

Original Paper

Seismic resolution enhancement via deep learning with teacher-student network and domain adaptation

Han-Peng Cai^{a,*}, Hao-Nan Zhang^a, Li-Yu Zhang^a, Suo Cheng^b

^a School of Resources and Environment, University of Electronic Science and Technology of China, Chengdu, 611731, Sichuan, China

^b PetroChina Tarim Oilfield Company and Research and Development Center for Exploration and Development Technology of Ultra-Deep Complex Oil and Gas Reservoirs, Institute of Tarim Oilfield Company, Korla, 841000, Xinjiang, China

ARTICLE INFO

Article history:

Received 1 July 2025

Received in revised form

27 January 2026

Accepted 26 April 2026

Available online 29 April 2026

Edited by Meng-Jiao Zhou

Keywords:

Deep learning

Teacher-student network

High-resolution processing

Domain adaptation

ABSTRACT

High-resolution processing of seismic signals is crucial for enhancing the ability to characterize underground geological structures in detail and improving the accuracy of thin-layer reservoir identification. Traditional seismic high-resolution algorithms suffer from issues such as poor robustness, low computational efficiency, and neglect of inter-channel structural relationships. Meanwhile, most mainstream deep learning-based high-resolution methods rely on end-to-end networks, lacking guidance from prior information and ignoring differences between data domains, resulting in insufficient generalization capabilities. Therefore, this paper proposes a domain-adaptive knowledge distillation deep learning network for seismic data, DAKD-Net (Domain-Adaptive Knowledge Distillation Network). This method establishes a teacher-student network model using forward simulation datasets as training samples. The teacher network module establishes physical constraints between low- and high-resolution data, extracting high-frequency prior information during the guidance phase to direct the student network module in restoring details without prior conditions. Domain adaptation then generalizes the model to real seismic data, enhancing its generalization capability and structural representation accuracy in actual work area data. Structurally, DAKD-Net adopts a U-net backbone to fully extract spatial structural information across multi-track seismic profiles. Its training mechanism enables prior knowledge transfer through the teacher-student network, allowing high-resolution data recovery without prior information. In application, domain adaptation fine-tuning strategies enhance the network's generalization capability and structural representation effectiveness in real-world work areas. Experimental results demonstrate that the proposed method outperforms traditional approaches and classical deep networks in both vertical resolution and complex structural detail recovery, exhibiting robust performance and practical applicability.

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

1. Introduction

In seismic exploration, the quality of seismic data directly impacts the analysis and interpretation of subsurface geological structures. High-resolution seismic data provides finer details of subsurface structures, enabling more precise reservoir characterization and stratigraphic structural analysis. However, in practical seismic exploration, only low-resolution seismic data is typically acquired. Such low-resolution data often fails to effectively reveal

subsurface structural and lithological information, leading to imprecise exploration results. Therefore, how to employ technical methods to process low-resolution seismic data with high precision and efficiency, thereby restoring high-resolution data, represents an urgent challenge in seismic data processing.

Traditional high-resolution seismic data processing methods include spectral whitening, deconvolution, and inverse Q filtering. Spectral whitening aims to compensate for frequency attenuation by expanding the amplitude spectrum of the signal, encompassing techniques such as wavelet spectral whitening (Chen and Zhou, 2000) and Hilbert spectral whitening (Yang et al., 2007). Deconvolution mitigates filtering effects in seismic records caused by wave propagation, encompassing methods like predictive deconvolution (Robinson, 1967), homomorphic deconvolution

* Corresponding author.

E-mail address: hanpengcai@uestc.edu.cn (H.-P. Cai).

Peer review under the responsibility of China University of Petroleum (Beijing).

Table 1
SSIM and PSNR corresponding to different loss weights α .

