ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

Identification and characterization of fractured cavities in carbonate reservoirs using Swin-UNet transformer and seismic attribute compression fusion

Yunhao Cui[1,2], Yuhua Chen[1,2]*, Chao Xu[1,2], Yaping Huang[3], Qiang Guo[3], Zhiqiang Lu[4], Zhanpeng Chen[1,2], and Yuwen Qian[1,2]

[1]Key Laboratory of Coalbed Methane Resource & Reservoir Formation Process, Ministry of Education, School of Resources and Geosciences, China University of Mining and Technology, Xuzhou, Jiangsu, China

[2]Department of Geo-information Science, School of Resources and Geosciences, China University of Mining and Technology, Xuzhou, Jiangsu, China

[3]Department of Geophysics, School of Resources and Geosciences, China University of Mining and Technology, Xuzhou, Jiangsu, China

[4]Research Institute of Exploration and Exploitation, Sinopec Northwest China Petroleum Bureau, Urumqi, Xinjiang Uyghur Autonomous Region, China

Abstract

Identifying and characterizing fractured cavities is essential for exploring carbonate reservoirs. However, characterizing the development and distribution of fractured cavities through post-stack seismic attribute analysis remains challenging. Recently, convolutional neural networks (CNNs), such as UNet and its enhanced versions, have enabled the quantitative identification of fractured cavities. Despite these advancements, the local receptive field and weight-sharing mechanisms of these CNNs limit their capability to capture long-range features within strike–slip fault systems. In addition, neural networks are inherently affected by data uncertainty. To address these challenges, a two-step methodology is proposed. The first step utilizes a Swin-UNet transformer (UNETR) model, enhanced with an attention gate, to interpret fractured cavities. The transformer in Swin-UNETR improves the detection of fractured cavities in strike–slip fault zones, whereas the attention gate enhances the recognition of small fractured cavities by increasing their response in the feature maps. This enhanced Swin-UNETR model overcomes the limitations in modeling long-range features. In the second step, the fractured-cavity identification results are combined with seismic attributes from conventional analysis. Principal component analysis is employed both to increase the relative weight of the neural network recognition results in the attribute fusion and to reduce the uncertainty associated with any single identification method. The methodology was validated in the Shunbei area, yielding horizontal segmentation and vertical zonation of fractured cavities, as well as their characterization through fixed-grid modeling. By combining deep learning-based feature extraction with seismic attributes, this approach improves the accuracy of fractured cavity identification and characterization in carbonate reservoirs.

Keywords: Fractured cavity identification and characterization; Seismic attribute compression fusion; Convolutional neural network

***Corresponding author:**
Yuhua Chen
(chenyuhua@cumt.edu.cn)

Citation: Cui Y, Chen Y, Xu C, *et al*. Identification and characterization of fractured cavities in carbonate reservoirs using Swin-UNet transformer and seismic attribute compression fusion. *J Seismic Explor*. 2026;35(1):119-138. doi: 10.36922/JSE025420090

Received: October 14, 2025

Revised: November 13, 2025

Accepted: December 12, 2025

Published online: January 5, 2026

1. Introduction

Fractured-cavity reservoirs are carbonate rock formations that have developed along fault zones subjected to multiphase tectonic stresses. The combined effects of these tectonic movements, the development of karst features over time, and the upwelling of hot fluids create complex fractured cavities in the carbonate layers. These cavities serve as primary reservoirs for oil and gas, playing a critical role in both hydrocarbon storage and migration. Therefore, accurately locating and characterizing these reservoirs is critical for successful oil and gas exploration and development.[1]

Domestic and international researchers primarily utilize post-stack seismic attributes, including coherence, curvature, and amplitude, to detect and characterize fractured cavities. These traditional methods approximate the locations of fractures and cavities but cannot precisely delineate their size or shape. With the rapid advancements in deep learning for visual and medical imaging, and given the structural similarity between three-dimensional (3D) post-stack seismic data and magnetic resonance imaging scans, deep learning has seen increasing adoption in seismic interpretation. For example, Wu *et al.*[2] used a convolutional neural network (CNN) with the UNet architecture to identify karst paleocaves from 3D seismic datasets. They also developed a method to create synthetic datasets that mimic the seismic reflections of paleocaves. High-quality seismic data annotations are indispensable for effective feature recognition using UNet. Building on this, Li *et al.*[3] adopted an enhanced UNet CNN, using field-annotated cavity contours and synthetic datasets of faults, fractures, and karst caves to improve fractured cavity identification in strike–slip fault zones. Despite these advances, uncertainty persists in seismic interpretation with deep learning. To address this, Gui *et al.*[4] combined UNet++ with multiscale convolutional kernels and seismic attributes to improve accuracy. Zhang *et al.*[5] created a Bayesian deep learning framework that identifies cave shapes and measures uncertainty in the results. However, the increased model complexity leads to prolonged training and inference durations.

While the aforementioned methodologies have exhibited promising results, they still suffer from inherent limitations. In contrast to paleocaves that develop independently, the formation of fractured cavities is influenced by strike–slip fault zones that extend over several kilometers. The intrinsic local receptive field and weight-sharing mechanisms of CNNs fundamentally constrain their capability to detect large-scale fractured cavity systems. Early convolutional layers primarily focus on localized regions of the input data, thus failing to capture correlations across distant regions, which inherently limits their efficacy in identifying long-range geological features (strike–slip fault systems). Moreover, fractured cavities exhibit considerable variability in size and morphology, with smaller cavities often displaying complex geometries. The application of neural network-based methods for fractured cavity identification also introduces both data-driven and model-induced uncertainties, warranting effective mitigation strategies.

To address the inherent locality limitations of CNNs, the present study employs a Swin-UNet transformer (UNETR) network augmented with an attention gate for the identification of fractured cavities. The Swin-UNETR architecture synergistically combines transformer models.[6] with UNet frameworks, thereby enhancing the network's capability to recognize complex fractured cavities over long spatial ranges. The encoder component utilizes a sliding window operation to reduce computational complexity during the processing of seismic input data. In addition, to mitigate uncertainties arising from neural network methodologies, conventional seismic attribute techniques are incorporated to constrain the identification outcomes. Specifically, principal component analysis is employed to integrate the fractured-cavity segmentation results—treated as a dominant attribute—with multiple seismic attribute datasets.[7] This integrative approach combines fractured-cavity information from all methods into a single fused attribute volume, enhancing overlapping features while suppressing inconsistent ones and thereby improving the overall accuracy of fractured-cavity identification.

2. Methodology

2.1. Field seismic data and geological background

The post-stack seismic datasets used in this study, outlined by the blue box in Figure 1, were obtained from the Shunbei Oil and Gas Field. The field is primarily located in the Shuntuoguole Low Uplift and extends southeastward to the Guchengxu Uplift. Geographically, it is adjacent to the Shaya Uplift to the north and the Katakunlun Uplift to the south, with the Manjiaer Depression and Awati Depression bounding it to the east and west, respectively. The field lies in a structurally low "saddle" zone, generally characterized by relatively higher elevations in the northern and eastern parts and lower elevations in the southern and western parts, forming an overall gently undulating geomorphology.[8]

During the early Caledonian era, the Shunbei area experienced stable tectonic subsidence in the craton, characterized by a weak extensional tectonic regime. Subsequently, from the mid- to late Caledonian era to the early Hercynian era, the region experienced the formation

and evolution of low uplifts under a regional compressional tectonic regime. This was followed by further tectonic adjustments and reworking during the mid-to-late Hercynian period, culminating in the formation of the present-day tectonic framework.

As shown in the comprehensive stratigraphic column of Figure 2, the stratigraphic sequence in the Shunbei area ranges from the Lower Paleozoic Cambrian to the Cenozoic Quaternary, with a notable absence of the Jurassic System. The Yuertusi formation of the Lower Cambrian is extensively developed as a source rock, distinguished by its high organic carbon content and considerable hydrocarbon generation potential. The focus of this study is the Ordovician carbonate strata, which are buried at depths ranging from 7200 to 8800 m below sea level. These Ordovician formations, arranged

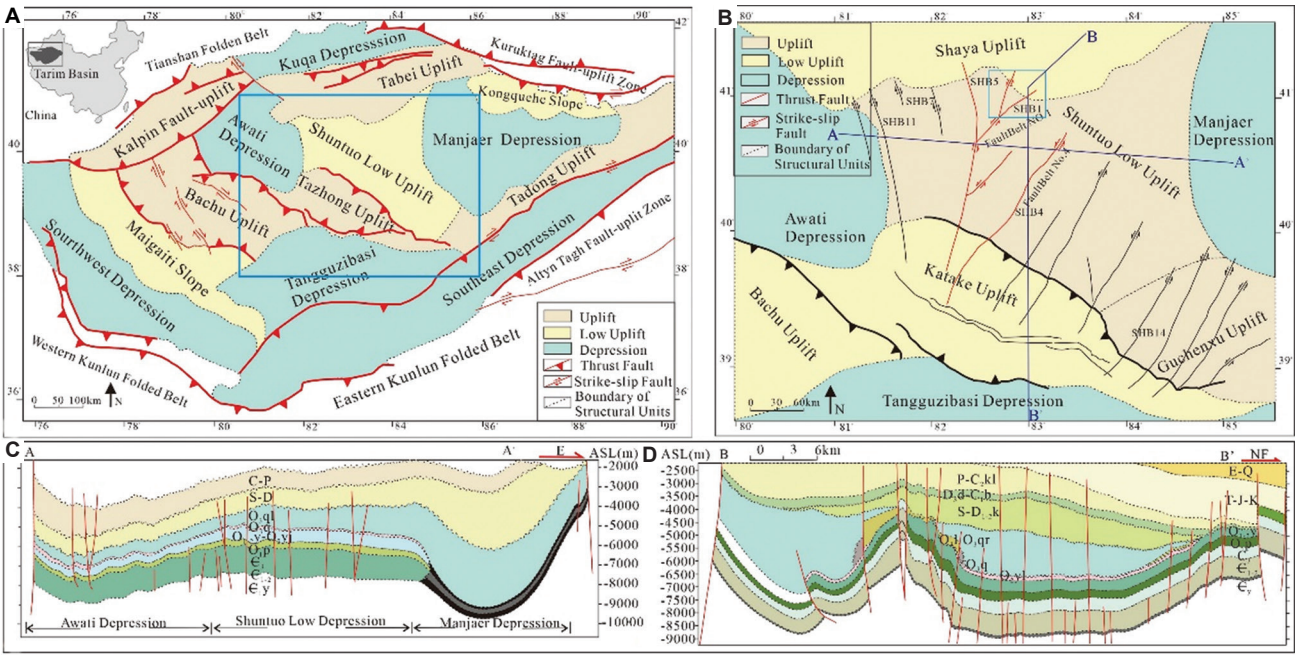

Figure 1. (A-D) Tectonic map and stratigraphic section of the central Tarim Basin, where the study area is located. Image modified with permission from Wang *et al.*[9] Copyright © 2024, IEEE.

