Available online at www.sciencedirect.com

ScienceDirect

Natural Gas Industry B 12 (2025) 174—185
www.keaipublishing.com/en/journals/natural-gas-industry-b/

Research Article

Seismic fault identification of deep fault-karst carbonate reservoir using transfer learning

Hanqing Wang[a,b], Han Wang[a,*], Kunyan Liu[a], Jin Meng[a], Yitian Xiao[a] & Yanghua Wang[b]

[a] *Petroleum Exploration and Production Research Institute, SINOPEC, Beijing, 102206, China*
[b] *Resource Geophysics Academy, Imperial College London, London SW7 2AZ, UK*

Received 12 August 2024; revised 30 October 2024; accepted 28 March 2025
Available online 19 April 2025

Abstract

Seismic fault identification is a critical step in structural interpretation, reservoir characterization, and well-drilling planning. However, fault identification in deep fault-karst carbonate formations is particularly challenging due to their deep burial depth and the complex effects of dissolution. Traditional manual interpretation methods are often labor intensive and prone to high uncertainty due to their subjective nature. To address these limitations, this study proposes a transfer learning−based strategy for fault identification in deep fault-karst carbonate formations. The proposed methodology began with the generation of a large volume of synthetic seismic samples based on statistical fault distribution patterns observed in the study area. These synthetic samples were used to pretrain an improved U-Net network architecture, enhanced with an attention mechanism, to create a robust pretrained model. Subsequently, real-world fault labels were manually annotated based on verified fault interpretations and integrated into the training dataset. This combination of synthetic and real-world data was used to fine-tune the pretrained model, significantly improving its fault interpretation accuracy. The experimental results demonstrate that the integration of synthetic and real-world samples effectively enhances the quality of the training dataset. Furthermore, the proposed transfer learning strategy significantly improves fault recognition accuracy. By replacing the traditional weighted cross-entropy loss function with the Dice loss function, the model successfully addresses the issue of extreme class imbalance between positive and negative samples. Practical applications confirm that the proposed transfer learning strategy can accurately identify fault structures in deep fault-karst carbonate formations, providing a novel and effective technical approach for fault interpretation in such complex geological settings.

Keywords: Seismic fault; Fault-karst carbonate; U-Net; Transfer learning; Attention mechanism

1. Introduction

With the increasing demand for fossil fuels and technological advancements in exploration and production, extensive extraction of oil and gas from depths greater than 5000 m is no longer a pipe dream, and this realization has begun to penetrate the public consciousness [1]. The carbonate reservoir, one of the most important reservoirs in the study of deep/ultra-deep oil and gas resources, has been both a focus of research and a source of difficulty. Different from shallow carbonate reservoirs, oil- and gas-rich areas in deep/ultra-deep carbonates are mostly located near deep faults, and the oil and gas are enriched in the deep faults fracture zone of connected oil sources; the nearby karst development areas have become an important field for exploring karst reservoirs [2−4]. Carbonate reservoirs impacted by both fracture and dissolution, named fault-karst carbonate reservoirs, have become the core of deep/ultra-deep carbonate exploration and development. To ease the development of such reservoirs, it is important to find as many faults as possible that are connected with deep source faults, providing support for well location and well trajectory [5].

Traditional seismic fault identification methods mainly fall into three categories, namely manual identification methods,

* Corresponding author.
E-mail address: wanghan2023.syky@sinopec.com (Han Wang).
Peer review under the responsibility of Sichuan Petroleum Administration.

https://doi.org/10.1016/j.ngib.2025.03.006

seismic attributes identification methods, and machine learning methods. The manual identification method primarily relies on experienced interpreters to draw fault lines on the profile according to the local characteristics of the fault, combined with the geological structure and stress distribution, and then explain the fault surface. Common indicators of faults include the coaxial axis being obviously off, the number of coaxes suddenly increasing, decreasing or disappearing, shape mutation, and the emergence of special waves [6]. However, this method has obvious defects. On one hand, it is inefficient and cannot meet the needs of large-scale interpretation; on the other hand, the accuracy is not reliable, as the interpretation results largely depend on the knowledge and experience of the interpreters. Especially in some areas where geological characteristics are not obvious, manual identification methods are extremely difficult to conduct [7].

The seismic attributes identification method, the mainstream method of fault recognition, has been widely utilized since 1995s. Using either continuous attributes or discontinuous attributes to identify faults have become the two main ways of identifying faults based on seismic attributes. Coherence attributes, curvature attributes, and variance attributes are the classic attributes of fault identification. Coherent attributes highlight the overall spatial distribution characteristics of faults, mainly through coherent changes. Classical techniques include C1 coherent cube technology based on correlation [8], C2 coherent cube technology based on semblance [9], and C3 coherent cube technology based on eigenstructure [10]. Curvature attributes are mainly used to identify faults through curvature changes, among which 2D surface curvature [11] and 3D volumetric curvature [12] are used for fault identification. The variance attribute has some advantages in the analysis of stratigraphic discontinuity, so it is also widely used in fault identification [13]. However, due to changes in lithology and special geological bodies, the values of the above attributes may be abnormal, and identification errors may occur [14]. Integrating attributes from multiple methods enhances fault identification accuracy, establishing this multi-attribute fusion as the preferred solution to overcome limitations of single-attribute analysis. The ant tracking algorithm is one of the representative technologies of attribute fusion, and it was one of the technologies proposed by Schlumberger in 2005 for fault tracking and interpretation [15]. It won the "Best Exploration Technology Award" of the World Petroleum Journal in 2005, but its parameter setting is relatively complex in application, and it still requires certain interpretation experience.

With the application of statistical methods in geological/seismic interpretation and the continuous improvement of computer processing power, machine learning methods have gradually shown unique advantages in automatic fault detection and interpretation. While traditional machine learning algorithms−exemplified by Support Vector Machines (SVM) [16,17]−established the foundational paradigm for automated fault detection, contemporary advances in deep learning have revolutionized seismic interpretation through data-driven feature extraction and end-to-end uncertainty quantification. Compared to conventional machine learning, deep learning can discover more feature details and therefore has an advantage in fault recognition. A paradigmatic advancement emerges in the evolution of convolutional neural network (CNN), which have redefined fault characterization through image segmentation paradigms—transitioning from 2D seismic section analysis to volumetric 3D interpretation frameworks with learnable spatial-contextual constraints. The advantage of using deep learning for fault recognition is that complex algorithms are no longer needed to extract features, thus improving the efficiency of fault identification. Huang et al. [18] developed a "big data" platform for seismic data interpretation and found that deep learning methods with high computational requirements, such as CNN, provided better results in identifying geological faults in 3D seismic data. Considering the difficulty of fault data labeling, Wei et al. [19] constructed CNN directly from unbalanced patch data through transfer learning, thereby improving the accuracy of fault identification. However, CNN exhibits two problems related to fault identification. First, it requires a large number of sample labels as the basis, which can be difficult to provide; second, the network structure of CNN is complex and the calculation cost is high. Wu et al. [20] innovatively applied the modified U-Net network to fault interpretation and created a fault specimen generation method that solved the problem of sample generation and labeling and realized end-to-end 3D seismic fault identification. Subsequently, a large number of scholars have improved and optimized that breakthrough, further improving the accuracy and efficiency of fault identification. In order to avoid the inaccuracies caused by manual selection of training labels by different interpreters, Liu et al. [21] used a random number of synthetic seismic datapoints with different positions and throws as the training and validation dataset, and they introduced residual units into U-Net to build a ResUNet fault identification model. To overcome the shortcomings of 2D and 3D neural networks for fault identification, Lin et al. [22] introduced the channel attention mechanism in the original U-Net architecture and built 2.5D Channel Attention U-Net (2.5D CA-NET). The model improved computational efficiency by using four adjacent seismic slices as inputs instead of the entire 3D seismic data. Based on the U-Net network and V-Net, Chen et al. [23] proposed an improved CUNet fault intelligent identification method that uses convolution operation to replace maximum pooling in downsampling. The 3D inverse convolution operation replaced the upsampling convolution operation, thus preserving the signal details and enhancing the extraction of fault image features. In deeply-buried fault-controlled karst carbonate reservoirs—a distinctive play characterized by syndepositional fault-karst coupling—conventional fault identification methods prove suboptimal due to the absence of consistent morphological templates, as structural discontinuities are inherently obscured by multistage karstification-induced geobody heterogeneity.

