# Using 3D Seismic Technology to Accurately Identify Faults and Reservoir Characteristics

Jiangyin Xue

Institute of Geology, No.4 Oil Production Plant, Daqing Oilfield Co, Ltd., 163000, Daqing, China

**Abstract.** With the accelerating development of the Chinese economy, the demand for oil and gas energy is increasing day by day. The accuracy and efficiency of fault interpretation are crucial for the exploration and development of oil and gas reservoirs. The presence of oil and gas in reservoirs will inevitably cause changes in the geophysical response characteristics, and different types of reservoirs have different response characteristics in logging curves such as sound, discharge, and electricity. 3D seismic technology is currently one of the most effective techniques for obtaining fault structural features and identifying small faults, but conventional seismic data processing techniques are unable to meet the accuracy requirements of current seismic exploration. With the development of computer technology, more and more scholars are applying deep learning (DL) algorithms to the identification of faults and reservoirs to further improve the recognition effect. This article is based on the DL algorithm and utilizes 3D seismic technology to design a method for automatic and accurate recognition of fault and reservoir features. Obtain labeled seismic amplitude faults, reservoir data, and seismic amplitude data to be identified from the 3D seismic amplitude data volume, and construct a CNN training model to identify the fault and reservoir data. The experimental results show that the method designed in this article can improve the efficiency and accuracy of fault and reservoir identification.

**Key words:** 3D seismic technology; Fault reservoir; Feature recognition.

## 1. Introduction

Earthquake fault identification is an important field in seismology and geology research, aimed at identifying fault lines in the crust in order to predict the location and intensity of earthquakes [1]. The accuracy of fault identification determines the accuracy of constructing maps. Therefore, how to effectively automatically pick up faults from seismic data has always been a hot topic in seismic data interpretation research [2]. In modern society, oil and natural gas are still indispensable resources, and fault interpretation is the key and foundation of earthquake interpretation, which is of great significance and value for academic research and industrial production [3]. By identifying faults and determining their distribution underground, the distribution and scale of oil and gas reservoirs can be accurately located. In the early days, it mainly relied on manual interpretation of faults on seismic profiles. In the process of fault tracking, researchers observed the direction of seismic faults manually, and then used numerical analysis methods to track faults line by line, thereby extending to three-dimensional space. This greatly depends on the experience of interpretation personnel, and the accuracy and efficiency of fault interpretation are difficult to meet practical production needs [4].

With the rapid development of computer technology, some new fault identification methods have been proposed one after another. The traditional method is to connect discontinuous points of the same phase axis into fault lines on a two-dimensional seismic profile. On this basis, people usually use various geophysical methods, such as coherent volume analysis, square analysis, edge detection, post stack target processing, forward model analysis, and full three-dimensional interpretation, To improve the accuracy and reliability of fault identification [5]. 3D seismic technology has been widely applied in coalfield exploration and is currently one of the most effective technical means to obtain structural characteristics of coal areas, identify small faults, reservoirs, and target coal seam distribution [6]. With the deepening of oil and gas exploration, fractured and vuggy reservoirs are constantly being developed. How to accurately and efficiently find favorable fractured and vuggy reservoirs using existing development well data is an urgent problem that needs to be solved. Analyzing seismic attributes is an essential step in reservoir prediction. The unconventional information about geometry, kinematics, and statistical characteristics derived from seismic data volumes using specific software is called seismic attributes [7]. Due to the irregular spatial distribution of fracture cavities, there are two problems in using seismic data to directly extract seismic waveforms for classification within the time window of interpretation layers: seismic waveforms from non reservoir segments have a serious impact on classification results; The seismic waveforms between different reservoirs have significant phase differences, resulting in unreliable classification results [8].

With the widespread application of 3D seismic exploration technology, more and more 3D seismic data volumes are being collected, and individual data volumes are becoming larger and larger. Artificial interpretation can no longer meet the requirements of high-resolution fine interpretation of faults. The automation and intelligence of fault interpretation are inevitable trends [9]. The quality of seismic data processing directly determines the accuracy of seismic interpretation. In actual 3D seismic exploration, due to various limitations, the results of 3D seismic data processing are not satisfactory, resulting in differences in the geological structures interpreted later in actual mine exposure, which poses hidden dangers to the safety production of the mine. In recent years, DL technology has been one of the important research directions in the field of machine learning, and its emergence has greatly promoted the development of AI technology. This article is based on the DL algorithm and utilizes 3D seismic technology to design a method for automatic and accurate recognition of fault and reservoir features. Obtain labeled seismic amplitude faults, reservoir data, and seismic amplitude data to be identified from the 3D seismic amplitude data volume; Construct a CNN training model to identify seismic faults and reservoir data.

