



OPEN ACCESS

EDITED BY

Paolo Capuano,
University of Salerno, Italy

REVIEWED BY

Randel Tom Cox,
University of Memphis, United States
Diana Núñez,
Complutense University of Madrid, Spain

*CORRESPONDENCE

Chenqiang Zhang,
✉ 3340952632@qq.com

RECEIVED 27 May 2025

ACCEPTED 18 July 2025

PUBLISHED 08 August 2025

CITATION

Lei Y, Zhang C, Wu W, Chen M, Wen X, He X,
Bai C, Qin S, Li Y and Wang L (2025) 3D fault
detection method using TransVNet.
Front. Earth Sci. 13:1635344.
doi: 10.3389/feart.2025.1635344

COPYRIGHT

© 2025 Lei, Zhang, Wu, Chen, Wen, He, Bai,
Qin, Li and Wang. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC
BY\)](#). The use, distribution or reproduction in
other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

3D fault detection method using TransVNet

Yang Lei¹, Chenqiang Zhang^{1*}, Wenjing Wu¹, Mingchun Chen^{1,2},
Xiaotao Wen², Xilei He², Chenggang Bai¹, Siping Qin¹, Ying Li¹
and Lijing Wang¹

¹Sinopec Geophysical Corporation Nanfang Branch, Chengdu, China, ²The College of Geophysics,
Chengdu University of Technology, Chengdu, China

Introduction: In seismic structural interpretation, fault detection plays a crucial role as it serves as the foundation and key step for identifying favorable oil and gas zones. Currently, many re-searchers are utilizing deep learning for automated fault detection. However, the accuracy and continuity of predictions generated by existing convolutional neural networks (CNNs) on real seismic data fail to meet practical production requirements.

Methods: To address this issue, we integrated the Transformer architecture into the V-Net framework, proposing a fault detection method based on the TransVNet network. This approach utilizes semantic segmentation technology to generate a fault probability volume by assessing the likelihood of each data point in the input dataset being part of a fault.

Results: For comparison, we referenced the classical U-Net network and the recently proposed TransUNet network, validating the feasibility of our method through theoretical seismic data. Subsequently, we applied the TransVNet network to actual seismic data. Without employing transfer learning, the fault detection results demonstrate that our proposed method exhibits superior fault detection capability, higher prediction accuracy, and better continuity compared to existing approaches.

Discussion: The method proposed in this article demonstrates that deep learning can be applied to fault detection in complex regions, which can enhance the accuracy and continuity of fault detection.

KEYWORDS

fault detection, seismic interpretation, convolutional neural network, TransVNet, deep learning

1 Introduction

The development of fault detection methods has evolved over time, leading to a range of conventional techniques such as coherence cube technology (Bahorich and Farmer, 1995; Gersztenkorn and Marfurt, 1999), curvature attributes (Al-Dossary and Marfurt, 2006; Yin et al., 2014), ant tracking (Colomi et al., 1992), and attribute fusion (Yuan et al., 2022), as well as deep learning-based approaches.

Early end-to-end fault detection networks relied on deep convolutional neural network (CNN) models (Xi et al., 2018). Over time, U-Net has become the dominant architecture for fault detection. The U-Net network, originally proposed by Ronneberger et al. (2015), significantly improved segmentation accuracy. Oktay et al. (2018) enhanced U-Net by introducing Attention Gate (AG) structures at the end of each skip connection, resulting

in the AG-UNet network (Oktay et al., 2018). Wu et al. (2019) applied the U-Net neural network to fault detection, achieving consistent resolution between input and out-put layers (Wu et al., 2019). The U-Net architecture has demonstrated strong advantages in fault detection, leading to numerous subsequent variants. Gao et al. (2022) proposed a nested residual U-Net with improved performance (Gao et al., 2022). While Shen et al. (2022) combined wavelet transform with CNN for enhanced results (Shen et al., 2022). Lyu et al. (2022) proposed the U-SegNet network specifically for buried-hill fault detection, which effectively reduce the case of fault miss and false detection (Lyu et al., 2022). Sun et al. (2023) applied the EAREsU-Net to reduce the semantic gap between the en-coder and decoder (Sun et al., 2023). Milletari et al. (2016) introduced the V-Net network, an extension of U-Net that better preserves detailed information during downsampling (Milletari et al., 2016).

All the aforementioned deep learning-based fault detection methods are improved upon the classical U-Net network. However, due to inherent limitations of the network itself, these U-Net-based approaches exhibit insufficient accuracy and poor continuity when applied to real-world data.

Recently, the Transformer model has achieved remarkable success in natural language processing (NLP) (Vaswani et al., 2017). After the Transformer was originally designed for one-dimensional sequences, researchers have begun exploring its adaptation to fault detection. Tang et al. (2023) proposed 2.5D Transformer U-Net, demonstrating significant improvements in result continuity (Tang et al., 2023). Wang et al. (2023) proposed a Transformer-assisted dual U-Net network, which improved the accuracy of fault detection results (Wang et al., 2023). Inspired by Cao et al. (2021), who introduced a Transformer-enhanced U-Net network (Cao et al., 2021), we integrated the self-attention-based Transformer model with the V-Net architecture to develop the TransVNet network. Subsequently, we tested the fault detection performance of our proposed TransVNet network, the well-established U-Net network and the recently introduced TransUNet network using synthetic seismic data. Finally, we applied all three methods to real-world data for validation. Based on the test results from both synthetic and field data, we demonstrate that the TransVNet network significantly enhances the accuracy and continuity of fault detection.