In order to solve the above problems, this paper proposes a transfer learning strategy for fault identification of deep fault-karst carbonate. First, a large number of synthetic samples were generated by using existing methods based on the statistical rules of fault distribution in the study area, the existing U-Net network was improved, and the attention block was added to generate a pretrained model. Then, combined with the manual fault verification and interpretation results, manual labeling was carried out to generate real sample labels, which were subsequently used as training sets to fine-tune the pretrained model and improve its interpretation accuracy. Through practical application, the results of this study prove that the new transfer learning strategy can accurately identify the fault structure in deep fault-karst carbonate, thereby providing a new technical method for this kind of stratigraphic fault interpretation.

2. Material and methods

As mentioned above, seismic fault identification of deep fault-karst carbonate reservoirs using transfer learning can be treated as a feature image segmentation issue. Therefore, this study revolved around constructing a deep learning strategy to accurately identify the fault developing regions from 3D post-stack seismic data. The identification result is expressed fault; a value of 1 indicates a fault developed zone, while a value of 0 indicates a non−fault developed zone. The identification workflow using deep learning is illustrated in Fig. 1. The entire procedure consists of three parts.

(1) **Pretraining the model**: 3D synthetic samples were input into the optimized 3D U-Net network for training, resulting in the creation of the pretrained model.
(2) **Fine-tuning the model:** Real samples were utilized to refine the fault identification model through partial parameter retention—maintaining frozen weights in the pretrained encoder while adaptively optimizing task-oriented decoder layers via stochastic gradient descent to obtain the finally fault identification model.
(3) **Fault identification and application:** The 3D seismic data to be identified were input into I the fault identification model to obtain automatic fault recognition.

2.1. Data processing

Traditional seismic fault interpretation is carried out on a two-dimensional seismic section, which facilitates marking faults and the construction of sample sets. However, faults are geological bodies with spatial distribution characteristics, and the construction of seismic bodies with three-dimensional fault labels as sample sets will improve the quality of sample data and the model's representation of the actual spatial distribution characteristics of faults. Due to the time-consuming and labor-intensive production of 3D seismic sample data with fault labels, a combination of forward modeling labeling (synthetic data) and manual labeling (real data) was adopted to construct the sample set used in this study.

2.1.1. Synthetic data building

For synthetic samples, we adopted the method proposed by Wu et al. [20,24], namely, we obtained fault models through forward modeling. Since the seismic forward simulation obtains seismic wave response characteristics under the assumption of geological characteristics, theoretical seismic data and fault labels can be obtained directly, which greatly reduces the cost of artificial sample construction. The fault parameterization scheme was adaptively calibrated to the fracture density distribution observed in the study area, generating 200 synthetic 3D seismic volumes with deterministic fault annotations. Due to the phenomenon of "blank" when adding stratum folds in the forward simulation process, combined with the actual conditions of hardware computing resources, a 200 × 200 × 200 sample set was randomly generated first, and then a 128 × 128 × 128 sample set was cut by focusing on the fault development area. The detailed steps of the construction process follow.

Step 1. Build a 1D horizontal reflection model, which can be expressed $r(x, y, z)$, the reflection coefficient of which is between -1 and 1 randomly. The resulting reflection model is shown in Fig. 2a (I).

Step 2. Extruding the reflection model, i.e., adding folds to the horizontal layer, is accomplished by a composite function consisting of 2D Gaussian functions and a linear scale function expressed $\frac{1.5z}{z_{\max}}$, as follows:

$$s_1(x,y,z) = a_0 + \frac{1.5z}{z_{\max}} \sum_{k=1}^{k=N} b_k e^{\frac{(x-c_k)^2+(y-d_k)^2}{2\sigma_k^2}}, \tag{1}$$

where a_0, b_k, c_k, d_k, and σ_k are parameters related to the respective fold properties. Different spatial fold structures can be obtained by assigning different values for combination. Through shift mapping and interpolation, the original reflection model is moved longitudinally, and the new reflection model with fold structure is obtained as $r(x, y, z + s_1(x, y, z))$, which is shown in Fig. 2a (II).

Step 3. To account for the structural complexity of natural fault systems, vertical displacement is mechanically regulated through the integration of planar shearing dynamics. The formula is as follows:

$$s_2(x,y,z) = e_0 + fx + gy, \tag{2}$$

where e_0, f, and g are parameters related to planar shearing. We also add $s_2(x, y, z)$ to the reflection model to obtain a new reflection model, $r(x, y, z + s_1(x, y, z) + s_2(x, y, z))$, which is shown in Fig. 2a (III).

Step 4. After obtaining a fold model that varies both laterally and longitudinally, faults are added using Gaussian or linear functions. For faults, fault distance, dip, and dip angle are three key parameters, and the corresponding fault structure can be obtained by setting corresponding values. The displacement of the fault should be contained within the predetermined range of the maximum sampling point. Fig. 2a (IV) shows the result of adding a normal fault using a linear function.