## 2. Methodology

### 2.1 Application of DL and 3D Seismic Technology in Fault and Reservoir Identification

The classification of reservoir types is one of the important contents of fine reservoir description, which is of great significance for accurately and quantitatively evaluating oil and gas reserves [10]. Faults in seismic images typically exhibit highly discontinuous or low continuous reflection patterns, and based on this characteristic, many methods have emerged to highlight faults by calculating attributes that can reflect reflection continuity or discontinuity. However, these attributes are also sensitive to noise and geological features. The purpose of seismic attribute analysis in the research area is to transform seismic attributes into information related to geological anomalies such as coal seams and structures that can serve seismic interpretation, thereby improving the ability of seismic data in coal seam prediction and micro amplitude structural interpretation. The traditional reservoir classification method requires a large number of characteristic parameter evaluations and selections, but usually only a few selected logging parameters can be used, and the effective information of other logging parameters cannot be fully utilized. It cannot overcome the limitations of nonlinear factors such as empirical formulas, and is greatly affected by human factors, which brings many disadvantages to actual production. Faults occur at the edges of rock layers with significant relative displacement, often exhibiting highly discontinuous features of the same phase axis in seismic images. Due to changes in near surface conditions and various complex human factors during the acquisition process, there are various types of interference waves in the collected seismic records. These interference waves reduce the seismic signal-to-noise ratio and resolution, affecting the entire process of data processing and causing difficulties in further improving the quality of seismic profiles [11].

DL learns the transformed features of the hidden nodes in the network from the input signal, and these features become increasingly abstract through layer by layer transformation, thereby solving complex nonlinear problems in practice [12]. Convolutional Neural Network (CNN) is an important product of computer technology development and one of the representative algorithms of DL [13]. CNN is a neural network model that specializes in processing image matrices, with its core being a convolutional layer that assigns network names. The main function of convolutional layers is to extract information from input data. Low level convolution extracts contour, texture, edge and other information, while deep level convolution extracts deep level semantic information, more complex and specific features, and can combine low-level semantic information. After feature extraction by convolutional layers and processing by other neural network layers, the model can obtain the prediction results of faults. CNN can directly transmit the original image to relevant detection devices and extract and display various features of the original image, making it easier for people to label and supervise. Reduce the number of weight parameters by reducing the number of connections; Another characteristic is weight sharing, which means that multiple positions in a feature map use the same weight to detect the same type of feature, which helps to further reduce the number of parameters. Therefore, the computational workload during network training is greatly reduced, and the complexity of the network model is also reduced. In terms of fault and reservoir identification, CNN can extract fault and reservoir features from seismic data through convolutional layers, and learn high-level feature representations through structures such as pooling and fully connected layers, thereby achieving accurate identification of faults and reservoirs. The CNN structure is shown in Figure 1.

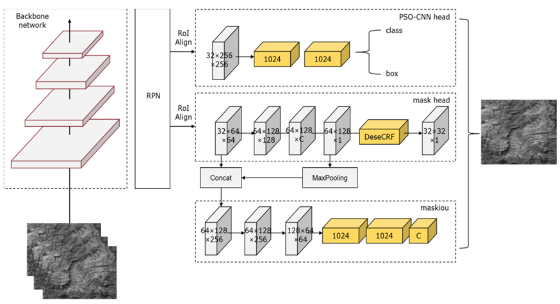


Figure 1 CNN structure

## 2.2 Algorithm Principle

The convolutional layer mainly completes feature extraction through convolution operation, which requires the use of convolution kernels to slide the window on the input image one by one and calculate to complete the convolution operation. In the convolution operation, each parameter of the convolution kernel is connected to the local pixels of the input data, and the convolution calculation is carried out by multiplying and adding each element to obtain the output result. This convolution operation can be used to learn local features, and finally further calculate the output of the convolutional layer by adding bias parameters. The calculation formula for convolution is shown in equation (1):

$$y_{mn} = f\left(\sum_{j=0}^{q-1}\sum_{i=0}^{p-1} x_{m+i,n+j} w_{ij} + b\right), (0 \le m < M, 0 \le n < N) \tag{1}$$

$x$ is usually represented as an RGB image or grayscale image, representing the input data of the convolutional layer, $y$ represents the output of the convolutional layer, $m \times n$ represents the size of the output of the convolutional layer, $w$ represents the convolutional kernel of $p \times q$ , $b$ is the bias parameter, and $f$ represents the activation function.