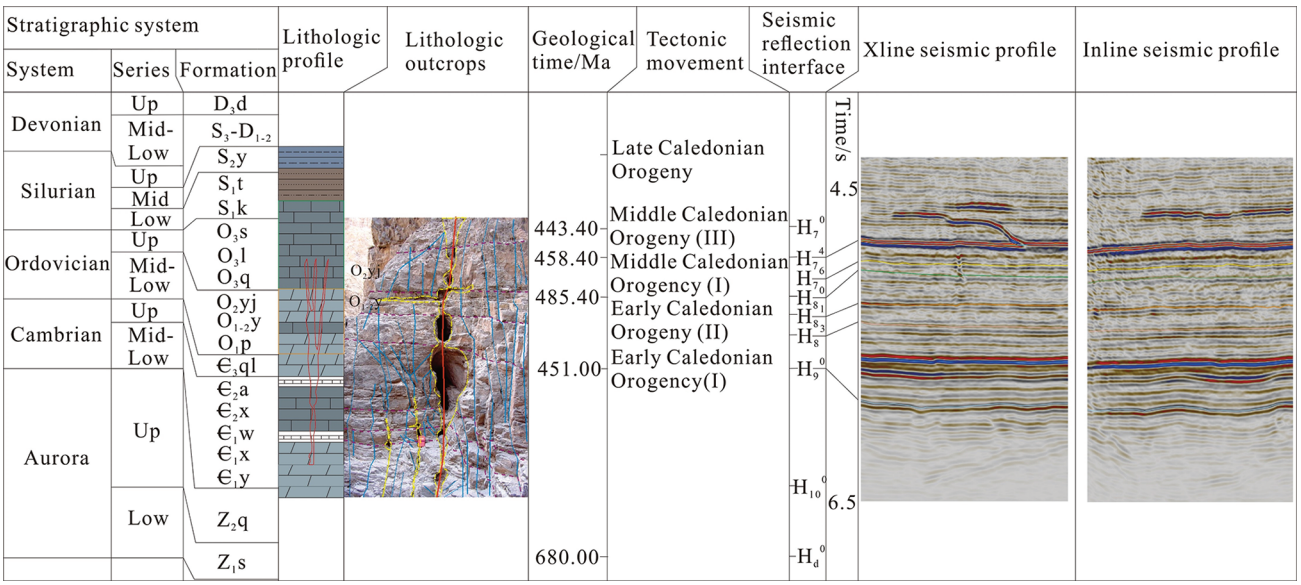

Figure 2. Tectonic map and stratigraphic section of the central Tarim Basin. Image modified from Zheng *et al.*[11]

stratigraphically in ascending order, include the Lower Ordovician Penglaiba formation (O_1p), the Middle-Lower Ordovician Yingshan formation ($O_{1-2}y$), the Middle Ordovician Yijianfang formation (O_2yj), the Upper Ordovician Qiaerbake formation (O_3q), the Lianglitage formation (O_3l), and the Santamu formation (O_3s). Within this sequence, the carbonate units in the upper portion of the Yingshan formation and the Yijianfang formation represent the principal targets for oil and gas exploration in the study area.[10]

A northeast-trending strike–slip fault system has developed in the study area, providing an effective conduit for hydrocarbon migration and promoting hydrocarbon accumulation. The reservoir space is primarily composed of tectonically generated fractured cavities.[12] Seismic data acquired along this fault zone during exploration are shown in Figure 3. The target area shown in the figure has dimensions of 2501 time slices × 1022 inlines × 865 crosslines. To better characterize the spatial distribution of multiscale carbonate fractured cavity reservoirs in the Ordovician and Cambrian strata, strong reflection axes were removed from the seismic data (Figure 4), thereby preventing the reflection information of the underlying fractured cavities from being masked by these strong reflection axes. The data were then divided into six segments, referred to as object area (OA)-crop1 through OA-crop6. These segments correspond to the southern, central, and northern sections of the study area.

2.2. Fractured cavities identification using Swin-UNETR

2.2.1. Training of datasets

As a data-driven supervised learning approach, deep learning requires feature-rich training datasets for effective model optimization. In carbonate fault-controlled reservoirs, fractured cavities manifest as clusters of strong reflections on a weak reflection background in seismic datasets. As small-scale geological anomalies within

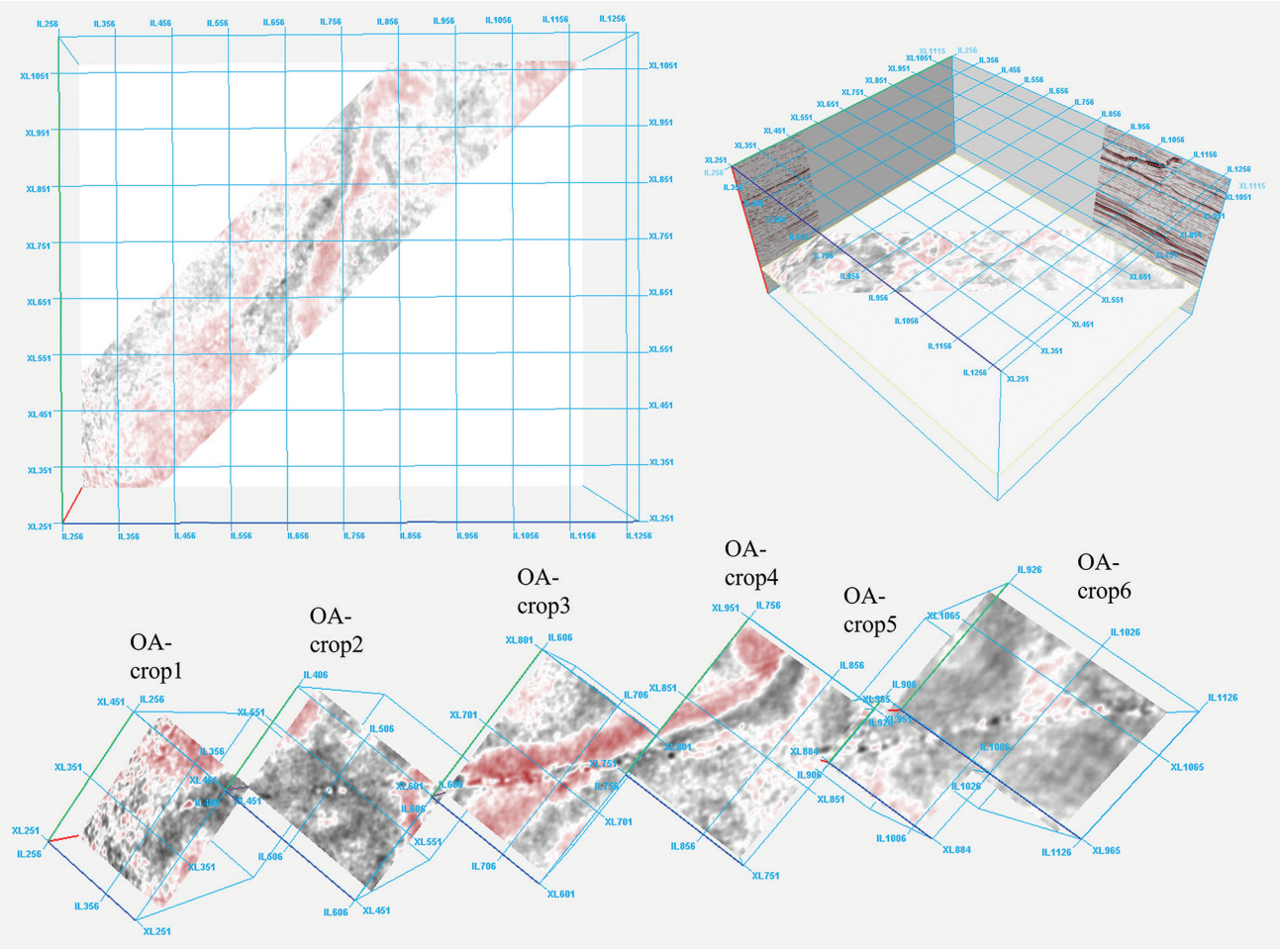

Figure 3. Three-dimensional exploration of the seismic data in the study area
Abbreviation: OA: Object area.

carbonate reservoirs, fracture cavities create diffraction effects during seismic wave propagation. This phenomenon arises from wave impedance contrasts, generating multiple vertical strong energy anomalies. Figure 5 shows that forward modeling reveals the beaded anomalous reflection characteristics of fractured cavities on post-stack seismic sections. These are typically shown by seismic attributes such as strong amplitude anomalies, high amplitude variation rates, weak to moderate coherence, low wave impedance, and low velocity. As shown in Figure 6, inversion of the seismic profiles indicates that fractured cavities in the OA exhibit nearly ellipsoidal outlines with significant variation in scale, ranging from 10 to 150 m.[13]

Based on the relationship between the morphology and size of fractured cavities and the beaded anomalous reflections observed on seismic profiles, a convolutional synthesis method using a seismic reflectivity model and Ricker wavelets was employed to construct training samples of fractured cavities. The process of establishing a seismic reflectivity model involved three key steps. First,

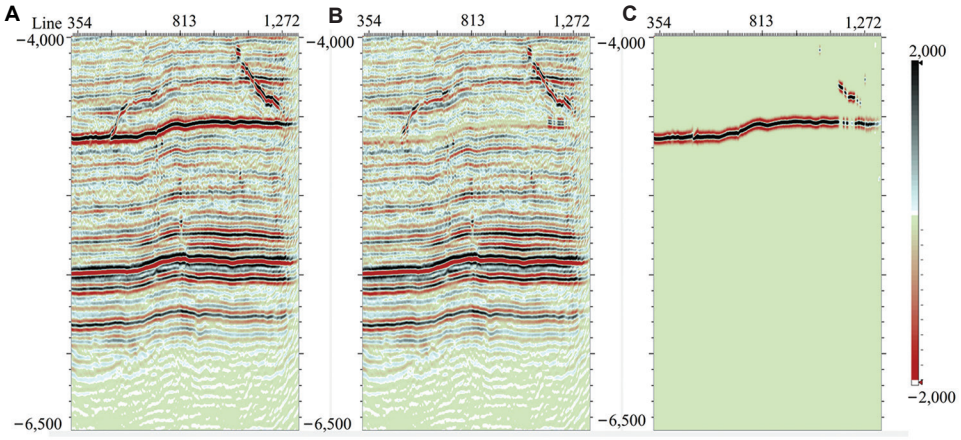

Figure 4. Object area data after strong axis removal processing. (A) Original seismic profile. (B) Seismic profile after removing the strong axis. (C) Removed H_7^4 strong axis.

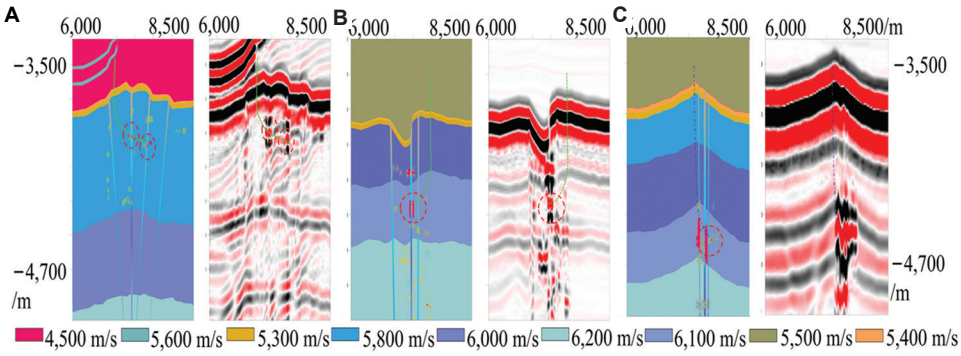

Figure 5. Forward modeling of the beaded reflection characteristics of the fractured cavities. (A) Compression section. (B) Pull-apart segment. (C) Translational segment.