α	SSIM	PSNR
0.5	0.8964	32.8518
0.7	0.9846	38.2939
0.9	0.9398	33.2817

(Oppenheim, 1965), and spectral simulation deconvolution (Zhao et al., 1996). While these techniques enhance high-frequency information, limitations persist in phase and frequency-domain processing. Consequently, researchers have successively proposed sparse deconvolution based on L_1 and L_p norms (Debye and Van Riel, 1990), high-resolution methods utilizing Hilbert-Huang transforms (Dragomiretskiy and Zosso, 2013), sparse pulse response spectrum inversion based on compressive sensing (Chen and Wang, 2015), generalized S-transform approaches (Tian et al., 2009), and deconvolution achieved by neural networks to derive inverse wavelets (Lu and Mu, 1996). The inverse Q filtering algorithm estimates the Q factor based on the Q -factor model of seismic wave propagation and compensates for amplitude loss and frequency/phase attenuation in seismic waves according to attenuation theory (Hale, 1981). Although stable and effective inverse Q filtering algorithms and time-frequency domain inverse Q filtering methods (Wang, 2006) can avoid the problem of directly calculating the Q factor, inverse Q filtering still suffers from high computational complexity and poor noise resistance. While these methods partially restore high-frequency information, most rely on processing single-trace data, neglecting the continuity of inter-trace geological structures and overall textural features. Their limited noise resistance and poor adaptability to complex geological regions hinder their ability to meet the demands of large-scale, high-precision seismic interpretation.

Classic deep learning networks include Generative Adversarial Networks (GANs) (Goodfellow et al., 2014), Convolutional Neural Networks (CNNs) (Yuan et al., 2021), and Residual Networks (Zhou et al., 2021). In recent years, deep learning techniques have achieved significant progress in geophysics, such as lithology identification (Yu et al., 2019), fault detection (Wu et al., 2019), and fracture prediction (Gao et al., 2021). The success of deep learning in image super-resolution has led to its introduction into high-resolution seismic processing. CNNs are among the most commonly used architectures in high-resolution techniques, reconstructing high-resolution images or signals by extracting spatial features from images. SRCNN was the first model to employ convolutional neural networks for high-resolution reconstruction (Dong et al., 2014); subsequently, CNNs were introduced into seismic data high-resolution processing (Yuan et al., 2019); and a deep convolutional neural network was proposed that simultaneously achieves high-resolution processing and denoising of seismic images (Li et al., 2021). Additionally, other models have been introduced for high-resolution tasks. The use of GANs was pioneered for high-resolution processing by proposing a high-resolution generative adversarial network model (Ledig et al., 2017). Noise suppression and resolution enhancement in seismic data were achieved using an elastic convolution dictionary learning-based approach (Lan et al., 2023). A diffusion model-based seismic data resolution enhancement method was also proposed recently (Zhang et al., 2024). Furthermore, a pseudo-reflection coefficient model using a U-net architecture was trained to achieve end-to-end high-resolution seismic data processing (Sun et al., 2021). A high-resolution seismic data processing method based on the S-transform and a complex-valued U-net was introduced, enabling the network to jointly learn time-

and frequency-domain information (Li et al., 2023). However, most existing deep learning approaches directly learn the mapping between low-resolution and high-resolution seismic data within an end-to-end framework, lacking constraints from seismic prior information. Simply increasing network depth or width often leads to insufficient generalization for real seismic data, making it difficult to deploy in diverse and complex field environments.

Teacher-student networks employ knowledge distillation strategies to transfer knowledge from a complex, high-performance teacher model to a smaller or simpler student model. This approach reduces computational costs and memory consumption while maintaining performance. First introduced in 2015, this technique has been widely applied in model compression, transfer learning, and multi-task learning scenarios (Hinton et al., 2015). Its core principle lies in the student model learning not only the actual data labels (hard labels) but also the additional information embedded in the teacher model's predictions (soft labels) for the input data. A teacher-student network-based prior method for facial recognition was proposed to enable the student network to capture prior knowledge without requiring prior estimation (Wang et al., 2022). Genuine prior information was leveraged during the teacher model training phase to enhance high-resolution performance. This prior information was then transferred to the student network, allowing it to directly reconstruct target images from low-resolution facial images. This approach effectively mitigates the impact of inaccurate prior estimation.