For a random event $X$ with $n$ possibilities, entropy can be used to represent the expected amount of information for all possibilities:

$$H(X) = -\sum_{i=1}^{n} p(x_i)\log(p(x_i)) \tag{2}$$

If $Y$ represents the label variable of $x$ , $y$ represents the label value of $x$ (0 and 1), then $(x, y)$ represents the samples in the training set, and $p\{Y = y \mid x\}$ represents the probability that $x$ 's label is $y$ . For each batch of $n$ samples $\{(x_i, y_i), i = 1, \cdots, n\}$ input from the training set, this paper uses the cross entropy function as its loss function, namely:

$$L = -\frac{1}{n}\sum_{i=1}^{n} \log p(Y = y_i \mid x_i) \tag{3}$$

Then use the Adam optimization algorithm to calculate the minimum value of the loss function. It can be implemented quickly and easily, with low memory requirements, suitable for unstable objective functions, and has great advantages compared to other types of random optimization algorithms.

When performing directional filtering on fault areas, it may cause blurring of the fault edges. Due to the discontinuity of geological information on both sides of the fault, if the data points on both sides of the fault are forcibly smoothed, the geological information on both sides of the fault will become continuous and the fault information will be destroyed. The idea of using Kuwahara filtering is to calculate the eigenvalues of all neighborhoods of the target point, and select the maximum value of the eigenvalues as the target value of the point. Identifying fault location through similarity information $C$ :

$$C = v_1 / \sum_i v_i \tag{4}$$

Among them, $v_i$ is the feature value obtained by singular value decomposition of local image data, arranged in descending order, so $v_1$ is the maximum feature value.

The seismic waveform is a comprehensive reflection of the amplitude, frequency, phase and other parameters of seismic signals. Applying neural network technology to classify and describe the spatial similarity of seismic wave waveforms, seeking similarities and differences, highlighting seismic wave waveforms of the same category, and ultimately obtaining the planar distribution pattern of seismic anomaly bodies. Assuming the sample is a point in the $n$ dimensional space $R^n$ , the feature vector $x_i$ of the $i$ th sample is represented as:

$$x_i = (x_i^1, x_i^2, \cdots, x_i^n) \tag{5}$$

Due to the input image being represented as a matrix in the computer world, only a few edges of the image may be detected and cropped out, resulting in a loss of a large amount of edge information. In addition, the input matrix needs to be filled before the convolution operation to ensure consistent size between the input and output images. Assuming the image size of the input convolutional layer is $H_{input} \times W_{input}$ , $p$ and $s$ represent the number of filling steps and step size, respectively. Assuming the size of the convolutional kernel is $k \times k$ , the size of the feature map after convolution operation is $H_{out} \times W_{out}$ . The formula for calculating the size of the feature map is shown in equations (6) and (7):

$$H_{out} = (H_{input} + 2 \times p - k)/s + 1 \tag{6}$$

$$W_{out} = (W_{input} + 2 \times p - k)/s + 1 \tag{7}$$

Precision represents how many of the predicted positive samples are true positive samples, reflecting the model's ability to accurately predict the accuracy of positive samples. The larger the Precision value, the better the performance. The formula is shown in equation (8):

$$\Pr ecision = \frac{TP}{TP + FP} \tag{8}$$

## 3. Result analysis and discussion

To verify the effectiveness of the method proposed in this article, experiments are required. The experimental data comes from logging data from 6 wells in a certain work area's exploration and production, with a total of 6000 sample points. One well data (1000 sample points) was selected as the test dataset for CNN model performance testing, while the remaining 5 well data (5000 sample points) were used as the training dataset for training CNN models. In CNN, it is necessary to strictly distinguish between training data and test data, and test data can only be used as a criterion for judging the accuracy of the CNN model. Generally, no analysis is done on the test dataset in the preliminary work. The testing environment is shown in Table 1.