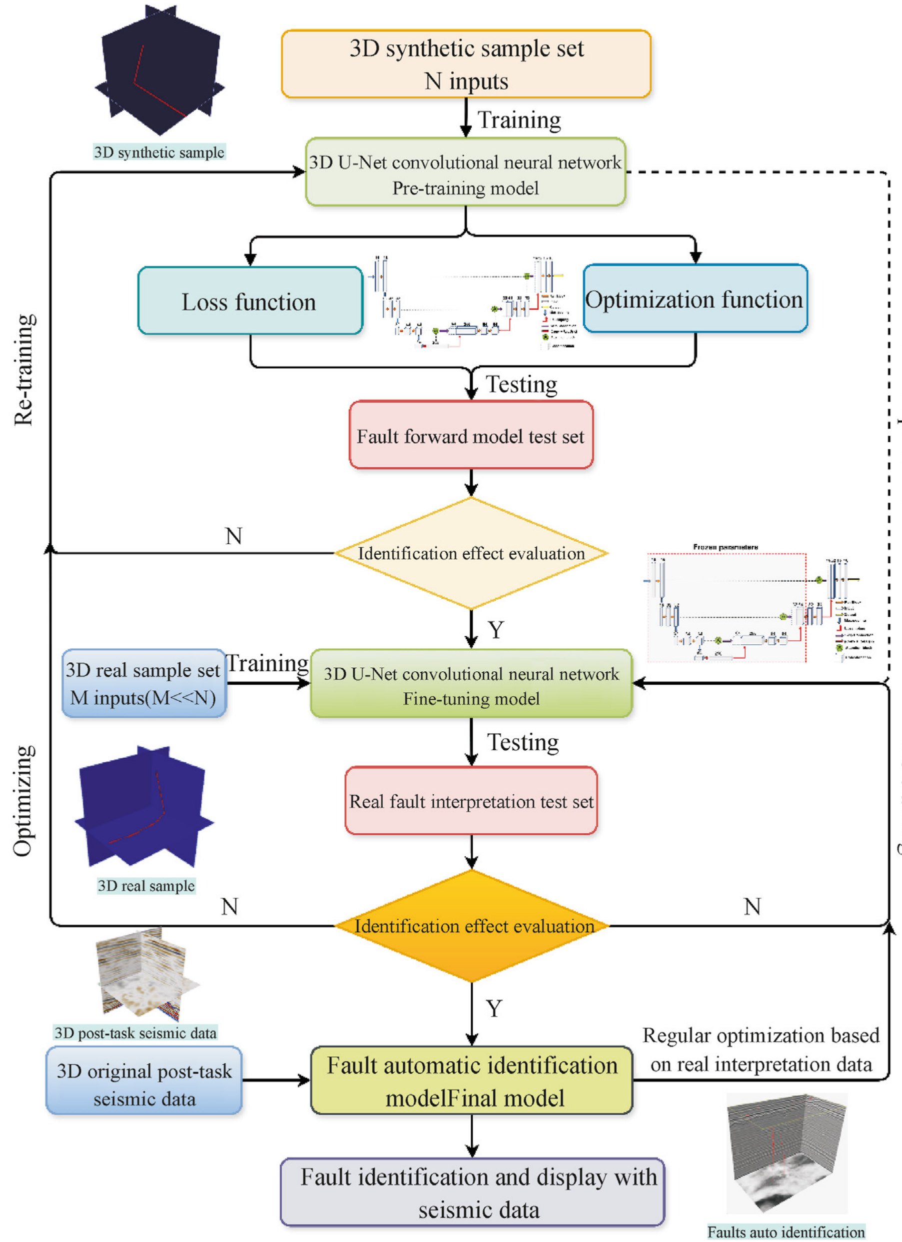

Fig. 1. Fault identification workflow using transfer learning.

Step 5. A 1D reflection model with faults is obtained and convolved with a Riker wavelet to obtain 3D seismic data with fault structure. Fig. 2a (V) shows the result of adding a normal fault using a linear function (Inline profile).

Step 6. In order to further improve the authenticity of synthetic seismic images, some random noise is added to the obtained data, as shown in Fig. 2a (VI).

Step 7. The generated 3D seismic volumes may exhibit localized data voids due to acquisition artifacts. To mitigate these discontinuities and optimize computational workflows, we implemented fault-centered cropping protocols, producing standardized seismic sub-volumes (128 × 128 × 128) with fault zones as geometric anchors. As demonstrated in Fig. 2a (VIII), the cropped inline profile preserves structural continuity while eliminating null-data regions."

Step 8. Finally, the amplitude of the seismic data is normalized, and the value is converted to between −1 and 1. The result is shown in Fig. 2a (VIII) (Inline profile).
Completing these eight steps creates seismic data with fault labels (Fig. 2b). The synthetic dataset generation process incorporates multiple stochastic parameters, which were calibrated against based on geostatistical theory. We generated 200 statistically independent 3D training volumes (128 × 128 × 128) that preserve geological plausibility while

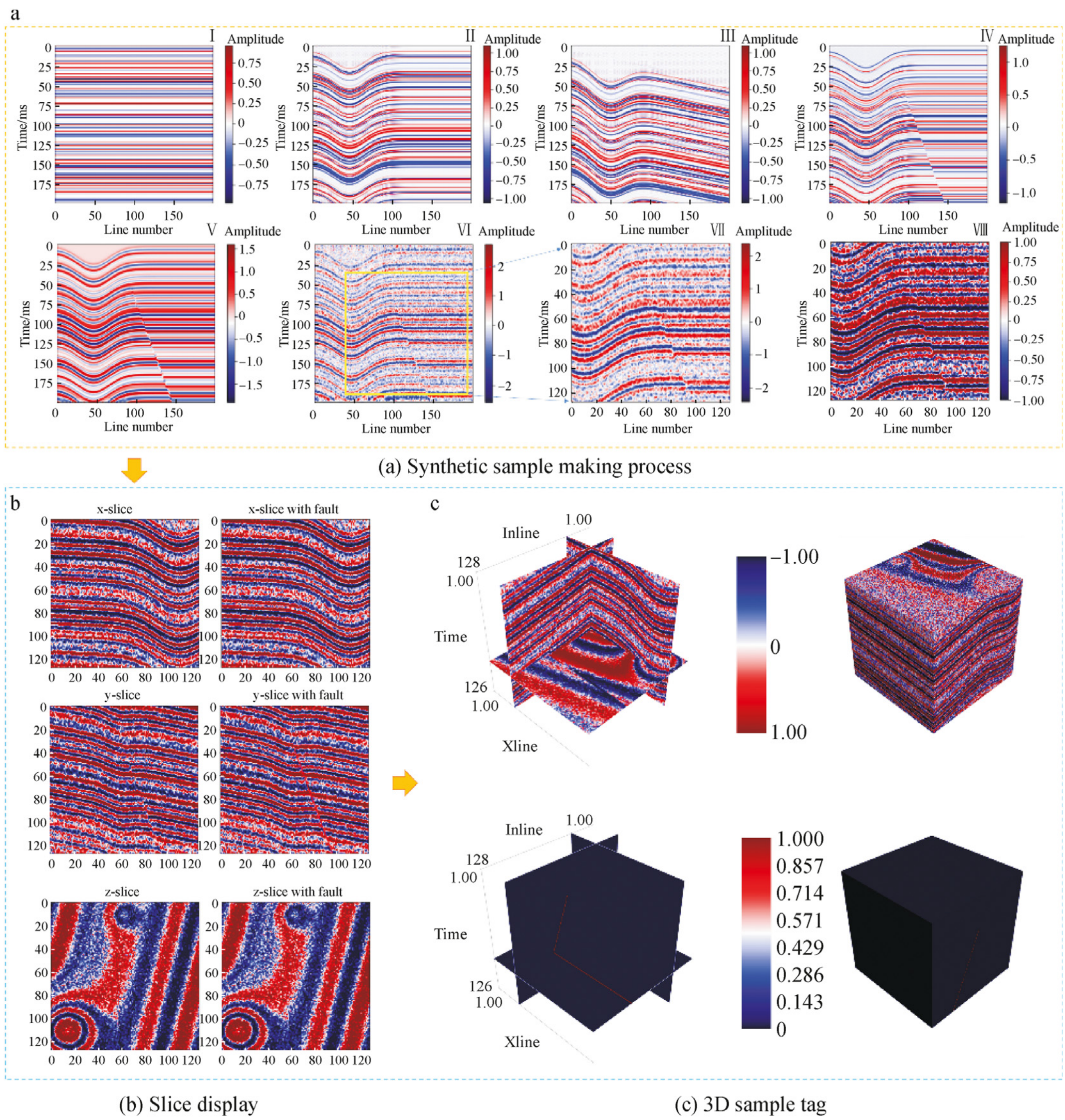

Fig. 2. 3D synthetic fault sample production flow.

ensuring dataset variability. To facilitate training, the data is usually binarized; namely, the fault value should be set to 1 and the non-fault value to 0, which produces a three-dimensional array containing only 0 and 1 (Fig. 2c).