Therefore, to address the limitations of traditional methods—underutilization of structural information between seismic channels and insufficient generalization capabilities of end-to-end networks—this paper proposes DAKD-Net, a high-resolution deep learning network for seismic data based on teacher-student networks and domain adaptation.

First, a training dataset is constructed by simulating seismic forward modeling: convolving with different frequency band wavelets yields a low-resolution input dataset, a high-resolution prior dataset formed by convolving with mid-frequency wavelets, and a high-resolution label dataset formed by convolving with high-frequency wavelets. During teacher network training, the “low-resolution input + high-resolution prior” is treated as joint input, with “high-resolution labels” as the supervision target, enabling precise reconstruction of high-frequency details. Simultaneously, soft label loss and feature space loss guide the student network to learn and transfer prior knowledge. During the real-data phase, domain adaptation strategies fine-tune the network's high-level feature modules, enabling efficient model transfer from synthetic domains to real seismic domains and enhancing the network's generalization capability for real seismic data. Experimental validation demonstrates that DAKD-Net not only excels in model data but also exhibits outstanding resolution enhancement and structural recognition capabilities in actual work areas.

2. Methods

The proposed DAKD-Net aims to automatically recover high-resolution features from low-resolution seismic data without requiring direct input of prior knowledge. Its overall architecture, as shown in Fig. 1, comprises three core modules: the teacher network module, the student network module, and the domain-adaptive transfer module. The network adopts a U-net-based architecture, utilizing a symmetric encoder-decoder structure with skip connections to extract and reconstruct cross-layer information, fully accounting for structural continuity and texture features across channels. The overall method comprises three stages.

Guided Training stage: By constructing synthetic seismic datasets, we establish physical mapping relationships between

