Submission code: ZKNMB6GL5J

See this and other abstracts on our website: https://home.sbgf.org.br/Pages/resumos.php

Fault Detection in 3D Seismic Data with Machine Learning Methods and Multiple Attributes

GUSTAVO CUSTODIO (Instituto de Pesquisas Tecnológicas (IPT)), Hugo Ferreira Saar (Instituto de Pesquisas Tecnológicas (IPT)), Aline Fernandes Heleno (Instituto de Pesquisas Tecnológicas (IPT)), Carlos Tadeu de Carvalho Gamba (Instituto de Pesquisas Tecnológicas (IPT)), Celso Luciano Alves da Silva (Instituto de Pesquisas Tecnológicas (IPT)), Cristina Maria Ferreira da Silva (Instituto de Pesquisas Tecnológicas (IPT)), Denis Bruno Virissimo (Instituto de Pesquisas Tecnológicas (IPT)), Elisa Morandé Sales (Instituto de Pesquisas Tecnológicas (IPT)), Felipe Silva Silles (Instituto de Pesquisas Tecnológicas (IPT)), Leonides Guireli Netto (Instituto de Pesquisas Tecnológicas (IPT)), Ney Ferreira de Souza Guerra (Instituto de Pesquisas Tecnológicas (IPT)), Otavio Coaracy Brasil Gandolfo (Instituto de Pesquisas Tecnológicas (IPT)), Rafael Andrello Rubo (Petrobras), Thiago Yuji Aoyagi (Instituto de Pesquisas Tecnológicas (IPT))

Fault Detection in 3D Seismic Data with Machine Learning Methods and Multiple Attributes

Abstract Summary

This study explores classical machine learning methods for fault detection in two seismic volumes: F3 (Netherlands) and Thebe (Australia). For each volume, we created a training dataset by extracting features through seismic attribute computation and a sliding window method. We compared multiple classifiers through a 10-fold cross-validation, with the Support Vector Machine (SVM) achieving the highest F1 score. The selected model was tested on both intra-volume (same volume, different region) and inter-volume scenarios. Results showed that models performed best when trained and tested on data from the same volume.

Introduction

Fault detection is of great importance in seismic interpretation. Since geological faults serve both as carrier beds and reservoir seals for oil and gas, there is a growing interest in identifying them, especially using artificial intelligence. Recent approaches for fault detection often make use of deep learning methods such as convolutional neural networks (CNNs), as highlighted in the review by Saar et al. (2025). Despite their high accuracy in fault detection, these methods require the use of expensive hardware.

In this work, instead of deep learning techniques, we apply classical machine learning methods that are less computationally expensive, specifically ensembles of classifiers such as XGBoost and Random Forests, to classify voxels within seismic volumes. Fault classification is performed on field data by performing training on two different datasets: the F3 offshore dataset from the Netherlands (Silva et al., 2019) and the Thebe dataset from Australia (An et al., 2021). These volumes had their faults manually annotated, which served as labeled data for evaluating the performance of the machine learning models.

The main objective of this work is to provide fault detection methods other than deep learning techniques that offer better performance for the available computational resources and datasets. We compare the performance of different classical machine learning algorithms and test the most suitable one to detect faults in not interpreted seismic volumes.

Method and Theory

Each seismic volume is divided into two different regions: one region is used for training and the other for testing. The goal is to assess how well the model performs on an uninterpreted region of the seismic volume. We used the following regions in each volume:

- F3: inlines 100 to 600, crosslines 300 to 1100, and time slices from 400 to 1800 ms.

- Thebe: inlines 501 to 900, crosslines 200 to 2000, and time slices from 2500 to 6000 ms.

For the F3 volume, crosslines 300 to 500 were used for training, and crosslines 501 to 1100 for testing. For the Thebe volume, crosslines 200 to 600 were used for training, and crosslines 601 to 2000 for testing. In order to prepare the training of the model, we randomly selected voxels from both fault and non-fault regions. Due to the imbalance between fault and non-fault voxels, with non-faults being more prevalent, models become biased toward prediction non-fault regions. To avoid this bias, we balanced the amount of positive and negative samples, using 3 negative samples (non-faults) for each positive sample (faults). Additionally, the dataset contains a numerical label indicating whether the corresponding voxel is in a fault region (1) or not (0).

Before training the model, we preprocessed the amplitude volume by enhancing contrast by clipping the first and last percentiles and extracting seismic attributes. We extracted 4 different attributes: chaos, gradient structure tensor, semblance and RMS (root mean square). We focused on edge detection attributes, as they highlight discontinuities in the seismic data, which enhances seismic faults (Kim et al., 2021). The attributes were computed using a 3x3x9 kernel, except for the RMS attribute, which was calculated using a 1x1x9 kernel. We used the `d2geo` Python library (Fitz-Gerald, 2018) for attribute computation.

