Petroleum Science 23 (2026) 3834–3853

Contents lists available at ScienceDirect

Petroleum Science

journal homepage: www.keaipublishing.com/en/journals/petroleum-science

Original Paper

Multivariate data-driven fracture identification and distribution pattern in tight sandstone reservoirs using an improved CNN-Attention-BiLSTM: A case study of the Permian Lower Shihezi Formation in the Hangjinqi area, Ordos Basin, China

Bao-Yu Liang [a,b,c], Lian-Bo Zeng [a,b,c,*], Shao-Qun Dong [c,d], Xue-Qun Tan [e], Ji-Bo Ren [b,c], Hong-Tao Li [e], Zi-Yi Yang [b,c], Shi-Qiang Liu [b,c], Zhen Wang [b,c]

[a] Hainan Institute of China University of Petroleum (Beijing), Sanya, 572024, Hainan, China
[b] College of Geosciences, China University of Petroleum (Beijing), Beijing, 102249, China
[c] State Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum (Beijing), Beijing, 102249, China
[d] College of Science, China University of Petroleum (Beijing), Beijing, 102249, China
[e] Petroleum Exploration and Production Research Institute, SINOPEC, Beijing, 102206, China

ARTICLE INFO

Article history:
Received 25 October 2025
Received in revised form
6 March 2026
Accepted 6 May 2026
Available online 9 May 2026

Edited by Xiu-Fang Hu

Keywords:
Fracture identification
Tight reservoirs
Full waveform sonic logs
Conventional logs
Deep learning

ABSTRACT

Natural fractures in tight sandstone reservoirs play an important role in hydrocarbon migration and accumulation. Fracture identification remains challenging due to the scarcity of labeled data and the complex logging responses of fractures. To address these problems, we propose a novel hybrid deep learning framework (CNN-Attention-BiLSTM) for labeled data balancing. First, labeled fracture classification based on full waveform sonic logs (FWS) characteristics is employed to screen unlabeled data, replacing sampling algorithms for data balancing. This approach provides conventional logging with more fracture labels that align with authentic geological information, thereby enhancing the reliability of fracture labels. Subsequently, one-dimensional convolution is applied to construct multi-dimensional fracture logging response patterns that characterize fracture development. A Channel Self-Attention (CSA) mechanism is introduced to assign optimal weights to response patterns across different dimensions, achieving an optimized pattern combination and thereby offering clearer response pattern guidance for subsequent identification models. A double-layer BiLSTM (DL-BiLSTM) is then utilized to mitigate the impact of sedimentary cycles on logging identification, while capturing both short- and long-term dependencies of fracture responses across different network layers. Ultimately, intelligent fracture identification is realized. The identification method is applied to the H1 member of the Lower Shihezi Formation in the Hangjinqi area, China. The test set accuracy is higher than 90%, and blind wells verification demonstrates an improvement of over 8% in accuracy compared to conventional methods. The identification results reveal that fractures are the most developed in H1-2 interval, followed by H1-1 and H1-3 intervals, while H1-4 interval is the least developed. The fracture distribution pattern is evidently controlled by both sedimentary rhythms and reservoir properties, resulting in complex storage and flow capabilities. The findings can provide guidance for the migration, accumulation and efficient development of tight sandstone gas.

© 2026 Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The increasing demand for oil and gas and the steady transformation of the global energy structure have made unconventional oil and gas increasingly prominent in the field of petroleum geology. Among them, tight sandstone gas has emerged as one of the most important unconventional hydrocarbon resources,

* Corresponding author.
E-mail address: lbzeng@cup.edu.cn (L.-B. Zeng).
Peer review under the responsibility of China University of Petroleum (Beijing).

https://doi.org/10.1016/j.petsci.2026.05.004
1995-8226/© 2026 Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

demonstrating significant potential for succession (Zou et al., 2018; Sun et al., 2019). The Hangjinqi area in the northern Ordos Basin is an important block of tight sandstone gas in the Ordos Basin, and the H1 member of the Permian Lower Shihezi Formation is one of the primary gas-rich layers. Previous studies indicate that the tight sandstone reservoirs in this formation exhibit complex pore structures and significant heterogeneity (Wang et al., 2020; Qin et al., 2022a). As effective storage spaces and main seepage channels of tight sandstone, natural fractures are crucial for gas migration, accumulation and subsequent fracturing stimulation (Wang et al., 2025). Fracture identification is extremely critical to clarify the fracture development characteristics and analyze the vertical distribution. The information it reflects can provide critical guidance for predicting the distribution of pore-fracture systems and establishing development patterns (Liu et al., 2018).

Core, image logs, FWS logs and conventional logs serve as the main technical means for fracture identification. Compared to other drilling data, conventional logs offer advantages such as lower acquisition difficulty, reduced cost, and detectable anomalous responses to open fractures underground. Therefore, they are widely applied for fracture identification. Consequently, researchers have attempted various methods to process conventional logging data for fracture detection. Examples include fracture indicator parameter method based on traditional logging curve calculation (Lyu et al., 2016; Dong et al., 2020a; Xie et al., 2023), feature enhancement method based on fracture sensitive curves (Zhang et al., 2021; He et al., 2023), fracture identification method based on R/S analysis and finite difference (Xiao et al., 2019; Li et al., 2023), etc. However, these approaches are limited by the fact that fractures represent relative anomalies against background values, exhibit strong nonlinearity, and display diverse response characteristics. Additionally, these methods often involve significant subjective human factors and face challenges in parameter determination, resulting in certain shortcomings in the accuracy and stability of fracture identification. In recent years, the application of various machine learning methods for fracture identification using conventional logs has become a major trend, leading to the development of a series of techniques such as Random Forest (RF), Support Vector Machine (SVM), eXtreme Gradient Boosting (XGBoost), Multi-kernel Fisher Discriminant Analysis (MKFD), and neural networks (Dong et al., 2016, 2020b; Lu et al., 2023). Previous studies have demonstrated the effectiveness of these methods in logging identification. Nevertheless, several key scientific challenges remain in the log-based identification of fractures in tight sandstones:

(1) The first is how to extract effective fracture response characteristics from the complex logging responses. According to existing core and image logs, approximately 95% of subsurface fractures are open and effective (Fig. 2). However, the identification of these open fractures in logging data is often interfered with by factors such as lithological variations, fluids, and gas-bearing properties (Tang et al., 2011; Lai et al., 2021b). For instance, the triple-porosity log responses induced by fracture-developed zones (Acoustic (AC): 255 μs/m, Density (DEN): 2.3 g/cm^3) resemble those of high-quality reservoirs. Deep mud invasion along densely developed fractures can lead to a sharp decrease in resistivity (Deep and Shallow Lateral Resistivities (LLD/LLS) < 10 Ω·m), while special minerals (e.g., metallic nodules) and water-bearing layers can also cause resistivity reduction. Moreover, there are notable differences in logging responses between fractures in medium-to-fine sandstone (AC: 235 μs/m, DEN: 2.46 g/cm^3) and fractures in coarse sandstone (AC: 260 μs/m, DEN: 2.25 g/cm^3). The aforementioned factors contribute to significant overlap in the value ranges of fracture-related log responses and some non-fracture-related log responses. Therefore, a critical challenge lies in how to eliminate the interference caused by absolute value anomalies resulting from these factors on fracture responses and thereby extract the effective logging response characteristics attributable to fractures.
(2) The second is how to achieve comprehensive identification of different fracture logging response types. Fractures vary in aperture and density, which can lead to significantly different log response patterns (Dong et al., 2020a; Bagheri and Falahat, 2022). For instance, in tight sandstone intervals where core and image logs indicate minimal fracture development, there may be no significant response in AC, which would exhibit low values. However, the separation between deep and shallow resistivity might be relatively pronounced, showing high-value anomalies. In contrast, if densely fractured zones are present around the wellbore, this could lead to a sharp increase in AC, resulting in high values, while both deep and shallow resistivity decrease with reduced separation, both manifesting as low values. Such completely opposite response characteristics will also cause huge interference to the intelligent identification of fractures using conventional logs.
(3) The third is how to achieve more scientific data balance processing. Under actual geological conditions, tight sandstone reservoirs are predominantly characterized by high-angle fractures. However, due to the limited lateral detection range of core and image logs, vertical wells have a low probability of intersecting these high-angle fractures. This results in a limited number of labeled fracture samples and a relative abundance of nonfracture samples. Current sampling methods primarily focus on the data level without adequately taking geological information into account. For sequential data of logging curves, random under-sampling and SMOTE oversampling methods may randomly discard samples or generate synthetic data points that lack sequential dependencies (Hemmatian et al., 2025), potentially leading to the loss or distortion of critical sequential information. Moreover, for multi-class samples, data distribution processed by these methods often contradicts actual geological conditions, resulting in distortion or duplication of the actual data distribution. The incomplete multi-category pattern information leads the model to learn virtual features. This weakens the modeling ability of the whole sequence.

However, the aforementioned machine learning methods still exhibit significant limitations when applied to these complex well-logging identification problems. For instance, SVM constructs a hyperplane (either original or kernel-mapped) in a static feature space, which fails to capture the relative anomalous responses of fractures within sequential features. In the context of fracture identification-characterized by multi-dimensional features, strong sequential dependencies, and non-linear data, the recognition and generalization capabilities of SVM are notably limited. XGBoost, while utilizing gradient boosting trees to iteratively fit well-logging curves and possessing strong non-linear modeling power, still constitutes a static modeling process for the data. It cannot account for crucial multi-scale sequential patterns and the extraction of logging combination patterns in fracture identification. Traditional Recurrent Neural Networks (RNN), although capable of attending to short-term fracture anomalies, are prone to issues like gradient vanishing and explosion, significantly impairing their ability to model long-term and complex sequential dependencies. While Long Short-Term Memory (LSTM) networks can mitigate long-term dependency problems to some extent, they still struggle with interference from sedimentary microfacies cycles (both normal and reverse) on fracture identification. Furthermore, constrained by a single sequential scale, their comprehensive characterization of fractures remains inadequate.

To address the above geological problems and existing model limitations, this paper specifically adopts the following methods:

Regarding the problem (1), this paper utilizes a dual-layer architecture of DL-BiLSTM. Based on the construction of sliding window sequence data, the first BiLSTM layer is employed to detect high-frequency relative anomalies in logging responses induced by fractures within short windows. On the other hand, the dual-layer structure enables the model to account for low-frequency trend shifts in long-term dependencies of the sequential data. This approach mitigates interference from sedimentary factors in fracture identification and captures precise multi-scale information of fracture responses.

The CSA mechanism is incorporated following one-dimensional convolution (Conv1D) processing to overcome the problem (2). This component is introduced to focus on the interrelationships among different response patterns. It adaptively learns and assigns higher weights to feature combinations that contribute most significantly to fracture classification, thereby achieving optimization and focusing of feature combinations. Consequently, an intelligent multi-class fracture identification model is established to resolve difficulties in recognition caused by significant disparities among different logging response types.

In addition, a method based on FWS classification thresholds is proposed to supplement fracture labels, replacing mathematical sampling methods for data balancing to address the problem (3). This approach integrates geologically reasonable labels into the sequential data, avoiding the sequential distortions introduced by conventional sampling methods and improving the geological rationality of the model.

Finally, a new method for fracture logging identification (CNN-Attention-BiLSTM) is established to achieve comprehensive identification of multiple fracture logging response types in tight sandstone reservoirs of the Hangjinqi area, northern Ordos Basin.

2. Geological settings

The Ordos Basin, a hydrocarbon-rich cratonic basin developed on the North China Platform, comprises six major secondary tectonic units: Yishan Slope, Yimeng Uplift, Jinxi Fault-Fold Belt, Weibei Uplift, Tianhuan Depression and Western Margin Thrust Belt (Fig. 1(a)) (Xu et al., 2018). The Hangjinqi area is located in the northern part of the Ordos Basin (Fig. 1(a)), spanning three major secondary tectonic units: Yimeng Uplift, Yishan Slope and Tianhuan Depression. Prior to the Late Paleozoic, the Hangjinqi area existed as a long-term inherited paleo-uplift (Wang et al., 2020; Xu et al., 2022). Its present structural configuration is a gentle slope dipping from northeast to southwest with low-amplitude fold deformation. Controlled by the regional tectonic stress field, the Hangjinqi area is characterized by three major basement-related faults: the Bo'erjianghaizi Fault, Wulanjilinmiao Fault, and Sanyanjing Fault, trending approximately EW and NE (Wu et al., 2017). Based on basement relief, cover development, and fault boundaries, the Hangjinqi area can be divided into six zones: Shiguhao, Haoraozhao, Gongkahan, Azhen, Shilijiahan, and Xinzhao (Wu et al., 2017; Qin et al., 2022a). The Jin 30 Block is located in the Xinzhao zone of southwestern Hangjinqi, at the junction of the three major tectonic units (Fig. 1(b)).

The Upper Paleozoic strata in the Jin 30 Block consist of the Upper Carboniferous Taiyuan Formation, Lower Permian Shanxi Formation, Middle Permian Lower Shihezi Formation and Upper Shihezi Formation, and Upper Permian Shiqianfeng Formation, from bottom to top. The depositional environment evolved from marine and deltaic facies to fluvial-dominated continental facies (Zhu et al., 2008; Li et al., 2019). As the main gas-bearing interval, the H1 member of Lower Shihezi Formation comprises a braided river to braided river delta system dominated by coarse clastic sediments (Anees et al., 2022a; Ashraf et al., 2022; Qin et al., 2022b). The reservoir is composed of shore sand bodies, deltaic braided channels, interdistributary channels, and channel-bar sand bodies (Anees et al., 2022b; Tan et al., 2023). It shows obvious positive cycle characteristics. The lower part is dominated by gray white gravel-bearing sandstone and medium sandstone, and gradually transitions to fine sandstone upwards. The uppermost part is gray black mudstone or carbonaceous mudstone (Fig. 1(c)) (Wang et al., 2006; Ashraf et al., 2022).

Since the deposition of the Lower Shihezi Formation, the study area has undergone three major tectonic movements: the Indosinian, Yanshanian, and Himalayan Orogenies. The Indosinian Movement, primarily driven by the collision between the Yangtze and North China Plates along with obstruction from the Siberian Plate, resulted in a dominantly N–S compressional stress field. By the middle stage of the Yanshanian Movement, influenced by the continuous activity of the Tianshan-Xingmeng Orogenic Belt and the sustained compression from the subduction of the Kula-Pacific Plate beneath the Aleutian-Japan Trench, the principal stress field gradually shifted to the NNW–SSE orientation, and further evolved to the NW–SE trend during the late Yanshanian. During the Himalayan Movement, due to the collision between the Indian and Asian Plates and the NW-ward subduction of the Pacific Plate, the study area was dominated by the NE–SW compressional stress field (Fig. 1(d)). These three tectonic phases have exerted significant modification on the study area (Xu et al., 2006; Jiang et al., 2016; Xu, 2020). Influenced by multiple phases of tectonic activities, the Jin 30 Block developed multiple sets of natural fractures, predominantly high-angle shear fractures (Fig. 2). Based on orientation, these can be classified into four dominant sets: NEE–SWW, NNE–SSW, approximately NS, and NWW–SEE trends (Fig. 1(d)). Previous studies on the development of micro-fractures in the Hangjinqi area have yielded substantial insights. It has been observed that tight sandstone reservoirs contain abundant grain edge fractures, along with intragranular and grain-penetrating fractures. The development of these micro-fractures increases with larger grain size, higher content of rigid clastic components, and lower content of interstitial material. Additionally, their formation is influenced by effective stress and tectonic stress. The presence of these micro-fractures significantly enhances reservoir permeability and contributes importantly to reservoir physical properties and connectivity (Guo et al., 2019; Zhao, 2022; Qin et al., 2022b, 2023). However, research on the characteristics, identification, and distribution patterns of macro-fractures remains notably insufficient, and a clear fracture development model to characterize the influence and control of fractures on hydrocarbon migration has yet to be established.