2.1.2. Real data building

For the real samples, we adopted the fault modeling method; the fault model was obtained through the fault fine interpretation and geostatistics method. Based on the fault interpretation results of the actual blocks, the fault plane model was generated and meshmatched according to the sampling rate of the seismic data, then binarized to obtain the real samples. To ensure spatial consistency with the actual fault interpretation, we generated 50 geostatistically constrained 3D training sub-volumes (128 × 128 × 128) through real fault model source from seismic cropping. What follows are the detailed steps of the entire construction process.

Step 1. The standard post-stack seismic data are processed with amplitude preserving, and the region with obvious fault characteristics verified by drilling, logging, and production data is selected as the sample construction area (Fig. 3a).

Step 2. Under the condition of stratification control, section interpretation is used to mark the fault development area by section with fault lines (Fig. 3b (I)).

Step 3. The point division of the fault lines is carried out, and the connection between fault lines is initially constructed according to geological understanding, as shown in Fig. 3b (II).

Step 4. Integrating borehole-to-seismic calibration data from 25 wells with geostatistical kriging simulations, we generated coherent fault surfaces that honor both structural dip

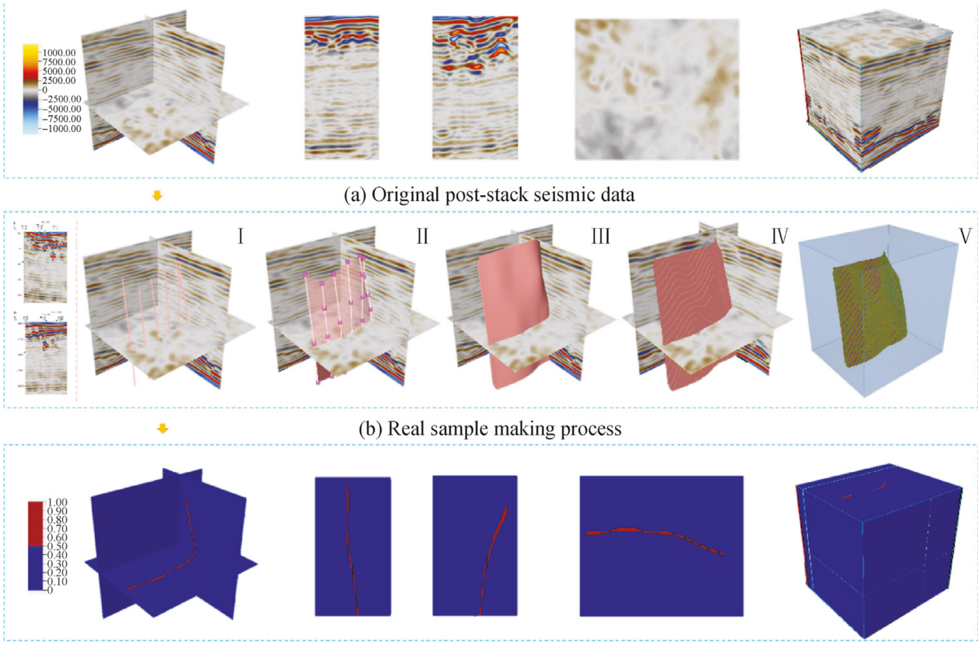

Fig. 3. 3D real fault sample production flow.

distributions and throw gradients, ultimately constructing the fault model shown in Fig. 3b (III).

Step 5. Taking the seismic grid as a reference, the fault model is meshed and the fault is reprocessed (Fig. 3b (IV)).

Step 6. The fault data matched with the seismic grid are written into the seismic grid, and the original data in the seismic grid are replaced (Fig. 3b (V)).

Step 7. The generated seismic data with fault labeling are binarized. The fault value is set to 1, and the non-fault value is set to 0, producing the output shown in Fig. 3c. As with the synthetic data sample, the final output of the real fault sample is a 3D array containing only 0 and 1, and multiple samples can be generated by clipping.

2.1.3. Building the datasets

Some prior studies show that images with more faults are more effective at training a CNN for fault segmentation than are those with fewer faults [21]. However, according to our experience, only when the dataset covers all possible fault feature patterns in the study area can we obtain a good identification result. Therefore, we selected fault samples with different degrees of fault development when constructing the sample set (i.e., a combination of more and fewer faults). The data distribution of the synthetic sample set is shown in Fig. 4, where fault density means the ratio of the number of grids where faults are located to the total number of grids. As the figure shows, the distribution of fault density was between 0 and 0.4, relatively concentrated between 0 and 0.2. The frequency decreased with an increase of fault complexity, and when fault density was greater than 0.1, the frequency decreased significantly (Fig. 4a). Correspondingly, the cumulative frequency rose gently, except for a slight fluctuation at 0.1 (Fig. 4b).

2.2. Fault identification model

2.2.1. Network architecture

To clarify the mapping relationship between 3D seismic data and buried fault locations, we constructed an end-to-end deep CNN [25]. The U-Net architecture originally developed

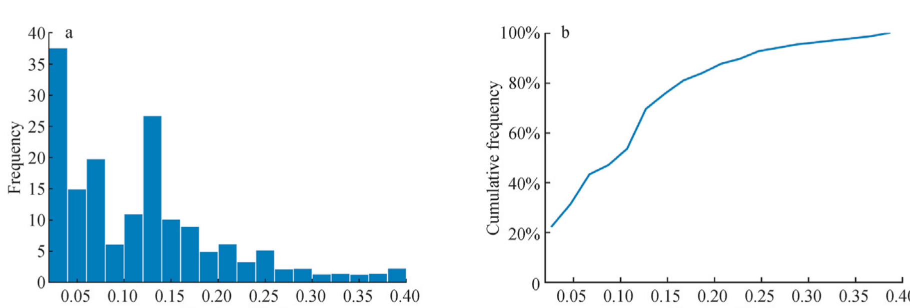

Fig. 4. Fault training set sample cumulative distribution: (a) frequency distribution of density of fault and (b) cumulative frequency distribution of density of fault.