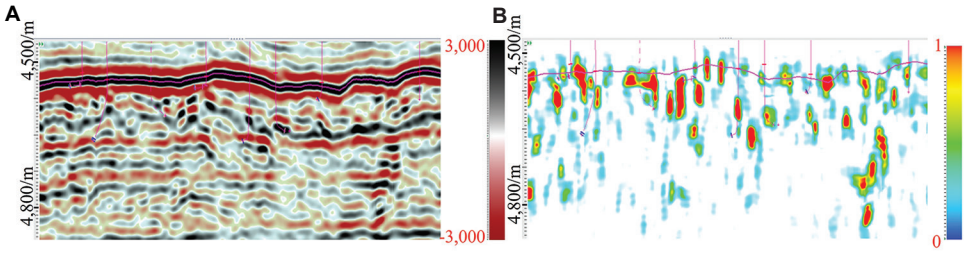

Figure 6. Amplitude curvature seismic attribute profile of fractured cavities. (A) Raw seismic profile. (B) Amplitude curvature seismic attribute profile.

building an initial 3D horizontal-layered reflectivity model (*X*, *Y*, *Z*); second, incorporating formation dip and fold structure models (X_s, Y_s, Z_s)into the horizontal layers; and finally, integrating the fractured cavities components ($\bar{X}$, $\bar{Y}$, $\bar{Z}$) into the composite model.

The geometric fold model (X_s, Y_s, Z_s) was generated by applying a vertical offset to the initial model (*X*, *Y*, *Z*). This vertical offset field was decomposed into large-scale fold surfaces S_1 (X, Y, Z), small-scale local folds S_2 (X, Y, Z), and multidirectional dip components S_3 (X, Y, Z).

$$S_1(X,Y,Z) = A_0 + \left(1 - \left|\frac{a(X-X_0)+b(Y-Y_0)}{X_{max}+Y_{max}}\right|\right) \sum_{k=1}^{N}\left(A_k \cos\left(\frac{k\pi E(X,Y)}{\lambda_k}\right) + B_k \sin\left(\frac{k\pi E(X,Y)}{\lambda_k}\right)\right)$$

$$E(X,Y) = \sqrt{\left(\frac{\cos\phi_k(X-X_0)-\sin\phi_k(Y-Y_0)}{L[x]_k}\right)^2 + \left(\frac{\sin\phi_k(X-X_0)-\cos\phi_k(Y-Y_0)}{L[y]_k}\right)^2} \quad (1)$$

The first large-scale displacement field, as shown in **Equation 1**, was defined by a Fourier series. In this context, the parameters X_0 and Y_0 represent the central position of the modified Foucault's equation. The Fourier coefficients A_k and B_k define the amplitude of the folds. The attenuation factors *a* and *b* control the amplitude variation of the folds in the X and Y directions, respectively. The elliptic coefficients $L[x]_k$ and $L[y]_k$ control the degree of fold curvature in the X and Y directions, respectively, with $L[x]_k$ corresponding to the ellipse's major axis. The parameter ϕ_k denotes the clockwise rotation angle of the ellipse, and the parameter λ_k controls the compression ratio of the wavelength.

$$S_2(X,Y,Z) = C\frac{Z}{Z_{max}}\sum_{k=1}^{N} Dk\exp\left[-\left(\frac{\cos\phi_k(X-E_k)-\sin\phi_k(Y-F_k)}{2\sigma[x]_k}\right)^2 - \left(\frac{\sin\phi_k(X-E_k)-sin\phi_k(Y-F_k)}{2\sigma[y]_k}\right)^2\right] \quad (2)$$

The second small-scale fold displacement field, as shown in **Equation 2**, was defined by a Gaussian equation; the Gaussian coefficients $\sigma[x]_k$ and $\sigma[y]_k$ represent the major and minor axis lengths of the ellipse, with $\sigma[x]_k$ denoting the major axis. The ϕ_k is the clockwise rotation angle of the fold, whereas E_k and F_k specify the central position of the Gaussian function. The fold amplitude is controlled by a separate parameter D_k, whereas an attenuation factor C adjusts its vertical variation. The vertical extent of the folds in the Z direction can be controlled by specifying appropriate lower and upper limits, for example, by setting $\sigma[x]_k = \sigma[x]_0 + dZ$ and $\sigma[y]_k = \sigma[y]_0 + eZ$.

$$S_3(X,Y,Z) = Gx + Hy + I \quad (3)$$

Finally, the third displacement field, as shown in **Equation 3**, used a linear equation to simulate the dip of strata in various directions, with parameters G and H controlling the dip angles in the X and Y directions, respectively.

$$\begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z + S_1(X,Y,Z) + S_2(X,Y,Z) + S_3(X,Y,Z) \end{bmatrix} \quad (4)$$

By superimposing the vertical displacement fields S_1 (X, Y, Z), S_2 (X, Y, Z), and S_3 (X, Y, Z), the spatial coordinates (X_s, Y_s, Z_s)of the fold structure model were derived **(Equation 4)**.

$$\begin{bmatrix} X_s' \\ Y_s' \\ Z_s' \end{bmatrix} = \begin{bmatrix} X_s \\ Y_s + S_1(X,Y,Z) \\ Z_s \end{bmatrix} \quad (5)$$

In addition, natural folds (X_s', Y_s', Z_s') are often complex, allowing an arbitrary combination of S_1 (*X*, *Y*, *Z*), S_2 (*X*, *Y*, *Z*), and S_3 (*X*, *Y*, *Z*). Vertical variation of fold surfaces can be achieved by incorporating a lateral displacement field **(Equation 5)** or by systematically adjusting the central coordinates of each fold equation.[14]

After generating the formation dip and fold reflectivity model, the next step simulated the structural characteristics of a fault-controlled fractured-cavity system. Field outcrop observations (Figure 2) show that the dominant features are fault-related fractures and near-cylindrical caves, which can be approximated as vertically elongated ellipsoids with major axes perpendicular to the ground surface. To effectively simulate the reflection characteristics of fractured cavities, a 3D, vertically elongated ellipsoid was first constructed, as shown in **Equation 6**.

$$f(p) = (p-c)^T R_1{}^T A R_1 (p-c) \quad (6)$$

The 3D domain of the ellipsoid is defined by **Equation 7**:

$$\begin{cases} if\ f(\mathrm{p}) \le 1 : inside\ chimney\ cube \\ if\ f(\mathrm{p}) > 1 : outside\ chimney\ cube \end{cases}$$

In the ellipsoid function (Equation VI), p = (X, Y, Z) denotes the coordinates of a point in the 3D reflectivity model, whereas c = (c_X, c_Y, c_Z)represents the center of the ellipsoid, randomly sampled within the model domains. Matrix A is a diagonal matrix specified by the ellipsoid's three radii, as shown in **Equation 8**:

$$A = \begin{bmatrix} \frac{1}{r_X^2} & 0 & 0 \\ 0 & \frac{1}{r_Y^2} & 0 \\ 0 & 0 & \frac{1}{r_Z^2} \end{bmatrix} \quad (8)$$

Among these parameters, the values of r_x, r_y, and r_z were randomly selected from predefined ranges [1, 12], [1, 12], and [10, 100], respectively, to construct cavities of varying sizes. To ensure predominantly vertically elongated ellipsoids, the range for r_z was set wider. In addition, constraints $r_x > 0.1r_z$ and $r_Y > 0.1r_z$ were imposed to avoid extremely elongated ellipsoids, which are geologically uncommon. Matrix R_1 **(Equation 9)** is a rotation matrix that rotates the ellipsoid around the X-axis and Y-axis.

$$R_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} \quad (9)$$

Among these, rotation angles α and β were randomly sampled from a narrow range of [−10°, 10°] to generate ellipsoids that are either slightly inclined or nearly vertical. By randomly sampling all parameters in **Equation 10**, numerous ellipsoids with diverse shapes, sizes, orientations, and positions can be created. Nevertheless, because natural fractured cavities are not ideal ellipsoids, random smooth perturbations were introduced to generate irregular ellipsoids that better resemble real-world features.

$$\begin{bmatrix} \bar{X} \\ \bar{Y} \\ \bar{Z} \end{bmatrix} = \begin{bmatrix} X_s + S_k(X,Y,Z) \\ Y_s + S_k(X,Y,Z) \\ Z_s + S_k(X,Y,Z) \end{bmatrix} \quad (10)$$

After defining the 3D tubular region of the fracture cave, the concave structure within the ellipsoid was further characterized. Concave structures typically exhibit a downward-bending morphology, so the reflection layer within the channel body was vertically offset. The vertical offset is defined in **Equation 11**.

$$S_k(X,Y,Z) = \begin{cases} 0, & if\ f(X,Y,Z) > 1 \\ \gamma\left(f(X,Y,Z) - 1\right) + \varepsilon(X,Y,Z), & \\ if\ f(X,Y,Z) \le 1 & \end{cases} \quad (11)$$

In **Equation 11**, $f(X, Y, Z)$ represents the ellipsoid function. γ is a positive scalar randomly selected from the range [10, 20], and $\epsilon(X, Y, Z)$ is a random perturbation field used to simulate fractures or faults within the fractured cavity that may cause displacement of the reflection layer. The perturbation field has a relatively small magnitude compared to the first term γ(f(X, Y, Z)−1). When the perturbation field approaches zero, $S_K(X, Y, Z) \approx \gamma(f(X, Y, Z)-1$, with its value being non-positive and decreasing smoothly from 0 at the fracture-cave boundary to the minimum at the center. At this point, the offset causes a vertical shear of the reflection layer, resulting in a smoothly downward-bending morphology. The resulting concave structure produces distinct annular features on the horizontal slice. If the perturbation is significant, the offset $S_K(X,Y,Z)$ exhibits disordered characteristics, leading to the formation of curved and displaced reflection layers within the channel body, including features such as holes or faults.[2]

After constructing a reflectivity model incorporating fold and fractured cavity features, synthetic seismic data were simulated by convolving the model with Ricker wavelets in the vertical direction.[15] The frequency of the Ricker wavelets was randomly selected within the range of 10–40 Hz. To enhance the realism of the synthetic data and align it more closely with real-world data, random noise was further introduced. Figure 7 presents the corresponding label data, where fractured cavity regions are marked as 1 and non-fractured cavity regions as 0. Using this method, a total of 800 data samples (each with a size of 128 × 128 × 128) were generated for model training. Although the training set was substantially smaller than the actual study area, it was sufficient to achieve accurate identification of the main fractured-cavity features.

2.2.2. Network architecture

The algorithmic structure proposed in this paper is illustrated in Figure 8. Overall, it adopts a fusion architecture that combines CNN and transformer models. The core of the encoder strikes a balance between feature extraction capability and computational efficiency by using the Swin transformer module[16] as its fundamental framework.

