IPTC-22145-EA

Machine Learning-Based Feature Importance Analysis of Seismic Attributes to Assist Fault Prediction

Fan Jiang and Philip Norlund, Halliburton Landmark

This paper was prepared for presentation at the International Petroleum Technology Conference held in Riyadh, Saudi Arabia, 21 - 23 February 2022.

Abstract

Summary

Deep learning approaches have recently shown success in automating feature extraction from seismic data and saving substantial interpretation time for geoscientists, especially for fault prediction. However, fault prediction results based on seismic data are characterized by the low resolution of fault probability that could extend the real fault plane out of its true range. Additionally, the low fidelity of the fault prediction results creates uncertainty locating the true fault. In this paper, we analyze thirty well-known seismic attributes and try to utilize the natural property of each attribute to help build and train a deep neural network to improve the resolution of fault prediction. We trained five machine learning classifiers separately to analyze the feature importance of each attribute to the predictability of fault planes. We then selected several of the most important attributes as additional inputs for a multi-channel convolutional neural network to improve the accuracy of fault prediction. Several synthetic and field data tests are tested to validate our approach.

Methods

Seismic attributes, a property extracted from seismic data, is used to analyze subsurface structures and reservoir characterization, helping geological interpretation in different exploration areas. They can be analyzed to enhance information that might be too subtle in a traditional amplitude image and increases the visibility of different subsurface objects. Seismic attributes can be divided into different classes, such as amplitude, frequency, phase, and structure. In most cases, time-based attributes are related to structure whilst amplitude-based attributes are designed to classify stratigraphy. Roden et al. (2015) described several methodologies to analyze combinations of seismic attributes of any kind for meaningful patterns that correspond to geological features. They used principal component analysis (PCA) and self-organizing maps (SOMs) to analyze multiple attributes in the interpretation workflow. Emujakporue and Enyenihi (2020) extracted and analyzed several seismic attributes from field data to obtain information about the structures, stratigraphy, and hydrocarbon potential from available seismic and a suite of well log data. They showed that different attributes provided different importance levels to enhance the visibility of the geometrical characteristics of seismic events and are sensitive to the lateral variation of azimuth, continuity, similarity, curvature, and other subsurface properties. In this paper, we generated synthetic seismic data (Wu et al., 2019) and calculated over thirty of the most popular attributes representing different categories. The ground-truth fault plane is delineated from the synthetic seismic data (Figure 1). We consider those attribute cubes as training data and the ground truth fault as a training label. Table 1 describes all attributes we generated in different categories and Figure 2 shows all the attributes we used for feature importance analysis.

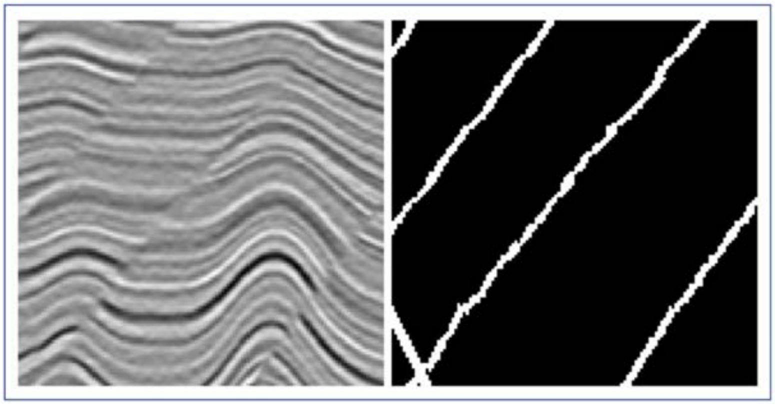

Figure 1—A synthetic seismic data (left) and its fault mask (right).