2 Methodology

2.1 TransVNet architecture

Figure 1 illustrates the architectures of the U-Net and TransUNet networks. Both networks consist of an encoder and a decoder component. Figure 2 shows the architecture of the TransVNet network. The main structure of TransVNet combines the Transformer and V-Net networks, while still maintaining the overall encoder-decoder framework.

The encoder section consists of three convolutional layers with downsampling operations, where a Transformer network is embedded after the third convolutional layer. During downsampling, the number of feature channels progressively doubles at each stage to enhance detection of subtle fault features. Key technical improvements include replacing pooling functions with convolutional layers to achieve data dimensionality reduction

and expanded receptive fields, as well as incorporating ResBlock with skip connections to preserve shallow-layer features through feature map stacking.

The decoder employs a progressive upsampling strategy involving three stages with decreasing convolutional operations (three convolutions in the first stage, two in the second, and one in the final stage). Attention mechanisms bridge the downsampling and upsampling paths, while cross-level feature concatenation enables multi-scale information fusion. This asymmetric convolutional configuration facilitates deep feature mining, combining CNN's local perception with Transformer's global modeling capabilities through skip connections.

For fault detection optimization, convolutional downsampling preserves high-resolution features, Transformer captures long-range spatial dependencies, and residual learning ensures fault continuity. The architecture synergistically enhances prediction accuracy by maintaining structural details during dimensionality reduction while modeling both local and global seismic patterns.

2.2 Loss function

In deep learning, the loss function plays a critical role. By minimizing the loss function, the model gradually converges and reduces the prediction error of the TransVNet network. Therefore, the choice of loss function significantly impacts the model's performance.

The network utilizes a weighted sigmoid cross-entropy loss function to address the severe class imbalance between positive (fault) and negative (non-fault) samples in seismic data, where non-fault points significantly outnumber fault points. By introducing a weighting coefficient to the standard sigmoid cross-entropy function, this approach mitigates the training bias caused by extreme sample imbalance, effectively reducing the model's tendency to favor the majority class (non-fault regions) (Lin et al., 2018). The mathematical formulation of the loss function is expressed as:

$$loss = q \times z \times \log(\text{sigmoid}(x)) + (1 - z) \times (-\log(1 - \text{sigmoid}(x))) \quad (1)$$

x represents the predicted fault probability, z denotes the ground truth label (1 for fault, 0 for non-fault), and q is the weighting coefficient that controls the balance between positive and negative samples.

This equation can be simplified from its original form (Equation 1) to explicitly high-light the role of the weighting factor in penalizing misclassified fault points more heavily than non-fault points. The weighting mechanism ensures that the model prioritizes accurate fault detection while maintaining numerical stability during optimization (Equation 2).

$$loss = (1 - z) \times x + (1 + (q - 1) \times z) \times \log(1 + e^{-x}) \quad (2)$$

2.3 Attention mechanism

This study incorporates a spatial attention mechanism into the network architecture to enhance feature fusion,