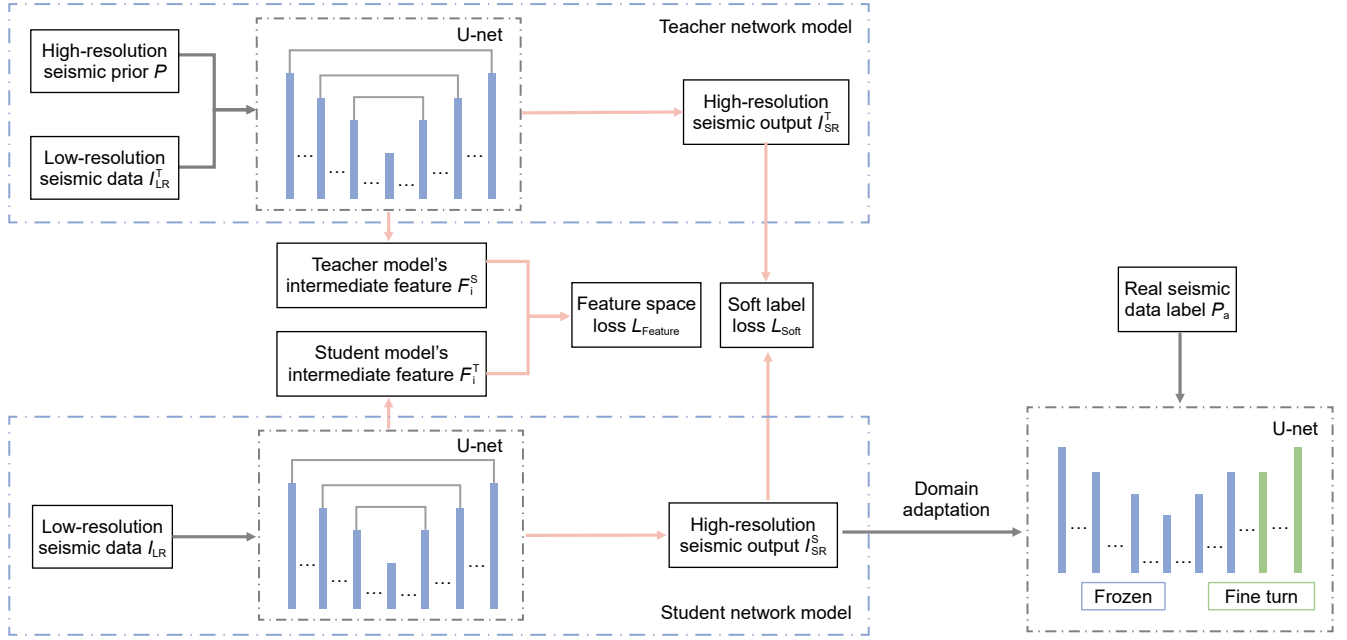


Fig. 1. Architecture of the domain-adaptive knowledge distillation seismic high-resolution network.

low- and high-resolution data. Prior data is introduced to guide the neural network in learning structural representations of high-frequency features. During this phase, the neural network extracts intermediate features across multiple scales, providing supervisory signals for subsequent stages through soft labels and feature space losses.

Autonomous Recovery stage: Remove prior inputs and guide the student network to recover high-frequency features solely from low-resolution data. During training, enhance the network's ability to restore fine-grained details through intermediate feature supervision and structural similarity loss in the guided training phase, enabling high-resolution reconstruction without prior conditions.

Domain Adaptive Transfer stage: Based on a small amount of actual seismic data, the lower-level feature extraction module is frozen, and only the parameters of the higher-level components are fine-tuned to achieve transfer from the synthetic domain to the actual seismic data domain, thereby improving the network's generalization capability in real-world data.

2.1. Guided training stage

Traditional CNN-based high-resolution methods often produce blurred reconstructions and lose fine details when lacking geological structure information, particularly in areas with complex geological features such as faults and tectonic boundaries. Therefore, incorporating seismic data as prior information enables the neural network to learn more accurate geological structures and details, which is crucial for enhancing high-resolution performance. By using low-resolution seismic data and high-resolution prior information as joint inputs, and feeding the concatenated data into the network, the process is defined as follows:

$$I_{SR}^T = f_{\text{Feature}}([I_{LR}, P]) \quad (1)$$

where P denotes the high-resolution seismic prior, I_{LR} represents the low-resolution seismic data, f_{Feature} represents feature extraction

function for teacher networks, I_{SR}^T is the high-resolution seismic output generated by this module, and $[,]$ denotes the concatenation operator. When synthesizing data, low-, medium-, and high-frequency seismic data are generated by convolving wavelets of different frequency bands with reflection coefficients. Among them, the medium-frequency wavelet's convolution data is selected as the high-resolution prior, mainly for the following reasons:

The convolution formula for synthetic data is as follows:

$$s_b(t) = (r * w_b)(t), \quad b \in \{\text{low, mid, high}\} \quad (2)$$

where $s_b(t)$ represents the amplitude of the seismic trace; the reflection coefficient r is convolved with wavelets of different frequency bands to generate S_{low} , S_{mid} , and S_{high} , corresponding to: $I_{LR} = S_{\text{low}}$, $P = S_{\text{mid}}$, $I_{HR} = S_{\text{high}}$. We select the medium-frequency convolution S_{mid} as the high-resolution prior. Physical and acquisition constraints: Absorption and scattering during seismic propagation cause high-frequency energy to decay rapidly according to the following equation:

$$A(f; t, Q) = \exp(-\pi f t / Q) \quad (3)$$

where A is the amplitude attenuation factor; f is the frequency; t is the propagation time; and Q is the quality factor. As frequency f increases or propagation time t increases, high-frequency amplitudes decay rapidly according to the above formula. Medium frequencies fall within the effective passband, where the wavelet and phase estimation are more stable. Using this as a prior can reduce label noise and over-sharpening risks. Frequency spectrum alignment: The medium-frequency spectrum covers low-frequency information while extending toward higher frequencies, located within the reliable passband of actual data, closer to the dominant frequency and bandwidth of the field wavelet. This helps reduce discrepancies between the synthetic domain and real domain distributions, making the teacher signal more reliable during distillation/transfer.

