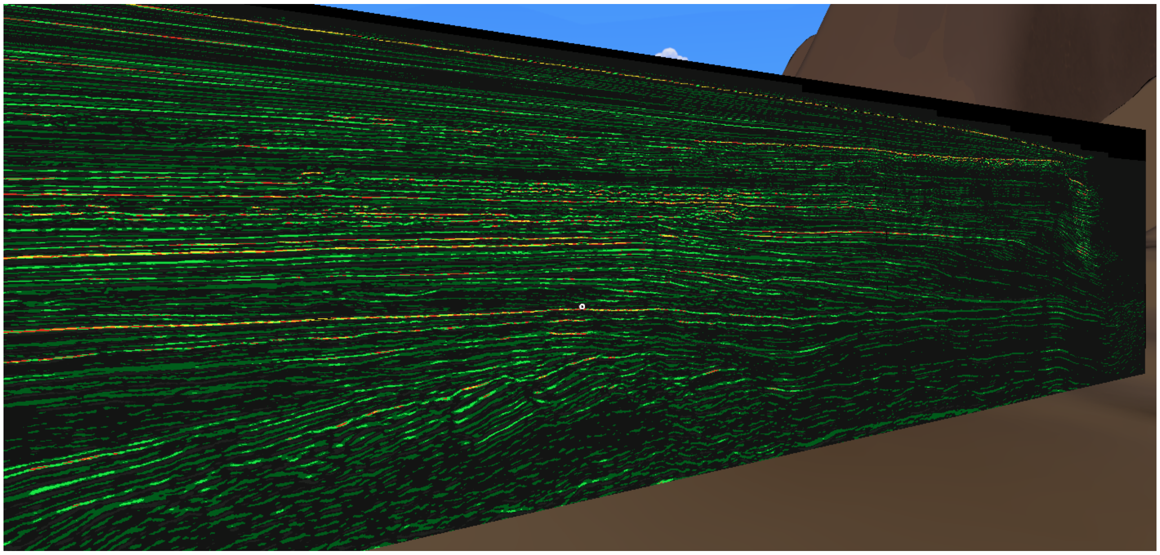

FIG. 7. A seismic data set (Australian data) visualized with SIERA in virtual reality as a three-dimensional volume.

fault detection images, most fault features have very high probabilities, and there are very limited noisy features. To verify the fault prediction result, we show in Figure 6 magnified views of the crossline, inline, and time slice images that are illustrated in Figure5. We observe that most faults are clearly labeled in the CNN fault probability images (Figure 6b, 6d, and 6f), and these faults can track by following the probability features.
We observe the CNN model works very well to detect faults in 3D field seismic, while this model is trained by using only 200 synthetic seismic images. The results show that the 3D fault prediction using the trained CNN model is highly efficient and also needs less time to compute the large CNN fault probability volume by using GPU.

## SEISMIC DATA VISUALIZATION

Visualization has strong advantages in complex data because it provides a visual encoding to get insight into revealing more information and feature presented in the data. It helps users perform effective analytics employing various human-computer interactions. SIERA can take whole seismic datasets and visualize them in VR as interactive 3D volumes by utilizing volumetric rendering techniques. These 3D volumes include numerous tiny voxels which each contain their own values for multiple assigned data variables, such as amplitude or ML certainty values. Each data volume contains a large number of voxels, which corresponds to the seismic data resolution. By reading the values of the data stored in voxels and applies end-user customizable color and transparency settings, SIERA presents a unique and desired visualization of the dataset volumetrically. Figure 7 shows the visualization of the Australian data. In VR, users can physically walk into these volumes and view the sub-surface structure from within, providing a more immersive way to view the data in close detail. Manipulation of the color and transparency for each of the many voxels which together make up a visualization's volume allows for the ability to