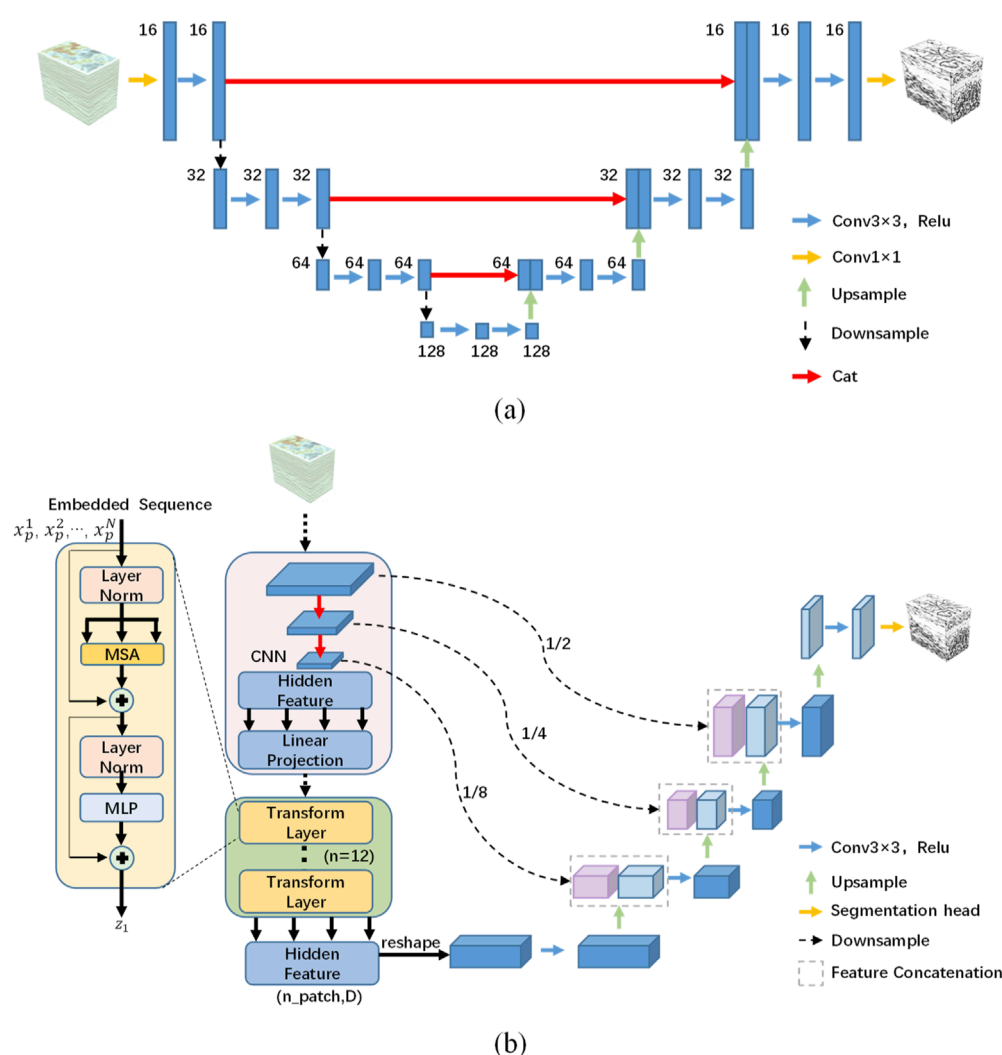


FIGURE 1
The network architectures of UNet (a) and TransUNet (b) both consist of encoder and decoder components.

directing the algorithm's focus toward target regions (fault structures) while suppressing background interference. By amplifying feature responses in critical areas and improving attention to potential fault zones, the mechanism specifically addresses the challenge of detecting low-order faults with subtle seismic signatures.

The attention module operates after merging shallow-layer features (retaining high-resolution spatial details) with upsampled deep-layer features (carrying semantic context). Through this configuration, the network learns to evaluate the spatial significance of geological structures and dynamically reweights feature maps by assigning higher importance to regions likely to contain faults. This adaptive weighting process strengthens fault-related features while diminishing irrelevant seismic patterns, effectively improving detection precision.

The implementation leverages the complementary strengths of multiscale features: shallow layers provide precise spatial localization, while deep layers encode broader geological context.

By fusing these through attention-guided recalibration, the model achieves enhanced sensitivity to discontinuous seismic reflections characteristic of faults, particularly those with weak expression or complex geometries.

Figure 3 schematically illustrates the implementation of the attention mechanism. In this process, g represents the input shallow-layer channel features, and x denotes the deep-layer channel features. The workflow begins by applying convolutional operations to both the shallow and deep channel features, followed by an elementwise summation to integrate their information. The combined result is then processed using a ReLU activation function to introduce non-linear transformations. Next, a $1 \times 1 \times 1$ convolutional kernel is applied to the output of the ReLU activation, refining the feature maps while preserving spatial relationships. The refined features are subsequently normalized through a softmax activation function, which generates spatial attention weights in the range $[0,1]$ to highlight critical regions. Finally, these attention weights are multiplied with the original deep-layer channel features

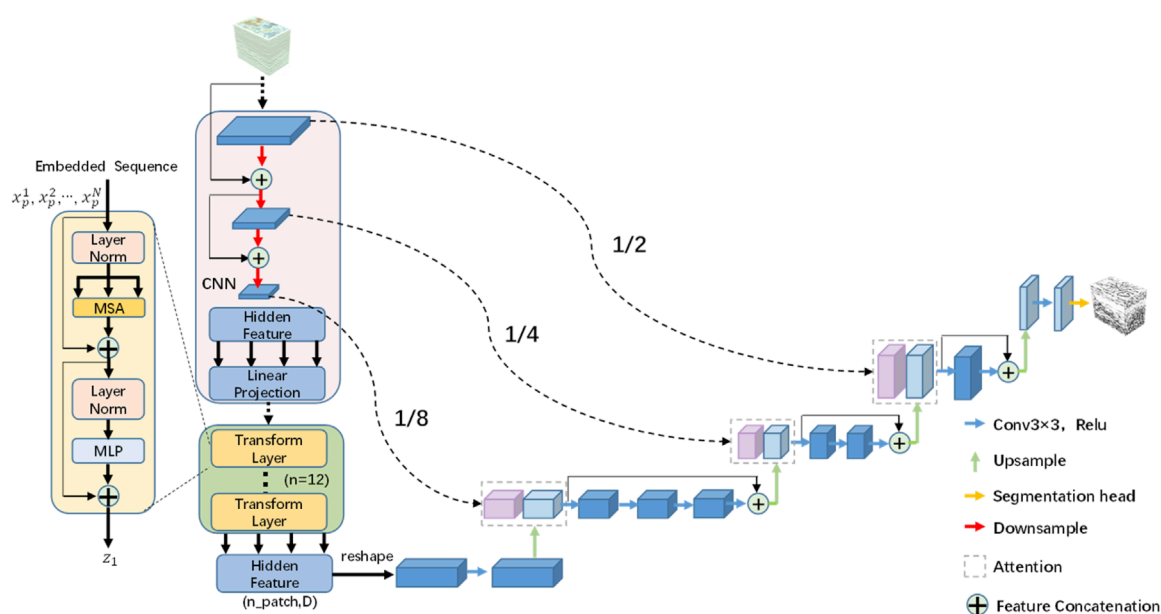


FIGURE 2
The network architectures of TransVNet.