3. Data and method

3.1. Data

This data is derived from core, image logs, FWS logs and conventional logs in Hangjinqi area of Ordos Basin, with a total of 4403 samples obtained. Among them, the core samples are derived from seven coring wells. Image logs are derived from the

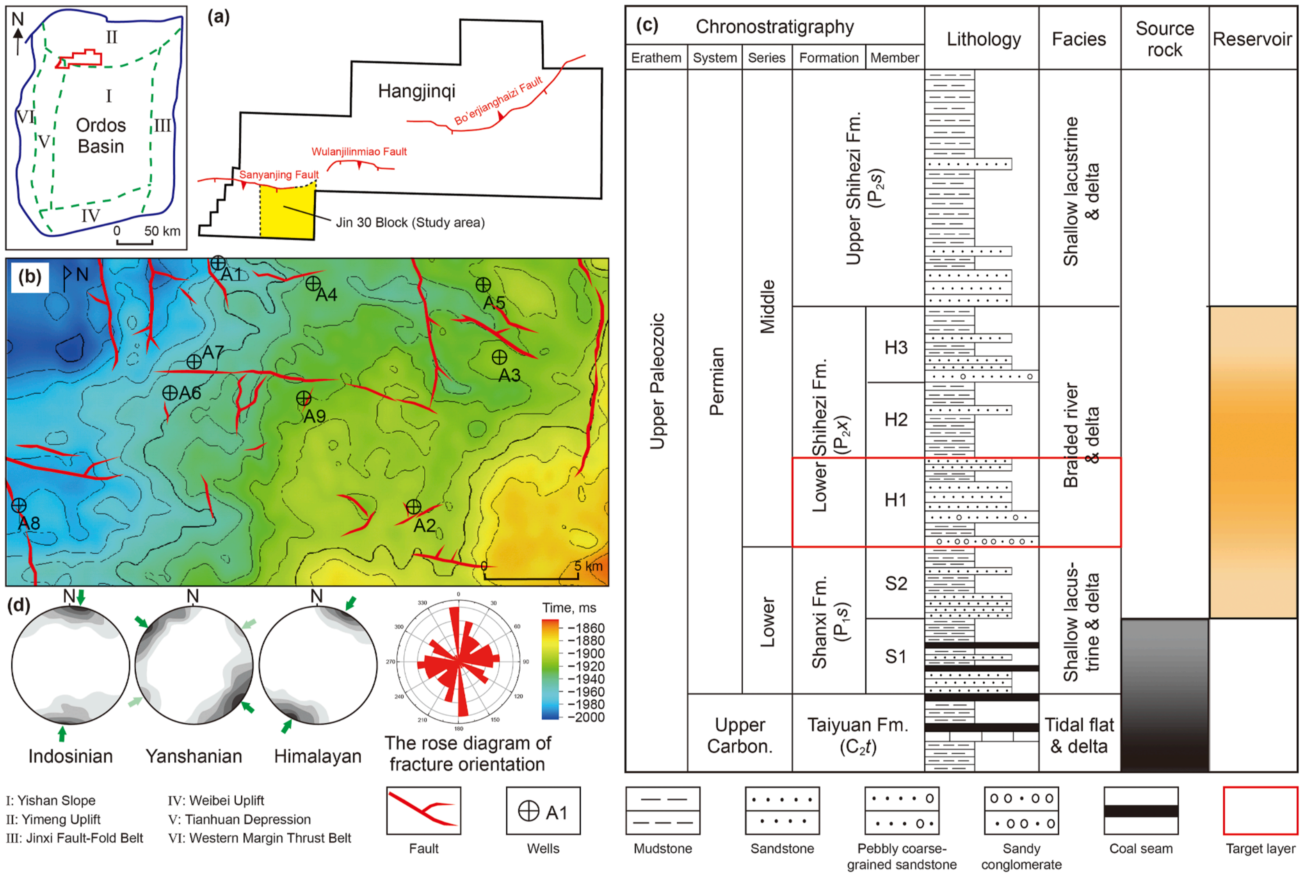

Fig. 1. Regional tectonic setting and stratigraphic column. **(a)** Tectonic location of the Hangjinqi area; **(b)** distribution of faults and wells in the study area; **(c)** the generalized stratigraphic column of the Upper Paleozoic in the study area; **(d)** in-situ stress orientation and rose diagram of fracture orientation. Modified from Wang et al. (2020); Xu (2020); Anees et al. (2022c); Li (2025).

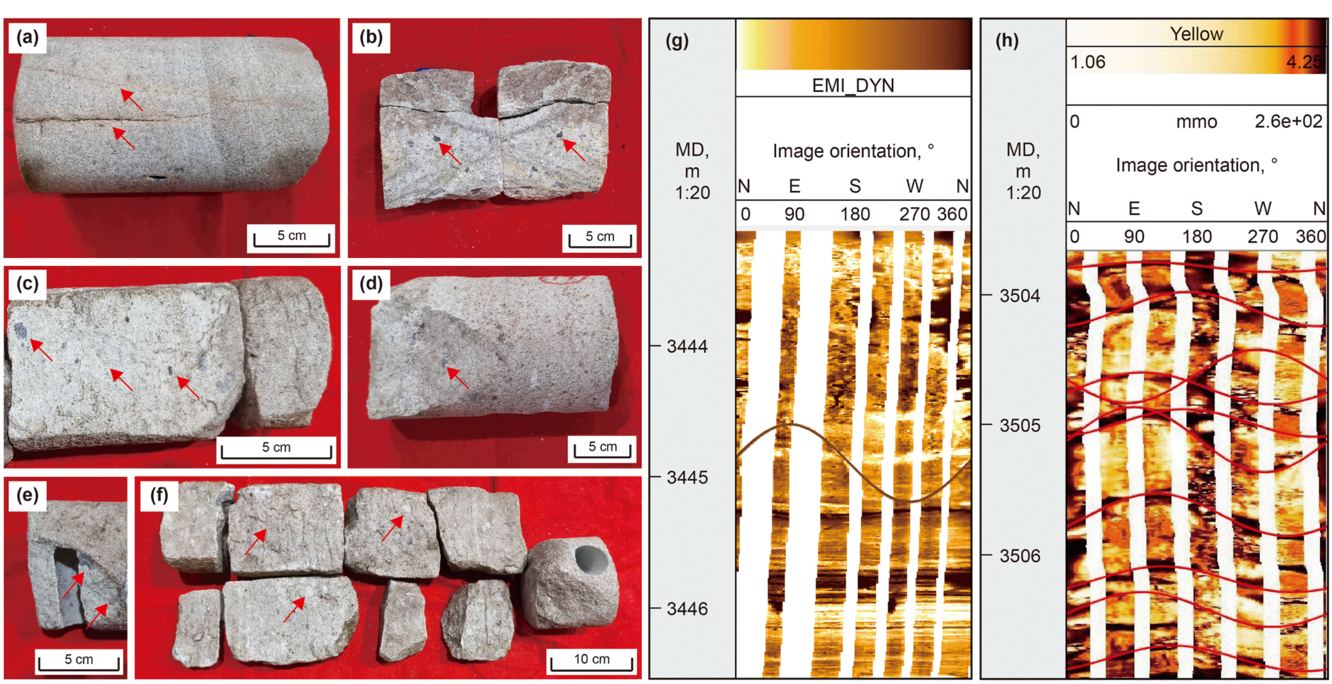

Fig. 2. Fractures observed in cores and image logs from the study area. **(a)** Vertical shear fracture, Well A1, 3559.7 m; **(b)** vertical shear fracture cutting through gravels, Well A2, 3463.5 m; **(c)** vertical shear fracture cutting through gravels, Well A3, 3461.5 m; **(d)** high-angle shear fracture cutting through gravels, Well A2, 3461.6 m; **(e)** high-angle shear fracture cutting through gravels, Well A3, 3461.3 m; **(f)** vertical shear fracture cutting through gravels, multiple steps are developed, Well A3, 3462.6 m; **(g)** high-angle filled shear fracture, Well A4, 3445 m; **(h)** networked fractures, Well A5, 3503.6–3507 m.

formation images around nine wells obtained by XRMI micro-resistivity image logs. Conventional logs are derived from sixteen wells, including seven coring wells and nine image log wells. FWS logs are derived from nine well logs, including P-wave, S-wave (including fast and slow S-wave), Stoneley wave and acoustic wave amplitude. Based on well log data, three types of conventional log responses were identified, and three corresponding fracture labels (Type-1, Type-2, and Type-3) were subsequently established. Type-1 and Type-2 correspond to the development of single or multiple parallel fractures, and Type-3 corresponds to the development of a densely networked fracture zone (Fig. 2).

3.2. Method

We propose a new intelligent method for fracture identification (CNN-Attention-BiLSTM). The principle and workflow are as follows.

3.2.1. CNN-attention-BiLSTM identification process

(1) Log data standardization for target wells in the study area. Calibrating FWS logging data with imaging log labels and core fracture labels, the labeled dataset was established (Fig. 3(a)). Fracture response characteristics were classified in FWS logs and multidimensional FWS features were optimized, then the XGBoost method was used to determine the classification threshold for fracture and nonfracture, thereby obtaining reliable unlabeled FWS data to be used as virtual labels. Therefore, data balancing is carried out to maintain the relative balance of the data volume between the fracture labels and the nonfracture labels (Fig. 3(b)). The detailed principle can be found in Section 3.2.2 and Fig. 4.

(2) Reconstruction of the sequential dataset using fracture labels from image logs, core data, and supplemented FWS labels (as shown in Fig. 3(c)), where different colors represent different fracture response types. Sequences for these multiple types were constructed.

(3) The constructed sequence data underwent Conv1D processing and dimensional transposition. The CSA was then utilized to weight the multi-dimensional response patterns. The output was subsequently subjected to dimension restoration and residual connection (Fig. 3(d)). Then it was fed into the DL-BiLSTM to establish the multi-class fracture identification model (Fig. 3(e)).

(4) The validation wells to be predicted were processed by constructing sequences following the same methodology. These sequences were then sequentially processed through convolution and pattern weighting. Finally, they were input into the trained CNN-Attention-BiLSTM model to obtain identification results for each fracture response type (Fig. 3(f)). The results were compared with actual fractures interpreted from core and image logs to evaluate the accuracy of the method.

(5) Two additional BiLSTM models were constructed using random under-sampling and SMOTE oversampling techniques for data balancing during sequence construction. Following the same workflow and parameters, these models were trained and validated on blind wells to compare the stability and accuracy of their identification results with those of the proposed method.

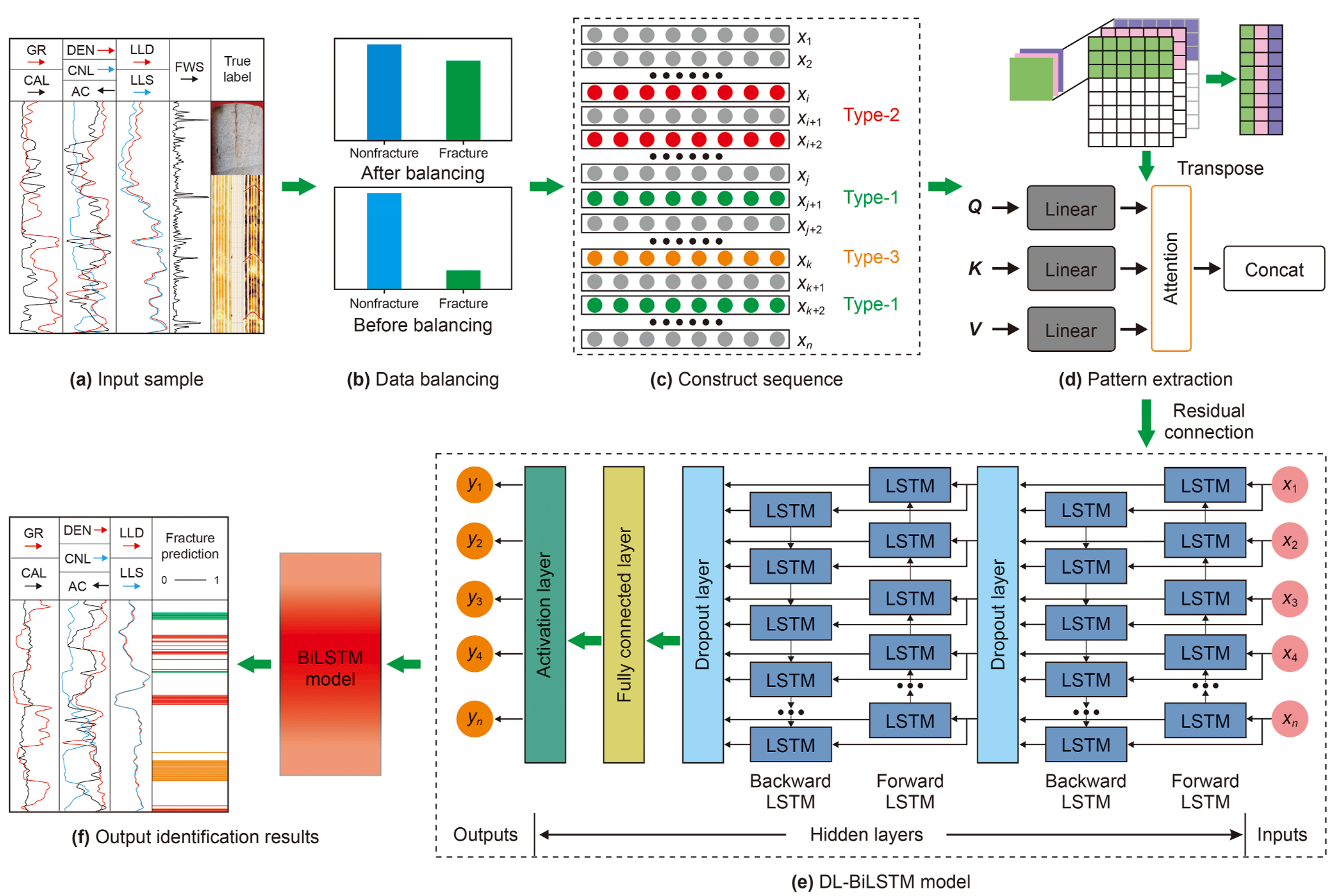

Fig. 3. Process of CNN-attention-BiLSTM for multi-type fracture identification.

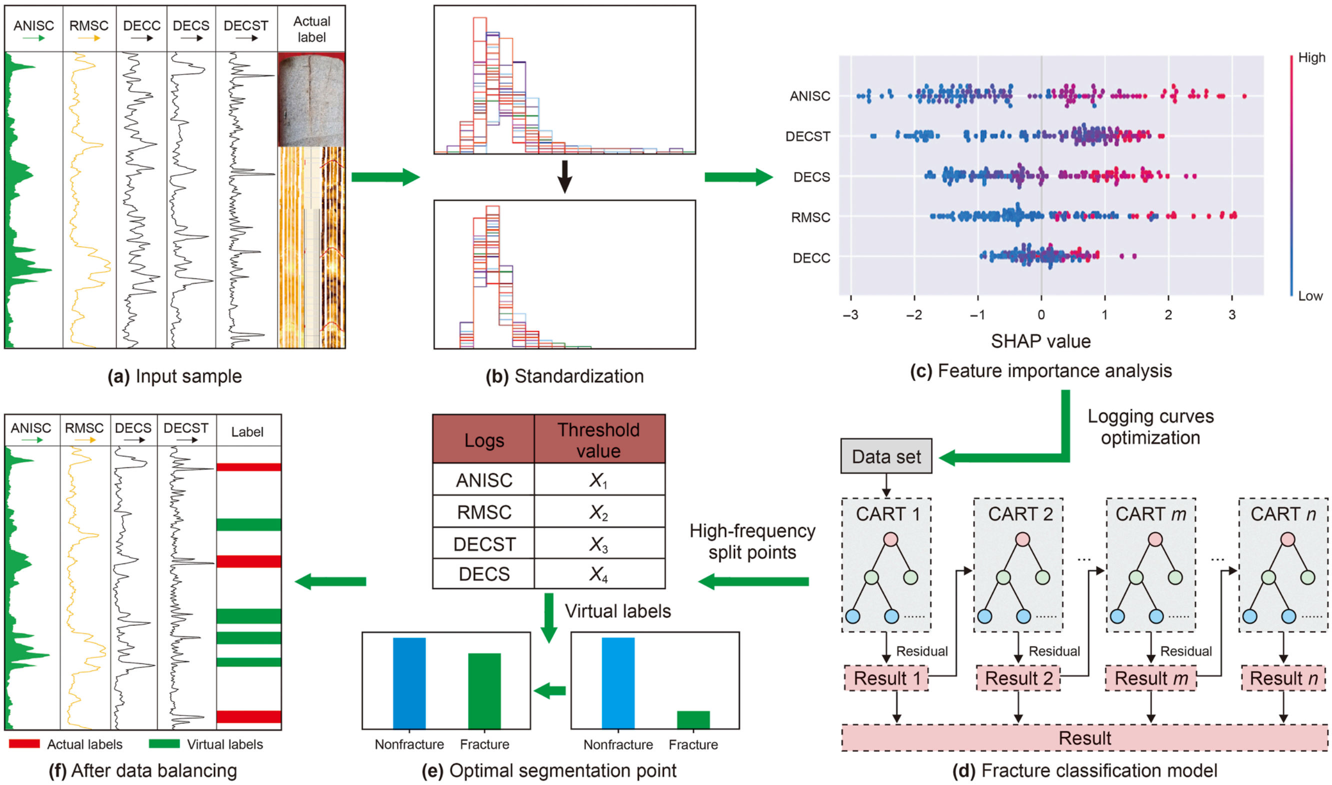

Fig. 4. Fracture data balancing process.