Sampling and imaging consistency: Under common survey spacing and imaging resolution, high-frequency components are more likely to approach or exceed the Nyquist limit, causing

aliasing. In contrast, the medium-frequency prior enhances details while maintaining axial continuity and lateral stability. The specific formula is as follows:

$$f_{\max} \leq \min\left(\frac{1}{2\Delta t}, \frac{v_{\min}}{2\Delta x \sin \theta_{\max}}\right) \quad (4)$$

where $f_{\max}$ is the maximum frequency, Hz; exceeding this value may cause aliasing in recording/imaging; Δt is the temporal sampling interval, s; $v_{\min}$ is the minimum propagation velocity; Δx is the spatial sampling interval, m; and $\theta_{\max}$ is the maximum propagation angle.

Resolution-stability trade-off: Compared to low-frequency priors (insufficient detail) and overly high-frequency priors (unstable, prone to artifacts), the medium-frequency prior provides a more balanced constraint between phase preservation, reflector boundary enhancement, and amplitude reliability. It serves as an anchor point for network high-resolution learning. Moreover, it serves as a physical constraint to prevent high-frequency hallucinations. In deep learning-based super-resolution methods, since the mapping from low-resolution to high-resolution is inherently an ill-posed problem, unconstrained networks often generate high-frequency textures that appear realistic but are geologically invalid. The mid-frequency prior acts as a physical anchor: it provides a reliable structural skeleton (e.g., the accurate phase and spatial location of major faults and layers) existing within the effective seismic passband. Consequently, the network is restricted to extending the spectrum strictly along these physically verified structural locations, rather than generating high-frequency events in void spaces. This mechanism ensures that the recovered high frequencies are structurally consistent with the underlying low-frequency geology, significantly reducing the probability of generating non-physical artifacts. Therefore, we adopt S_{mid} as the high-resolution prior to guide spectral extension and structural detail recovery, while the final network output further extends the frequency band and enhances resolution during real-domain training.

The key to high-resolution seismic data lies in its spectral information. The mid-frequency wavelet can directly provide the detailed frequency features required for spectral extension, containing rich structural details of the seismic data. Using the prior information synthesized by the mid-frequency wavelet can also provide more generalized frequency features, avoiding the model's over-reliance on specific geological interpretations. Furthermore, as mentioned above, the mid-frequency wavelet serves as a physical constraint to provide seismic skeleton information, preventing the generation of high-frequency “hallucinations”. In summary, the mid-frequency wavelet data is selected as the prior information.

The specific architecture of the teacher network constructed based on the U-net architecture is shown in Fig. 2, comprising the following key components.

1. **Input Layer:** The low-resolution seismic data I_{LR} and high-resolution prior information P are concatenated as joint input to the network.
2. **Feature Extraction Module:** Convolutional layers are used to extract deep features. Features F_{P} are extracted from the high-resolution prior via a convolutional path, while features F_{LR} are extracted from the low-resolution seismic data using a combination of convolutional and upsampling layers. Due to the limited spatial resolution of the low-resolution input, direct use in downstream processing may result in insufficient structural representation. Thus, upsampling is applied to compensate for spatial resolution loss and ensure compatibility during feature fusion.

3. **Feature Fusion Module:** A concatenation layer is used for feature fusion. The depth features extracted from high-resolution prior F_{P} and low-resolution data F_{LR} are concatenated to generate a joint feature tensor F_{joint} . This joint representation captures both the reflection patterns from the low-resolution input and the structural details provided by the high-resolution prior.
4. **Encoder-Decoder Structure:** The fused feature tensor F_{joint} is fed into the U-net architecture. The encoder progressively down-samples the features to capture multi-scale contextual information, while the decoder restores spatial resolution using transposed convolutions. Skip connections bridge encoder and decoder stages to preserve spatial details.
5. **Residual Learning Module:** To further enhance high-frequency detail reconstruction, a residual connection is added before the final output. The predicted high-resolution result is added to the original low-resolution input, enabling the network to focus on learning the missing high-frequency components.
6. **Output Layer:** The network outputs the predicted high-resolution seismic data I_{SR}^{T} while storing intermediate features from each layer for use in the teacher network's guidance of the student network phase.