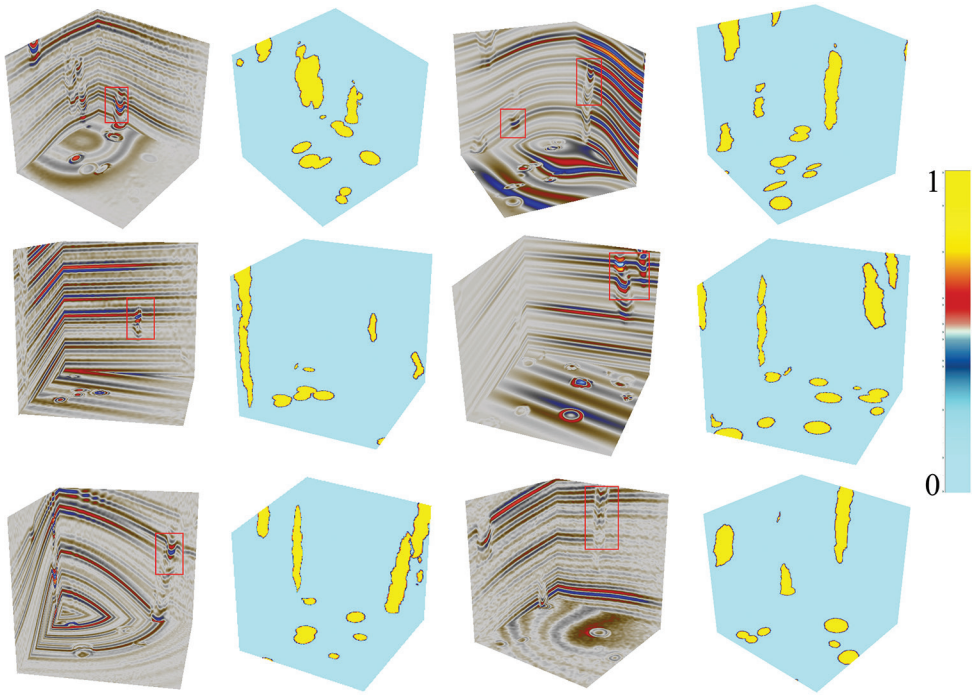

Figure 7. Three-dimensional fractured cavity data were synthesized by convolving the reflection model with the Ricker wavelet

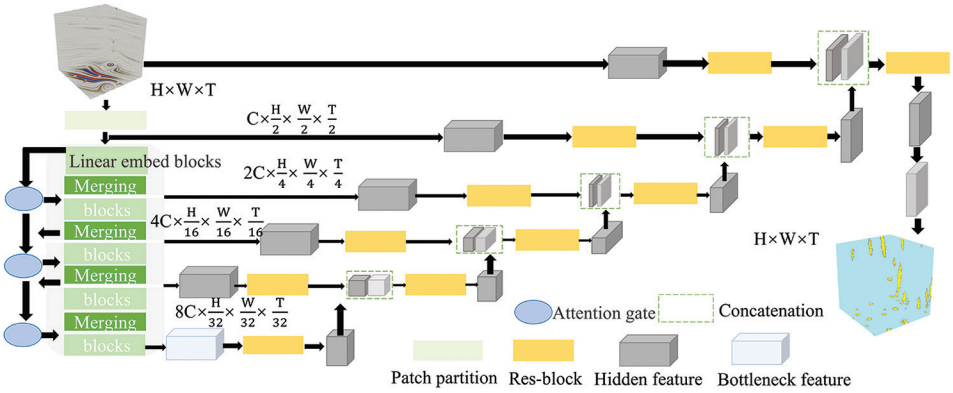

Figure 8. Overview of the network architecture

The self-attention mechanism in the standard transformer operates globally. This leads to a quadratic increase in computational load and memory requirements as the size of input seismic data grows, which limits its application to high-resolution 3D seismic datasets. The Swin transformer restricts attention calculations to local windows of fixed size, significantly reducing computational demands. In addition, it introduces a sliding window mechanism that enables cross-boundary feature interactions between windows, ensuring that the model can capture long-range contextual information. This sliding window strategy reduces the network's processing complexity from cubic to linear in relation to the input data size, greatly enhancing its adaptability to large-scale seismic datasets.

For the decoder, inspired by the UNETR design,[17] multiple convolutional layers with residual connections and transposed convolutional layers are used to upsample the feature maps generated by the encoder.

In the encoder stage, the input seismic data ($x \in R^{H \times W \times D}$) passes through a patch partition layer, where each dimension is divided by a factor ($h \times w \times d$) to form several non-overlapping data blocks of size $\frac{H}{h} \times \frac{W}{w} \times \frac{D}{d}$. Each sub-block corresponded to the seismic reflection characteristics of a local underground region. Next, the linear embedding layer in the Swin transformer stage maps the 3D data into a one-dimensional sequence space of

dimension *C*, where $C = h \times w \times d$. As illustrated in Figure 9, the core of each Swin transformer module consisted of a window-based multi-head self-attention (W-MSA) module and a shifted W-MSA module. The input features first underwent layer normalization and then entered the W-MSA module to extract local intra-window self-attention features. The output features, after layer normalization, were passed through a multi-layer perceptron with a ReLU activation function. Subsequently, the features underwent layer normalization again and entered the shifted W-MSA module, enabling information transfer between local windows and extracting contextual feature information. Residual connections were applied between every two layers of normalization operations. The self-attention mechanism used here was calculated using **Equation 12**:

$$Attention(Q,K,V) = Softmax\left(\frac{QK^T}{\sqrt{d}}\right)V \tag{12}$$

Among them, Q, K, and V represent the query, key, and value, respectively; d denotes the size of the query and key vectors.

The calculation formula for the l-th layer of the Swin transformer module is as follows **(Equations 13-16)**:

$$z_1^l = W - MSA\left(LN\left((z_0^{l-1})\right)\right) + z_0^{l-1} \tag{13}$$

$$z_0^1 = MLP\left(LN\left(z_1^l\right)\right) + z_1^l \tag{14}$$

$$z_1^{l+1} = SW - MSA\left(LN\left(z_0^1\right)\right) + z_0^l \tag{15}$$

$$z_0^{l+1} = MLP\left(LN\left(z_1^{l+1}\right)\right) + z_1^{l+1} \tag{16}$$

To automatically focus on fracture bodies of various shapes and sizes, an attention gate mechanism[18] was introduced, as shown in Figure 10. The feature map x^lrepresents the features to be processed in the current module, whereas the gating signal g is the high-level feature from the previous layer. The attention gate module takes both the feature map x^l and the gating signal g as inputs, applies 1×1×1 convolutions to each to unify their dimensions, and then performs element-wise addition. After activating the combined features, the weights corresponding to areas of fractured cavity development increased through iterative optimization, thereby enhancing the model's prediction accuracy.

The decoder is a CNN-based neural network that connects to the encoder via multi-resolution features through skip connections, forming a U-shaped architecture. The decoder's decision followed the Swin-UNETR model,[19] as illustrated in Figure 8. At each decoder stage, the input features first pass through a residual module composed of two 3 × 3 × 3 convolutional layers with instance normalization, which processes the features. These processed features were then concatenated with the upsampled features from the subsequent stage. After concatenation, the features were reshaped. The features processed at each stage were upsampled using a deconvolution layer; when the spatial dimensions doubled, they were concatenated with the corresponding features from the encoder at the same stage. The concatenated features were further processed through a residual module.

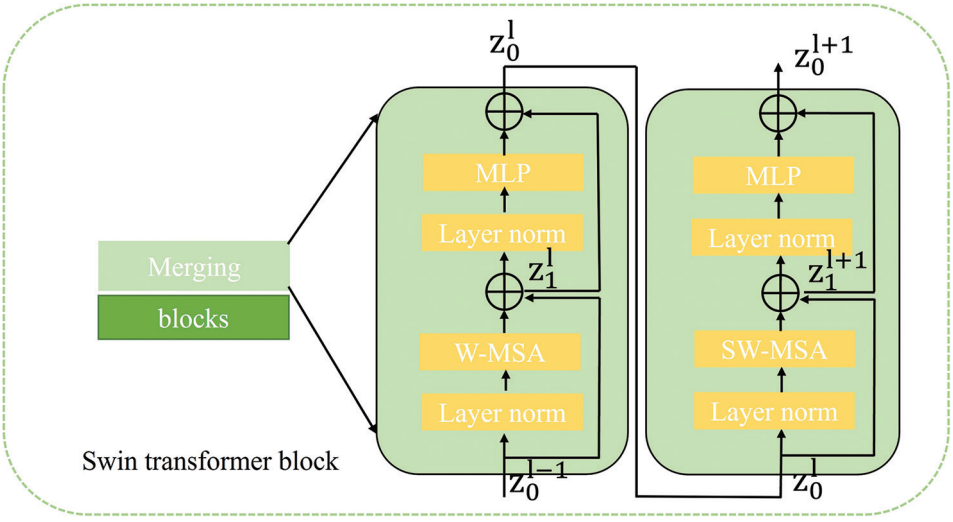

Figure 9. The structure of the Swin transformer module
Abbreviations: MLP: Multi-layer perceptron; SW-MSA: Shifted window-based multi-head self-attention; W-MSA: Window-based multi-head self-attention.

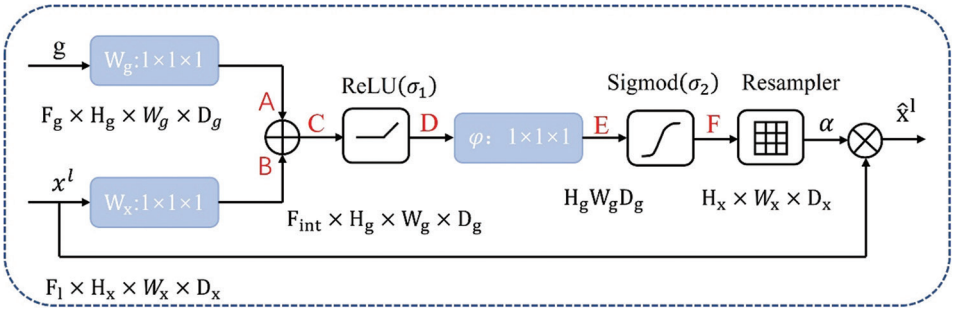

Figure 10. The structure of the attention gate module

Finally, a residual module serves as the segmentation head; after a 1×1×1 convolutional layer and a sigmoid activation function, the fractured-cavity segmentation map is produced.

2.3. Seismic attribute compression and fusion based on principal component analysis

To use multi-attribute seismic data to constrain the neural-network segmentation results, we preserved the correlations between different attributes (i.e., overlapping fractured-cavity information), eliminated redundant information, and introduced new variables (principal components) that comprehensively summarize fractured-cavity features across all original attributes. To extract the dominant features of fractured cavities via attribute compression, this study utilized principal component analysis. This algorithm independently assigns weights to the calculation factors of various seismic attributes, distributing weights according to the similarities and differences among attributes, thus enabling effective fusion of diverse seismic attributes.