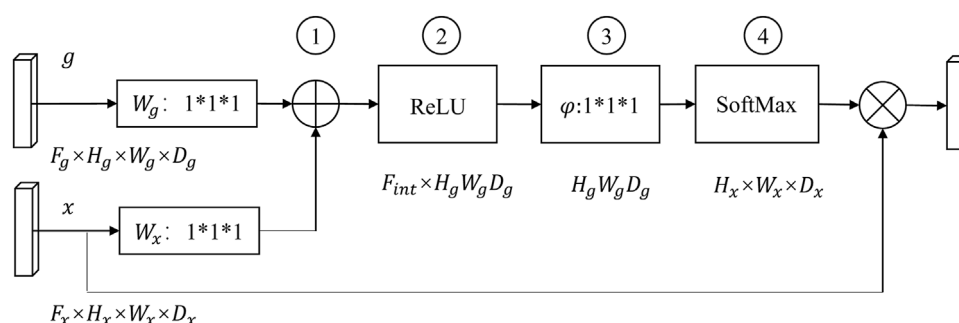


FIGURE 3
Spatial attention mechanism.

(x), adaptively enhancing fault-related patterns and suppressing irrelevant background noise. This sequential integration enables the network to prioritize geologically significant structures by dynamically reweighting multi-scale features, thereby improving fault detection accuracy while maintaining computational efficiency through compact 3D operations.

3 Model training

3.1 Model testing

The dataset was constructed using the toolkit developed by Wu et al. (2019), which generates synthetic fault volumes through forward modeling. Following Wu et al.'s finding that 200 fault data pairs suffice to train a reliable neural network for fault segmentation, we employed 220 synthetic 3D seismic

records ($128 \times 128 \times 128$ samples each) with corresponding fault labels. Of these, 200 volumes were used for training and 20 for testing.

Figure 4 displays representative seismic data samples and their fault labels from the training dataset. The horizontal axis represents CDP (Common Depth Point), and the vertical axis represents time. Prior to training, all input data underwent amplitude normalization to mitigate overfitting caused by amplitude variations across different seismic surveys.

For validation, a seismic profile extracted from one test volume was analyzed. Figure 5 compares the fault detection results between U-Net, TransUNet and TransVNet. The U-Net outputs exhibit poor continuity with numerous misdetection regions, notably failing to detect faults on the right section. The TransUNet model demonstrates higher resolution capabilities, but produces more erroneous results and similarly fails to detect faults on the right side. Obvious issues have been circled in Figure 5. In contrast,

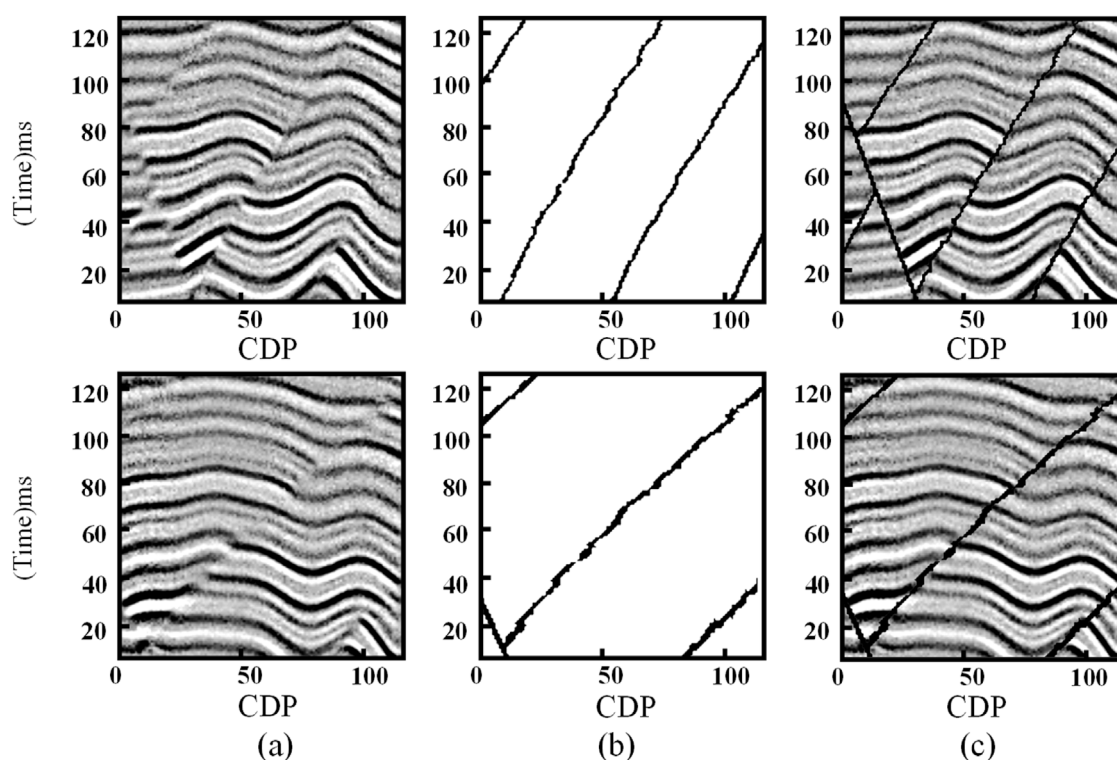


FIGURE 4
Seismic data (a) and fault label (b) and fault-overlaid seismic profile (c).