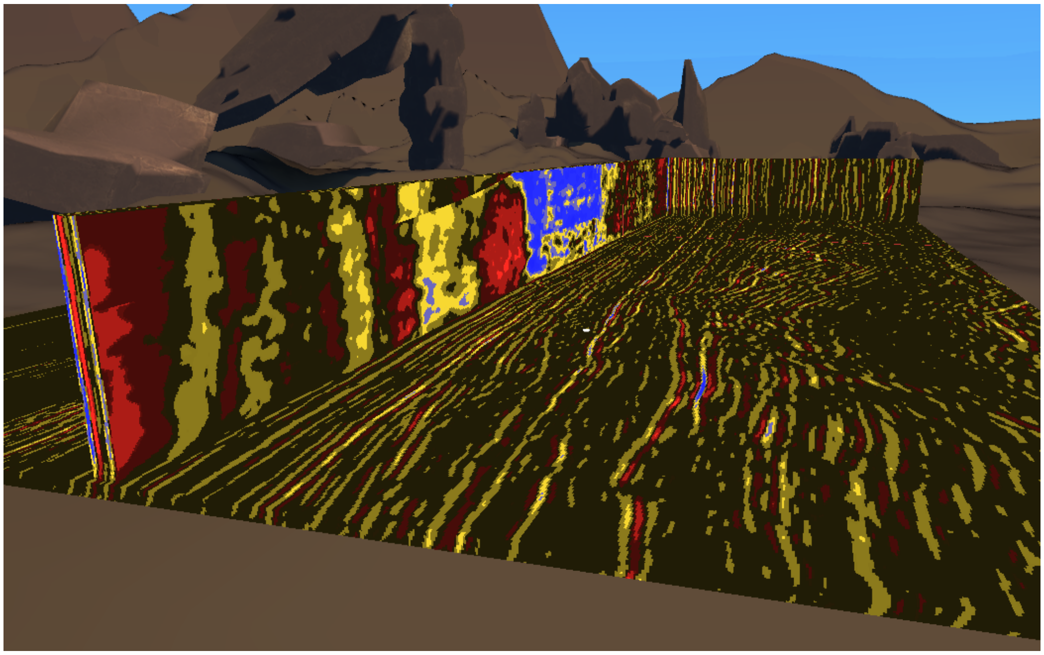

FIG. 8. Data visualization with colors chosen to correspond with specific data ranges.