To constrain the teacher network module to generate high-resolution seismic data, a loss function combining L_1 loss and Structural Similarity (SSIM) loss is adopted. The SSIM loss emphasizes the preservation of textural details and structural similarity within seismic data. By incorporating the SSIM component, the model is encouraged to maintain fine-grained features such as seismic reflections and geological boundaries. This helps enhance the fidelity of the reconstructed seismic signals. The L_{Voxel} loss is defined as:

$$L_{\text{Voxel}} = \|I_{\text{SR}}^{\text{T}} - I_{\text{HR}}\|_1 \quad (5)$$

where I_{HR} denotes the input high-resolution seismic data.

SSIM is defined as:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (6)$$

Here, x and y represent two seismic data samples, μ_x and μ_y denote their respective means, σ_x^2 and σ_y^2 are the standard deviations, and σ_{xy} is the covariance between x and y . Constants C_1 and C_2 are stabilization terms used to avoid division by zero. A higher SSIM value indicates a greater structural similarity between the predicted high-resolution result and the high-resolution ground truth. Therefore, L_{SSIM} is defined as $1 - \text{SSIM}(x, y)$, aiming to minimize structural discrepancies between seismic data, which is defined as:

$$L_{\text{SSIM}} = 1 - \text{SSIM}(I_{\text{SR}}^{\text{T}}, I_{\text{HR}}) \quad (7)$$

The final loss function for the teacher network module is defined as:

$$L_{\text{Teacher}} = \alpha \times L_{\text{Voxel}} + (1 - \alpha) \times L_{\text{SSIM}} \quad (8)$$

where α is the weight of L_{Voxel} and L_{SSIM} . During the training process, the weights of the L_1 loss and SSIM loss are dynamically adjusted. In the early stages, greater emphasis is placed on the loss to ensure numerical accuracy. In the later stages, the weight of the SSIM loss is gradually increased to enhance structural similarity. This strategy ensures a balanced trade-off between numerical precision and structural consistency in the model's output.