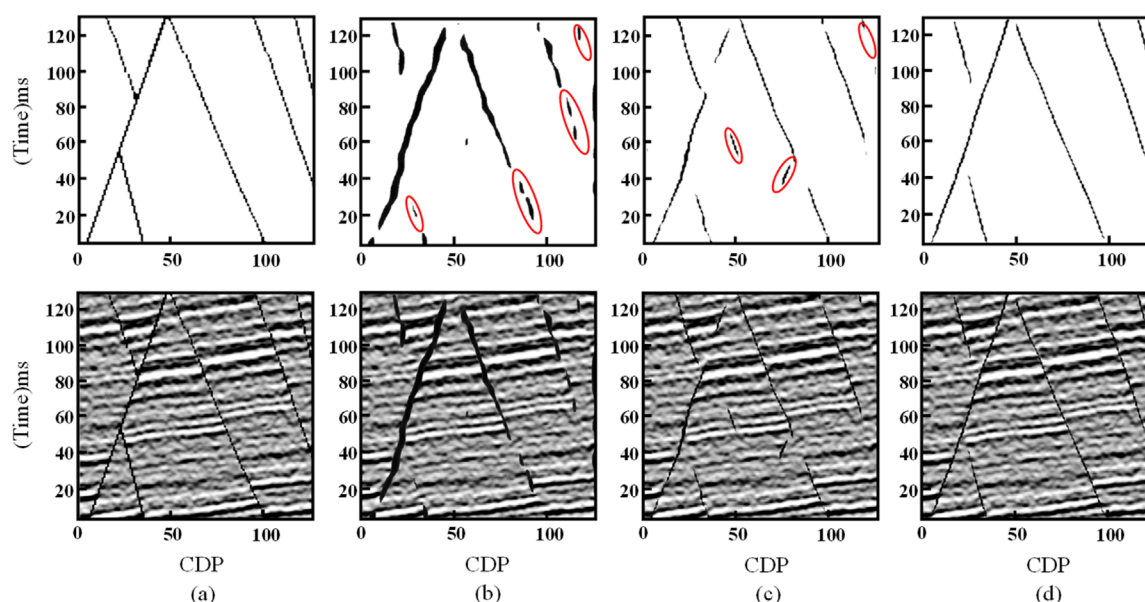


FIGURE 5
Fault labels (a) and fault detection results between U-Net (b), TransUNet (c) and TransVNet (d).

TransVNet addresses these limitations effectively, producing accurate fault predictions with significantly improved spatial continuity. The enhanced performance demonstrates TransVNet's

superior capability in maintaining structural coherence while reducing false positives, particularly for subtle or discontinuous fault features.

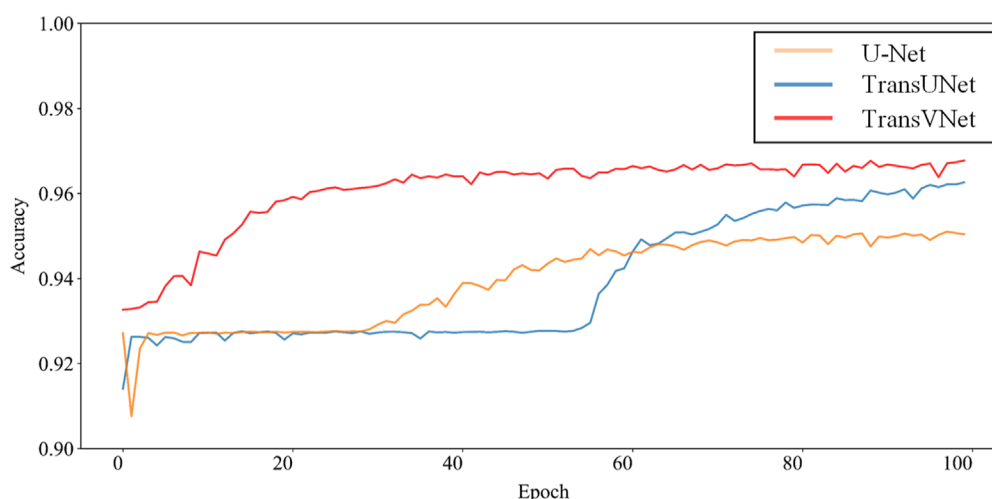


FIGURE 6
Accuracy curve of U-Net, TransUNet and TransVNet.

3.2 Accuracy curve

Figure 6 displays the accuracy curves of the three networks after 100 training epochs. As evident from the figure, the accuracy curve of the TransVNet model reaches higher values compared to those of the U-Net and TransUNet models. This clearly demonstrates the superior performance of TransVNet over both U-Net and TransUNet.

TABLE 1 Evaluation metrics values of the 3 methods.

Method	DC	Sensitivity	Specificity
U-Net	0.4409	0.7117	0.8971
TransUNet	0.5609	0.4815	0.9840
TransVNet	0.7122	0.6070	0.9934

3.3 Evaluation metrics

To quantitatively evaluate the performance of the models, we employed three evaluation metrics to assess the prediction results of the three models: Dice Coefficient (DC), Sensitivity, and Specificity. The DC is a similarity measure function used to calculate the similarity between predicted results and ground truth. Sensitivity measures the model's ability to correctly identify true faults, while Specificity evaluates its capability to correctly recognize non-fault regions.