For the analysis of a specific seismic attribute, a seismic attribute observation data matrix was constructed based on its characteristic features. Suppose the number of samples of the seismic attribute data is *n,* and the number of seismic attribute features is *p*, then the seismic attribute observation data matrix is **(Equation 17)**:

$$X = \begin{bmatrix} X_{11} & \cdots & X_{1P} \\ \vdots & \ddots & \vdots \\ X_{n1} & \cdots & X_{nP} \end{bmatrix} \tag{17}$$

Since the numerical ranges and dimensions of different seismic attribute data vary, it is necessary to convert multiple seismic attribute datasets to a common observation range; in other words, attribute standardization must be performed, as shown in **Equation 17**.

$$X = \frac{X_{ij}}{X_{1j}}\left(i = 1,2,\ldots n; j = 1,2,\ldots n, p\right) \tag{18}$$

The standardized seismic attribute data matrix X was then analyzed, and its correlation matrix R_2 was calculated. The matrix R_2 is symmetric, and the correlation coefficient after standardization is simplified as follows **(Equation 19)**:

$$R_2 = \begin{bmatrix} r_{11} & \cdots & r_{1p} \\ \vdots & \ddots & \vdots \\ r_{n1} & \cdots & r_{np} \end{bmatrix} \tag{19}$$

The seismic attribute features were obtained by calculating the correlation coefficient matrix R_2, where $\overline{x_i}$ and $\overline{x_j}$ are the means of the i-th and j-th attributes and are denoted as $r_{ij} = \frac{1}{n-1}\sum_{i=1}^{n}(x_{li} - \overline{x_i})(x_{li} - \overline{x_j})\left(i, j = 1,2,\ldots p\right)$, respectively. The characteristic equation $|R_2 - \lambda I| = 0$ was considered, and *k* principal components were selected. The variance contribution rate of the k-th principal component is given by β_k, and the proportion of total variance explained by the first *k* principal components is expressed by the cumulative contribution rate *M* **(Equation 20)**:

$$\beta_k = \left(\sum_{i=1}^{p} i\right)^{-1} M = \sum_{i=1}^{k} \lambda_k \left(\sum_{i=1}^{p} i\right)^{-1} \tag{20}$$

In practice, the cumulative contribution rate *M* of the features was chosen to exceed 90%.[20]

2.4. Experiments and analysis

Figure 11 illustrates the flowchart of the fractured-cavity identification and characterization methodology adopted in this study. All code utilized in this study was developed using Python version 3.8. The computational environment consisted of a Windows 10 operating system and an NVIDIA GeForce RTX 4080 Ti GPU with 16 GB of memory.

Model training was conducted on a dataset composed of 800 input samples, each with dimensions of 128 × 128 × 128, processed in batches of four. The training procedure spanned 25 epochs.

The key hyperparameters of Swin-UNETR were set as follows: an initial learning rate of 0.0001 was adopted, the Adam optimizer was selected for parameter updating, and the optimization objective was defined as a weighted combination of Dice loss and cross-entropy loss. The evaluation metrics included loss, Dice coefficient, and intersection over union, with the training and validation results presented in Figure 12.

3. Results and discussion

A significant advantage of data-driven models lies in their ability to generalize. In this study, we applied our methodology to the designated study area and systematically characterized the segmental and hierarchical features of the identified fractured cavities on both horizontal planes and vertical profiles.

Within the study region, the Ordovician system comprises approximately 2,000 m of mudstone and sandstone, whereas the stratigraphic sequence from the Middle Ordovician to the Cambrian encompasses over 3,000 m of marine carbonate formations. Multiple seismic reflection horizons have been delineated by geologists within these strata. Notably, the Ordovician carbonate reservoirs are regarded as the primary targets for exploration in the central Tarim Basin.

Guided by the geometric and kinematic frameworks of strike–slip faulting, which is predominantly characterized by horizontal displacement, our analysis focuses on specific horizons: H_7^4 (demarcating the boundary between the Middle and Upper Ordovician), H_7^6, H_7^8, H_8^0 (the interface

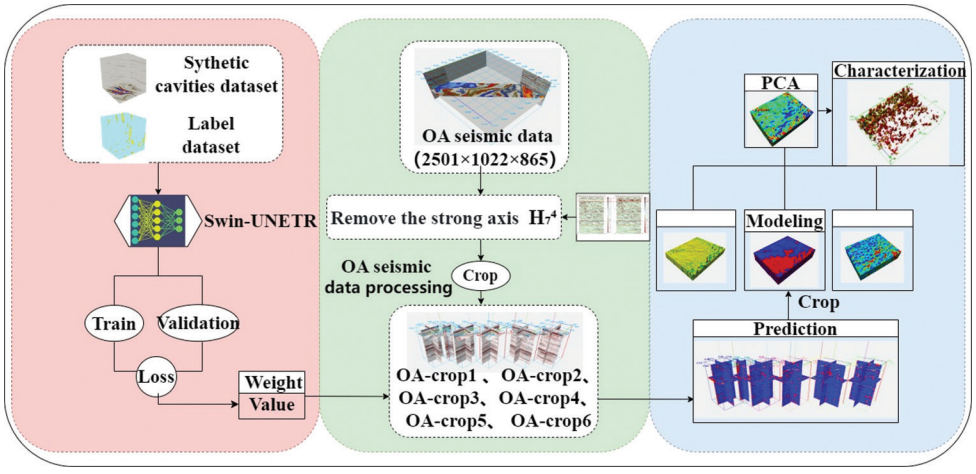

Figure 11. The flowchart of the fractured cavities identification and characterization method
Abbreviations: OA: Object area; PCA: Principal component analysis.

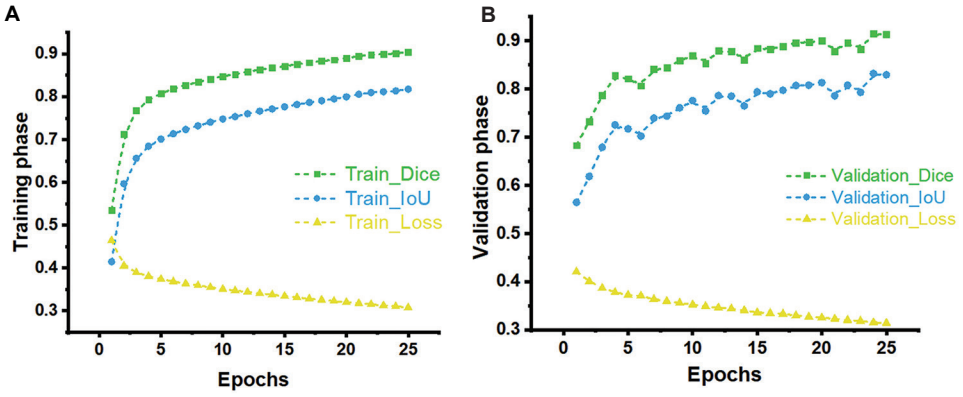

Figure 12. Evaluation metric curves. (A) Training and (B) validation phases.
Abbreviation: IoU: Intersection over Union.

between the Lower Ordovician and Upper Cambrian), H_8^1 (separating the Upper and Lower Cambrian), and H_9^0 (the basal Cambrian horizon). These horizons served as the basis for presenting the identification outcomes of fractured cavities within horizontal stratigraphic levels.

Within the study area, fractured cavities are primarily controlled by the strike–slip fault zone (OA) rather than dissolution processes. To clarify this control mechanism, scanning coherence detection was first performed on the post-stack seismic data of the study area to obtain the seismic dip gather. Subsequently, elastic impedance inversion was conducted based on this dip gather, generating key petrophysical parameters including elastic modulus and Poisson's ratio. Finally, the maximum principal stress field was derived through systematic calculations. Figure 13 presents the maximum principal stress field integrated with geological data. The fault zone can be categorized into three main structural segments: Compression segments, pull-apart segments, and translation segments, which alternate sequentially along the fault strike. Figures 14-19 illustrate that strata in compression segments are uplifted by stress. In these areas, independent, fractured cavities with circular outlines frequently develop at stress-concentrated locations, often referred to as cave-like features. When compressive and strike–slip stresses are superimposed, such cave-like anomalies are widely distributed.

In pull-apart segments, pull-apart and strike–slip stresses act synergistically. Regions subjected to strong

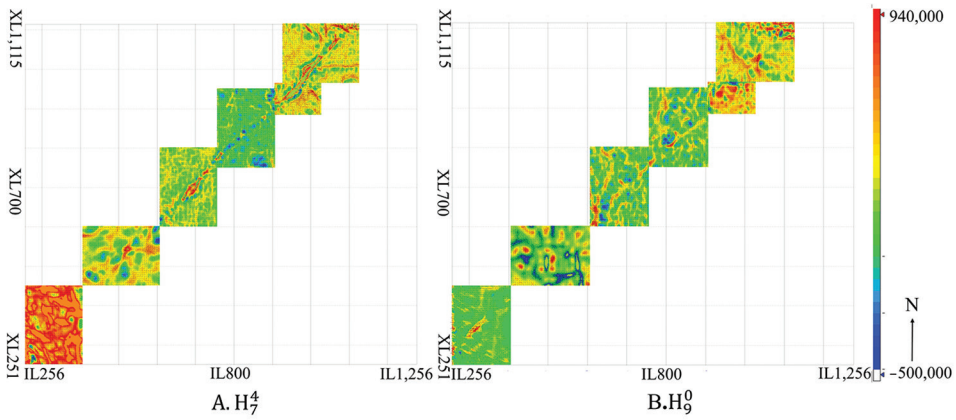

Figure 13. Maximum principal stress plane maps of the H_7^4 and H_9^0 horizons in the object area

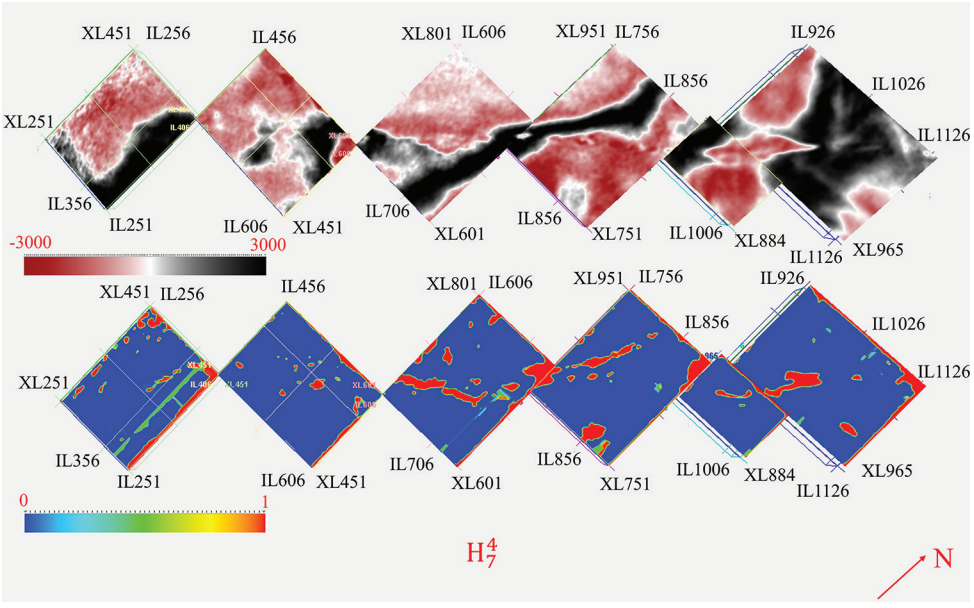

Figure 14. Original seismic data and fractured cavities identification results of the H_7^4 horizon in the object area