Similarly to the work of Gomes et al. (2023), we extracted features using a 5x5 window on each inline separately. As shown in Figure 1, this method extracts voxels from a seismic volume within a specific region defined by the window and rearranges the extracted voxel values into a vector to be input as a single sample to the model. Therefore, each row in the dataset has W x T features, where W is the number of points extracted from the window and T is the number of attributes used, including the amplitude. We constructed two training datasets, each containing 125 features. Both datasets are composed of 25% points from fault regions and 75% from non-fault regions. In the F3 dataset, this corresponds to 10,000 faulted and 40,000 non-fault points. In the Thebe dataset, it totals 50,000 fault and 150,000 non-fault points.

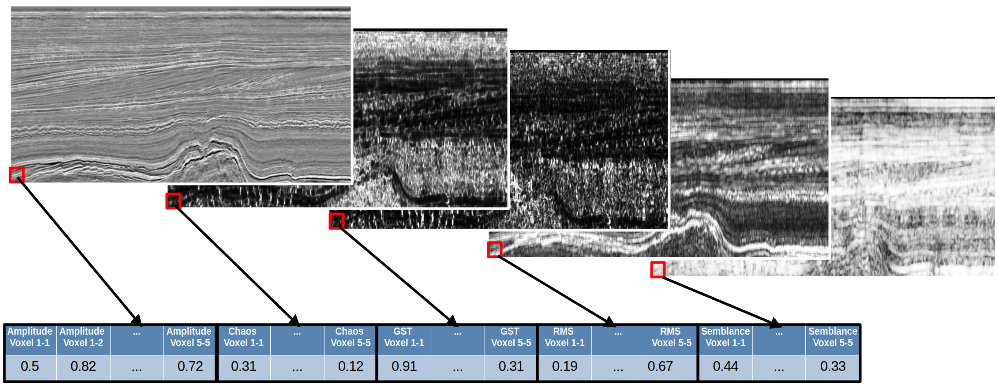


Figure 1: A 5x5 window extracting features from multiple attributes. For each attribute, the window extracts voxel values from the same region and concatenates these values in a single row.

We evaluated multiple classification algorithms on these datasets, including XGBoost (Chen et al., 2015), Random Forest (Breiman, 2001), and SVM (Cortes and Vapnik, 1995), selecting the one with the highest F1 score for testing. The performance was assessed using 10-fold cross-validation.

For each model, we apply z-score normalization to the features as shown in Eq. (1):

$$Z = \frac{x - \mu}{\sigma} \tag{1}$$

where x represents the value of a voxel, μ is the mean, and σ is the standard deviation of the training dataset. We removed outliers by clipping z-scores to the -3 to 3 range, covering 99.7% of the data and replacing extremes with the limits.

We implemented two different approaches for training: we used part of the F3 dataset for training and tested it on another F3 region, as well as on the Thebe dataset. We also performed a cross-testing approach between the volumes, where a model is trained using the training region of Thebe and applied on the test region of F3.

Results

We performed a 10-fold cross-validation using three different classifiers in both the F3 and Thebe datasets. Table 1 depicts the comparison between different classifiers across different metrics: accuracy, recall, precision and F1 score. We used the F1 score as the decisive criteria to select the classification model used for segmenting faults in the test volumes. Clearly, we can see that the SVM achieved the best performance among the three classifiers.

Table 1: Comparison in performance of classification algorithms across different metrics.

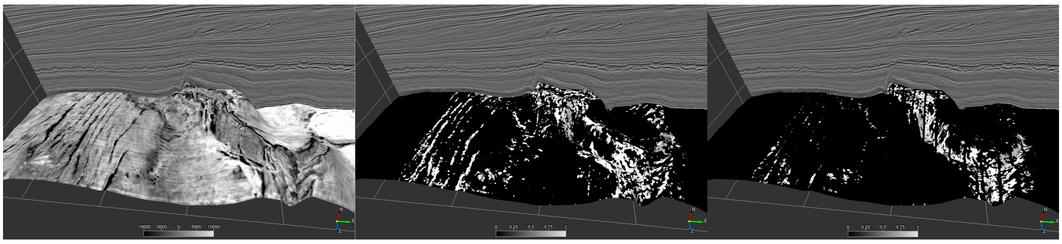
Dataset	Model	Accuracy	Recall	Precision	F1
F3	Random Forest	0.88	0.46	0.79	0.52
	XGBoost	0.89	0.55	0.78	0.65
	SVM	**0.90**	**0.57**	**0.82**	**0.67**
Thebe	Random Forest	0.88	0.46	0.80	0.58
	XGBoost	0.89	0.59	0.77	0.67
	SVM	**0.90**	**0.61**	**0.81**	**0.69**