Equations 3–5 demonstrate the calculation formulas for these metrics:

$$DC = \frac{2TP}{2TP + FP + FN} \quad (3)$$

$$sensitivity = \frac{TP}{TP + FN} \quad (4)$$

$$specifcity = \frac{TN}{TN + FP} \quad (5)$$

where TP (True Positives) denotes regions correctly identified as faults in both predictions and ground truth, FP (False Positives) represents areas misclassified as faults in pre-dictions but actually non-fault in reality, and FN (False Negatives) corresponds to true fault regions missed by the predictions, TN (True Negative) denotes regions correctly identified as non-faults in both predictions and ground truth. A higher IoU value indicates superior precision

TABLE 2 Evaluation metrics values of the 3 methods.

Method	Parameters	Training times	Prediction times
U-Net	351,161	118	107
TransUNet	109,340,785	5,138	181
TransVNet	305,193,073	341,734	254

in fault detection. Table 1 presents the evaluation metrics of fault identification results obtained using three methods: U-Net, TransUNet, and TransVNet. For evaluation metrics computation, a threshold of 0.7 is applied to the fault probability volume, where voxels with probabilities exceeding this value are classified as fault regions.

Analysis of Table 1 reveals that TransVNet achieves the highest Dice Coefficient (DC) value of 0.7122, indicating its fault identification results most closely match the ground truth labels. Furthermore, the numerical results show that both TransUNet and TransVNet exhibit specificity values approaching 1, demonstrating that the incorporation of Transformer architecture significantly enhances non-fault recognition capability.

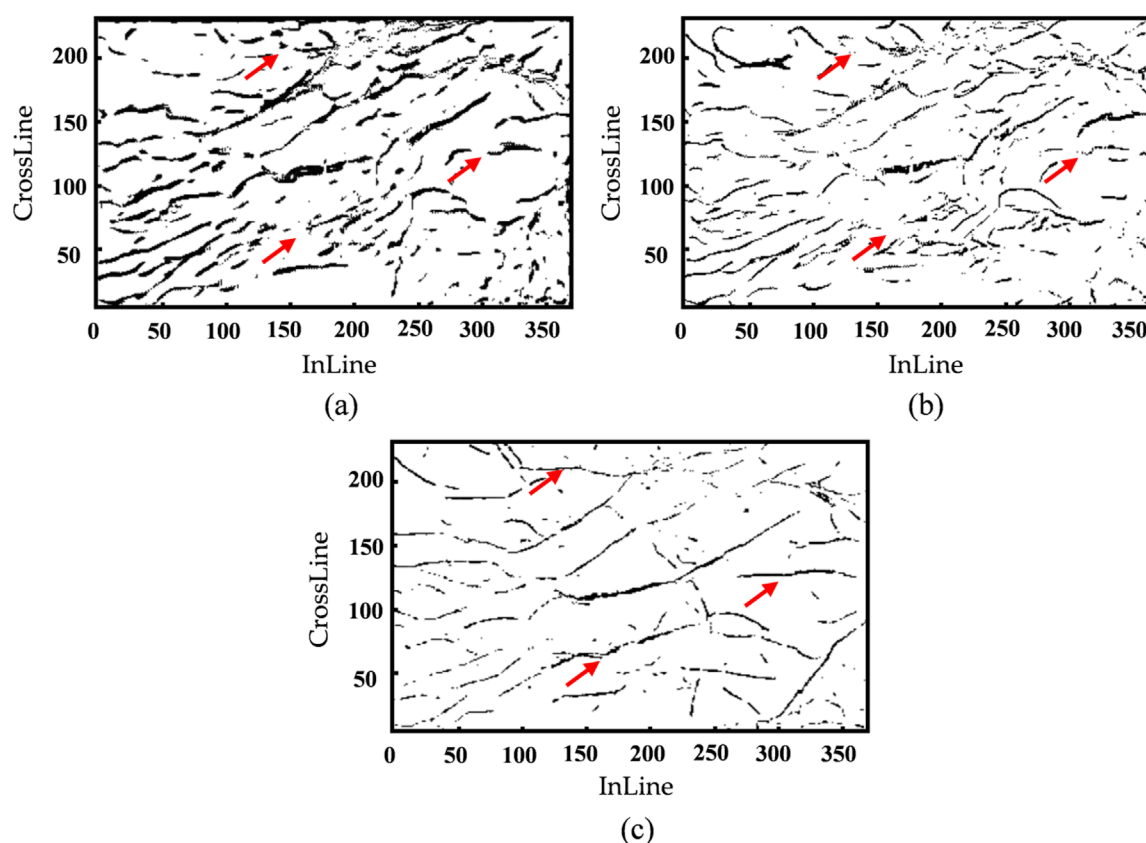


FIGURE 7
Fault detection results of U-Net (a), TransUNet (b) and TransVNet (c).

We also compared the number of training parameters, training and prediction times of the three methods. Table 2 shows the parameters of the three methods. Training time refers to the time consumed in training 1 epoch.

By analyzing Table 2, we can find that the training parameters of TransVNet are much more than those of the other two methods, and its training time is also much longer than the other two methods. This is a drawback of TransVNet. However, the prediction times of the three methods do not differ much, which indicates that the prediction time has little to do with the model.

4 Practical application

The study utilizes real seismic data from a specific block in an oilfield, with data dimensions of InLine = 373, CrossLine = 227, and Time sample = 2,901. The area is influenced by a strike-slip stress field, resulting in fragmented strike-slip faults, complex structural patterns, and well-developed low-order faults that serve as critical pathways for hydrocarbon migration. Accurate fault characterization is essential for determining reservoir locations and understanding hydrocarbon migration pathways in this fault-controlled system.