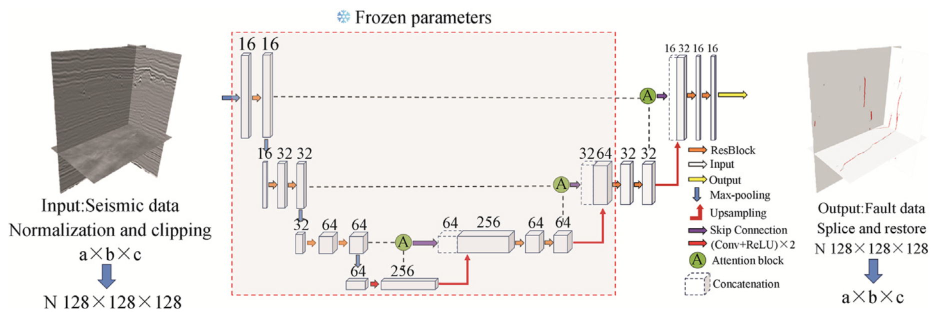

Fig. 5. An improved convolutional neural network (3D U-Net) for fault identification.

by Ronneberger et al. [26] served as the model for the network, as it has been shown to be successful in resolving the nonlinear fault identification problem. In contrast to the original U-Net structure, the CNN uses a 3D convolutional kernel, replacing the original convolution layer with residual blocks (ResBlock) and adding attention blocks to improve the model's capacity to identify fault features. Fig. 5 displays the network structure's specifics. The network receives an input of 3D seismic data, which is divided into an expansion path for precise fault position estimation (right side of Fig. 5) and a contraction path for extracting geological characteristics (left side of Fig. 5). After being divided into 128 × 128 × 128 patches, the raw seismic data is fed into the CNN network. Every stage in the left contracting path consists of two ResBlock layers, which are then followed by a downsampling max-pooling operation with a stride of 2. Two 3 × 3 × 3 convolutional layers and a ReLU activation are employed at the deepest layer. Every step on the right expanding path has two ResBlock layers and an upsampling layer, and every upsampling procedure includes an attention block. Concatenation is used to combine the high-resolution features with the upsampled outputs from the expansion path.

Transfer learning takes the trained model as the initialization for new network training. Since most of the data and tasks are related, fine-tuning the network by transfer learning can help speed up learning efficiency and economize many computing resources. In this study, after training the pretrained model with synthetic data, we froze the downsampling process and part of the upsampling processes, leaving only an upsampling layer and the output layer to update weights (the area in the red dotted box in Fig. 5).

We used ResBlocks in place of the original U-Net structure's convolutional layers to enable a deeper network with a lower chance of vanishing gradient. Fig. 6 (a) displays the structure's details. The input was sent via a second convolution layer (Conv), batch normalization, Leaky Relu, and another batch normalization. The final output was obtained by adding the result to the input and then passing it via a Leaky Relu activation. Additionally, we added attention blocks to the network (Fig. 6b). The attention block computed attention coefficients (α), which were then used to scale an input feature (x). The gating signal (g) was gathered from a coarser scale. After using a Conv to make x and g identical in size, they were put together. So, a Relu, a Conv, a Sigmoid, and a resampling layer are required to obtain the attention coefficient. Trilinear was used for the grid resampling of attention coefficients.

2.2.2. Loss function

When fault locations are designated as positive samples and other locations as negative samples, the fault detection problem can be thought of as a binary segmentation problem. A typical loss function in binary segmentation problems is binary cross entropy (BCE), which performs well in general image segmentation tasks in which the distribution of positive and negative samples is balanced [27]. In defect detection, however, there are significantly fewer positive samples than

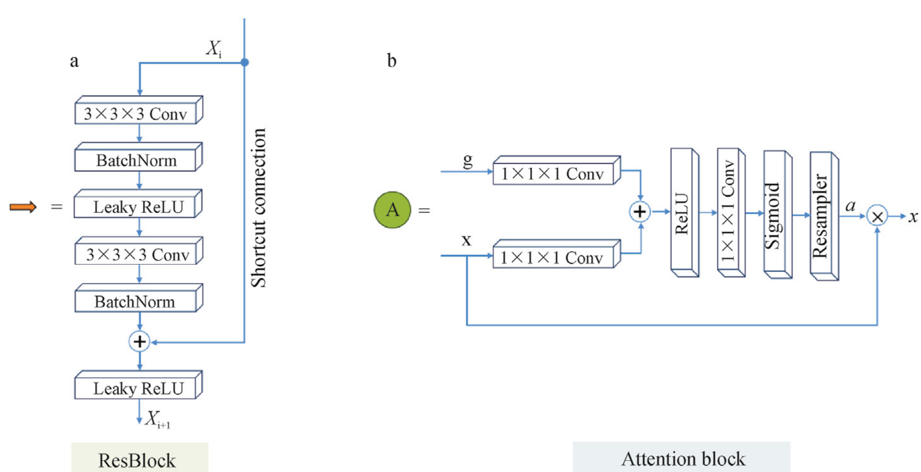

Fig. 6. The structures of the ResBlock (a) and the attention block (b).

negative samples. As a result, when the network is trained using BCE loss, it is simpler for the network to obtain negative predictions in order to minimize the objective function. To optimize this problem, we selected the Dice loss function [28]:

$$L_{Dice} = 1 - Dice = 1 - \frac{2|X \cap Y|}{|X| + |X|}, \tag{3}$$

where the Dice coefficient is used to evaluate the similarity of predictions and labels. X and Y denote the prediction and ground truth, respectively. $|X \cap Y|$ is approximately the addition of dot product between the pixels of the predictions and the ground truth. |X| and |Y| are the pixels in their respective corresponding images. We define the loss function for fault detection, based on the Dice loss, as follows:

$$L_{fault} = 1 - \frac{2\sum_{i=1}^{N} y_i y_i'}{\sum_{i=1}^{N} y_i + \sum_{i=1}^{N} y_i'}, \tag{4}$$

where y_i and y_i' denote the prediction and ground truth, respectively.

2.2.3. Model evaluation

The loss curves of the model trained using transfer learning algorithms and the model trained directly with real data are compared in Fig. 7. The loss curve of the training dataset for the model that was trained directly using manual calibration of real data consistently declines; however, the validation dataset's loss curve does not converge or decrease, indicating a major overfitting situation (Fig. 7a). The loss curves of the training and validation datasets for the model utilizing transfer learning techniques both gradually decline and converge (Fig. 7b).

The identification effects of the ant tracking algorithm (Fig. 8a), the unoptimized U-Net network (Fig. 8b), and our model (Fig. 8c) were compared, and untrained seismic data were used to test the model in order to confirm its influence on fault identification. A single ant tracking color bar was utilized to compare the recognition effects of the three approaches. While the ant tracking algorithm is capable of detecting both small and large fractures, it is not the best method for locating local trunk fault. After initial optimization, the 3D U-Net network exhibited overfitting, meaning it was not accurate enough for local fault characterization. Compared with the above two methods, our model depicted the main fault more precisely, but the characterization of local details still requires further strengthening (Fig. 8).