3.2.2. FWS for fracture data balancing

FWS logs can effectively record raw data such as P-wave, S-wave and Stoneley wave. The comprehensive and rich data information enables its strong application value in lithology analysis, gas reservoir evaluation, fracture zone detection, rock mechanics and in-situ stress analysis, etc. (Che et al., 2020; Lai et al., 2021b; Zhao et al., 2025). Compared to conventional logs, FWS logs offer superior resolution and effectiveness in fracture identification due to their unique advantages in P-wave and S-wave measurement, S-wave anisotropy and Stoneley wave attenuation, etc. (Zeng et al., 2010; Tang et al., 2017; Che et al., 2020; Lai et al., 2021a; Wang et al., 2022; Zhao et al., 2025). Additionally, the radial detection depth of logs is greater than that of both core and image logs, making FWS logging theoretically well-suited for indicating fracture development.

This study utilizes FWS logging data to supplement labeled data, aiming to obtain a more reasonable and balanced data distribution for subsequent fracture identification. Among all the curves, multiple FWS curves were utilized, including decay of energy based on cross-correlation (compressional-wave), decay of energy based on cross-correlation (shear-wave), decay of energy based on cross-correlation (Stoneley-wave), acoustic anisotropy and ratio of mean shear-wave velocity to mean compressional-wave velocity, denoted as DECC, DECS, DECST, ANISC, and RMSC, respectively. DECC, DECS, and DECST are composite attenuation indices calculated from the cross-correlation of full-waveform sonic for compressional, shear, and Stoneley waves. They exhibit high sensitivity in reflecting acoustic energy absorption, scattering, and structural discontinuities within fractured formations. ANISC is an anisotropy parameter extracted from the characteristics of fast and slow shear waves, used to characterize directional differences induced by fractures in the formation. RMSC is derived from the ratio of compressional-wave to shear-wave velocities and indirectly reflects fracture development by recording formation heterogeneity caused by fracture zones. The methodological workflow is illustrated in Fig. 4. First, the original logging data are standardized to eliminate the influence of the data distribution differences of the same curve between wells on the model (Fig. 4(a) and (b)). Subsequently, SHAP (SHapley Additive exPlanations) feature importance values are introduced to reflect the direction and magnitude of each feature's impact on prediction outcomes. The sign of the SHAP value for a feature indicates whether it pushes the model prediction toward fracture or nonfracture, positive values favor fracture classification, while negative values favor nonfracture. The absolute value reflects the strength of the feature's influence on the prediction for a given sample. Values close to zero indicate negligible contribution. Based on the contribution of features to fracture identification, feature curves with no significant effect are excluded, thereby achieving feature optimization (Fig. 4(c)). The optimized dataset is then input into an intelligent identification model for fracture classification (Fig. 4(d)). Here, the XGBoost method is employed to iteratively construct multiple decision trees. Each tree learns the residuals of the preceding trees, and model identification performance is enhanced through hyperparameter optimization, feature importance analysis, and model interpretation techniques (Gu et al., 2021). This process continues until the model's learning outcomes stabilize. Regularization terms are incorporated during hyperparameter tuning to prevent overfitting, and the split point values of each decision tree generated during model development are recorded. Based on the high-frequency split points of the decision trees, the classification thresholds of each feature under optimal classification conditions are obtained as the thresholds for distinguishing between fractured and non-fractured intervals (Fig. 4(e)). Finally, these threshold values are applied to screen unlabeled FWS data, selecting well segments that meet the criteria as virtual labeled data to supplement the labeled data from core and image logs (Fig. 4(f)). This approach achieves a balance between fractured and non-fractured data in the training dataset.

3.2.3. The multi-kernel Conv1D and CSA for multidimensional feature extraction and weighting

The length variation of fracture development zones is distinctly manifested in well logging responses. Consequently, logging interpretation often reveals multi-scale response patterns, such as long-term cycle skipping and short-term spike anomalies. To capture these, the approach processes the established multi-feature sequence data through a parallel architecture of three Conv1D branches coupled with CSA. The convolutional layers generate a richer set of combined logging patterns, while the attention mechanism self-supervisedly identifies global dependencies and internally computes the weights for these pattern combinations. This allows local patterns to be aggregated into more discriminative global features. These features are then fed into a subsequent DL-BiLSTM, providing it with clearer guidance on fracture-induced log responses. This facilitates the model's ability to focus on temporal dependencies on the basis of multi-log response patterns, ultimately establishing a more robust identification model.

The convolutional kernels in our parallel branches are sized at 3, 5, and 7, respectively, enabling the simultaneous capture of features at different scales. The raw input data is processed through multiple kernels in each branch to extract local features from the multidimensional log curves, thereby yielding features from a greater diversity of pattern dimensions (Fig. 5(a)). The operation of the Conv1D is defined by the following formula (LeCun et al., 1998):

$$\boldsymbol{y}_{t,f} = \sigma\left(\sum_{c=1}^{C}\sum_{\tau=-\left[\frac{k}{2}\right]}^{\left[\frac{k}{2}\right]} \omega_{f,c,\tau}\cdot x_{t+\tau,c} + b_f\right) \tag{1}$$

where $x_{t+\tau,c}$ represents the value at the depth point $t+\tau$ on the log curve c; $\omega_{f,c,\tau}$ represents the weight of the convolutional kernel corresponding to the f-th output feature; b_f represents the bias term; $\tau = -[k/2], \dots, [k/2]$ represents the sliding step of the convolutional kernel along the depth; C represents the dimensionality of the input log curves; $\boldsymbol{y}_{t,f}$ represents the convolutional output value at depth point t for the f-th output feature; σ represents the activation function.

Self-Attention mechanisms and their extended variants have been demonstrated to enhance the classification performance of neural networks (Bahdanau et al., 2014; Lan et al., 2025). By enabling each element in a sequence to compute its correlation with all other elements, the mechanism effectively captures global dependencies within the data. It primarily consists of three components (Fig. 5(c)): Query ($\boldsymbol{Q}$), Key ($\boldsymbol{K}$), and Value ($\boldsymbol{V}$) (Vaswani et al., 2017). The specific formulas are as follows:

$$\boldsymbol{Q} = \boldsymbol{X}\boldsymbol{W}_Q, \boldsymbol{K} = \boldsymbol{X}\boldsymbol{W}_K, \boldsymbol{V} = \boldsymbol{X}\boldsymbol{W}_V \tag{2}$$

$$\text{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}) = \text{softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d_k}}\right)\boldsymbol{V} \tag{3}$$

where $\boldsymbol{W}_Q$, $\boldsymbol{W}_K$, and $\boldsymbol{W}_V$ represent the weight matrices for the Query, Key, and Value linear transformations, respectively; $\boldsymbol{Q}$, $\boldsymbol{K}$, and $\boldsymbol{V}$ incorporate combinations of fracture logging response patterns residing in distinct latent spaces; T represents the transpose of the keys matrix; $\sqrt{d_k}$ denotes the square root of the dimensionality of $\boldsymbol{K}$ and the softmax function is applied to normalize the attention weights.

It should be noted that the identification of fracture intervals relies on the combined patterns from multiple log curves. Thus, to address the complex response patterns of fractures, it is necessary to perform feature weighting on these combined patterns from different log series prior to inputting them into the BiLSTM model. This preprocessing reduces the interference of non-informative features during the BiLSTM training, thereby enhancing the model's sensitivity to characteristic fracture responses. Therefore, Conv1D processing and CSA are employed to achieve this objective. Technically, the input matrix generated by the Conv1D is transposed and mapped into multiple distinct projection spaces ($\boldsymbol{Q}$, $\boldsymbol{K}$, $\boldsymbol{V}$). This operation shifts the Self-Attention computation from the temporal (depth) axis to the feature (logging response pattern) dimension (Fig. 5(b)). Then, the features extracted from these projection spaces are then concatenated and transformed, enabling the model to learn dependencies between different patterns (Fig. 5(c)). By employing this method, the model achieves a

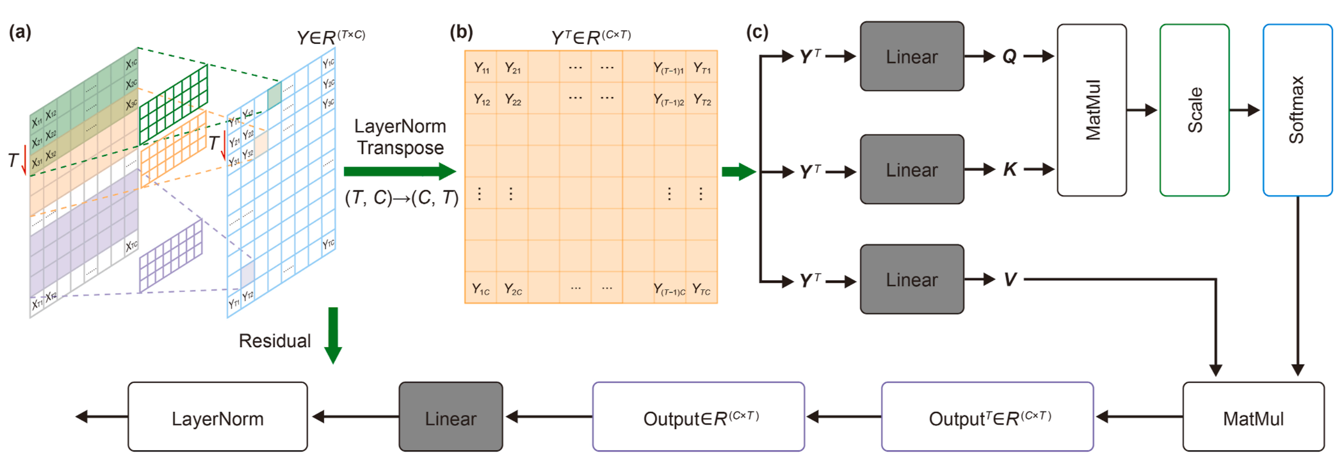

Fig. 5. Process of Conv1D processing and Channel Self-Attention mechanism. **(a)** Multi-core 1D convolution computation; **(b)** dimensional transpose of the well logging data matrix; **(c)** input and output of the Channel Self-Attention module.

more comprehensive understanding of both local and global features across different scales, ultimately enhancing its capacity to model complex data.

3.2.4. The principle of DL-BiLSTM

LSTM is a special type of Recurrent Neural Network (RNN) designed to address the issues of gradient vanishing and gradient explosion during the training of long sequence data (Yang et al., 2023). Compared to traditional RNNs, it demonstrates stronger robustness in addressing sequence recognition and prediction problems and has been widely applied to fracture identification using conventional logs (Dong et al., 2024). The structure of an LSTM unit is shown in Fig. 6(a). It introduces three gating mechanisms (the input gate, forget gate, and output gate) along with a cell state (S_t), to achieve selective memorization, forgetting, and output of the logs. This enables selective transmission of effective logging responses throughout the sequence. The detailed computational process is given by Equations (4)–(9) (Graves, 2013):

$$\boldsymbol{f}_t = \sigma\left(\boldsymbol{W}_f \cdot [\boldsymbol{y}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_f\right) \tag{4}$$

$$\boldsymbol{i}_t = \sigma(\boldsymbol{W}_i \cdot [\boldsymbol{y}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_i) \tag{5}$$

$$\tilde{\boldsymbol{C}}_t = \tanh(\boldsymbol{W}_c \cdot [\boldsymbol{y}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_c) \tag{6}$$

$$\boldsymbol{C}_t = \boldsymbol{f}_t \odot \boldsymbol{C}_{t-1} + \boldsymbol{i}_t \odot \tilde{\boldsymbol{C}}_t \tag{7}$$

$$\boldsymbol{o}_t = \sigma(\boldsymbol{W}_o \cdot [\boldsymbol{y}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_o) \tag{8}$$

$$\boldsymbol{y}_t = \boldsymbol{o}_t \odot \tanh(\boldsymbol{C}_t) \tag{9}$$

where $\boldsymbol{f}_t$, $\boldsymbol{i}_t$, and $\boldsymbol{o}_t$ represent the forget gate, input gate, and output gate respectively; $\boldsymbol{x}_t$ and $\boldsymbol{y}_t$ represent input and output respectively; $\boldsymbol{C}_t$ and $\tilde{\boldsymbol{C}}_t$ represent cell state and candidate memory respectively; $\boldsymbol{W}_f$, $\boldsymbol{W}_i$, $\boldsymbol{W}_C$, and $\boldsymbol{W}_o$ are weight matrices associated with the respective gates and cell state; $\boldsymbol{b}_f$, $\boldsymbol{b}_i$, $\boldsymbol{b}_C$, and $\boldsymbol{b}_o$ respectively represent the bias terms in LSTM; tanh represents the hyperbolic tangent activation function.

However, conventional logging identification of fractures largely relies on contextual depth information to detect anomalous responses indicative of fractures. Therefore, BiLSTM provides a more effective solution. It builds a bidirectional structure by adding a backward layer on the basis of the LSTM architecture. Using the cyclical characteristics reflected by the (Gamma) GR curve as a constraint, simultaneously capturing both historical and future information, it further eliminates the influence of sedimentary rhythms on the fracture identification. This design better accommodates the need for global context awareness in fracture identification from log data (Shan et al., 2021; Zhang et al., 2024). In this paper, we establish a dual-layer BiLSTM model to further utilize the advantage that different BiLSTM layers can capture information at varying scales (as shown in Fig. 6(b)). Building upon the aforementioned well log response patterns, this method further enhances the modeling capacity for both short-term anomalies and long-term trends (Fig. 10(a) and (d) show the long-term series of long fracture interval depth, while Fig. 10(b) shows the short-term series of short fracture interval depth). This enables a more comprehensive understanding of the dependencies within sequential data, thereby improving the accuracy of fracture identification. Additionally, by incorporating regularization layers that randomly drop out units during the training phase, the model avoids over-reliance on specific neurons. Combined with early stopping, this approach enhances the model robustness against overfitting. Finally, an activation layer is incorporated to fulfill multi-class classification requirements.