When applying deep learning for fault detection, transfer learning is typically used to enhance model performance by adapting

synthetic-trained models to real seismic data. However, effective transfer learning requires highly accurate fault labels for the real training data. Due to the prevalence of low-order faults in this area, which are challenging to delineate precisely, transfer learning was not implemented to avoid introducing errors from imperfect labels.

Figure 7 compares fault detection results from U-Net (a), TransUNet (b) and TransVNet (c) through travel time slices along the map view of T6 horizon at depth (The green line in Figure 8 is the T6 horizon). As evidenced by Figure 7, the fault identification results obtained using the U-Net network exhibit significant noise interference and poor fault continuity, with relatively low resolution. Similarly, the TransUNet network's results also suffer from substantial noise artifacts and discontinuous fault representations. In contrast, the TransVNet network effectively addresses the issue of unclear fault traces and significantly enhances the continuity of identified faults. The most notable differences between the three methods' performances are indicated by red arrows in the figure.

Figure 8 displays the fault identification results along InLine and CrossLine sections obtained using different methods. The cross-section analysis reveals that the U-Net network produces results with blurred fault boundaries, poor continuity, and low resolution. Similarly, the TransUNet network also suffers from unclear fault identification and discontinuous results. Yellow arrows in the figures highlight these significant errors in both methods.

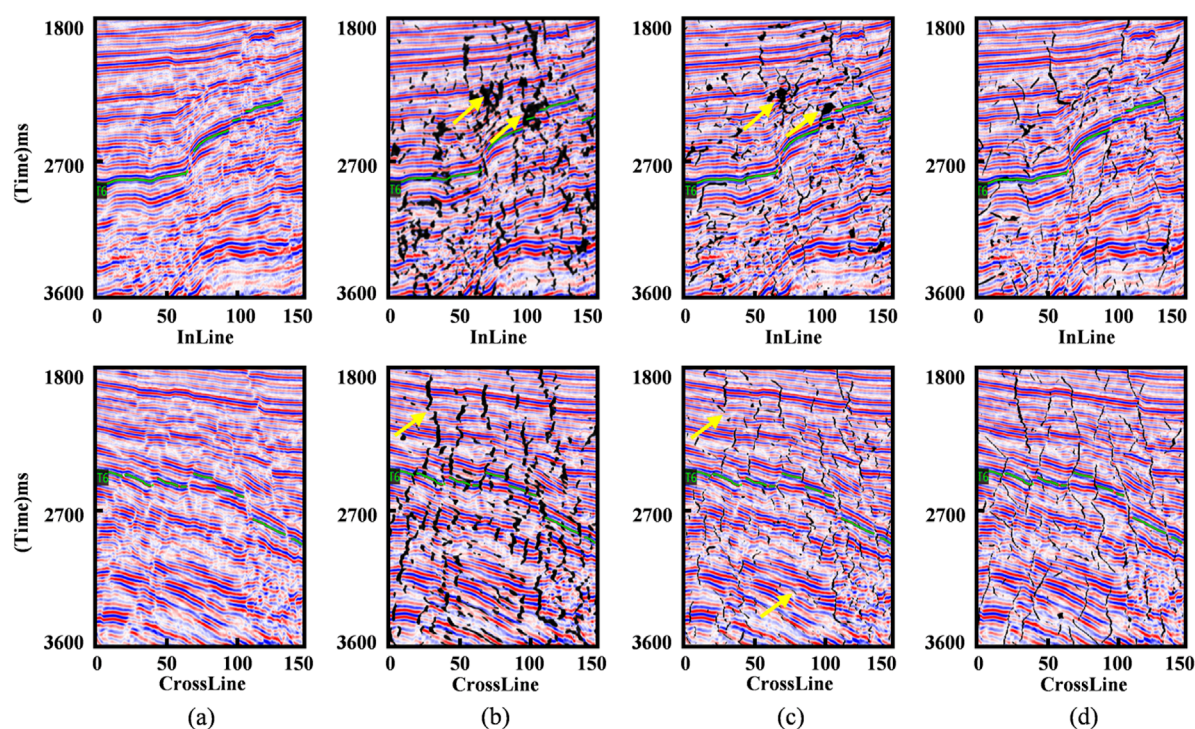


FIGURE 8
Seismic profile (a) and fault detection results of U-Net (b), TransUNet (c) and TransVNet (d).

In contrast, the TransVNet network delivers fault identification results with superior continuity, effectively eliminating most unclear interpretations while maintaining high accuracy. The actual data testing thus demonstrates that compared to both TransUNet and U-Net networks, the TransVNet network significantly improves both the continuity and accuracy of fault identification results.

5 Conclusion

This study proposes an intelligent fault detection method based on the TransVNet convolutional neural network. Through experimental validation and field applications, the following innovative findings have been established:

- (1) Architectural Innovation: The deep integration of Transformer networks with V-Net architecture significantly enhances deep learning performance in fault detection. Our hybrid design synergizes local geometric feature extraction (via 3D convolutions) and global modeling (via self-attention mechanisms).
- (2) Operational Superiority: The proposed method generates fault predictions with enhanced continuity and improved accuracy. Field tests demonstrate its capability to resolve low-order faults in complex strike-slip systems, effectively addressing practical requirements for reservoir characterization in oilfield exploration and development.
- (3) However, the training time of the TransVNet network model is much longer than that of the other two methods. In the future, further optimization can be carried out in this issue.