3. Results

To further test the robustness of the model, block X of the TH Oilfield served as the research area for fault identification. The TH Oilfield is a vast Paleozoic marine carbonate oil field located in the northern Tarim Basin, south of the Aksukule Uplift and in the midst of the Shaya Uplift. It is primarily composed of Ordovician karst fracture-caved reservoirs. The Yingshan Formation and the Yijianfang Formation of the Middle and Lower Ordovician Series are the principal intervals for oil and gas production. The reservoir depth is relatively large, often ranging from 5300 to 6200 m. The

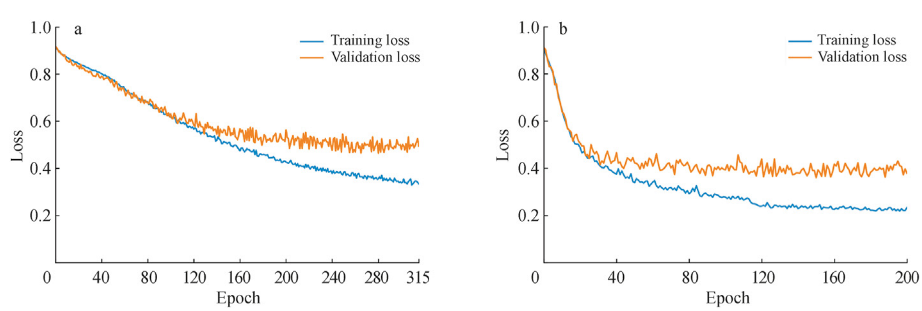

Fig. 7. Loss function optimization effect: (a) the loss curves of the model trained directly with real data; (b) the loss curves of the model trained with transfer learning strategies.

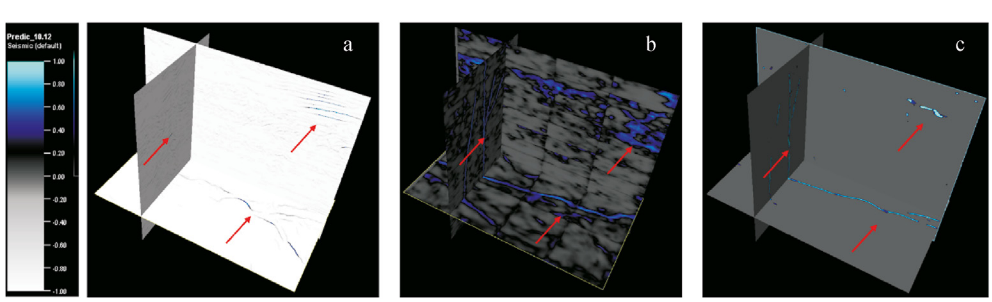

Fig. 8. Fault identification blind testing and comparative analysis: (a) fault identification based on ant tracking; (b) fault identification based on 3d-Unet; and (c) fault identification based on our method.

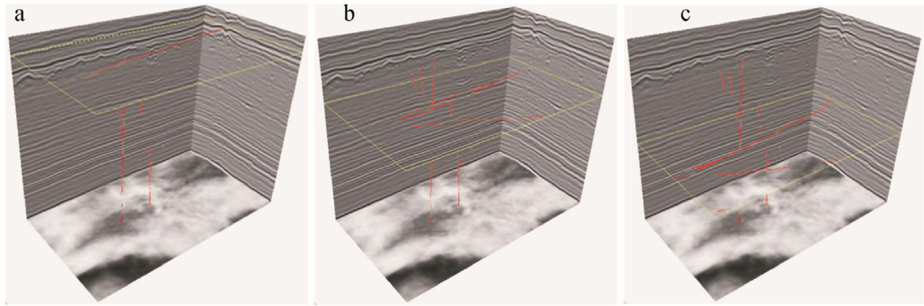

Fig. 9. Fault identification in block X.

primary reservoir space and gas flow channels are made up of a vast number of dissolution pores, fractures, caverns, and faults, with clear scale variations beneath the formation and multi-stage overlaid karst.

3.1. 3D fault identification

There are many different ways that faults can occur. Currently, identifying faults is the first step toward identifying potentially productive locations for partial wells. However, automatic problem identification is desperately needed to assist production, as manual interpretation is labor intensive and time consuming. The fault identification results based on seismic data in block X are shown in Fig. 9. Through the constraint of single well data, according to seismic interpretation, experts proposed that there are 5 main fault systems in the study area. It can be seen from the 3D interpretation results that all faults in the study area were identified using our model.

3.2. Fault characteristics of different profiles

According to the Inline profile, the seismic profiles of Inline = 2406, Inline = 2436, and Inline = 2306 were randomly extracted (Fig. 10). From the identification results, it can be seen that the fault has obvious corrosion characteristics of a fault-karst reservoir.

Xline = 6959, Xline = 7059, and Xline = 7159 seismic profiles were extracted at random based on the Xline profile. The fault displayed a strong “homologous fault” feature on the main fault, and the interpretation results indicated that the fault did not exhibit karst features. Instead, it displayed the flower structure typical of a strike-slip fault. An additional investigation and well drilling plan can focus on the related secondary faults (Fig. 11).

Seismic slices with Time = 3660, Time = 3830, and Time = 4247 were randomly extracted according to the Time slice (Fig. 12). It can be seen from the identification results that the main fault zone has good continuity on the plane, while the secondary fault develops intermittently and has obvious strike-slip characteristics.

4. Discussion

4.1. Sample building

Although creating synthetic data is simple, the training model's capacity for generalization is limited. Nevertheless, a modest sample size is insufficient to increase the training model's identification accuracy. As a result, distinct technological approaches should be used for various issues. In this research, we used the transfer learning approach to enhance model recognition accuracy by fusing synthetic and real data. The sample building process involves many difficulties, particularly when creating real samples. The fault modeling approach used in this work was based on the well-seismic constraint. The fault was then directly modeled using fault interpretation, and the fault was matched using the seismic grid sampling rate. Since faults are interpreted as collections of fault lines, when creating fault models, it is essential to connect these lines using geological reasoning and integrate geostatistics to produce a regional surface. As a result, the information between the lines is really retrieved by interpolation, which raises a fundamental question about the ideal grid size. Denser meshes result in finer interpolation; however, this comes at a great cost in terms of processing. A balance must be achieved in grid resolution, as excessively coarse grids result in significant discrepancies between synthetic and actual fault geometries. As a general rule, faults explanation should try to make the faults as accurate as possible before

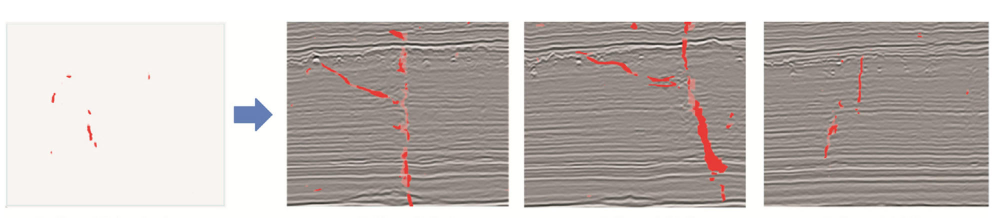

Fig. 10. Fault identification of the seismic Inline profile.