4. Results

4.1. Data balancing using FWS logs

First, FWS logs are calibrated by using core and image log data to classify the FWS response types of fractures. The FWS response characteristics of fractures in the study area are categorized into two distinct types. Type-1 is primarily characterized by significant DECS, with DECST being either significant or subtle, increased ANISC, and the increased RMSC (3486–3488 m, 3494–3495 m and 3497–3499 m in Fig. 7). This response pattern is likely induced by single or multiple high-angle fractures near the borehole, which significantly alter DECS signatures and increase RMSC. But fracture density in such intervals is relatively low, and permeability varies considerably. The Stoneley wave may or may not exhibit clear attenuation. Type-2 displays prominent DECS and DECST, increased ANISC, and cycle skipping in RMSC (3477.5–3481.5 m in Fig. 7). This pattern is interpreted to result from densely connected

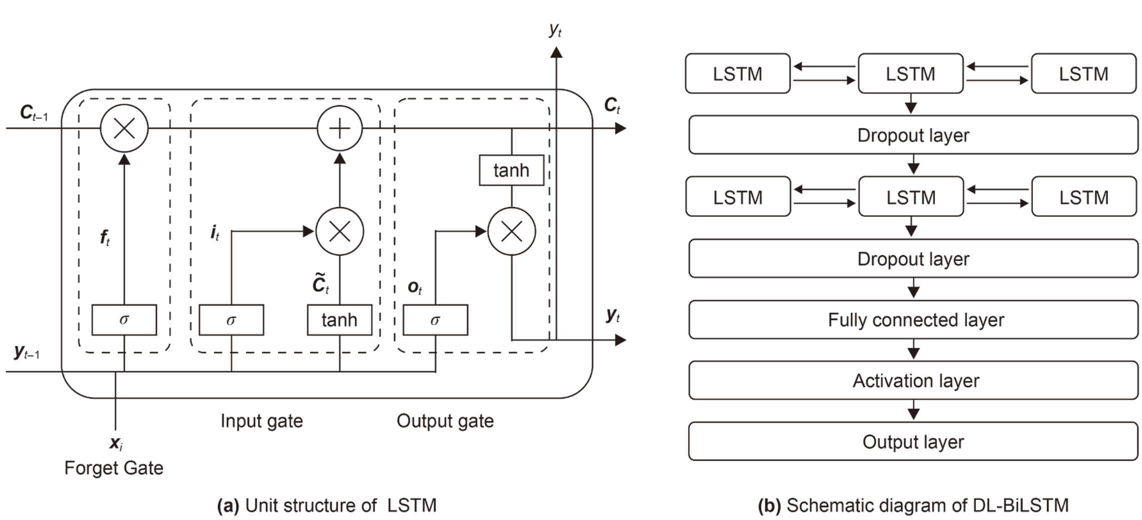

Fig. 6. Principle of DL-BiLSTM unit structure.

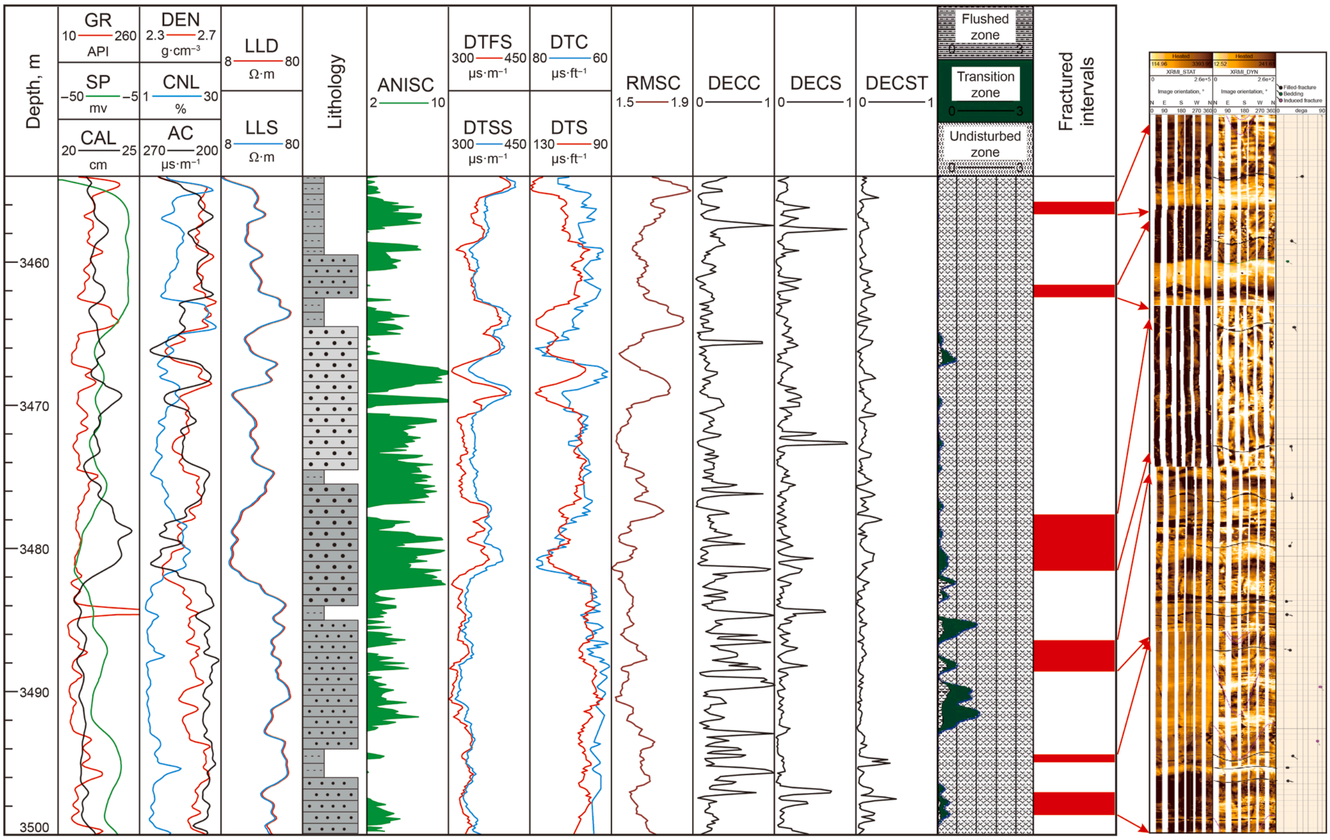

Fig. 7. FWS logging response characteristics of fractures in tight sandstone within the study area.

fracture networks, forming localized high-permeability zones. These zones cause significant alterations in both shear and Stoneley waves, while substantial compressional wave variations lead to RMSC cycle skipping. The developed fracture intervals show strong correspondence in FWS log responses, with good consistency across different logging measurements. Therefore, integrating multiple acoustic logging curves enables accurate identification of fracture intervals, thereby reducing interpretational ambiguity and enhancing recognition accuracy.

The SHAP importance analysis was introduced to quantify the magnitude and direction of each feature's contribution to model predictions, and the optimization and iteration of logging curves were carried out. Based on global feature importance ranking by mean SHAP values (Fig. 8), the results show that the contribution of the DECC feature is nearly zero. Consequently, DECC was removed from the dataset to iteratively optimize the classification model.

Following the procedure outlined in Section 3.2.2, the XGBoost method was employed to establish the fracture classification model. The classification performance of the model can be evaluated using precision, recall, accuracy and F1 score. The specific formulas are as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{10}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{11}$$

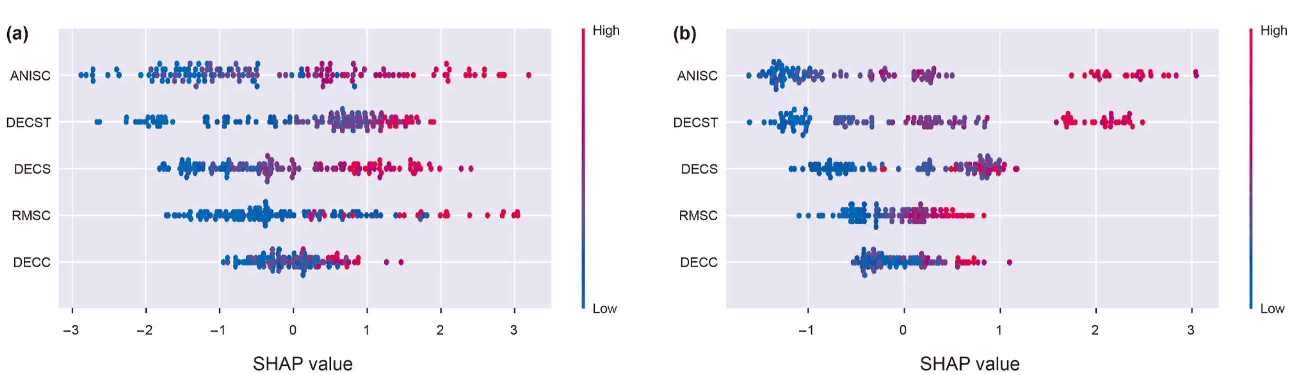

Fig. 8. SHAP feature importance plot of FWS logs. **(a)** Response characteristics for Type-1 fractures; **(b)** response characteristics for Type-2 fractures.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{12}$$

$$\text{F1 score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{13}$$

Precision reflects the probability of correctly identifying true fractures and the probability of misclassifying nonfractures as fractures. Recall represents the probability of detecting actual fractures. Accuracy represents the overall correctness in identifying both fractured and non-fractured intervals. The F1 score is the harmonic mean of precision and recall. It is noteworthy that during the identification process, response deviations arise from the significantly greater radial detection depth of FWS logs compared to core and image logs. This necessitates allowing a certain degree of misidentification where some non-fractured intervals may be classified as fractured. Given the greater emphasis on accurately identifying actual fractures, recall is considered more representative than precision and accuracy in this context. As shown in Fig. 9, the recall rate for Type-1 fractures reached 89.1%, and the recall rate for Type-2 fractures reached 87.0%. Meanwhile, precision, accuracy, and F1 score all exceeded 85%, indicating highly satisfactory identification performance. Finally, the split thresholds of each significant feature in the decision tree are statistically analyzed, and their high-frequency split points are selected as the optimal classification thresholds (Table 1).

Table 1
Feature classification thresholds by FWS log.

Name	Type-1	Type-2
ANISC	3.341	4.754
DECST	0.166	0.313
DECS	0.368	0.114
RMSC	1.656	1.619

4.2. Conventional logging response characteristics and sensitivity of fractures

Fractures identified on core, image logs, and FWS logs were used to calibrate conventional logs, allowing for classifying conventional logging response types of fractures. The fracture responses in the study area were classified into three types. Type-1 is mainly characterized by a significant difference between Deep and Shallow Lateral Resistivities (LLD/LLS), while Acoustic (AC), Density (DEN), Compensated Neutron (CNL) and Caliper (CAL) show little noticeable change (as shown in Fig. 10(a)). Type-2 mainly exhibits caliper expansion, insignificant variation in AC, DEN, CNL, and significant difference between deep and shallow lateral resistivities (as shown in Fig. 10(b)). The log responses of the first two types are likely caused by single or multiple high-angle fractures around the borehole. These fractures do not interconnect to form a network (Fig. 10(c)), which result in negligible changes in porosity logs (AC, DEN, CNL), minimal to severe caliper expansion, and significant resistivity responses due mainly to mud invasion (Deng et al., 2006). This response pattern corresponds to Type-1 of the FWS logging response characteristics. Type-3 is mainly characterized as caliper expansion, increased AC, decreased DEN, increased CNL, and decreased deep and shallow resistivity (LLD, LLS) with no significant positive separation or even negative separation between them (as shown in Fig. 10(d)). This pattern may result from intensively network-like fracture zones causing severe borehole breakouts and large-scale mud invasion, leading to dramatic changes in caliper, porosity logs, and resistivity (Fig. 10(e)) (Deng et al., 2006). The combined conventional logging response characteristics for fracture identification are summarized in Table 2. Based on different fracture-induced log patterns, corresponding labels (Type-1, Type-2, and Type-3) are established.

Cross-plots for fracture identification were constructed using eight conventional logging curves (Fig. 11). The left and bottom margins of the plot display probability density distribution curves of the log data. Blue-filled areas represent nonfracture, while other colors indicate different fracture types. The probability density curves reflect the sensitivity of each logging curve to fracture detection: a larger separation between peaks and less overlap indicate better fracture distinction performance. Complete overlap between two curves suggests that the corresponding log has no significant response to fractures. The results demonstrate that LLD/LLS shows the strongest response to fractures, followed by CAL, AC, DEN, and CNL, while GR and spontaneous potential (SP) exhibit relatively poor performance. The criteria for fracture classification on these curves are provided in Table 3. It should be noted that since the responses caused by fractures are relative anomalies, their value ranges are also dependent on the background values. The cross plot does not exhibit distinct fracture classification boundaries. The method proposed in this paper is

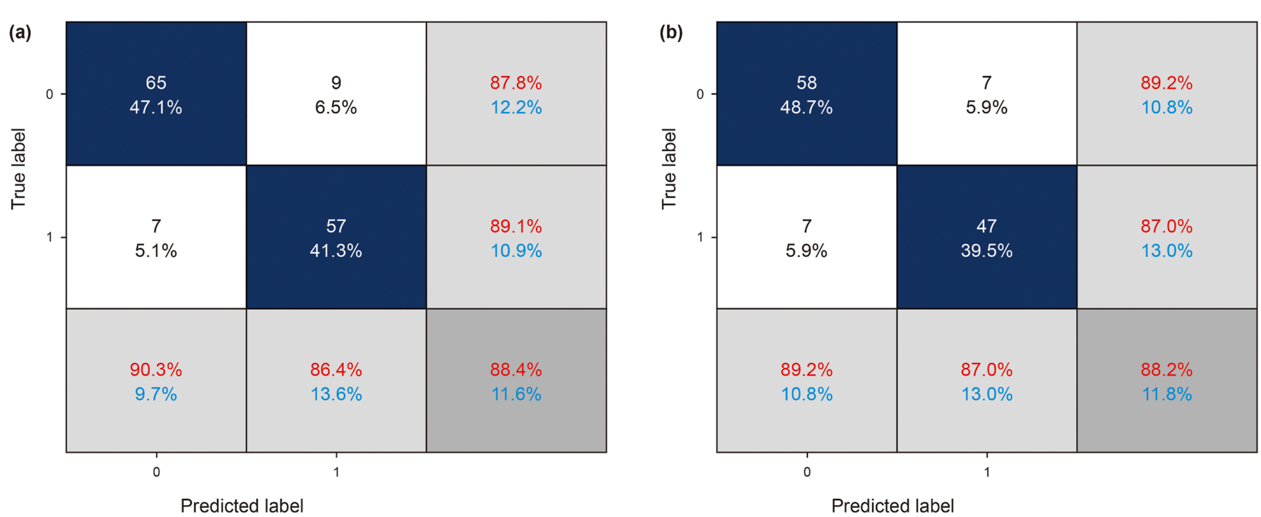

Fig. 9. Confusion matrix of XGBoost test samples of Type-1 (left) and Type-2 (right). 1 represents fracture, 0 represents nonfracture, blue represents identified samples, white represents misidentified samples, light gray represents precision (lower left) and recall (upper right) respectively, and dark gray represents accuracy.

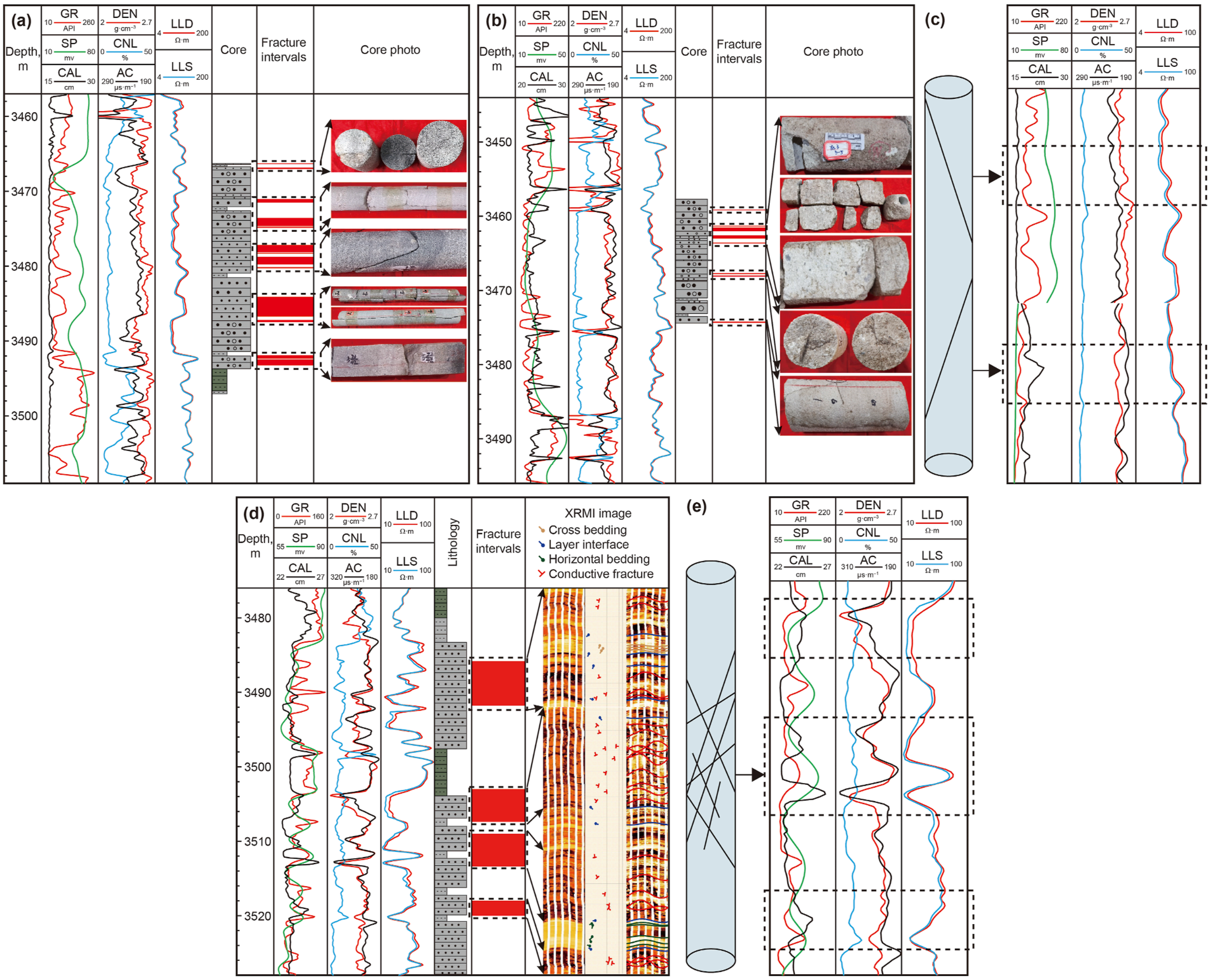

Fig. 10. Conventional logging response characteristics of fracture in tight sandstone reservoirs. **(a)** Response characteristics of Type-1, Well A7; **(b)** response characteristics of Type-2, Well A3; **(c)** fracture morphology and response pattern of Type-1 and Type-2; **(d)** response characteristics of Type-3, Well A5; **(e)** fracture morphology and response pattern of Type-3.