These advancements provide geoscientists with a robust tool for high-precision fault interpretation, particularly in structurally complex basins where accurate fault delineation critically influences drilling success rates and reservoir management strategies.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

YL: Methodology, Writing – review and editing, Writing – original draft. CZ: Writing – review and editing, Methodology. WW: Writing – review and editing, Data curation. MC: Writing – review and editing, Methodology. XW: Conceptualization, Writing – review and editing. XH: Validation, Writing – review and editing. CB: Validation, Writing – review and editing. SQ: Writing – review and editing, Investigation. YL: Writing – review and editing, Methodology. LW: Writing – review and editing, Validation.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. Funding was provided by

the China Magnetotelluric Array, National Science and Technology Major Project (2024ZD1000200 and 2024ZD1000206).

Conflict of interest

Authors YL, CZ, WW, MC, CB, SQ, YL, and LW were employed by Sinopec Geophysical Corporation Nanfang Branch.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Al-Dossary, S., and Marfurt, K. J. (2006). 3D volumetric multispectral estimates of reflector curvature and rotation. *Geophysics* 71 (5), 41–51. doi:10.1190/1.2242449
- Bahorich, M., and Farmer, S. (1995). 3-D seismic discontinuity for faults and stratigraphic features: the coherence cube. *AAPG Bull.* 14 (10), 1053–1058. doi:10.1190/1.1437077
- Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., and Tian, Q. (2021). *SwinUnet: Unet-like pure transformer for medical image segmentation*. ArXiv, abs/2105.05537.
- Colorni, A., Dorigo, M., and Maniezzo, V. (1992). *Distributed optimization by ant colonies*.
- Gao, K., Huang, L., and Zheng, Y. (2022). Fault detection on seismic structural images using a nested residual U-net. *IEEE Trans. Geosci. Remote Sens.* 60, 1–15. doi:10.1109/tgrs.2021.3073840
- Gersztenkorn, A., and Marfurt, K. J. (1999). Eigenstructure-based coherence computations as an aid to 3-D structural and stratigraphic mapping. *Geophysics* 64 (5), 1468–1479. doi:10.1190/1.1444651
- Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollar, P. (2018). *Focal loss for dense object detection*. arXiv:1708.02002v2.
- Lyu, F., Zhou, H., Liu, J., Zhou, J., Tao, B., and Wang, D. (2022). A buried hill fault detection method based on 3D U-SegNet and transfer learning. *J. Petroleum Sci. Eng.* 218, 110917. doi:10.1016/j.petrol.2022.110917
- Milletari, F., Navab, N., and Ahmadi, S. A. (2016). “V-net: fully convolutional neural networks for volumetric MedISAI image segmentation,” in 2016 Fourth International Conference on 3D Vision, Stanford, CA, USA, 25–28 October 2016 (IEEE), 565–571.
- Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., et al. (2018). *Attention U-Net: learning W-here to look for the pancreas*. ArXiv, abs/1804.03999.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). *U-Net convolutional networks for BiomedISAI image segmentation*. Cham: Springer.
- Shen, S., Li, H., Chen, W., Wang, X., and Huang, B. (2022). Seismic fault interpretation using 3-D scattering wavelet transform CNN. *IEEE Geosci. Remote Sens. Lett.* 19, 1–5. doi:10.1109/lgrs.2022.3183495
- Sun, Q., Wang, X., Ni, H., Gong, F., and Du, Q. (2023). Fault identification of U-Net based on enhanced feature fusion and attention mechanism. *Electronics* 12, 2562. doi:10.3390/electronics12122562
- Tang, Z., Wu, B., Wu, W., and Ma, D. (2023). Fault detection via 2.5D transformer U-Net with seismic data pre-processing. *Remote Sens.* 15, 1039. doi:10.3390/rs15041039
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). *Attention is all you need*. arXiv:1706.03762.
- Wang, Z., You, J., Liu, W., and Wang, X. (2023). Transformer assisted dual U-net for seismic fault detection. *Front. Earth Sci.* 11, 1047626. doi:10.3389/feart.2023.1047626
- Wu, X. M., Liang, L. M., Shi, Y. Z., and Fomel, S. (2019). FaultSeg3D: using synthetic data sets to train an end to end convolutional neural network for 3D seismic fault segmentation. *Geophysics* 84 (3), IM35–IM45. doi:10.1190/geo2018-0646.1
- Xiong, W., Ji, X., Ma, Y., Wang, Y., Albinhassan, N. M., Ali, M. N., et al. (2018). Seismic fault detection with convolutional neural network. *Geophysics* 83 (5), O97–O103. doi:10.1190/geo2017-0666.1
- Yin, X. Y., Gao, J. H., and Zong, Z. Y. (2014). Curvature attribute based on dip scan with eccentric window. *Chin. J. Geophys.* 57 (10), 3411–3421. doi:10.1190/segam2014-0219.1
- Yuan, S., Jiao, X., Luo, Y., Sang, W., and Wang, S. (2022). Double-scale supervised inversion with a data-driven forward model for low-frequency impedance recovery. *Geophysics* 87 (2), R165–R181. doi:10.1190/geo2020-0421.1

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.