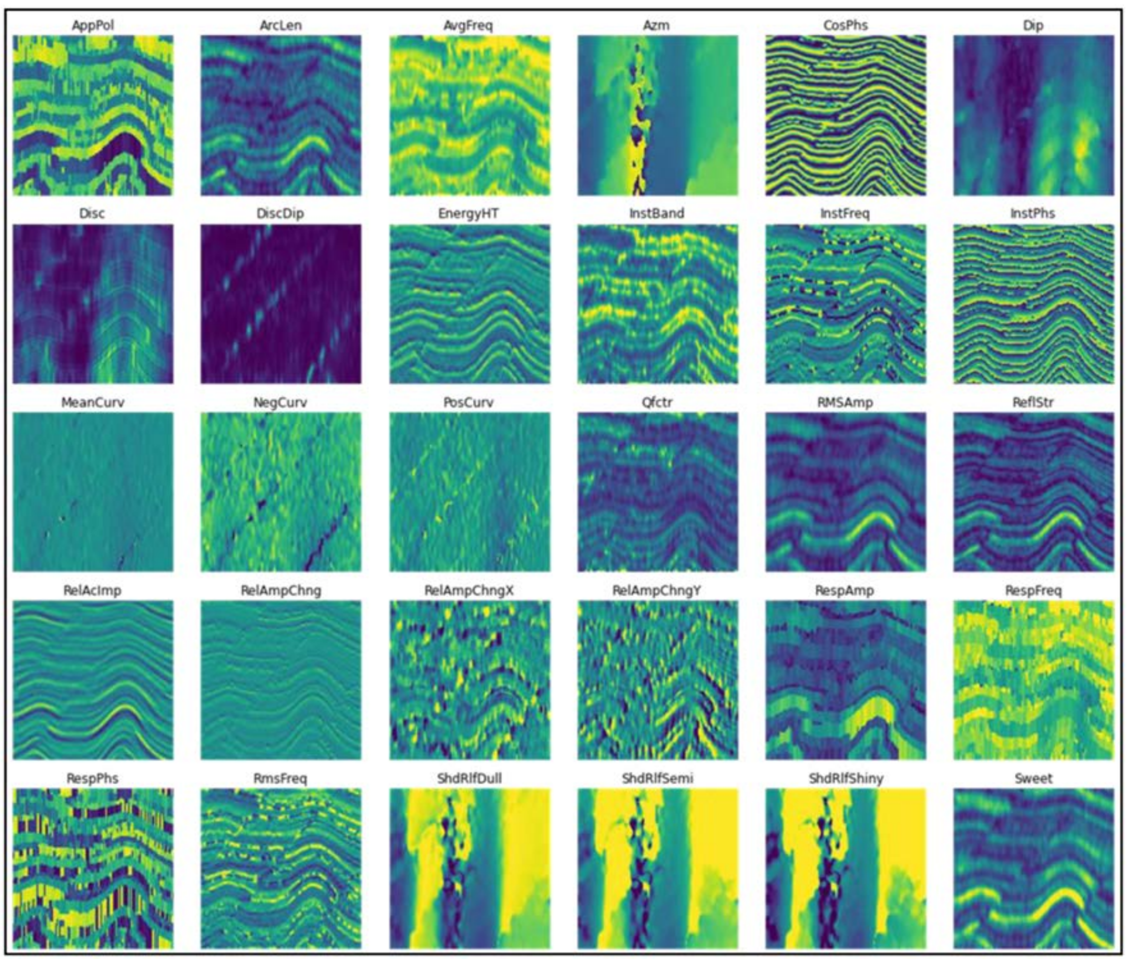

Figure 2—Thirty seismic attributes, derived from seismic in Figure 1, used for feature importance analysis.

Table 1—Seismic attributes used for feature importance analysis.

Seismic Attributes				
Amplitude	Frequency	Phase	Structure	Others
Reflection Strength	Average Frequency	Apparent Polarity	Azimuth	Dull Surface
Relative Amp Change	Instantaneous Frequency	Cosine of Phase	Dip	Shiny Surface
Response Amplitude	Response Frequency	Instantaneous Phase	Discontinuity	Semi-Shiny Surface
RMS Amplitude	RMS Frequency	Response Phase	Discontinuity Along Dip	Arc Length
Quality factor	Thin Bed Indicator		Mean Curvature	Energy half-time
			Most Negative / Positive Curvature	Relative Acoustic Impedance
			Relative Amp Change in X / Y direction	Sweetness

We then use the above generated attributes as training data and labels to feed into several ensemble learning classifiers. Ensemble learning is a model that makes predictions based on several different models (Rokach, 2010). By combing individual models, the ensembled model could perform more flexibly and is less data sensitive. Bagging and Boosting are considered as two of the most popular ensemble methods. Bagging represents a workflow to train a bunch of individual models in a parallel way, where each model is trained by a random subset of the data, such as the Random Forest and Decision Tree method. Boosting describes a training process to train a series of individual models sequentially. Each model learns from mistakes made by the previous model, such as Adaptive Boosting and Gradient Boosting method. The reason we selected multiple ensemble classifiers is to avoid any bias from specific ensemble classifiers which could be sensitive to the seismic attributes and affect the accuracy of importance analysis. Figure 3 describes that workflow to use several ensemble classifiers to analyze the predictability towards each of the attributes in terms of fault prediction. After passing each attribute through the classifier, the importance factor in each classifier is measured by observing the effect on model accuracy by randomly shuffling each predictor variable. The mechanism is designed to measure how effective each attribute is at reducing uncertainty with different classifiers.