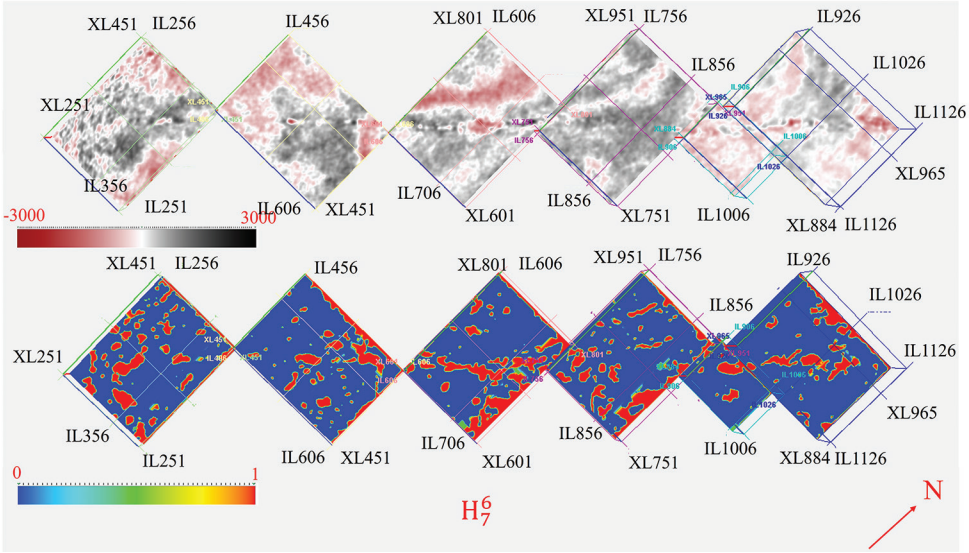

Figure 15. Original seismic data and fractured cavities identification results of the H_7^6 horizon in the object area

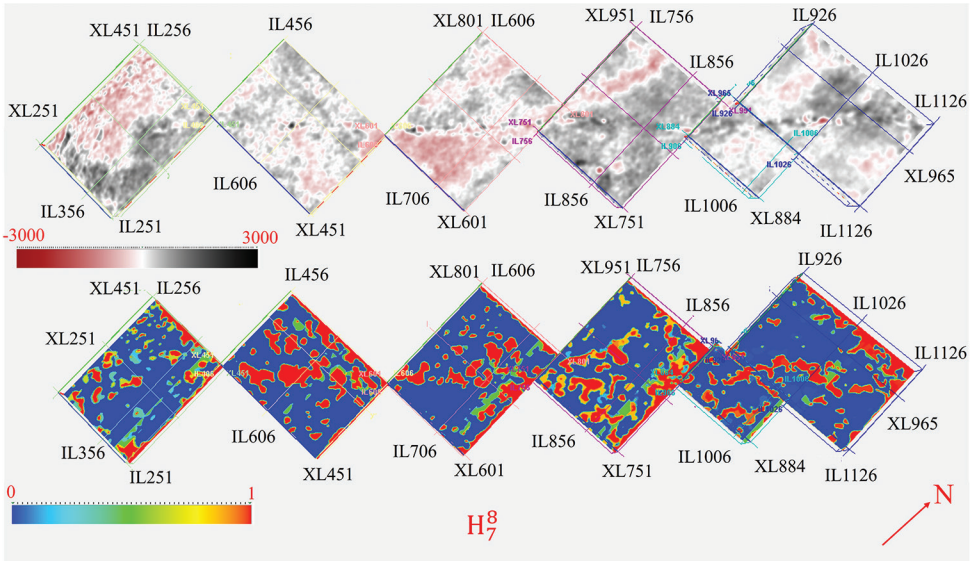

Figure 16. Original seismic data and fractured cavities identification results of the H_7^8 horizon in the object area

pull-apart stress host elongated fractured cavities termed fault-related bodies that are aligned with the fault trend. Adjacent cave-like features near these fault-related bodies may merge at their boundaries, forming larger and more irregular cavity systems. In translation segments, only strike–slip stress is present. Here, fractured cavities with regular morphologies develop along the fault, whereas cave-like features are scarce.

When the stress regime transitions from pull-apart to compression, multiple fault-related bodies typically form. Under compressive stress, these bodies converge toward the central region. Cave-like features often develop at the edges of compression zones, with their cores either contacting fault-related bodies or lying between several such bodies. When the stress regime shifts from translation to compression, a single fault-related body generally emerges within the uplifted strata. This body penetrates the core of the cave-like feature but does not connect its opposite sides.

Furthermore, the fractured cavities exhibited a hierarchical structure in the vertical profile of the study area. As shown in Figures 14-19 , the density and

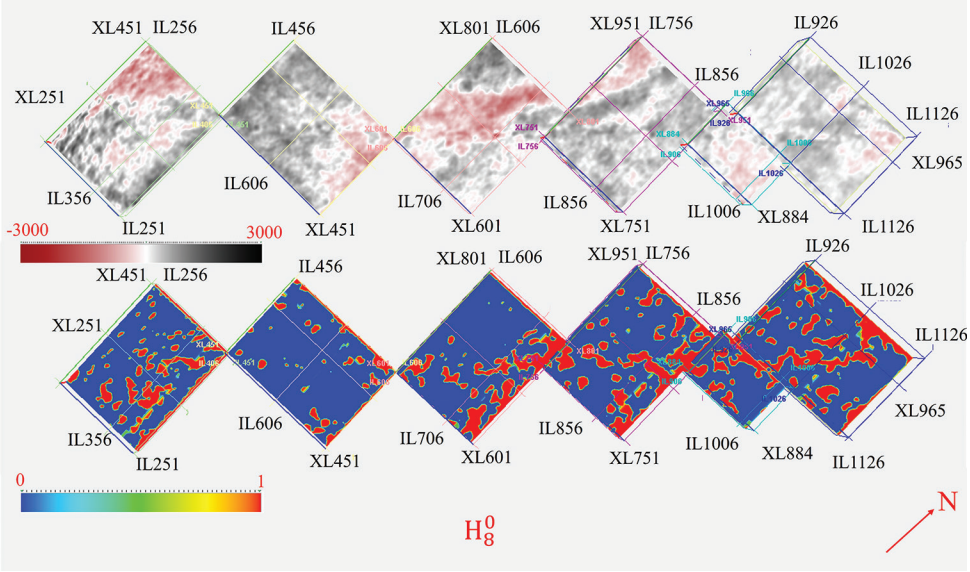

Figure 17. Original seismic data and fractured cavities identification results of the H_8^0 horizon in the object area

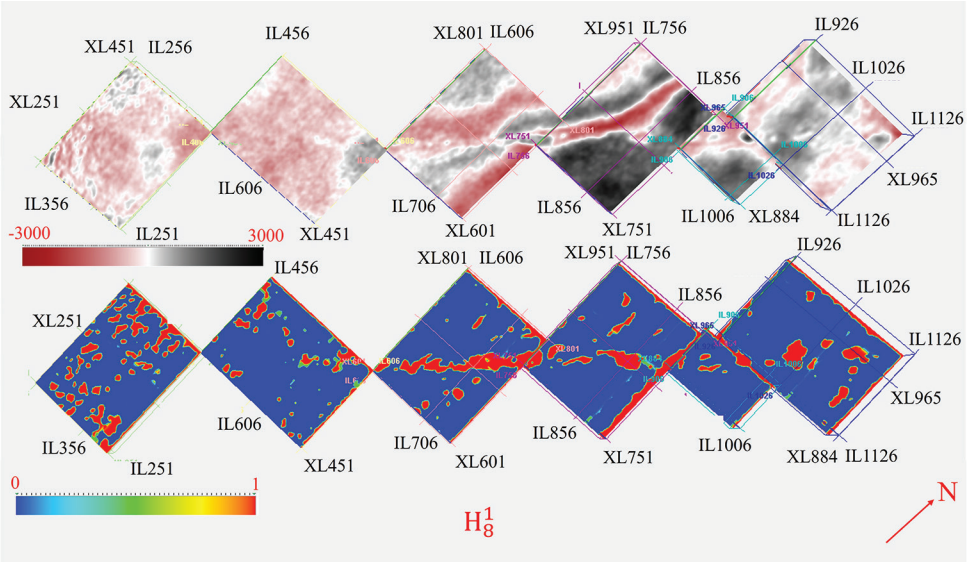

Figure 18. Original seismic data and fractured cavities identification results of the H_8^1 horizon in the object area

complexity of fractured cavity development gradually increased from stratigraphic levels H_7^4–H_7^8, whereas transitioning from dense to sparse from H_7^8 to H_9^0. This variation is presumably linked to differential dissolution sensitivity caused by lithological heterogeneity in the carbonate strata, such as the interbedding of dolomite and limestone. Figure 20 further demonstrates that fractured cavities developed between H_7^4 and H_7^8 can be distinctly categorized from those between H_7^8 and H_9^0, based on differences in their structural characteristics and abundance.

The first type of fractured cavities is characterized by vertical elongation in the profile, manifesting as steep and upright structures. Larger cavities tend to exhibit a roughly funnel-shaped morphology, whereas smaller ones display nearly circular outlines. Extensive fault-related fractured cavities formed within the strike–slip fault zone are almost continuous, extending from shallow to deep intervals. In addition, smaller, cave-like fractures occur on both sides of the fault, extending through the entire target stratum. Conversely, the second type of fractured cavities develops predominantly in a near-horizontal

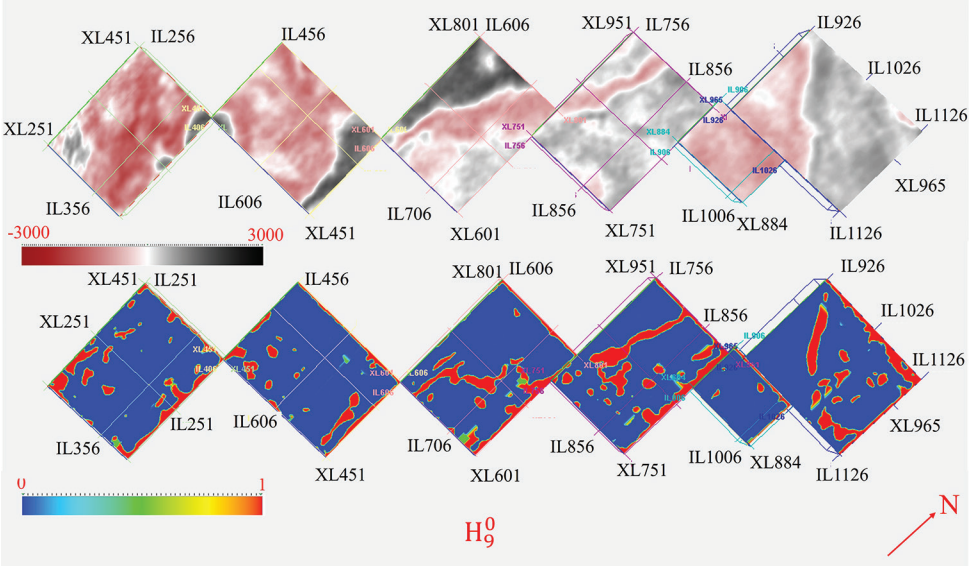

Figure 19. Original seismic data and fractured cavities identification results of the H_9^0 horizon in the object area