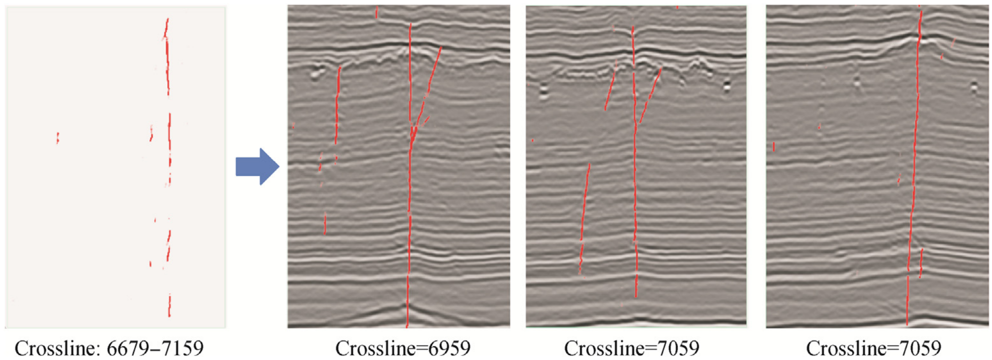

Fig. 11. Fault identification of the seismic Xline profile.

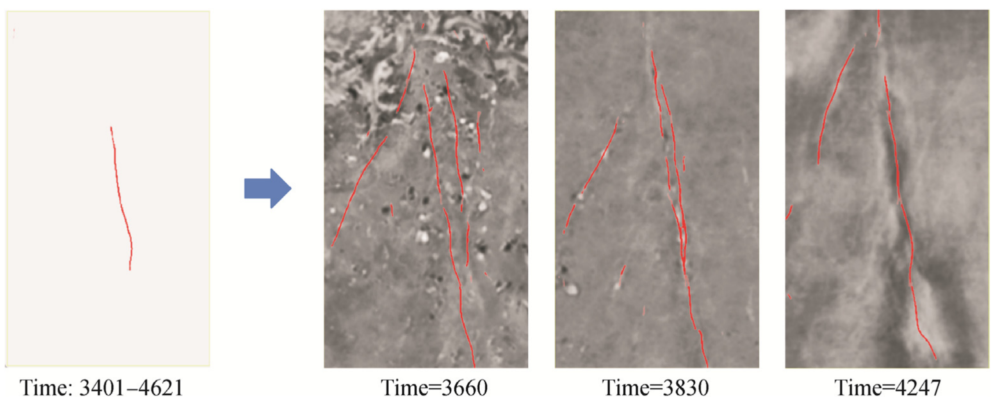

Fig. 12. Fault identification of the seismic Time slice.

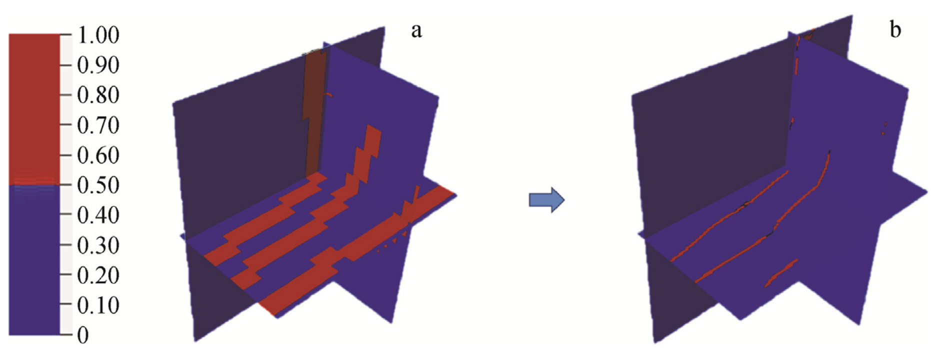

Fig. 13. The influence of grid on label quality: (a) bad label; (b) good label.

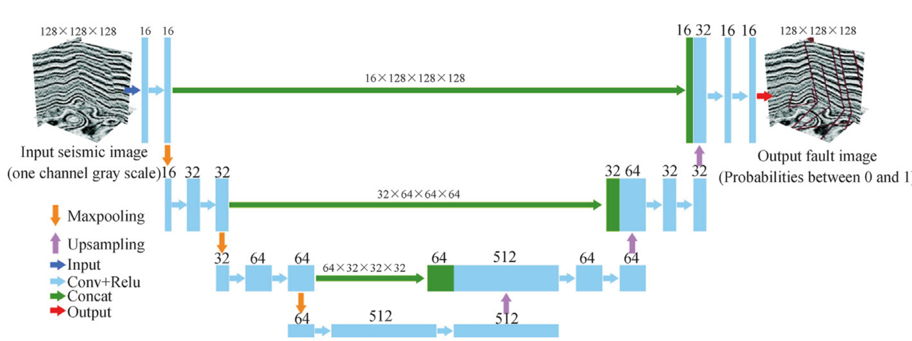

Fig. 14. A simplified end-to-end convolutional neural network (U-Net) for 3D fault detection (Sourced from Wu et al. [20]).

putting them into the seismic grid. Then, when writing them into the seismic grid, they merely need to be united with it. If the label is created according to the seismic grid sampling rate at the beginning, the subsequent label samples are not useable. Fig. 13 shows the difference between a "bad label" and a "good label."

4.2. Model optimization

Optimizing models and algorithms is another efficient method of raising the accuracy of fault recognition, in addition to enhancing the quality of the data. Since its proposal, the U-Net network topology has found extensive application in computer vision, particularly in the domain of medical image segmentation. Subsequently, it was brought into the field of geophysics. Choosing a local window or cube on each image pixel point for fault recognition is a prerequisite for early fault recognition using the CNN method. This process is computationally demanding, particularly for 3D fault recognition. In 2019, Wu et al. [20] introduced a CNN architecture optimized for fault identification; they streamlined the U-Net framework by reducing its complexity and reformulated fault detection as a tractable end-to-end binary image segmentation task. To increase computational efficiency, they streamlined the original U-Net by lowering the number of convolutional layers and the characteristics of each layer (Fig. 14). This inspired our model optimization strategy, specifically, modular refinement of the network architecture through integration of field-acquired seismic constraints while preserving the U-Net framework. For instance, we developed a hybrid loss function combining weighted cross-entropy and geological continuity constraints, and implemented multi-scale residual connections with self-attention gating modules in this study. Additionally, these architectural innovations can be deployed as preliminary validation benchmarks, analogous to established U-Net variants like V-Net (for volumetric medical segmentation) and U-Net++ (with nested dense connections) in their respective domains.

In particular, it should be noted that although the transfer learning strategy proposed in this paper is universal, the identification effect may not be satisfactory due to the problem of sample characteristics in the actual operation process. We mainly carried out fault identification for carbonate formations, and the applicability to other types of formations, such as sandstone and other clastic rock formations, still needs to be examined. However, it is undeniable that with continuous improvement of the quantity and quality of synthetic and real samples, as well as continuous optimization of algorithms, automatic fault recognition based on deep learning will achieve more ideal results.