Table 1 Experimental Environment Configuration

| Name | Version |
|---|---|
| Processor | 2.6GHz Intel Core i7 |
| Operating system | Windows 10 |
| PyCharm | 2021 |
| Python | 3.6 |
| Mayavi | 4.7.4 |
| Segyio | 1.9.6 |

This article introduces ROC curve and AUC value as model evaluation indicators. The ROC curve, also known as the receiver operating characteristic curve, is a comprehensive indicator that reflects the sensitivity and specificity of continuous variables. The abscissa of the ROC curve is the pseudo positive class rate (FPR), which refers to the proportion of predicted positive but actually negative samples to all negative class samples; The y-axis represents the true class rate (TPR), which refers to the proportion of predicted and actually positive samples to all positive class samples. The AUC value refers to the area under the ROC curve, usually 0.5<AUC<1.0. The closer the ROC curve approaches the (0,1) point, the higher the corresponding AUC value, and the better the model classification results. It can be seen that the AUC value can more intuitively reflect the quality of the ROC curve, which can be used to evaluate the performance of the model. Figure 2 shows the comparison of AUC values between the model established by our method and the traditional method. It can be seen from Figure 2 that the model performance of our method is much higher than that of the traditional method, and the fault classification performance is the best.

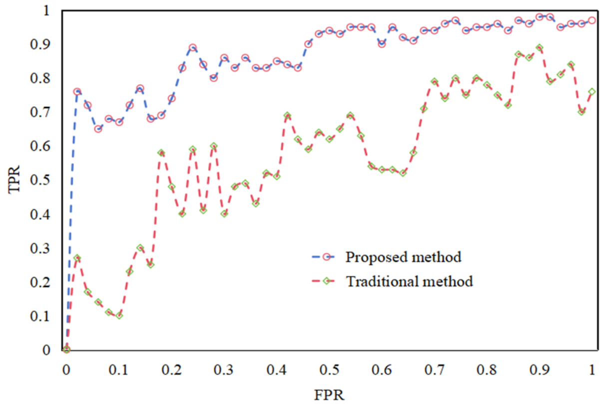


Figure 2 Comparison of AUC values for different methods

In order to verify the fault recognition effect of the method designed in this article, it will be compared with traditional methods for training and testing, as shown in Figure 3.

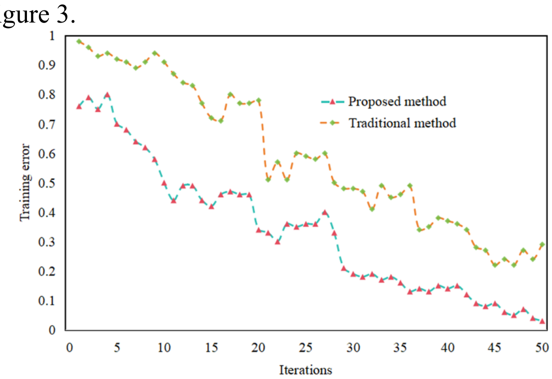


Figure 3 Comparison of training errors using different methods

From Figure 3, it can be seen that the prediction results of the model in this article are more accurate in feature extraction, eliminating many interference options and allowing for intuitive fault description. Multiple training adjustments were made in terms of training error. Although there was a significant initial error, as the number of iterations increased, the error value also gradually decreased and tended to stabilize. From this, it can be seen that the design method in this article is relatively reasonable, and CNN can also be applied to automatic identification of seismic faults without the problem of horizontal interference, effectively alleviating the dependence on manual experience.

Table 2 shows the accuracy of fault and reservoir identification corresponding to different training sample sizes in actual records. It can be seen that in actual seismic data, as the number of training samples decreases, the accuracy of prediction results gradually decreases, which is consistent with the conclusions obtained from model testing.

Table 2 The proportion of actual data training samples and recognition accuracy

| Training sample proportion/% | Recognition accuracy/% |
|---|---|
| 100 | 93.7 |
| 60 | 72.1 |
| 40 | 58.3 |
| 20 | 38.9 |