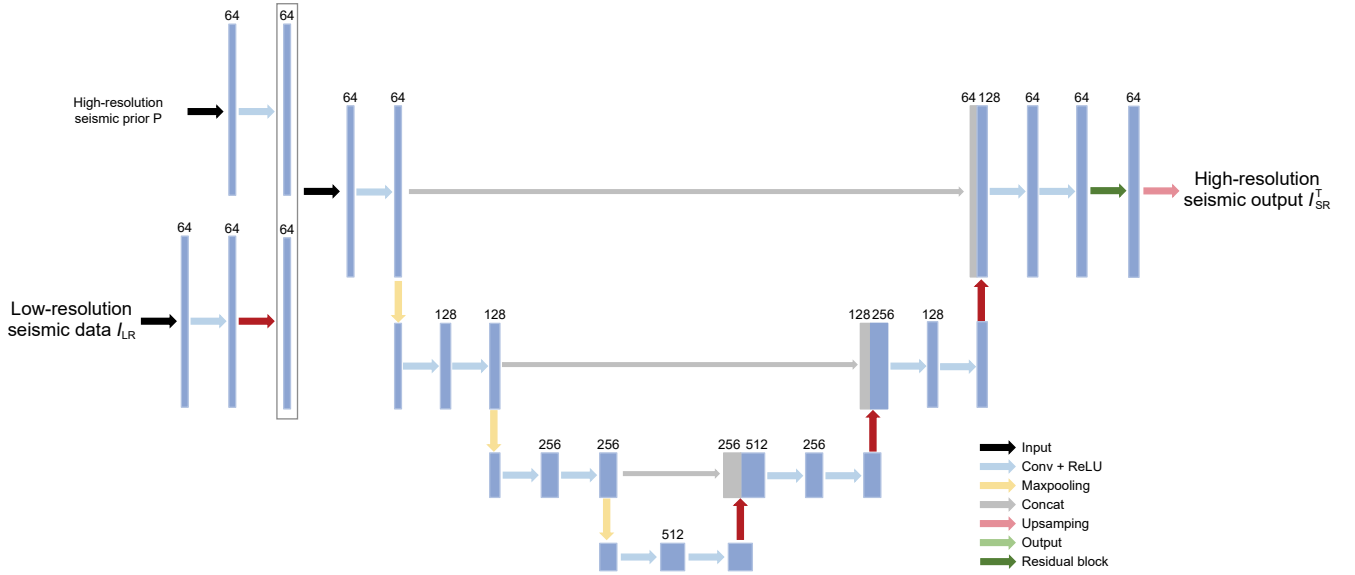


Fig. 2. Teacher network module of DAKD-Net.

2.2. Autonomous recovery stage

To enhance the network's high-frequency recovery capability under conditions without prior information input, DAKD-Net adopts a “guided-then-autonomous” training mechanism. During the guided training phase, the network jointly uses low resolution seismic data and precise prior data at the input end, and trains using a dataset synthesized through forward modeling, effectively obtaining geological structural information. However, in actual seismic exploration, it is challenging to obtain complete and accurate prior information about faults. Therefore, in the autonomous recovery phase, prior information is removed, and only low-resolution seismic traces are used as input. During training, the network continuously receives soft label supervision and feature space distillation information from the guided training phase, achieving generalization and transfer from “with prior information” to “without prior information”. This mechanism enables the network to independently reconstruct high-resolution seismic data during the inference stage while maintaining structural consistency. The two-stage collaborative training strategy enhances the network's ability to express complex geological features through constraints in the structural space.

The structure of the student network is shown in Fig. 3. It is similar to that of the teacher network, but with a key difference in the input—the student network only receives low-resolution seismic data as input and does not require additional prior information. Furthermore, it outputs features at each decoding stage and compares them with the corresponding outputs from the teacher network, ensuring that detailed knowledge from the teacher network is learned at different scales.

Through knowledge distillation, the intermediate features output by the teacher network are transferred to the student network. The student network compares its own intermediate-layer features with those of the teacher network and minimizes the differences between them. The loss function of the student network consists of multiple loss terms, including L_1 loss, SSIM loss, soft label loss, and feature space loss, defined as follows:

$$L_{\text{Student}} = \alpha \times L_{\text{Pixel}} + \beta \times L_{\text{Feature}} + \gamma \times L_{\text{Soft}} + \kappa \times L_{\text{SSIM}} \quad (9)$$

where α , β , γ , and κ are the weights of L_{Pixel} , L_{Feature} , L_{Soft} , and L_{SSIM} , respectively. To effectively transfer high-resolution seismic prior information to the student network, soft label loss L_{Soft} and feature space loss L_{Feature} are introduced as knowledge distillation losses. Since relying solely on pixel-level differences may be insufficient to capture important structural features in seismic data, SSIM loss L_{SSIM} is also incorporated to enhance the structural fidelity of the generated data. The SSIM loss measures the local structural similarity of seismic data, ensuring that the generated high-resolution data not only closely matches the high-resolution label data numerically but also maintains high consistency in seismic waveform characteristics and geological structure information.

L_{Pixel} is the pixel-wise L_1 loss function applied on the I_{SR}^S and I_{HR} , which is defined as:

$$L_{\text{Pixel}} = \|I_{\text{SR}}^S - I_{\text{HR}}\|_1 \quad (10)$$

where I_{SR}^S denotes the high-resolution prediction produced by the guided learning module, and I_{HR} represents the high-resolution seismic ground truth.