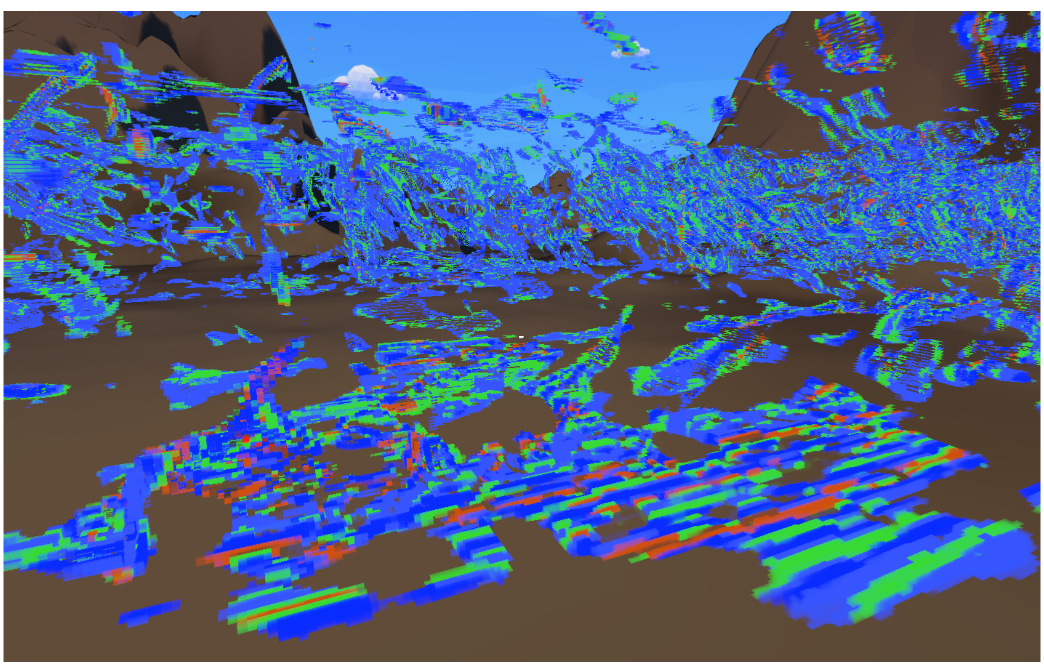

FIG. 9. Color-manipulated data volume with full-transparency applied to low ML certainty values, allowing one to view only the higher-certainty potential internal structures identified by ML.

emphasize important perspectives of the data by highlighting these with color (Figure 8), and the efficient filtering out of data which one does not want to view through adjusting transparency (Figure 9). Figure 9 shows that transparency manipulation is especially useful for allowing to see internal sub-surface structures that may not be visible when viewing an opaque visualization in its entirety. The color gradient mapped to an original dataset variable and the transparency gradient mapped to ML certainty values let a user see and directly compare how ML results relate back to aspects of the original data.

## CONCLUSIONS

We have presented a CNN to detect faults from 3D seismic images, in which the fault detection is considered as a binary segmentation problem. We simplify the original U-Net by reducing the number of convolutional layers and features at each layer to save GPU memory and computational time, but still preserves high performance in the 3D fault detection tasks. The neural network is trained by using only 200 pairs of 3D synthetic seismic and fault volumes. After training with only the synthetic data sets, the neural network can accurately detect faults from 3D field seismic volumes. To improve geoscientists abilities to analyze their data and to collaborate on its interpretation, virtual reality (VR) has significant potential. In this paper, we use the SIERA tool to provides a way to more intuitively and immersively interact with the three-dimensional nature of seismic data and ML results. This tool allows for analyzing several large data volumes simultaneously and scale them to sizes impractical with traditional techniques for better analysis and the creation of completely customizable and unique data visualizations through the use of voxel color and transparency manipulation.

## ACKNOWLEDGMENTS

We thank the sponsors of CREWES for continued support. This work was funded by CREWES industrial sponsors, NSERC (Natural Science and Engineering Research Council of Canada) through the grants CRDPJ 461179-13 and CRDPJ 543578-19. Partial funding also came from the Canada First Research Excellence Fund. This work was also funded by IBM Center for Advanced Studies Alberta (IBM CAS). We are thankful for fruitful discussions with D. Emery.

## REFERENCES

Aqrawi, A. A., and Boe, T. H., 2011, Improved fault segmentation using a dip guided and modified 3d sobel filter, *in* SEG Technical Program Expanded Abstracts 2011, Society of Exploration Geophysicists, 999–1003.

Bakker, P., 2002, Image structure analysis for seismic interpretation: Citeseer.

Di, H., Shafiq, M., and AlRegib, G., 2018, Patch-level mlp classification for improved fault detection, *in* SEG Technical Program Expanded Abstracts 2018, Society of Exploration Geophysicists, 2211–2215.

Douglas, N., Yim, D., Kartal, B., Hernandez-Leal, P., Maurer, F., and Taylor, M. E., 2019, Towers of saliency:

A reinforcement learning visualization using immersive environments, *in* Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces, 339–342.

Gersztenkorn, A., and Marfurt, K. J., 1999, Eigenstructure-based coherence computations as an aid to 3-d structural and stratigraphic mapping: Geophysics, **64**, No. 5, 1468–1479.

Guitton, A., 2018, 3d convolutional neural networks for fault interpretation, *in* 80th EAGE Conference and Exhibition 2018, vol. 2018, European Association of Geoscientists & Engineers, 1–5.

Guo, B., Li, L., and Luo, Y., 2018, A new method for automatic seismic fault detection using convolutional neural network, *in* SEG Technical Program Expanded Abstracts 2018, Society of Exploration Geophysicists, 1951–1955.