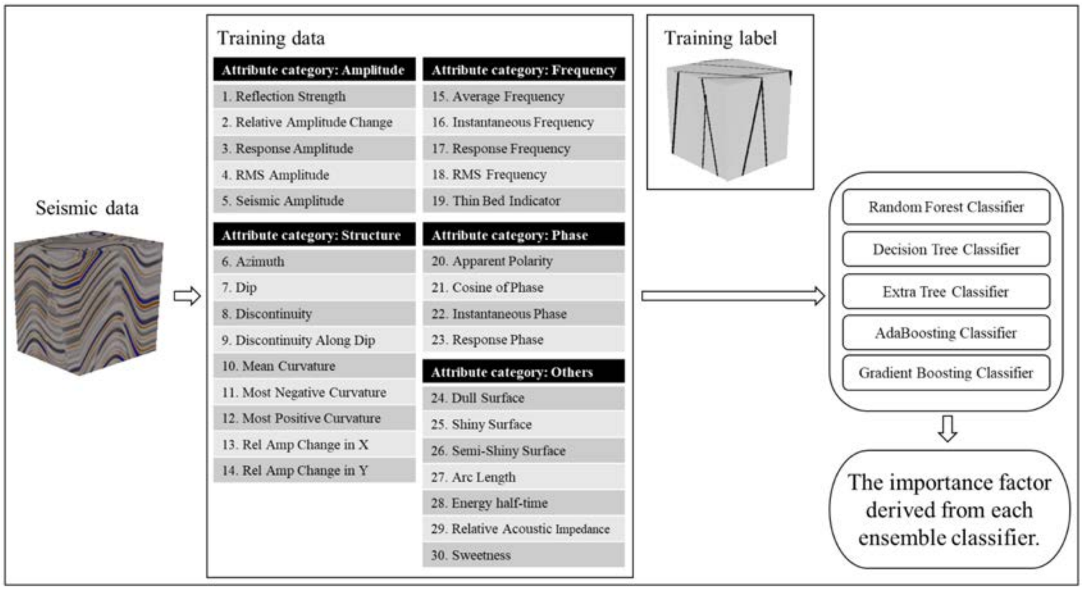

Figure 3—A workflow to analyze the importance of each seismic attribute for the predictability of fault interpretation.

Results

After training the model with different classifiers, we obtained the importance factors for each attribute. These importance factors represent the relative predictive strengths of the feature, or attribute, relative to the fault plane. Figure 4 shows a combined result from several ensemble classifiers. It describes the relationship between the importance factor and seismic attribute. Among all attributes, discontinuity-along-dip always provides the highest rank for all classifiers, followed by other attributes, such as positive-curvature and relative-amplitude-change. The attribute discontinuity-along-dip accounts for reflection dip and produces cleaner images than standard discontinuity (Hale, 2009), which highlights faults, channels, and diapirs. Other attributes, like most-positive-curvature, record the most positive rate of change of the reflection dip and azimuth, highlights reflection bumps in seismic reflection, and are closely related to the attribute most-negative-curvature, which highlights reflection sags (Jiang and Norlund, 2020). In a geological environment, normal faults often exhibit positive curvature on the up-thrown side and negative curvature on the down-thrown side. This is an indicator for other structure-related attributes to delineate fault images. Another attribute which ranks generally higher than others in this test is the relative amplitude change. This attribute serves as a directional high-resolution discontinuity attribute that reveals details in faults and channels along with time or depth and exhibits similar effects to the coherence amplitude gradient. In Figure 4, we normalized the bar of discontinuity-along-dip attribute for visualization purposes. Those high-ranking attributes represent a quantitative measure of a fault's character of interest which means they could help evaluate geological structure and improve the accuracy of fault interpretation. The higher the bar, the more important the attribute contributes to the fault interpretation. Although we implemented five different ensemble classifiers, the result shows a similar ranking across different attributes.