Table 2
Conventional logging combination characteristics of fractures.

Type	Conventional logging characteristics
Type-1	LLD/LLS↑
Type-2	CAL↑, LLD/LLS↑
Type-3	CAL↑, AC↑, DEN↓, CNL↑, LLD↓, LLS↓, LLD/LLS↓

designed to specifically identify these relative anomalies associated with fractures. Therefore, the classification criteria are intended to indicate the trends in logging responses within fracture developed intervals, rather than serving as strict boundaries for fracture classification. The contribution of each logging curve was obtained using the Random Forest algorithm. The contribution is calculated based on the principle of Gini impurity reduction (Equation (14)), i.e., in each decision tree, when a feature is used for a node split, it reduces the class Gini impurity of that node. The sum of the reductions in impurity before and after the split is considered as the feature's contribution at that node (Equations (15) and (16)). In the Random Forest, the contributions of each feature across all trees are averaged and normalized to obtain the final contribution metric (Equation (17)) (Breiman, 2001). Therefore, the more frequently and effectively a feature is used to improve node purity, the higher its contribution. Finally, based on the impurity reduction calculation, the obtained contributions are as follows in order: LLD/LLS (0.214), CAL (0.189), AC (0.132), LLD & LLS (0.115), DEN (0.103), CNL (0.098), GR (0.091), and SP (0.058).

$$\text{Gini\_impurity} = 1 - \sum_{k=1}^{K} P_k^2 \tag{14}$$

$$\Delta I\left(X_j\right) = I_{\text{parent}} - \left(\frac{N_{\text{L}}}{N} I_{\text{L}} + \frac{N_{\text{R}}}{N} I_{\text{R}}\right) \tag{15}$$

$$\text{Importance}_{\text{tree}}\left(X_j\right) = \sum_{t \in \text{Nodes}(X_j)} \Delta I_t \tag{16}$$

$$\text{Importance}_{\text{RF}}\left(X_j\right) = \frac{1}{T} \sum_{t=1}^{T} \text{Importance}_{\text{tree}_t}\left(X_j\right) \tag{17}$$

In the formula, P_K represents the proportion of samples

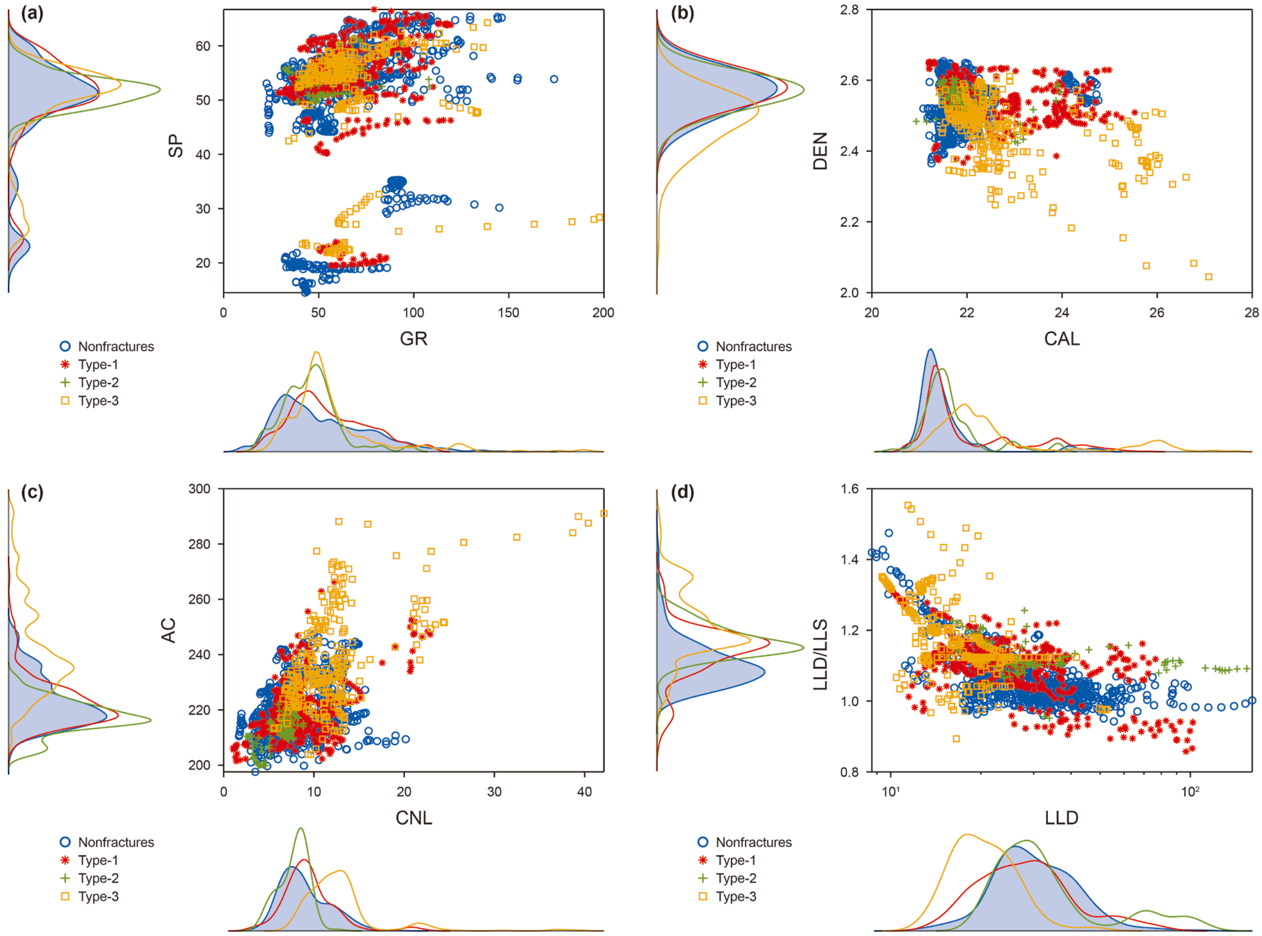

Fig. 11. Cross-plots of conventional logging sensitivity for fracture and nonfracture.

Table 3
Feature classification thresholds by conventional logs.

Name	Type-1	Type-2	Type-3
GR, API	/	/	/
SP, mv	/	/	/
CAL, cm	/	>21.4	>21.7
DEN, $g\cdot cm^{-3}$	/	/	<2.53
CNL, %	/	/	>9.2
AC, $\mu s\cdot m^{-1}$	/	/	>225.8
LLD, Ω·m	/	/	<25
LLS, Ω·m	<25	<25	<20
LLD/LLS	>1.07	>1.09 or <1	>1.12 or <1

belonging to the k-th class at the node; X_j represents the j-th feature used for splitting; t and T represent all nodes in a decision tree that utilize feature X_j and the total number of decision trees, respectively; $\Delta I\left(X_j\right)$ and ΔI_t represent the weighted average impurity of the child nodes and the impurity reduction brought by the split at the t-th node, respectively; N, N_L and N_R represent the number of samples at the parent node, left child node, and right child node, respectively; I_{parent}, I_L and I_R represent the impurity of the parent node, left child node, and right child node, respectively; $\mathrm{Importance}_{\mathrm{tree}_t}\left(X_j\right)$ and $\mathrm{Importance}_{\mathrm{RF}}\left(X_j\right)$ represent the importance of feature X_j in the t-th decision tree and final aggregated importance in the entire Random Forest model, respectively; $\mathrm{Nodes}\left(X_j\right)$ represents the set of all nodes in a single tree that use feature X_j as the splitting feature.

4.3. Validation for fracture identification using the CNN-attention-BiLSTM model

This paper proposes a deep learning intelligent identification method (CNN-Attention-BiLSTM) based on FWS logging-supplemented fracture labels for data balancing. Conventional logging curves were calibrated using core samples, image logging data, and virtual labels. The least contributive curve (SP) was excluded based on the contribution ranking of logging curves, while GR, CAL, DEN, CNL, AC, LLD, LLS, and LLD/LLS were selected as input features. Following the process described in section 3.2.1, a total of 4403 sequential samples were generated to construct a CNN-Attention-BiLSTM model for identifying three response types of high-angle fractures in tight sandstone reservoirs. It should be noted that although the GR log does not play a critical role in fracture identification itself, it serves as an important constraint on lithological information and must be input together with other curves for model construction and prediction. This allows the synchronous variations in AC, DEN, and other logs, which are controlled by lithology and sedimentary rhythms, to be identified and excluded by the model. For example, certain unconsolidated mudstone intervals may exhibit a significant increase in AC, a decrease in DEN, and substantial resistivity changes (as seen in the upper part of Fig. 10(a) and the middle part of Fig. 10(b)), or long-term variations in curves caused by progradation or retrogradation rhythms. This approach mitigates the influence of lithological effects and sedimentary rhythms on fracture identification. Furthermore, the Bayesian hyperparameter optimization method based on TPE (Tree-structured Parzen Estimator) is used to obtain the optimal parameters of the model (Table 4). It conducts directional exploration of hyperparameters by combining previous hyperparameter validations and acquires the optimal model scheme with fewer iterations. To ensure robustness, 5-fold cross-validation was employed: the original training dataset was randomly divided into five equally sized subsets. In each fold, four subsets were used for training and the remaining one for validation. This process was repeated five times, and the final performance metrics were averaged across all results.

The classification performance of the model can also be evaluated using Precision, Recall, Accuracy, and F1 score. The significance of these four metrics is as follows: Precision indicates the proportion of detected fractures/nonfracture labels that are true fractures/nonfracture labels; Recall indicates the proportion that true fracture or nonfracture labels are correctly identified; Accuracy indicates the overall probability of correct identification for both fracture and nonfracture within the entire dataset, i.e., the ratio of correctly identified fracture and nonfracture samples to the total number of samples; the F1 score is the harmonic mean of Precision and Recall, serving as a comprehensive metric for evaluating the model's identification performance. As shown in Fig. 12, the average precision values for nonfracture, Type-1, Type-2, and Type-3 are 96.8%, 92.6%, 94.0% and 94.7%, respectively. The average recall values are 96.5%, 94.4%, 91.0% and 93.4%, respectively. The overall accuracy is 96.5%, 94.4%, 91.0% and 93.4%, respectively. The average F1 scores are 96.5%, 92.4%, 93.6% and 93.5%, respectively.

For intelligent logging interpretation, generalization capability is an extremely important metric for evaluating model performance. Thus, this paper further validates the performance of the proposed model using two blind wells. As shown in Fig. 13, the second-to-last column represents fractures identified from core and image logging data, while the last column represents fractures predicted by the model. Green, red, and Orange represent Type-1, Type-2, and Type-3, respectively. The model successfully identified fractured intervals at depths of 3486–3492 m, 3503–3508 m, 3509–3513 m, and 3518–3520 m in Well A5 (Fig. 13(a)), as well as the interval at 3461–3470 m in Well A2 (Fig. 13(b)). Overall, the fracture identification results from the blind wells exhibit strong consistency with the actual fractured intervals, achieving an accuracy exceeding 90%. These findings demonstrate that the proposed model offers high accuracy in fracture identification for the H1 member of Lower Shihezi Formation in the Hangjinqi area, Ordos Basin.

4.4. Vertical distribution patterns of fractures

To investigate the vertical distribution patterns of fractures, we selected four wells with similar distances to the fault and comparable structural relief to form a cross-section (Fig. 14(a)), and the location of the section is shown in Fig. 1. Furthermore, fracture development in 79 wells across the study area was predicted. A sliding window method was employed, averaging data from five points (the target point plus two above and two below) to calculate fracture development probability. Subsequently, the ratio of fracture development intensity in each layer to the overall fracture development intensity in the well was computed. The fracture distribution in Fig. 14(a) and (b) reveals that H1-2 interval is the most intensely fractured vertically, followed by the H1-1 and H1-3 intervals, while the H1-4 interval shows the least fracturing. Notably, within the cross-section, Well A7 exhibits the highest fracture development, dominated by interbedded sandstone and mudstone layers vertically. Wells A2 and A9 show moderate fracture development, dominated by thick sandstone layers vertically. In contrast, Well A8 exhibits the lowest fracture development, with thick mudstone layers and poorly developed sandstones vertically. Accordingly, we statistically analyzed the thickness and frequency of sandstone layers in each layer across the 79 wells (Fig. 14(c)). It was found that although the sand-to-strata ratio of H1-2 interval is similar to that of H1-1 (both exceeding 0.6), H1-2 interval has a higher average number of sandstone layers, indicating more frequent sandstone-mudstone interbedding. In contrast, H1-3 and H1-4 intervals, which show significantly lower fracture development, have lower overall sand-to-strata ratios and fewer sandstone layers. This inter-well and inter-layer distribution pattern demonstrates that vertical fracture development is strongly controlled by the thickness of mechanical stratigraphy, with the fracture development intensity in the order: interbedded sandstone and mudstone layers, thick sandstone and thick mudstone, thick mudstone and poorly developed sandstones. Previous studies also suggest that mechanical stratigraphy thickness is a key factor controlling fracture development within individual layers; a thinner mechanical stratigraphy promotes more intensive fracturing (Zeng et al., 2020). In tight sandstone sequences, mechanical stratigraphy often correlates strongly with lithological stratigraphy. Under equivalent tectonic stress conditions, thin sandstone layers are more prone to fracturing due to stress concentration, resulting in higher fracture density.

Table 4
Parameters of the CNN-Attention-BiLSTM model.

Parameters	Value	Parameters	Value
LSTM layer	2	Dropout 1	0.11
Dropout layer	2	Dropout 2	0.32
Fully connected layer	1	Batch size	16
Dense units	64	Kernel size	3, 5, 7
LSTM units 1	256	Filters	64
LSTM units 2	112	Learning rate	0.002

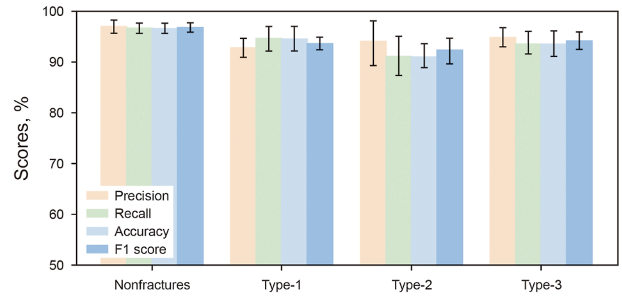

Fig. 12. Distribution of accuracy in 5-fold cross-validation for the CNN-attention-BiLSTM model.