5. Conclusions

(1) The pretrained model can be obtained by training synthetic data, and then the final recognition model can be obtained by fine-tuning the pretrained model with real data. This approach, based on transfer learning, can effectively improve fault identification.
(2) Based on the CNN framework, overfitting is mitigated through stochastic dropout regularization ($p = 0.5$), identification accuracy is enhanced via multi-scale residual learning modules replacing standard convolutions, and training efficiency is optimized by integrating channel-spatial attention mechanisms, collectively advancing fault detection performance
(3) Extremely imbalanced positive and negative samples in fault identification can be effectively solved by optimizing the loss function and substituting the Dice loss function for the weighted cross-entropy function.

CRediT authorship contribution statement

Hanqing Wang: Conceptualization, Investigation, Formal analysis, Writing−review & editing. **Han Wang:** Investigation, Formal analysis, Writing−review & editing, Funding acquisition. **Kunyan Liu:** Data curation, Resources. **Jin Meng:** Data curation. **Yitian Xiao:** Supervision. **Yanghua Wang:** Supervision.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to gratefully acknowledge the financial support provided by the China Postdoctoral Science Foundation (Grant No. 2024M763650) and the Excellent Young Scientists Fund Program of SINOPEC Petroleum Exploration and Production Research Institute (Grant No. yk2024010). We also thank Dr. Miao Tian (SINOPEC Petro E&P Research Inst) for technical support related to creating real fault labels.

References

[1] Ma Yongsheng, Cai Xunyun, Yun Lu, et al. Practice and theoretical and technical progress in exploration and development of Shunbei ultra-deep carbonate oil and gas field, Tarim Basin, NW China. Petrol Explor Dev 2022;49(1):1−20.
[2] Zhao Wenzhi, Hu Suyun, Liu Wei, et al. Petroleum geological features and exploration prospect of deep marine carbonate rocks in China onshore: a further discussion. Nat Gas Ind B 2014;1(1):14−23.
[3] Lu Xinbian, Wang Yan, Tian Fei, et al. New insights into the carbonate karstic fault system and reservoir formation in the Southern Tahe area of the Tarim Basin. Mar Petrol Geol 2017;86:587−605.
[4] Ukar E, Baqués V, Laubach SE, et al. The nature and origins of decametre-scale porosity in Ordovician carbonate rocks, Halahatang oilfield, Tarim Basin, China. J Geol Soc 2020;177(5):1074−91.

[5] Li Hongbin, Wang Guiwen, Li Yuhang, et al. Fault-Karst systems in the deep ordovician carbonate reservoirs in the Yingshan Formation of tahe oilfield Tarim Basin, China. Geoenergy Sci Eng 2023;231, 212338.
[6] Ghosh T, Saha RK, Singh SK, et al. LRSTV: a low-rank total variation-based seismic fault preserving denoising algorithm. J Appl Geophys 2023;210, 104948.
[7] Di Haibin, Shafiq MA, Wang Zhen, et al. Improving seismic fault detection by super-attribute-based classification. Interpretation 2019;7(3):SE251−S267.
[8] Bahorich M & Farmer S. 3-D seismic discontinuity for faults and stratigraphic features: the coherence cube. Lead Edge 1995;14(10):1053−8.
[9] Marfurt KJ, Kirlin RL, Farmer SL, et al. 3-D seismic attributes using a semblance-based coherency algorithm. Geophysics 1998;63(4): 1150−65.
[10] Gersztenkorn A & Marfurt KJ. Eigenstructure-based coherence computations as an aid to 3-D structural and stratigraphic mapping. Geophysics 1999;64(5):1468−79.
[11] Roberts A. Curvature attributes and their application to 3 D interpreted horizons. First Break 2001;19(2):85−100.
[12] Al-Dossary S & Marfurt KJ. 3D volumetric multispectral estimates of reflector curvature and rotation. Geophysics 2006;71(5):P41−51.
[13] Geng Zhi, Wang Hanqing, Fan Meng, et al. Predicting seismic-based risk of lost circulation using machine learning. J Petrol Sci Eng 2019; 176:679−88.
[14] Sun Fengyuan, Gao Jinghuai, Zhang Bing, et al. Coherence algorithm with a high-resolution time−time transform and feature matrix for seismic data. Geophys Prospect 2020;68(4):1113−25.
[15] Silva CC, Marcolino CS & Lima FD. Automatic fault extraction using ant tracking algorithm in the Marlim South Field, Campos Basin. In: 2005 SEG international exposition and annual meeting. Houston, Texas: United States of America; 2005.
[16] Di Haibin, Shafiq MA & AlRegib G. Seismic-fault detection based on multiattribute support vector machine analysis. In: 2017 SEG international exposition and annual meeting, Houston, Texas, United States of America; 2017.
[17] Ren Ke, Zou Guangui, Zhang Shaomin, et al. Fault identification and reliability evaluation using an SVM model based on 3-D seismic data volume. Geophys J Int 2023;234(1):755−68.
[18] Huang Lei, Dong Xishuang & Clee TE. A scalable deep learning platform for identifying geologic features from seismic attributes. Lead Edge 2017;36(3):249−56.
[19] Wei Xiong, Xu Ji, Ma Yue, et al. Seismic fault detection with convolutional neural network. Geophysics 2018;83(5):97−103.
[20] Wu Xinming, Liang Luming, Shi Yunzhi, et al. FaultSeg3D: using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation. Geophysics 2019;84(3):35−45.
[21] Liu Naihao, He Tao, Tian Yajun, et al. Common-azimuth seismic data fault analysis using residual UNet. Interpretation 2020;8(3):25−37.
[22] Lin Lei, Zhong Zhi, Cai Zhongxian, et al. Automatic geologic fault identification from seismic data using 2.5 D channel attention U-Net. Geophysics 2022;87(4):IM111−24.
[23] Chen Jixing, Liu Pengfei, Yuan Zhaolin, et al. CUNet deep learning method and its application in fault recognition. Progress in Geophysis 2024;39(2):561−71 (in Chinese).
[24] Wu Xinming & Hale D. 3D seismic image processing for faults. Geophysics 2016;81(2):IM1−1.
[25] Wang Han, Wang Hanqing, Wu Xinwei, et al. Low-order faults identification of carbonate reservoir based on 3D attention-based U-net. In: 85th EAGE annual conference & exhibition (including the workshop programme), Oslo, Norway; 2024.
[26] Ronneberger O, Fischer P & Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention−MICCAI 2015: 18th international conference, Munich, Germany, 2015, proceedings, part III 18 (pp. 234-241). Springer International Publishing.
[27] Milletari F, Navab N & Ahmadi SA. V-net: fully convolutional neural networks for volumetric medical image segmentation. In: Fourth international conference on 3D vision (3DV); 2016. Stanford, CA, USA. IEEE.
[28] Sudre CH, Li W, Vercauteren T, et al. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In: Deep learning in medical image analysis and multimodal learning for clinical decision support: third international workshop, DLMIA 2017, and 7th international workshop, ML-CDS 2017, held in conjunction with MICCAI 2017, québec city, QC, Canada, september 14, proceedings 3 2017 (pp. 240-248). Springer International Publishing.