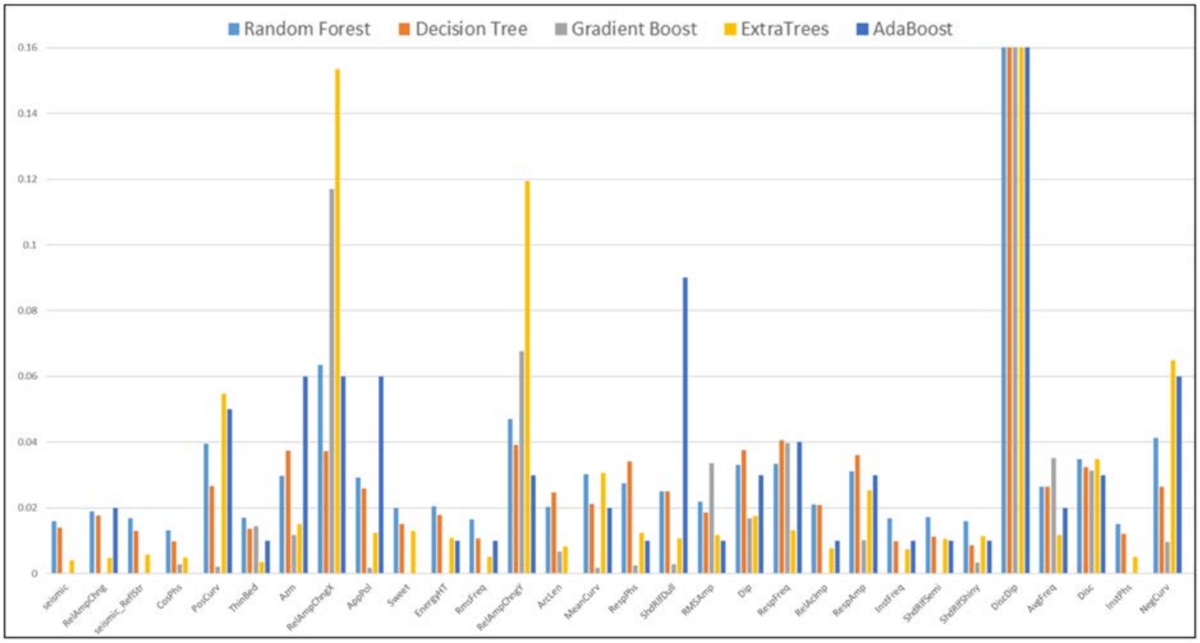

Figure 4—The comparison of importance index between different attributes among all five different ensemble classifiers.

We then consider those attributes as additional channels, along with the original seismic, to feed into a deep learning neural network (Jiang and Norlund, 2021) to train the model and predict faults in a dataset from the northwest of Australia. The prediction result (Figure 5) is improved, including better continuity of fault segments, and reveals several missing fault segments which the seismic-only model missed. We conclude that seismic attributes have a lot of potentials to help in seismic feature extraction and pattern segmentation by implementing deep learning neural networks.