Hale, D., 2013, Methods to compute fault images, extract fault surfaces, and estimate fault throws from 3d seismic images: Geophysics, **78**, No. 2, O33–O43.

Huang, L., Dong, X., and Clee, T. E., 2017, A scalable deep learning platform for identifying geologic features from seismic attributes: The Leading Edge, **36**, No. 3, 249–256.

Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., and Fei-Fei, L., 2014, Large-scale video classification with convolutional neural networks, *in* Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, 1725–1732.

Kaufman, A., 1994, Trends in volume visualization and volume graphics.

Krizhevsky, A., Sutskever, I., and Hinton, G. E., 2012, Imagenet classification with deep convolutional neural networks, *in* Advances in neural information processing systems, 1097–1105.

Kwon, O.-H., Muelder, C., Lee, K., and Ma, K.-L., 2016, A study of layout, rendering, and interaction methods for immersive graph visualization: IEEE transactions on visualization and computer graphics, **22**, No. 7, 1802–1815.

Lawton, B., loans, H., Abou Gharib, P., Maurer, F., Guarido, M., Trad, D., and Fathalian, A., 2020, Siera: The seismic information extended reality analytics tool, *in* Proceedings of the 2020 ACM International Conference on Interactive Surfaces and Spaces.

LeCun, Y., Bengio, Y., and Hinton, G., 2015, Deep learning: nature, **521**, No. 7553, 436–444.

Marfurt, K. J., Kirlin, R. L., Farmer, S. L., and Bahorich, M. S., 1998, 3-d seismic attributes using a semblance-based coherency algorithm: Geophysics, **63**, No. 4, 1150–1165.

Marfurt, K. J., Sudhaker, V., Gersztenkorn, A., Crawford, K. D., and Nissen, S. E., 1999, Coherency calculations in the presence of structural dip: Geophysics, **64**, No. 1, 104–111.

Randen, T., Pedersen, S. I., and Sønneland, L., 2001, Automatic extraction of fault surfaces from three-dimensional seismic data, *in* SEG Technical Program Expanded Abstracts 2001, Society of Exploration Geophysicists, 551–554.

Ronneberger, O., Fischer, P., and Brox, T., 2015, U-net: Convolutional networks for biomedical image segmentation, *in* International Conference on Medical image computing and computer-assisted intervention, Springer, 234–241.

Usher, W., Klacansky, P., Federer, F., Bremer, P.-T., Knoll, A., Yarch, J., Angelucci, A., and Pascucci, V., 2017, A virtual reality visualization tool for neuron tracing: IEEE transactions on visualization and computer graphics, **24**, No. 1, 994–1003.

Van Bemmel, P. P., and Pepper, R. E., 2000, Seismic signal processing method and apparatus for generating a cube of variance values, uS Patent 6,151,555.

Van den Oord, A., Dieleman, S., and Schrauwen, B., 2013, Deep content-based music recommendation, *in* Advances in neural information processing systems, 2643–2651.

Wu, X., 2017, Directional structure-tensor-based coherence to detect seismic faults and channels: Geophysics, **82**, No. 2, A13–A17.

Wu, X., and Hale, D., 2016, 3d seismic image processing for faults: Geophysics, **81**, No. 2, IM1–IM11.

Wu, X., Liang, L., Shi, Y., and Fomel, S., 2019, Faultseg3d: Using synthetic data sets to train an end-to-end convolutional neural network for 3d seismic fault segmentation: Geophysics, **84**, No. 3, IM35–IM45.

Wu, X., Shi, Y., Fomel, S., and Liang, L., 2018, Convolutional neural networks for fault interpretation in seismic images, *in* SEG Technical Program Expanded Abstracts 2018, Society of Exploration Geophysicists, 1946–1950.

Zhao, T., and Mukhopadhyay, P., 2018, A fault detection workflow using deep learning and image processing, *in* SEG Technical Program Expanded Abstracts 2018, Society of Exploration Geophysicists, 1966–1970.