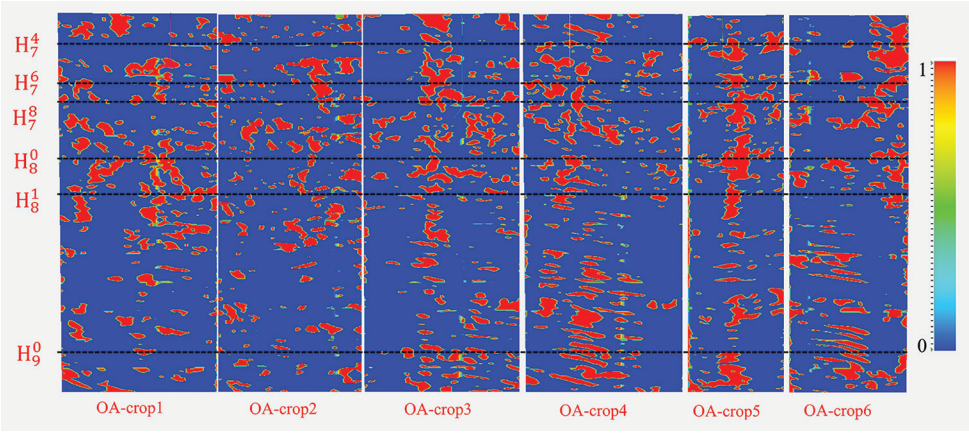

Figure 20. Fractured-cavity identification profile of the object area (OA)

orientation, exhibiting a flattened morphology. These cavities are distributed at regular intervals, have smaller sizes than the first type, and are less extensively developed. Nonetheless, they also display significant vertical connectivity within the fault zone.

To determine the outline of fault-controlled bodies, we employed a method that fuses different seismic attribute volumes after threshold truncation. The post-stack seismic data OA-crop1 is segmented into three parts: OA-crop1-1 (H_7^4–H_7^6), OA-crop1-2 (H_7^6–H_7^8), and OA-crop1-3 (H_7^8–H_8^0). Amplitude curvature and coherence detection attributes were extracted from these segments. Following outlier processing and standardization of the selected dominant attributes, attribute compression was performed using principal component analysis. Finally, the amplitude curvature, coherence detection attributes, and image segmentation results were merged.

Figures 21-23 show, from left to right, the network-based segmentation results, the calculated amplitude curvature and coherence attributes, and the fused attribute result. The figures demonstrate that the fused attribute volume integrates all the high-value features from the first three attribute volumes. A color-based comparison reveals that the fused attribute volume exhibits enhanced responses at locations where fractured cavities are consistently identified by all three attribute volumes. In regions where discrepancies occurred among the three attribute volumes, the fractured cavity features from the first volume were assigned greater weight. This fusion procedure combined the three attribute volumes into a single, integrated

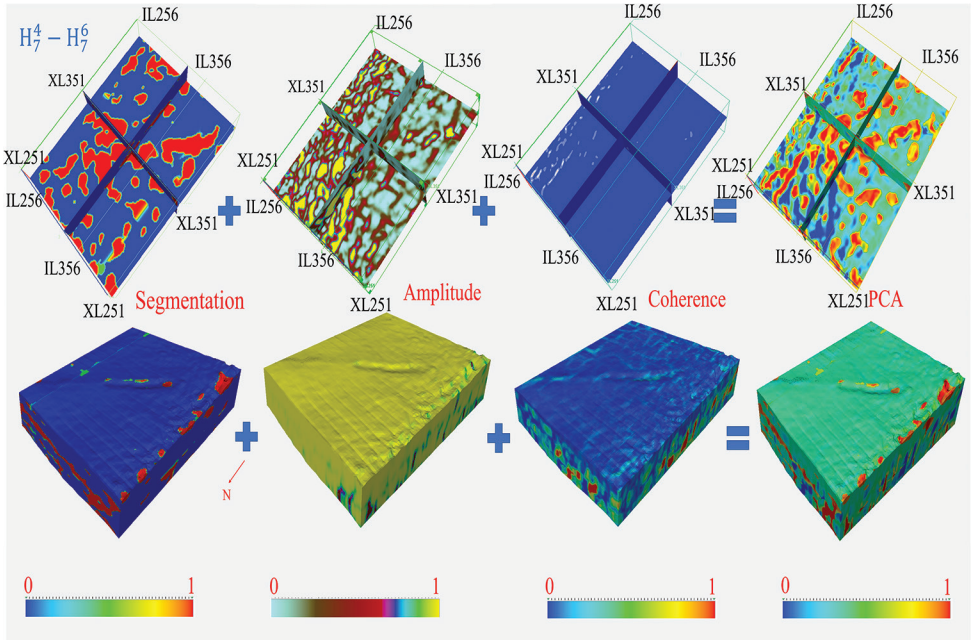

Figure 21. Seismic attribute fusion result map of the OA-crop1 H_7^4–H_7^6 horizon in the object area
Abbreviations: OA: Object area; PCA: Principal component analysis.

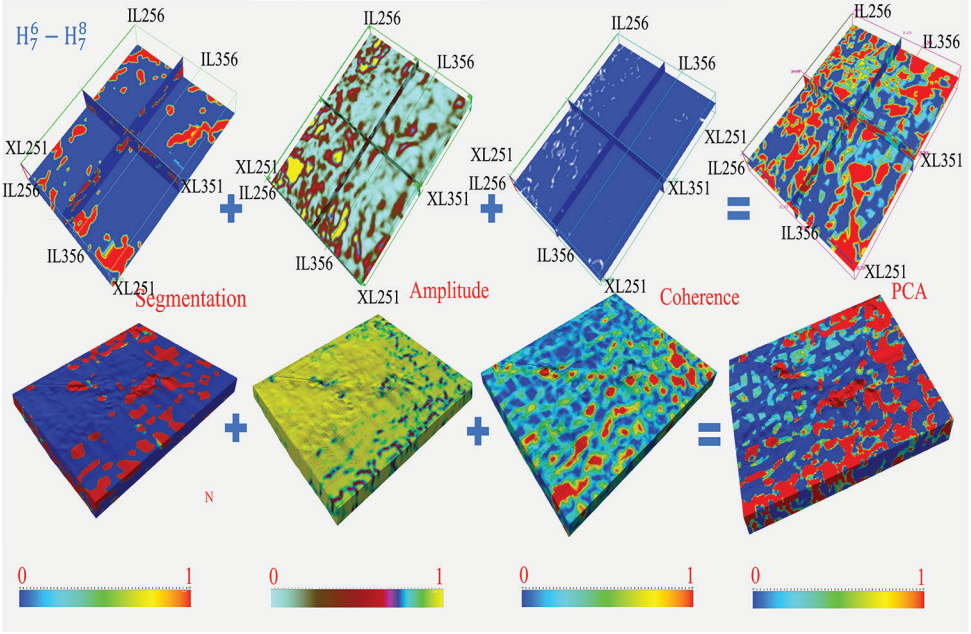

Figure 22. Seismic attribute fusion result map of the OA-crop1 H_7^6–H_7^8 horizon in the object area
Abbreviations: OA: Object area; PCA: Principal component analysis.

volume, effectively reducing the uncertainty associated with fractured cavities.

The fractured cavities were then modeled and characterized using a 4 × 4 × 1 grid in the local object area, where the color gradient from light to dark represents an increasing probability of fractured-cavity occurrence.[21] Figures 24-26 show the final fractured cavity body results.

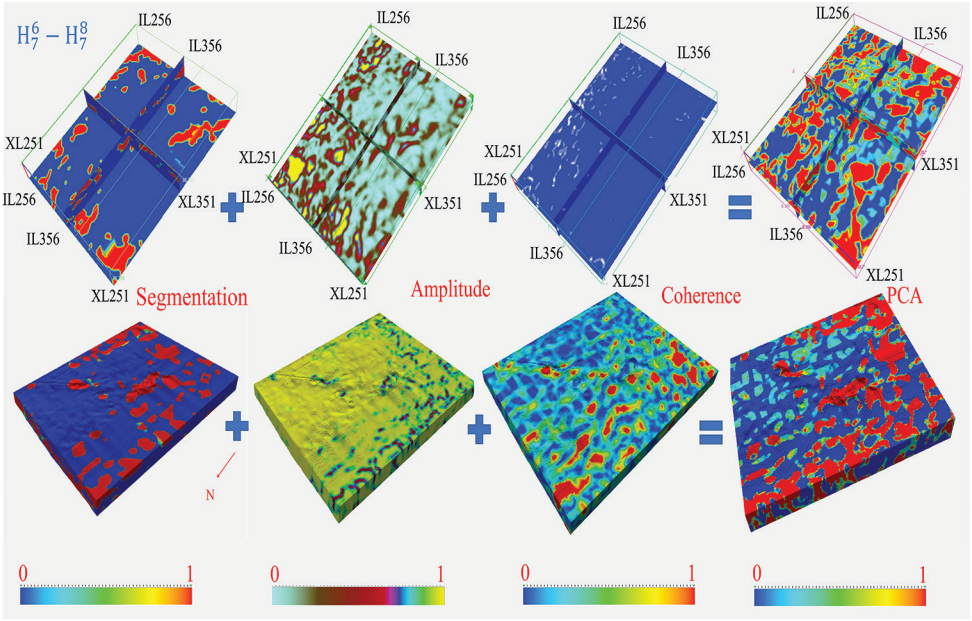

Figure 23. Seismic attribute fusion result map of the OA-crop1 H_7^8–H_8^0 horizon in the object area
Abbreviations: OA: Object area; PCA: Principal component analysis.