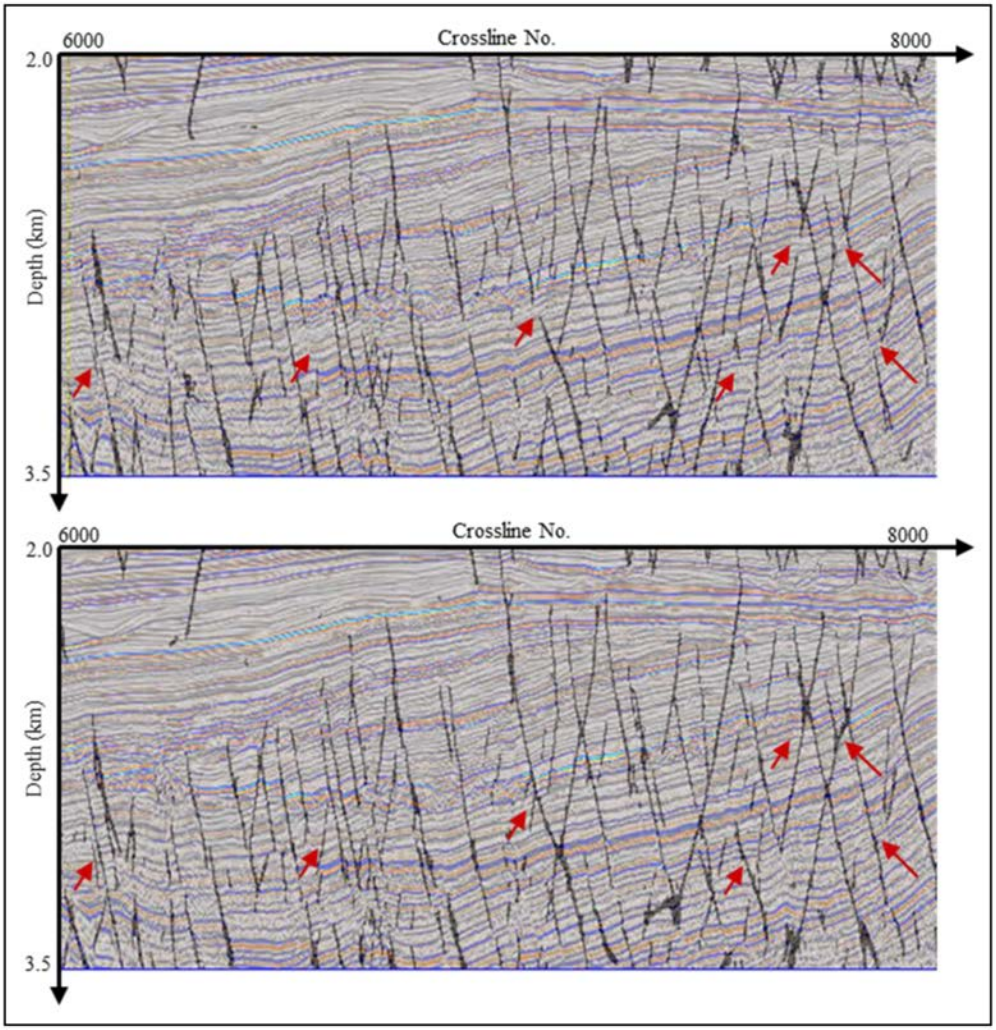

Figure 5—(Top) Fault probability map generated by a machine learning model trained by seismic-only data; (Bottom) Fault probability map generated by a machine learning model trained by five attributes simultaneously. Those attributes are seismic amplitude, discontinuity-along-dip, most-positive-curvature, relative-amplitude-change, and sweetness.

Discussions and Conclusions

The interpretation of faults in 3D seismic data is an important component of hydrocarbon exploration and development workflows. Machine learning-based fault interpretation provides an opportunity to leverage seismic attributes to help in the process of seismic feature engineering to extract geological characteristics and properties, prepare the input data in the forms of structured columns, and improve the performance of different machine learning models. Some techniques, such as ensemble classifiers, one-hot encoding, feature split, are shown to provide great improvement in different machine learning applications to decrease uncertainty in geophysical interpretation. Our proposed method of analyzing various seismic attributes, could also be applied to assist other interpretation projects, such as salt body detection (Jiang et al., 2020), lithology discrimination (Walker et al., 2005), and stratigraphic feature characterization (Chopra and Marfurt, 2008).

In this paper, we trained five ensemble machine learning classifiers to analyze the feature importance of each seismic attribute to the predictability of fault planes, then selected the attributes with the highest rank, along with seismic data, to feed into a multi-channel convolutional neural network to train the model. The selected attributes highlight the structure-property of normal faults, which often exhibit positive curvature on the up-thrown side and negative curvature on the down-thrown side. Our importance ranking system also confirms the effectiveness of structure-related attributes that could contribute more to the geophysical interpretation life cycle. The field test from Australia shows that the model trained by multiple attributes improved significantly on the predicted fault probability map over the seismic-only model with better imaging of fault planes. This could also simplify the fault extraction workflow to extract continuous and cleaner fault planes for the next velocity model building process.

Acknowledgments

We thank Geoscience Australia for providing the seismic data and their permission to use it in this research.

References

Chopra, S. and Marfurt, K., 2008, Seismic attributes for stratigraphic feature characterization, *SEG Extended Abstract*.

Dave, H., 2009, Structure-oriented smoothing and semblance: CWP Report 635.

Emujakporue, G. O. and Enyenihi, E.E., 2020, Identification of seismic attributes for hydrocarbon prospecting of Akos field, Niger Delta, Nigeria, Springer Nature Applied Sciences.

Jiang, F. and Norlund, P., 2020, Seismic attribute-guided automatic fault prediction by deep learning, EAGE Annual Meeting, *Extended Abstract*.

Jiang, F. and Norlund, P., 2021, Assisted fault interpretation by multi-scale dilated convolutional neural network, EAGE Annual Meeting, Extended Abstract.

Jiang, F., Norlund, P., and Wei, Z., 2020, Analysis of seismic attributes to assist in the classification of salt by multi-channel convolutional neural networks, the First EAGE Digitalization Conference and Exhibition, Extended Abstract.

Roden, R., Smith, T., and Sacrey, D., 2015, Geologic pattern recognition from seismic attributes: Principle component analysis and self-organizing maps, *Interpretation, SAE59-83*.

Rokach, L., 2010, Ensemble-based classifiers. *Artificial Intelligence Review*. **33**: 1–39.

Walker, R., Wong, C., Malcotti, H., Perez, E., and Sierra, J., 2005, Seismic multi-attribute analysis for lithology discrimination in Ganso Field, *Oficina Formation, Venezuela, The Leading Edge*, v**24**, 1081–1192.

Wu, X., Liang, L., Shi, Y. and Fomel, S., 2019, FultSeg3D: Using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation, *Geophysics*, **84**, IM35–IM45.