5. Discussion

5.1. Logging response pattern extraction using the CNN-attention-BiLSTM model

This paper employs a combination of Conv1D and CSA to extract the logging response patterns of fractures, thereby providing superior pattern guidance for conventional fracture identification. As illustrated in Fig. 15, we utilize Principal Component Analysis (PCA) to project the high-dimensional features of each layer onto a two-dimensional plane. This visualization allows us to examine the classification effectiveness at various stages of the multi-layer progressive model, with a comparative baseline provided by a double-layer BiLSTM model trained on raw data. A higher clustering degree of the data points indicates a

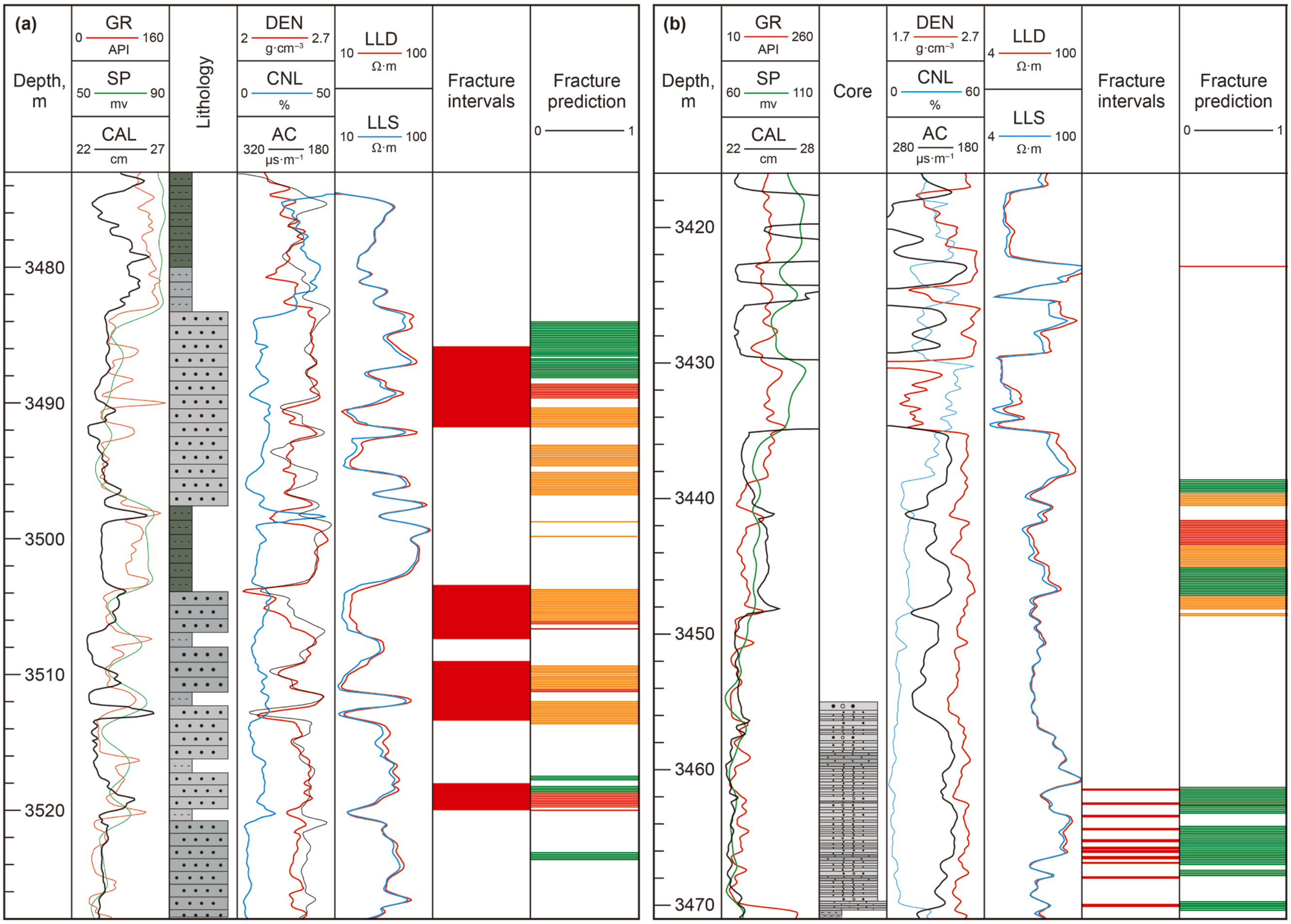

Fig. 13. Validation of fracture identification results in blind wells: **(a)** Well A5, **(b)** Well A2.

stronger feature abstraction capability at that specific layer. The results demonstrate that the Conv1D layer, building upon the raw curves, extracts multi-dimensional logging response patterns. Compared to the raw data, the feature distribution already begins to show partial class boundaries (Fig. 15(a) and (d)). Subsequently, CSA performs a weighted refinement of these patterns, enabling the model to learn discriminative features across channels, which results in a more distinct clustering (Fig. 15(b)). Ultimately, the CNN-Attention-BiLSTM model achieves a clear identification from the sequential data (Fig. 15(c)–(e) and (f)). In contrast to directly inputting raw data into a BiLSTM model, our approach more effectively learns the combined response patterns of fractures, playing a significant role in enhancing identification accuracy.

5.2. Comparison of identification results among FWS-augmented labels, random under-sampling, SMOTE oversampling, and transformer

This study conducted fracture identification using the proposed method, while also processing the dataset with the SMOTE and random under-sampling techniques. Corresponding DL-BiLSTM models were strictly established following the same workflow. Furthermore, we utilized the latest Transformer algorithm to construct a corresponding identification model, and accuracy calculations and blind well validation of the four methods were performed. The performance of the four methods was then compared. The proposed method achieved higher accuracy than the SMOTE oversampling method and significantly higher accuracy than the Transformer algorithm and random under-sampling (Fig. 16).

Well logging identification is inherently a modeling process based on limited sample data. When fracture labels are scarce, models tend to overfit the known data distribution, making it difficult for methods such as SMOTE oversampling to avoid the randomness of labels caused by an insufficient number of actual labels. This leads to the model being well in learning observed patterns but performing poorly in generalizing to unseen data (Saleh et al., 2025; Xu et al., 2025). Furthermore, the SMOTE algorithm carries the risk of generating a substantial amount of homogeneous fracture information. Its synthesized fracture data distribution fails to adequately represent the actual fracture distribution, which makes it difficult for the model to identify different fracture response patterns. The CNN-Attention-BiLSTM method incorporates Conv1D and CSA, enhancing the capability to extract well logging response patterns and optimize feature weights. And, it utilizes DL-BiLSTM to capture multi-scale sequential information of fractures, thereby substantially improving the model's performance in addressing complex fracture identification challenges from well logs. Consequently, based on the above discussion, the model validation process should not only focus on test set performance but more importantly on blind well verification. Using four different models, we obtained identification results from two blind wells. We found that the proposed method demonstrates significant improvement compared to the

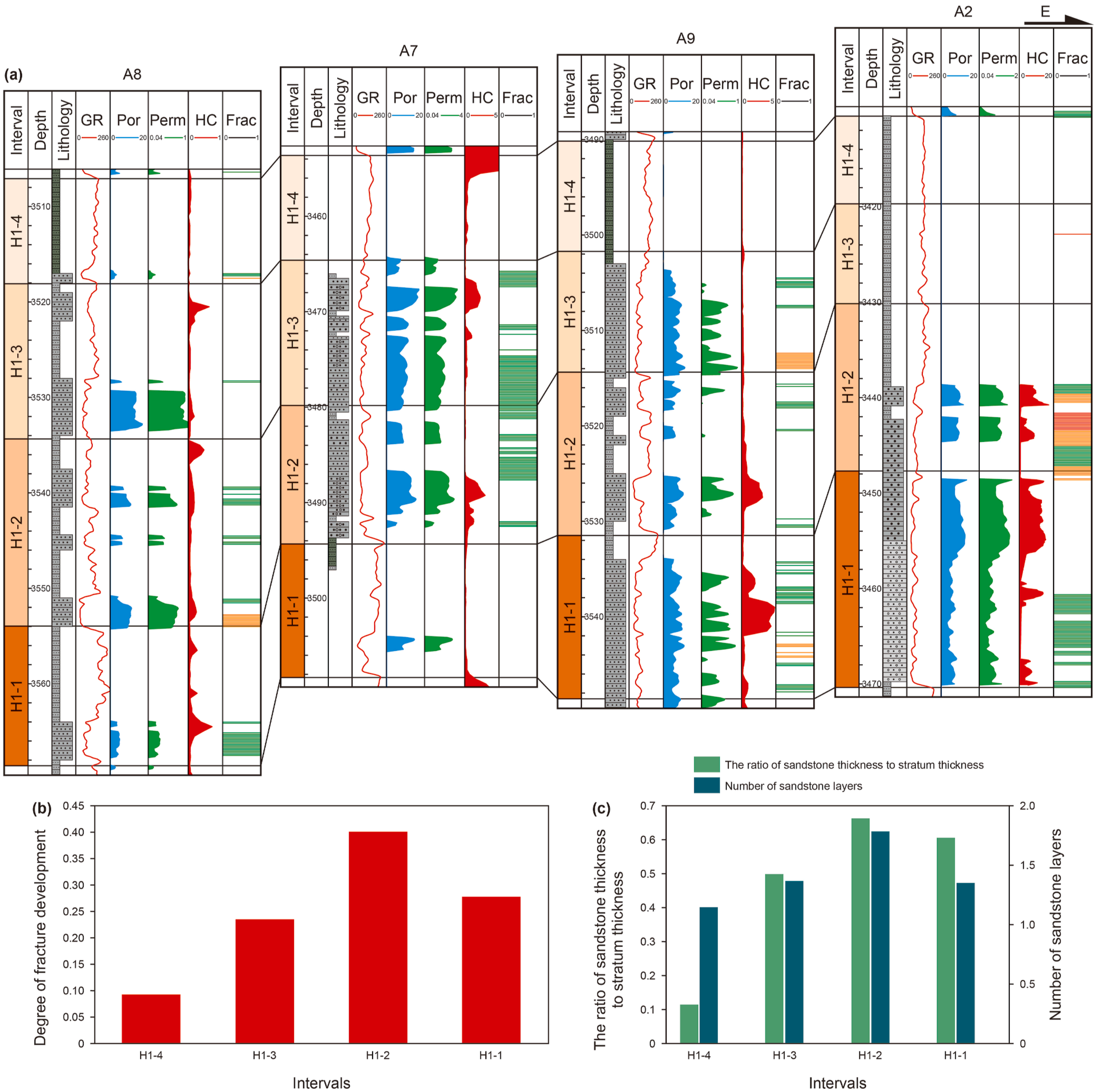

Fig. 14. Results of fracture identification. **(a)** Well cross-section composed of four wells with similar structural relief. Por: porosity, Perm: permeability, HC: total hydrocarbon content, Frac: fractures; **(b)** fracture development in the four intervals of the H1 member; **(c)** the ratio of sandstone thickness to stratum thickness, and the number of sandstone layers in the four intervals of the H1 member.

other three methods (Fig. 17). The fourth-to-last column in the figure shows the results obtained by the proposed method, the third-to-last column shows the results from the SMOTE algorithm, the second-to-last column shows the results from the random under-sampling algorithm, and the last column shows the results from the Transformer algorithm. It can be observed that the proposed method exhibits substantially better identification performance at intervals such as 3485–3492 m and 3509–3513 m in Well A5 (left), and 3462–3467 m in Well A2 (right). The ratio of the identified fracture length to the total length is used as the identification accuracy of the fracture length interval. The method proposed in this paper achieved improvements of 8.7%, 13.1% and 20.5% compared to Transformer, SMOTE and random under-sampling, respectively. Furthermore, this method also demonstrates superior performance in aligning fracture type identification with well logging response patterns. Therefore, compared to the other methods, the proposed method exhibits more robust identification capability and stronger generalization ability, which indicates better cross-well applicability.

5.3. *Influence of varying radial investigation depths of core, image logs, and different logging series on fracture identification*

In fracture identification using well logs, different logging series exhibit varying capabilities in reflecting radial fracture development due to their distinct depths of investigation. Within

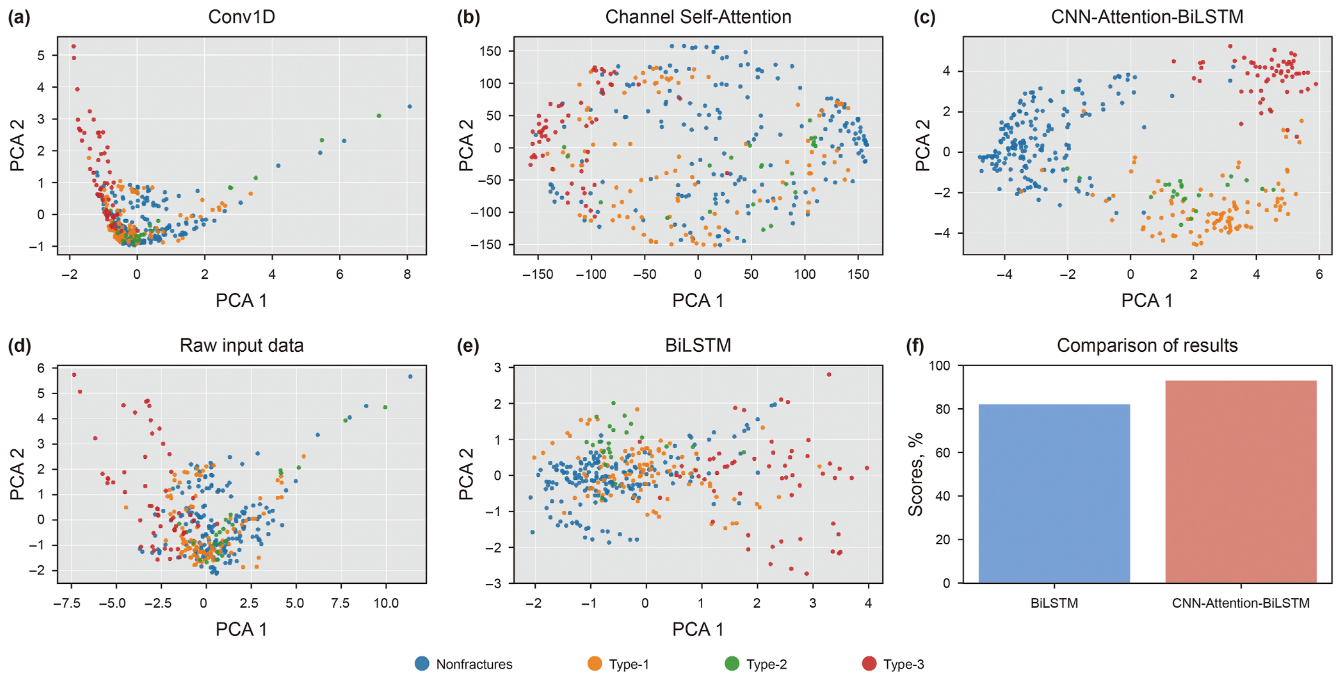

Fig. 15. Performance comparison of pattern extraction and identification between BiLSTM and CNN-attention-BiLSTM.