After selecting the SVM as the best classifier, we used it to detect faults on the F3 and Thebe test volumes. We used the SVM trained using data from the F3 training dataset to detect faults in the F3 test volume (F3-F3) and used the model trained with data from the Thebe dataset to detect in both the F3 (Thebe-F3) and Thebe (Thebe-Thebe) test volumes. Figure 2 shows the results of the SVM model for all training and test variations, with a morphological opening filter applied to remove noise from each seismic volume. We see in Figure 2a how the F3-F3 variation was able to detect the same faults as the Thebe-F3 variation, but with much less noise than the later. The noise was expected, given that the model trained using Thebe data was not adjusted for noisy data like the F3 volume. The Thebe-Thebe variation shown in Figure 2b detected less false positives than the F3, which was again expected, since the Thebe volume contains less noise.

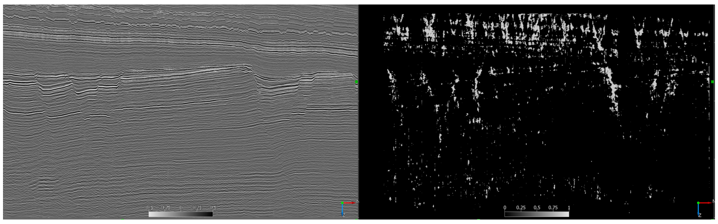
Conclusions

The best results were those obtained from training and testing on the same-origin volume. Thebe had less tendency to detect than F3, which was expected due to the F3 volume containing more noisy data. The visual results for the Thebe dataset were superior to those of F3, possibly also due to Thebe's lower noise levels.

The results show that, despite the higher precision of Deep Learning techniques, classical machine learning algorithms still show good performance in seismic fault detection. Using edge detec-



(a) Results for F3 volume in a deep horizon slice using the SVM model. From left to right: original seismic volume, detected faults for the F3-F3 variation and detected faults for the Thebe-F3 variation.



(b) Results for Thebe volume using the SVM model trained with Thebe voxels. Original seismic volume (left) and detected faults (right).

Figure 2: Visual results obtained for fault identification across different datasets.

tion attributes, these models can be trained on a specific region of a seismic volume and applied to other regions. Despite the high noise levels, it is also possible to apply them across different volumes.

References

An, Y., J. Guo, Q. Ye, C. Childs, J. Walsh, and R. Dong, 2021, Fault recognition dataset 2021: `https://www.fault-analysis-group.ucd.ie/Publications/206_Fault_recognition_Dataset.html`. (Accessed: 13-May-2025).

Breiman, L., 2001, Random forests: Machine learning, **45**, 5–32.

Chen, T., T. He, M. Benesty, V. Khotilovich, Y. Tang, H. Cho, K. Chen, R. Mitchell, I. Cano, T. Zhou, et al., 2015, Xgboost: extreme gradient boosting: R package version 0.4-2, **1**, 1–4.

Cortes, C., and V. Vapnik, 1995, Support-vector networks: Machine learning, **20**, 273–297.

Fitz-Gerald, B., 2018, D2geo: A framework for computing seismic attributes with Python: `https://github.com/dfitzgerald3`. (Accessed: 13-May-2025).

Gomes, J., V. Carneiro, and A. Cerqueira, 2023, Seismic fault classification using multilayer perceptron neural network algorithm: Presented at the Proceedings of the ACM Conference on Example.

Kim, M., J. Yu, N.-K. Kang, and B.-Y. Kim, 2021, Improved workflow for fault detection and extraction using seismic attributes and orientation clustering: Applied Sciences, **11**.

Saar, H. F., C. M. F. da Silva, A. F. Heleno, E. M. Sales, G. T. Custodio, T. Y. Aoyagi, F. S. Silles, C. L. A. da Silva, L. Guireli Netto, N. F. de S. Guerra, O. C. B. Gandolfo, C. T. de C. Gamba, D. B. Virissimo, and R. A. Rubo, 2025, Identifying AI Techniques for Detecting Geological Features in Seismic Data: Literature Review Method with GPT-4. Accepted for presentation at the 86th EAGE Annual Conference & Exhibition.

Silva, R. M., L. Baroni, R. S. Ferreira, D. Civitarese, D. Szwarcman, and E. V. Brazil, 2019, Netherlands dataset: A new public dataset for machine learning in seismic interpretation: arXiv preprint arXiv:1904.00770.