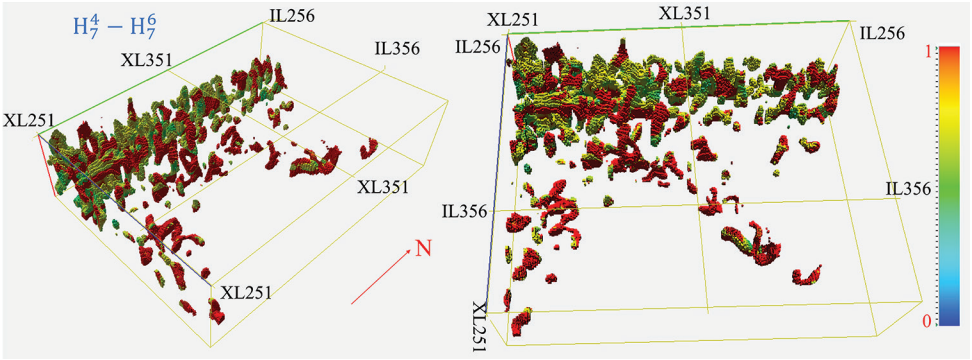

Figure 24. Fractured-cavity body map for the H_7^4–H_7^6 horizon in object area (OA)-crop1 of the OA

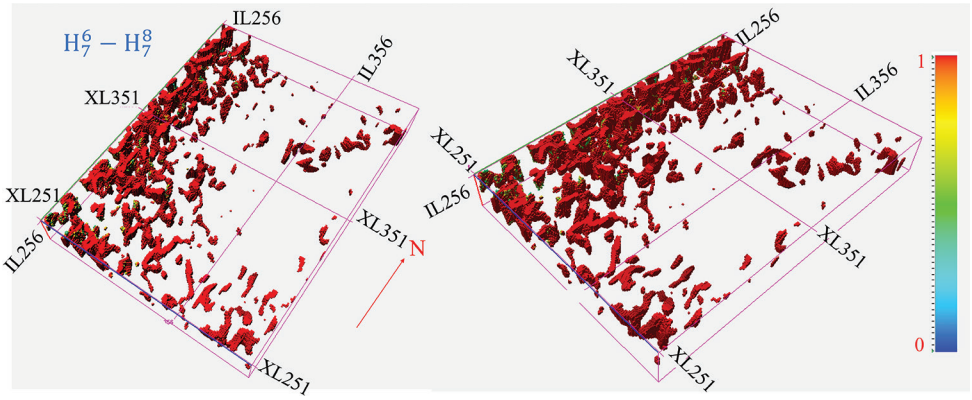

Figure 25. Fractured-cavity body map for the H_7^6–H_7^8 horizon in object area (OA)-crop1 of the OA

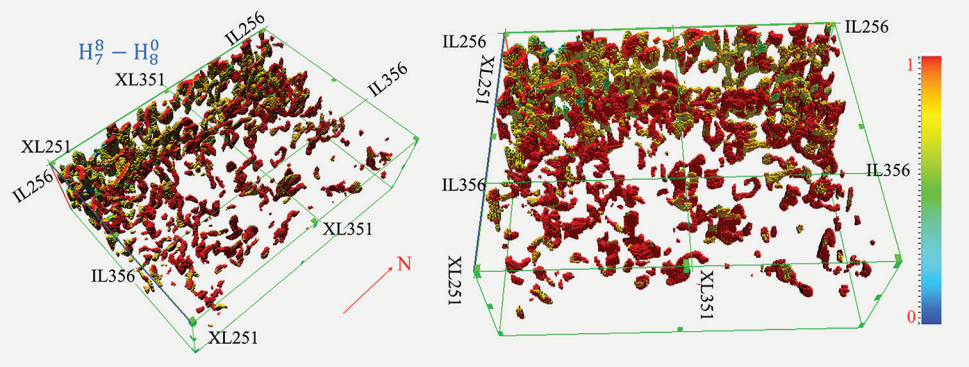

Figure 26. Fractured-cavity body map for the H_7^8–H_8^0 horizon in object area (OA)-crop1 of the OA

4. Conclusion

Fractured cavities in reservoirs within carbonate formations have complex geological features. These features make it hard to precisely locate and map fracture caves using conventional methods. This study introduces a novel two-step methodology to identify and delineate fractured-cavity reservoirs in response to these challenges.

First, the Swin-UNETR architecture was employed to automatically detect reflective signatures of fractured cavity reservoir strata in the Shunbei field. An attention gate mechanism was incorporated into the Swin-UNETR model, enhancing its capability to interpret the dense and complex features associated with fractured cavity labels in 3D seismic data. Then, we applied a feature extraction strategy that fuses multiple seismic attributes using principal component analysis. This approach reduced uncertainty, leading to more accurate and robust identification results.

Application of this methodology to Shunbei field data has demonstrated its accuracy in recognizing the characteristic beaded reflections in seismic records. The identified fractured-cavity reservoirs exhibit both planar and vertical spatial attributes, revealing segmented patterns in plain view. Fractured cavity development varied significantly across different tectonic segments. In compressional zones, stress-induced strata uplift resulted in nearly circular, isolated fractured cavities in areas of high stress concentration. In regions where compressive and strike–slip stresses interact, extensive irregular cave-like anomalies were developed. Pull-apart segments, influenced by both pull-apart and strike–slip stresses, hosted elongated fractured cavities that were aligned with the strike–slip direction. These cavities often had adjacent cave-like features whose edges interfaced with fault-related bodies, resulting in irregularly expanded fractured cavities. In translational segments that are subjected solely to strike–slip stress, fault-related fractured cavities exhibited regular edges aligned with the strike–slip direction, accompanied by sparse surrounding cave-like features.

Analysis of the OA profile revealed a stratified structural framework. The first category of fractured cavities extends vertically in profile and is characterized by steep, upright morphologies. Large-volume fractured cavities display roughly funnel-shaped geometries, whereas smaller cavities tend to be nearly circular. Within the strike–slip fault zone, substantial fault-related fracture caves demonstrated near-continuity across both shallow and deep strata, flanked by smaller cave-like features that penetrated the entire target formation. The second category predominantly developed in a near-horizontal direction, presenting flatter morphologies with discrete intervals between individual cavities. These features had smaller outlines and were more spatially constrained than the first category. Nevertheless, they exhibited strong vertical connectivity both within and beneath the fault zone. Finally, fractured cavities were delineated with a fixed grid size to standardize their spatial representation.

This integrated methodological framework provides a robust tool for the detailed characterization of fractured cavity reservoirs, thereby facilitating a deeper understanding of their spatial distribution and geological evolution within complex carbonate strata.

Acknowledgments

None.

Funding

This work was supported by the National Natural Science Foundation of China (Grant no. 42274180) and the Graduate Innovation Program of China University of Mining and Technology (Grant no. 2025WLJCRCZL003).

Conflict of interest

Qiang Guo is an Editorial Board Member of this journal, but was not in any way involved in the editorial and peer-review process conducted for this paper, directly or indirectly. The authors declare they have no competing interests.

Author contributions

Conceptualization: Yunhao Cui, Chao Xu
Formal analysis: Zhiqiang Lu
Investigation: Yunhao Cui, Chao Xu, Zhanpeng Chen, Yuwen Qian
Methodology: Yunhao Cui, Yaping Huang, Qiang Guo
Writing–original draft: Yunhao Cui
Writing–review & editing: Yuhua Chen

Availability of data

Data will be made available upon request to the corresponding author.

References

1. Ma Y, Cai X, Yun L, *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. *Pet Explor Dev*. 2022;49(1):1-20.

   doi: 10.1016/s1876-3804(22)60001-6

2. Wu X, Yan S, Qi J, Zeng H. Deep learning for characterizing paleokarst collapse features in 3-D seismic images. *JGR Solid Earth*. 2020;125(9):e2020JB019685.

   doi: 10.1029/2020jb019685

3. Li Z, Li H, Liu J, Deng G, Gu H, Yan Z. 3D seismic intelligent prediction of fault-controlled fractured-vuggy reservoirs in carbonate reservoirs based on a deep learning method. *J Geophys Eng*. 2024;21(2):345-358.

   doi: 10.1093/jge/gxae005

4. Gui Z, Zhang J, Zhang Y, Sun C. Characterization of fault-karst reservoirs based on deep learning and attribute fusion. *Acta Geophys*. 2024;73(2):1335-1347.

   doi: 10.1007/s11600-024-01420-5

5. Zhang G, Lin C, Ren L, *et al*. Seismic characterization of deeply buried paleocaves based on Bayesian deep learning. *J Nat Gas Sci Eng*. 2022;97:104340.

   doi: 10.1016/j.jngse.2021.104340

6. Vaswani A, Shazeer N, Parmar N, *et al*. *Attention is All You Need*. [arXiv Preprint]; 2023.

   doi: 10.48550/arXiv.1706.03762

7. Chen L, Liu X, Zhou H, Zhang H, Lyu F, Mo Q. Identification of carbonate cave reservoirs based on variational bayesian principal component analysis. *IEEE Trans Geosci Remote Sens*. 2023;61:1-10.

   doi: 10.1109/tgrs.2023.3333962

8. Li B, He Y, Chen W, Shang H, Wang L. Geological modeling of carbonate fracture-cavity reservoir: Case study of Shunbei fault zone No. 5. *Front Earth Sci*. 2025;13:1559030.

   doi: 10.3389/feart.2025.1559030

9. Wang S, Si X, Cai Z, Sun L, Wang W, Jiang Z. Fast global self-attention for seismic image fault identification. *IEEE Trans Geosci Remote Sens*. 2024;62:1-11.

   doi: 10.1109/tgrs.2024.3436066

10. Ma Q, Duan T. Multi-level ultra-deep fault-controlled karst reservoirs characterization methods for the Shunbei field. *Front Earth Sci*. 2023;11:1149678.

    doi: 10.3389/feart.2023.1149678

11. Zheng D, Zhang L, Shen F. Structural heterogeneities in fault-controlled cavity carbonate reservoirs: Insights into fracture and fault patterns. *J Struct Geol*. 2025;193:105358.

    doi: 10.1016/j.jsg.2025.105358

12. Li H, Lin C, Ren L, Zhang G, Chang L, Dong C. Quantitative prediction of multi-period tectonic fractures based on integrated geological-geophysical and geomechanics data in deep carbonate reservoirs of Halahatang oilfield in northern Tarim Basin. *Mar Pet Geol*. 2021;134:105377.

    doi: 10.1016/j.marpetgeo.2021.105377

13. Ding Y, Du Q, Fu L, Jian S. Reliability analysis of seismic attribute in the detection of fault-karst. *Interpretation*. 2020;8(4):SP61-SP70.

    doi: 10.1190/int-2019-0299.1

14. Wang S, Cai Z, Si X, Cui Y. A three-dimensional geological structure modeling framework and its application in machine learning. *Math Geosci*. 2023;55(2):163-200.

    doi: 10.1007/s11004-022-10027-9

15. Wu X, Liang L, Shi Y, Fomel S. FaultSeg3D: Using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation. *Geophys*. 2019;84(3):IM35-IM45.

    doi: 10.1190/geo2018-0646.1

16. Liu Z, Lin Y, Cao Y, *et al*. *Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows*. [arXiv Preprint]; 2021. p. 10012-10022.

    doi: 10.48550/arXiv.2103.14030

17. Hatamizadeh A, Tang Y, Nath V, *et al*. *UNETR: Transformers for 3D Medical Image Segmentation*. [arXiv Preprint]; 2021. p. 574-584.

doi: 10.48550/arXiv.2103.10504

18. Oktay O, Schlemper J, Folgoc LL, *et al*. *Attention U-Net: Learning Where to Look for the Pancreas*. [arXiv Preprint]; 2018.

    doi: 10.48550/arXiv.1804.03999

19. Hatamizadeh A, Nath V, Tang Y, Yang D, Roth H, Xu D. *Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images*. [arXiv Preprint]; 2022.

    doi: 10.48550/arXiv.2201.01266

20. Mishra S, Sarkar U, Taraphder S, *et al*. Multivariate statistical data analysis- principal component analysis. *Int J Livest Res*. 2017;7:60-78.

    doi: 10.5455/ijlr.20170415115235

21. Song S, Mukerji T, Hou J, Zhang D, Lyu X. GANSim-3D for conditional geomodeling: Theory and field application. *Water Resour Res*. 2022;58(7):e2021WR031865.

    doi: 10.1029/2021wr031865