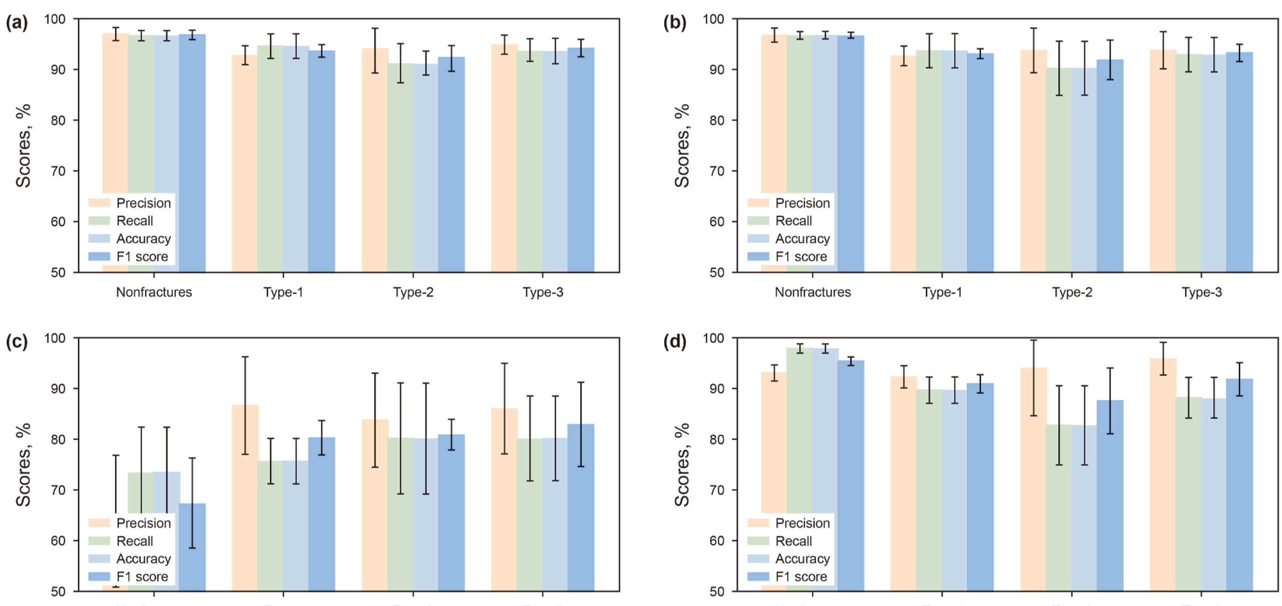

Fig. 16. Distribution of 5-fold cross-validation accuracy for different methods: **(a)** the proposed model, **(b)** SMOTE oversampling method, **(c)** random under-sampling method, **(d)** Transformer method.

FWS logging series, the investigation depth of the compressional wave is approximately 1 m or greater, the shear wave ranges from 0.3 m to 0.6 m, while the Stoneley wave is limited to only 0.2–0.3 m (He and Wang, 2012; Lai et al., 2025). In fracture identification using conventional logs, aside from deep lateral resistivity, which can exceed 1 m, the investigation depth of other logging series is only about 0.1–0.5 m (Lai et al., 2024). Furthermore, based on the analysis of the aforementioned logging response patterns, the mud invasion profile shows that fracture-developed intervals correspond to a flushed zone width of 0–0.5 m and a transition zone width of 0.5–1.2 m (Fig. 7). In intervals without fractures (3464–3467 m), both FWS and conventional logging responses tend to reflect characteristics of high-quality reservoirs, with relatively shallow flushed and transition zones (Fig. 7). Therefore, inconsistencies in identification scale resulting from varying logging investigation depths must be considered during the establishment of virtual labels and the logging-based identification process. During the supplementation of fracture labels using FWS, we introduced SHAP values to characterize the importance of each logging curve. According to the importance analysis, the compressional wave contributes the least to the classification outcome. And the depth of investigation of the compressional

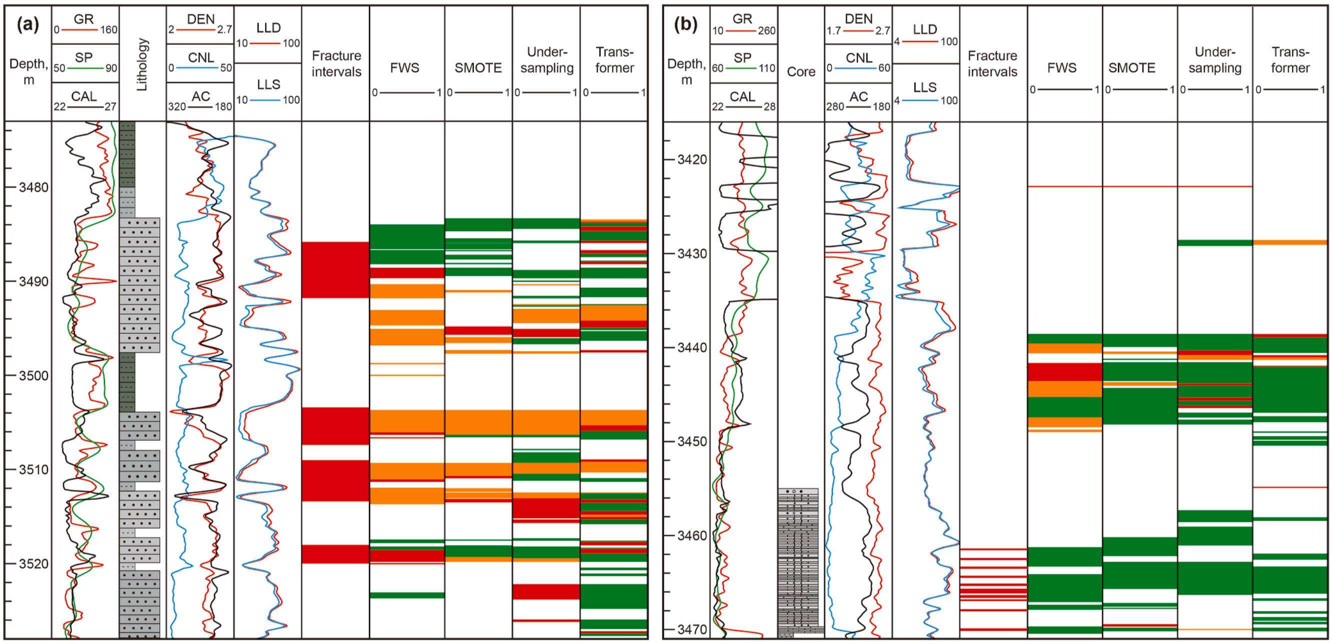

Fig. 17. Comparison of fracture identification performance by different methods. **(a)** Well A5; **(b)** Well A2.

wave in array acoustic logs differs significantly from that of other logging series. Additionally, the compressional wave is prone to interference from bedding planes (and fractures). Consequently, the DECC curve was ultimately excluded, and only other features were retained for the screening and expansion of labeled data, ensuring consistency in the scale of investigation.

The absence of fractures in core and image logs does not necessarily indicate the lack of fractures in adjacent formations near the wellbore (Lyu et al., 2016). In fact, due to the limited radial investigation depth of core and image logs, the probability of vertical wells encountering high-angle fractures in the study area is relatively low. Conventional logging, with its larger radial investigation depth, provides an indication of fracture development around the wellbore. Therefore, the blind wells likely contain more identified fractures than those detected by core and image logs (Lyu et al., 2016). For Well A5, compared to the actual fractures identified by image logs, the model not only accurately identified the actual fractured intervals but also identified two additional fractured intervals (3493–3498 m and 3523.5–3524 m). The interval 3493–3498 m shows clear Type-3 fracture response characteristics in conventional logs: caliper expansion, increased AC, decreased DEN, and decreased deep and shallow resistivity (LLD, LLS) (Fig. 13(a)). The interval 3522.5–3524 m shows clear Type-1 fracture response characteristics in conventional logs: significant separation between deep and shallow resistivities with minimal changes in the three porosity curves and caliper (Fig. 13(a)). Similarly, for Well A2 (Fig. 13(b)), the model not only accurately identified fractures actually observed in cores but also successfully detected fracture intervals in the section without core and image logs (3438–3448 m), which exhibited clear characteristic responses. Moreover, the three types of fractures identified within this interval showed great consistency in their respective response features with the logging response patterns illustrated in Fig. 10. Thus, the additional fractured intervals identified by the method, beyond the actual fracture labels, all display clear logging responses. This demonstrates both the reliability of the model's identification performance and generalization capability, and also reveals more authentic information about fracture development near the wellbore. The method significantly supplements fracture information provided by core and image logs, offering greater reference value for fracture characterization.

5.4. Fracture development patterns and geological significance

The method proposed in this paper, based on the traditional intelligent fracture identification, establishes a deep connection between the logging response type of fractures and their development intensity. The results show that fractures are do not only developed in large sets of tight sandstone, but also distributed in various forms across different regions. In the sandstone sections with good porosity and permeability (Well A7 in Fig. 14), fractures often appear as single fractures or multiple parallel fractures, with relatively less development (Types-1 and 2). In contrast, in sandstone intervals with poor porosity and permeability (Wells A2 and A9 in Fig. 14), the rock exhibits high brittleness, leading to the development of dense fracture networks (Type-3). On vertical profiles, the relationship between fracture development and gas logging responses is not simply a clear correspondence but presents multiple scenarios. In thick-bedded coarse-grained sandstone layers, which serve as high-gas-bearing zones, fractures are relatively less developed. Conversely, fractures are more commonly developed in the overlying and underlying intervals, connecting the upper and lower gas layers (Wells A2 and A7 in Fig. 14). In tight sandstone intervals, however, high-gas-bearing zones show a clear correspondence with fracture development intervals (Well A9 in Fig. 14). Therefore, based on the prediction results, this study statistically analyzed the average fracture development intensity within the single-layer thickness of each well. The reservoirs were classified into tight sandstone and high-quality sandstone using a porosity threshold of 10% and a permeability threshold of 1 mD from logging interpretation (Fig. 18(b)). The statistical results indicate that in tight sandstone, fracture development to some extent promotes high production, showing a positive correlation between cumulative well productivity and average fracture development intensity. In high-quality sandstone, although the average fracture development intensity is

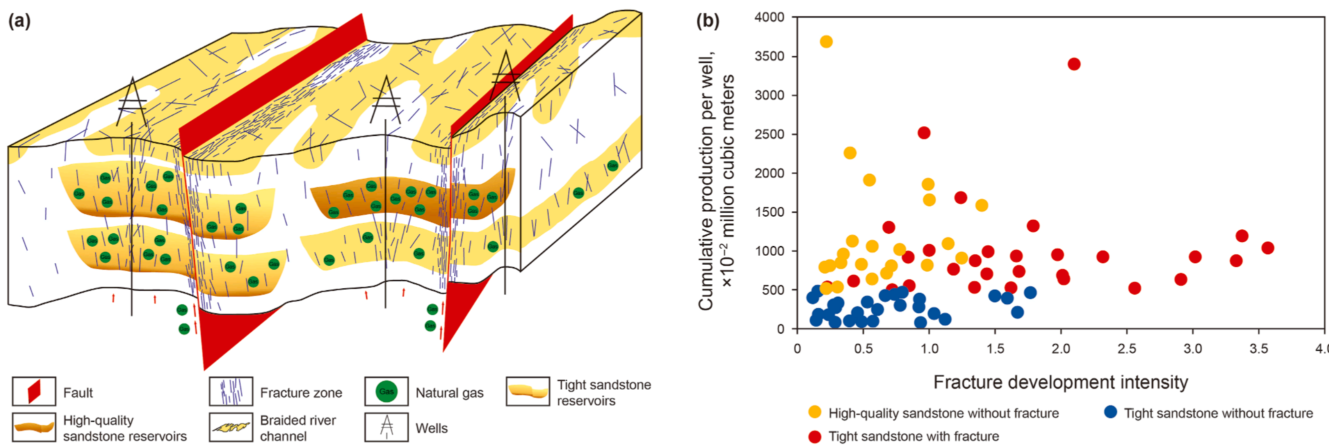

Fig. 18. Fracture development pattern and its relationship with productivity.

relatively low, higher productivity is still achieved, indicating a less direct relationship with fractures. Combined with vertical fracture identification results, it can be further inferred that in intervals dominated by high-quality sandstone reservoirs, fractures may primarily serve to connect upper and lower gas layers, acting as preferential migration pathways. In intervals dominated by tight sandstone, fractures play a more significant role in enhancing matrix reservoir properties, thereby facilitating the migration and accumulation of natural gas. Based on the above analysis, the model illustrating the relationship between fracture development and tight sandstone reservoir matching is established (Fig. 18(a)).

The intelligent fracture identification method proposed in this paper provides valuable guidance for understanding the development patterns, migration and accumulation mechanisms, and reservoir connectivity. It also offers a robust basis for selecting perforation zones in subsequent operations, demonstrating significant scientific and practical value.

6. Conclusions

(1) To address the similarities and differences in the response mechanisms and sensitivity of fractures in the Hangjinqi area of northern Ordos Basin, this study used image logs, FWS logs, and conventional logs to clarify distinct response types and characteristics of natural fractures. An intelligent fracture identification method based on CNN-Attention-BiLSTM was proposed. The method effectively leverages a combination of Conv1D and Channel Self-Attention to capture multi-dimensional logging response patterns and optimally assign weights to them, achieving clearer response patterns. Furthermore, it utilizes a DL-BiLSTM network to mitigate the impact of sedimentary cycles on log-based identification and captures both short- and long-term sequential dependencies of fracture responses through different network layers. This approach successfully overcomes the challenges associated with complex responses under multi-logging modes in fracture identification. Compared with the latest Transformer method, the proposed method demonstrates a significant improvement in the classification accuracy of minority-class fracture samples in the test set. The approach exhibits superior performance in identifying fractured intervals in blind wells, along with an 8.7% improvement in identification accuracy.

(2) Compared to the SMOTE oversampling and random undersampling methods, the approach based on FWS for supplementing fracture labels for data balancing not only maintains high classification accuracy on the test set but also effectively addresses the insufficient model generalization caused by artificially generated or eliminated label data. The results further demonstrate that the proposed method significantly improves the identification of fractured intervals, achieving accuracy increases of 13.1% and 20.5%, respectively, compared to the other two methods.

(3) The cross-well profile reveals that fractures are the most developed in the H1-2 interval, followed by the H1-1 and H1-3 intervals, while the H1-4 interval is the least developed. The fracture distribution pattern is clearly controlled by sedimentary rhythms, demonstrating that the more frequent interbedding of lithologies in the vertical profile corresponds to the more intensive fracture development. The density of fracture development also varies. Interconnected fractures tend to develop in tight sandstone intervals, where they serve to enhance reservoir quality. In contrast, high-quality sandstone intervals typically contain isolated fractures, which serve as conduits connecting upper and lower gas zones.

CRediT authorship contribution statement

Bao-Yu Liang: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Data curation, Conceptualization. **Lian-Bo Zeng:** Project administration, Methodology. **Shao-Qun Dong:** Project administration, Conceptualization. **Xue-Qun Tan:** Resources. **Ji-Bo Ren:** Writing – original draft, Methodology. **Hong-Tao Li:** Resources. **Zi-Yi Yang:** Software, Formal analysis. **Shi-Qiang Liu:** Investigation, Data curation. **Zhen Wang:** Writing – review & editing, Supervision.

Data availability

The authors do not have permission to share data.

Declaration of competing interest

Lian-Bo Zeng is an editorial board member for Petroleum Science and was not involved in the editorial review or the decision to publish this article. All authors declare that there are no competing interests.

Acknowledgements

This work was financially supported by the National Natural Science Foundation of China (No. 42002134).

References

Anees, A., Zhang, H.C., Ashraf, U., et al., 2022a. Identification of favorable zones of gas accumulation via fault distribution and sedimentary facies: insights from Hangjinqi area, northern Ordos Basin. Front. Earth Sci. 9, 822670. https://doi.org/10.3389/feart.2021.822670.

Anees, A., Zhang, H.C., Ashraf, U., et al., 2022b. Sand-ratio distribution in an unconventional tight sandstone reservoir of Hangjinqi area, Ordos Basin: acoustic impedance inversion-based reservoir quality prediction. Front. Earth Sci. 10, 1018105. https://doi.org/10.3389/feart.2022.1018105.

Anees, A., Zhang, H.C., Ashraf, U., et al., 2022c. Sedimentary facies controls for reservoir quality prediction of lower Shihezi member-1 of the Hangjinqi area, Ordos Basin. Minerals 12 (2), 126–144. https://doi.org/10.3390/min12020126.

Ashraf, U., Anees, A., Shi, W.Z., et al., 2022. Estimation of porosity and facies distribution through seismic inversion in an unconventional tight sandstone reservoir of Hangjinqi area, Ordos Basin. Front. Earth Sci. 10, 1014052. https://doi.org/10.3389/feart.2022.1014052.

Bagheri, H., Falahat, R., 2022. Fracture permeability estimation utilizing conventional well logs and flow zone indicator. Petrol. Res. 7 (3), 357–365. https://doi.org/10.1016/j.ptlrs.2021.11.004.

Bahdanau, D., Cho, K., Bengio, Y., 2014. Neural machine translation by jointly learning to align and translate arXiv: 1409.0473 [cs.CL]. https://doi.org/10.48550/arXiv.1409.0473.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32. https://doi.org/10.1023/A:1010933404324.