## 4. Conclusion

The improvement of fault segmentation accuracy is of crucial significance for the process of oil and gas exploration and development. For a long time, the application of 3D seismic exploration technology in dynamically reflecting geological changes has been relatively mature. The seismic fault and reservoir interpretation methods used rely heavily on manual experience. Such calculation methods are prone to errors due to manual errors, and are time-consuming and difficult. DL utilizes a large amount of fault and reservoir data for training, autonomously learning effective features of faults and reservoirs. Its powerful non-linear fitting ability and feature extraction ability can further improve the quality of fault detection. By applying CNN to automatic identification of seismic faults, it can effectively reduce human and horizontal interference, obtain higher fault interpretation, and provide effective reference for seismic data analysis. In order to better meet the urgent demand for efficient, high-precision, and high-resolution fault interpretation in current oil and gas exploration and development, this paper designs a method for automatic and accurate recognition of fault and reservoir characteristics based on DL algorithm and 3D seismic technology. Obtain labeled seismic amplitude faults, reservoir data, and seismic amplitude data to be identified from the 3D seismic amplitude data volume; Construct a CNN training model to identify seismic fault data. The experimental results show that the method designed in this article can improve the efficiency and accuracy of fault and reservoir identification. In the future, the precision of seismic interpretation should be improved, combining on-site and experimental methods to explore its productivity characteristics at a deeper level, and further improve the method system for predicting seismic faults and reservoirs.

## Reference

1. Gawad E A, Fathy M, Reda M, et al. Petroleum geology: 3D reservoir modelling in the central Gulf of Suez, Egypt, to estimate the hydrocarbon possibility via petrophysical analysis and seismic data interpretation[J]. Geological Journal, vol. 56, no. 10, pp. 5329-5342, 2021.
2. Salaun N, Toubiana H, Mitschler J B, et al. High-resolution 3D seismic imaging and refined velocity model building improve the image of a deep geothermal reservoir in the Upper Rhine Graben[J]. The Leading Edge, vol. 39, no. 12, pp. 857-863, 2020.
3. Bradaric A D, Andersen T, Lecomte I, et al. Recognition and characterization of small-scale sand injectites in seismic data: implications for reservoir development[J]. Journal of the Geological Society, vol. 2022, no. 2, pp. 179, 2022.
4. Zhang J, Wei K, Qin J. Resilience and Economic Loss Assessment of Highway Bridges in Deep Reservoir under Near-Fault Ground Motions[J]. Journal of Bridge Engineering, vol. 26, no. 3, pp. 04021007, 2021.
5. Yi B Y, Yoon Y H, Kim Y J, et al. characterization of thin gas hydrate reservoir in u11eung basin with stepwise seismic inversion[J]. Energies, vol. 14, no. 14, pp. 4077, 2021.
6. Pang Y, Cai L, He W, et al. Response to discussion of "Seismic assessment of deep water bridges in reservoir considering hydrodynamic effects using endurance time analysis"[J]. Ocean Engineering, vol. 221, no. 5, pp. 108538, 2021.
7. Teillet T, Fournier F, Zhao L, et al. Geophysical pore type inversion in carbonate reservoir: Integration of cores, well logs, and seismic data (Yadana field, offshore Myanmar)[J]. Geophysics, vol. 86, no. 3, pp. 1-69, 2021.
8. Maleki M, Davolio A, Denis José Schiozer. Quantitative integration of 3D and 4D seismic impedance into reservoir simulation model updating in the Norne Field[J]. Geophysical Prospecting, vol. 67, no. 1, pp. 167-187, 2019.
9. Karbalaali H, Javaheriani A, Dahlke S, et al. Seismic channel edge detection using 3D Shearlets- a study on synthetic and real channelized 3D seismic data[J]. Geophysical Prospecting, vol. 66, no. 7, pp. 1272-1289, 2018.
10. Nantanoi S, Rodriguez-Pradilla G, Verdon J. 3D seismic interpretation and fault slip potential analysis from hydraulic fracturing in the Bowland Shale, UK[J]. Petroleum Geoscience, vol. 2022, no. 2, pp. 28, 2022.
11. Protasov M I, Gadylshin K G, Tcheverda V A, et al. 3D diffraction imaging of fault and fracture zones via image spectral decomposition of partial images[J]. Geophysical Prospecting, vol. 67, no. 5, pp. 1256-1270, 2019.
12. Santos A B, Porsani M J, Góis, Manuelle S. 3D Seismic survey design using mixed-radix based algorithm inversion[J]. Geophysical Prospecting, vol. 67, no. 7, pp. 1720-1731, 2019.
13. Naibai M A, Shenglin J, Ruizhao Y, et al. Seismic response characteristics and identification of fault-karst reservoir in Shunbei area, Tarim Basin[J]. Shiyou Diqiu Wuli Kantan/Oil Geophysical Prospecting, vol. 54, no. 2, pp. 398-403, 2019.