Che, X.H., Zhao, T., Qiao, W.X., et al., 2020. Fracture identification and evaluation based on multi-pole acoustic logging. Oil Gas Geol. 41 (6), 1263–1272. https://doi.org/10.11743/ogg20200614 (in Chinese).

Deng, S.G., Wang, X.C., Fan, Y.R., 2006. Response of dual laterolog to fractures in fractured carbonate formation and its interpretation. Earth Sci. 31 (6), 846–850. https://doi.org/10.3321/j.issn:1000-2383.2006.06.013 (in Chinese).

Dong, S.Q., Wang, Z.Z., Zeng, L.B., 2016. Lithology identification using kernel fisher discriminant analysis with well logs. J. Petrol. Sci. Eng. 143, 95–102. https://doi.org/10.1016/j.petrol.2016.02.017.

Dong, S.Q., Zeng, L.B., Lyu, W.Y., et al., 2020a. Fracture identification and evaluation using conventional logs in tight sandstones: a case study in the Ordos Basin, China. Energy Geosci. 1 (3–4), 115–123. https://doi.org/10.1016/j.engeos.2020.06.003.

Dong, S.Q., Zeng, L.B., Lyu, W.Y., et al., 2020b. Fracture identification by semi-supervised learning using conventional logs in tight sandstones of Ordos Basin, China. J. Nat. Gas Sci. Eng. 76, 103131. https://doi.org/10.1016/j.jngse.2019.103131.

Dong, S.Q., Zeng, L.B., Wang, L.T., et al., 2024. Fracture identification in shale reservoir using a deep learning method: Chang 7 reservoirs, Triassic Yanchang Formation. Geoenergy Sci. Eng. 238, 212853. https://doi.org/10.1016/j.geoen.2024.212853.

Graves, A., 2013. Generating sequences with recurrent neural networks arXiv: 1308.0850 [cs.NE]. https://doi.org/10.48550/arXiv.1308.0850.

Gu, Y.F., Zhang, D.Y., Lin, Y.B., et al., 2021. Data-driven lithology prediction for tight sandstone reservoirs based on new ensemble learning of conventional logs: A demonstration of a Yanchang member, Ordos Basin. J. Petrol. Sci. Eng. 207, 109292. https://doi.org/10.1016/j.petrol.2021.109292.

Guo, P., Ren, D.S., Xue, Y.H., 2019. Simulation of multi-period tectonic stress fields and distribution prediction of tectonic fractures in tight gas reservoirs: A case study of the Tianhuan Depression in western Ordos Basin. China Mar Petrol. Geol. 109, 530–546. https://doi.org/10.1016/j.marpetgeo.2019.06.026.

He, S.B., Meng, K., Wang, C., et al., 2023. Fracture identification using conventional logs in ultra-low permeability sandstone reservoirs: A case study of the Chang 6 member of the Ordos Basin, China. Minerals 13 (2), 297–308. https://doi.org/10.3390/min13020297.

He, X., Wang, X.M., 2012. Radial detection depths of borehole stoneley modes. J. Acoust. Soc. Am. 131 (4_Suppl. ment), 3369. https://doi.org/10.1121/1.4708703.

Hemmatian, J., Hajizadeh, R., Nazari, F., 2025. Addressing imbalanced data classification with cluster-based reduced noise SMOTE. PLoS One 20 (2), e0317396. https://doi.org/10.1371/journal.pone.0317396.

Jiang, L., Qiu, Z., Wang, Q., et al., 2016. Joint development and tectonic stress field evolution in the southeastern Mesozoic Ordos Basin, west part of North China. J. Asian Earth Sci. 127, 47–62. https://doi.org/10.1016/j.jseaes.2016.06.017.

Lai, J., Chen, K.J., Xin, Y., et al., 2021a. Fracture characterization and detection in the deep cambrian dolostones in the Tarim Basin, China: insights from borehole image and sonic logs. J. Petrol. Sci. Eng. 196, 107659. https://doi.org/10.1016/j.petrol.2020.107659.

Lai, J., Dang, W.L., Su, Y., et al., 2025. Acoustic logging: Geological and engineering applications. Geol. Rev. 71 (3), 1046–1072. https://doi.org/10.16509/j.georeview.2024.09.091 (in Chinese).

Lai, J., Su, Y., Xiao, L., et al., 2024. Application of geophysical well logs in solving geologic issues: Past, present and future prospect. Geosci. Front. 15 (3), 101779. https://doi.org/10.1016/J.GSF.2024.101779.

Lai, J., Wang, G.W., Pang, X.J., et al., 2021b. Well logging geology: its past, present, and future. — on the occasion of publishing the second edition of well logging geology. Geol. Rev. 67 (6), 1804–1828. https://doi.org/10.16509/j.georeview.2021.09.021 (in Chinese).

Lan, W.Z., Tao, Y.H., Liang, B., et al., 2025. Shear wave velocity prediction based on bayesian-optimized multi-head attention mechanism and CNN-BiLSTM. Comput. Geosci. 195, 105787. https://doi.org/10.1016/j.cageo.2024.105787.

LeCun, Y., Bottou, L., Bengio, Y., et al., 1998. Gradient-based learning applied to document recognition. Proc. IEEE 86 (11), 2278–2324. https://doi.org/10.1109/5.726791.

Li, H.T., 2025. Microfacies analysis of braided river deposits based on seismic-to-well ties: A case study of the Jin-77 well block in the dongsheng gas field, Ordos Basin. Oil Gas Geol. 46 (1), 91–107. https://doi.org/10.11743/ogg20250107 (in Chinese).

Li, Y., Gao, X.D., Meng, S.Z., et al., 2019. Diagenetic sequences of continuously deposited tight sandstones in various environments: A case study from upper Paleozoic sandstones in the Linxing area, eastern Ordos Basin, China. AAPG Bull. 103 (11), 2757–2783. https://doi.org/10.1306/03061918062.

Li, Z.L., Fan, C.Y., Sun, B., et al., 2023. Characteristics, logging identification and major controlling factors of bedding-parallel fractures in tight sandstones. Geoenergy Sci. Eng. 228, 211956. https://doi.org/10.1016/j.geoen.2023.211956.

Liu, J.S., Ding, W.L., Dai, J.S., et al., 2018. Quantitative multiparameter prediction of fault-related fractures: A case study of the second member of the Funing Formation in the Jinhu Sag, Subei Basin. Pet. Sci. 15 (3), 468–483. https://doi.org/10.1007/s12182-018-0240-3.

Lu, G.Q., Dong, S.Q., Huang, L.L., et al., 2023. Fracture intelligent identification using well logs of continental shale oil reservoir of Fengcheng Formation in Mahu Sag, Junggar Basin. Earth Sci. 48 (7), 2690–2702. https://doi.org/10.3799/dqkx.2022.409 (in Chinese).

Lyu, W.Y., Zeng, L.B., Liu, Z.Q., et al., 2016. Fracture responses of conventional logs in tight-oil sandstones: A case study of the Upper Triassic Yanchang Formation in southwest Ordos Basin, China. AAPG Bull. 100 (9), 1399–1417. https://doi.org/10.1306/04041615129.

Qin, S., Shi, W.Z., Wang, R., et al., 2022a. Characteristics of tight sandstone reservoirs and their controlling factors of He-1 member in Hangjinqi Block, Ordos Basin. Earth Sci. 47 (5), 1604–1618. https://doi.org/10.3799/dqkx.2022.007 (in Chinese).

Qin, S., Wang, R., Shi, W.Z., et al., 2022b. Diverse effects of intragranular fractures on reservoir properties, diagenesis, and gas migration: insight from Permian tight sandstone in the Hangjinqi area, north Ordos Basin. Mar. Petrol. Geol. 137, 105526. https://doi.org/10.1016/j.marpetgeo.2022.105526.

Qin, S., Wang, R., Shi, W.Z., et al., 2023. Intragranular fractures within tight sandstone reservoirs in the Hangjinqi area, northern Ordos Basin, China: Morphology, intensity and controlling factors. Geoenergy Sci. Eng. 231, 212376. https://doi.org/10.1016/j.geoen.2023.212376.

Saleh, K., Mabrouk, W.M., Metwally, A., 2025. Machine learning model optimization for compressional sonic log prediction using well logs in Shahd SE field, Western Desert, Egypt. Sci. Rep. 15, 14957. https://doi.org/10.1038/s41598-025-97938-9.

Shan, L.Q., Liu, Y.C., Tang, M., et al., 2021. CNN-BiLSTM hybrid neural networks with attention mechanism for well log prediction. J. Petrol. Sci. Eng. 205, 108838. https://doi.org/10.1016/j.petrol.2021.108838.

Sun, L.D., Zou, C.N., Jia, A.L., et al., 2019. Development characteristics and orientation of tight oil and gas in China. Petrol. Explor. Dev. 46 (6), 1015–1026. https://doi.org/10.11698/PED.2019.06.01.

Tan, X.F., Huang, Y.T., Lei, T., et al., 2023. Sedimentary characteristics of sandy braided river deposits and factors controlling their deposition: A case study of the lower Shihezi Formation in the northern Ordos Basin, China. Geoenergy Sci. Eng. 228, 211932. https://doi.org/10.1016/j.geoen.2023.211932.

Tang, J., Zhang, C.G., Xin, Y., 2017. A fracture evaluation by acoustic logging technology in oil-based mud: A case from tight sandstone reservoirs in Keshen area of Kuqa Depression, Tarim Basin, NW China. Petrol. Explor. Dev. 44 (3), 418–427. https://doi.org/10.1016/S1876-3804(17)30048-4.

Tang, X.M., Bolshakov, A., Patterson, D., 2011. Integrated acoustic evaluation of reservoir fractures: From borehole out into the Formation. Petrophysics 52 (3), 199–206.

Vaswani, A., Shazeer, N., Parmar, N., et al., 2017. Attention is all you need. Adv. Neural Inf. Process. Syst. 30, 5998–6008.

Wang, Q.F., Deng, J., Huang, D.H., et al., 2006. Tectonic constraints on the transformation of Paleozoic framework of uplift and depression in the Ordos area. Acta Geol. Sin. Engl. Ed. 80 (6), 944–953. https://doi.org/10.1111/j.1755-6724.2006.tb00315.x.

Wang, R., Shi, W.Z., Xie, X.Y., et al., 2020. Clay mineral content, type, and their effects on pore throat structure and reservoir properties: Insight from the Permian tight sandstones in the Hangjinqi area, north Ordos Basin, China. Mar. Petrol. Geol. 115, 104281. https://doi.org/10.1016/j.marpetgeo.2020.104281.

Wang, S., Wang, G.W., Li, D., et al., 2022. Comparison between double caliper, imaging logs, and array sonic log for determining the in-situ stress direction: A case study from the ultra-deep fractured tight sandstone reservoirs, the Cretaceous Bashijiqike Formation in Keshen8 region of Kuqa Depression, Tarim Basin, China. Pet. Sci. 19 (6), 2601–2617. https://doi.org/10.1016/j.petsci.2022.08.035.

Wang, S., Wang, G.W., Wang, H.Z., et al., 2025. Logging evaluation of acoustic anisotropy and its relationship with "sweet spots" in lacustrine shale oil reservoirs: The Fengcheng Formation of the Mahu Sag, China. Pet. Sci. 22 (8), 3133–3151. https://doi.org/10.1016/j.petsci.2025.05.014.

Wu, X.Q., Ni, C.H., Liu, Q.Y., et al., 2017. Genetic types and source of the upper Paleozoic tight gas in the Hangjinqi area, northern Ordos Basin, China. Geofluids 2017, 4596273. https://doi.org/10.1155/2017/4596273.
Xiao, Z.K., Ding, W.L., Liu, J.S., et al., 2019. A fracture identification method for low-permeability sandstone based on R/S analysis and the finite difference method: A case study from the Chang 6 reservoir in Huaqing oilfield, Ordos Basin. J. Petrol. Sci. Eng. 174, 1169–1178. https://doi.org/10.1016/j.petrol.2018.12.017.
Xie, J., Hao, X.F., Zhang, Y.P., et al., 2023. Fracture identification and characteristics of carbonate underground gas storage: An example from the eastern area of Sulige gas field, Ordos Basin, China. Sci. Rep. 13, 22446. https://doi.org/10.1038/s41598-023-50072-w.
Xu, L.M., Zhou, L.F., Zhang, Y.K., et al., 2006. Characteristics and tectonic setting of tectono-stress field of Ordos Basin. Geotect. Metallogenia 30 (4), 455–462. https://doi.org/10.3969/j.issn.1001-1552.2006.04.007 (in Chinese).
Xu, N.N., Zhang, S.P., Wang, Y.S., et al., 2022. Diagenesis and pore genesis of tight sandstone in the Permian lower Shihezi Formation, northern Ordos Basin. Acta Sedimentol. Sin. 40 (2), 422–434. https://doi.org/10.14027/j.issn.1000-0550.2021.027 (in Chinese).
Xu, Q.H., Shi, W.Z., Xie, X.Y., et al., 2018. Inversion and propagation of the Late Paleozoic Porjianghaizi Fault (North Ordos Basin, China): Controls on sedimentation and gas accumulations. Mar. Petrol. Geol. 91, 706–722. https://doi.org/10.1016/j.marpetgeo.2018.02.003.
Xu, X.Y., 2020. Study on Faults Structure and Reservoir Control in Ordos Basin. Doctoral Dissertation. China University of Petroleum (East China), Qingdao.
Xu, Z.H., Zheng, B.Y., Liu, B., et al., 2025. Stratigraphic correlation of well logs using geology-informed deep learning networks. Processes 13 (5), 1288. https://doi.org/10.3390/pr13051288.
Yang, W.B., Xia, K.W., Fan, S.R., 2023. Oil logging reservoir recognition based on TCN and SA-BiLSTM deep learning method. Eng. Appl. Artif. Intell. 121, 105950. https://doi.org/10.1016/j.engappai.2023.105950.
Zeng, L.B., Ke, S.Z., Liu, Y., 2010. Research Methods for Fracture Characterization in Low-Permeability Oil and Gas Reservoirs. Petroleum Industry Press, Beijing.
Zeng, L.B., Lyu, P., Qu, X.F., et al., 2020. Multi-scale fractures in tight sandstone reservoirs with low permeability and geological conditions of their development. Oil Gas Geol. 41 (3), 449–454. https://doi.org/10.11743/ogg20200301 (in Chinese).
Zhang, H.Y., Wu, W.S., Chen, Z.X., et al., 2024. Well logs reconstruction of petroleum energy exploration based on bidirectional long short-term memory networks with a PSO optimization algorithm. Geoenergy Sci. Eng. 239, 212975. https://doi.org/10.1016/j.geoen.2024.212975.
Zhang, S.L., Yan, J.P., Cai, J.P., et al., 2021. Fracture characteristics and logging identification of lacustrine shale in the Jiyang depression, Bohai Bay Basin, eastern China. Mar. Petrol. Geol. 132, 105192. https://doi.org/10.1016/j.marpetgeo.2021.105192.
Zhao, L., 2022. Development characteristics of microfractures in tight sandstone reservoir and its influence on physical properties: A case study of Shiligiahan zone in Hangjinqi. Petrol. Reserv. Eval. Dev. 12 (2), 285–291. https://doi.org/10.13809/j.cnki.cn32-1825/te.2022.02.003 (in Chinese).
Zhao, X., Wang, G.W., Li, Y.F., et al., 2025. Development characteristics and logging identification of natural fractures in shale oil reservoirs of the Yingxiongling area, Qaidam Basin. Nat. Gas Geosci. 36 (4), 713–733. https://doi.org/10.11764/j.issn.1672-1926.2024.09.009 (in Chinese).
Zhu, H.T., Chen, K.Y., Liu, K.Y., et al., 2008. A sequence stratigraphic model for reservoir sand-body distribution in the Lower Permian Shanxi Formation in the Ordos Basin, northern China. Mar. Petrol. Geol. 25 (8), 731–743. https://doi.org/10.1016/j.marpetgeo.2008.03.007.
Zou, C.N., Yang, Z., He, D.B., et al., 2018. Theory, technology and prospects of conventional and unconventional natural gas. Petrol. Explor. Dev. 45 (4), 575–587. https://doi.org/10.11698/PED.2018.04.04.