The high-resolution prediction results generated by the teacher network are referred to as soft labels, while the actual high-resolution data are referred to as hard labels. Soft labels can include varying degrees of fuzziness and uncertainty, making the learning objectives more broad and thereby improving adaptability and generalization capabilities for unknown seismic data, providing simpler but more effective supervision for the network. In order to make the output of the autonomous recovery stage closer to the results of the guided training stage, soft ground truth loss L_{Soft} is used, which is implemented based on the L_1 loss, as shown in the following formula:

$$L_{\text{Soft}} = \|I_{\text{SR}}^S - I_{\text{SR}}^T\|_1 \quad (11)$$

where I_{SR}^S represents the high-resolution prediction results of the autonomous recovery stage, and I_{SR}^T represents the prediction results of the guided training stage, namely the soft ground truth, which is generated from I_{LR} and P . The soft ground truth contains

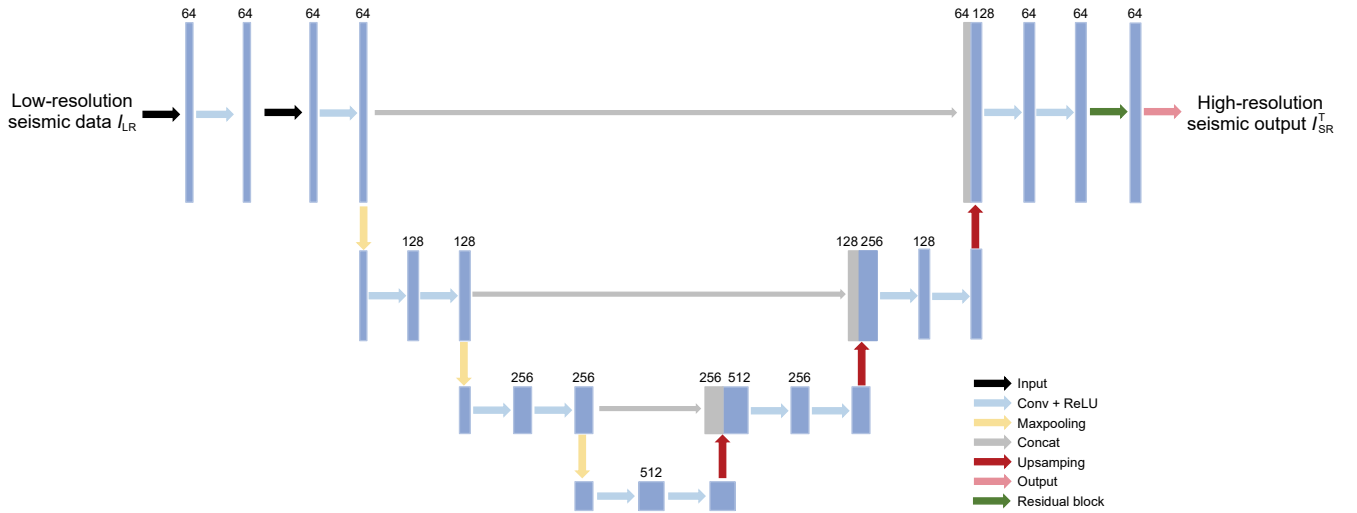


Fig. 3. Student network module of DAKD-Net.

rich high-resolution prior information and is easier to learn compared to the ground truth targets.

To explicitly constrain the intermediate features between the guided training phase and the autonomous recovery phase, the L_{Feature} feature space loss is used to calculate the difference between the output features of the two in the intermediate layer, as shown in the following formula:

$$L_{\text{Feature}} = \frac{1}{n} \sum_i^n \|F_i^S - F_i^T\|_1 \quad (12)$$

where F_i^S and F_i^T represent the intermediate layer features of the guided training stage and the autonomous recovery stage, respectively. These features are typically the output tensors of intermediate layers, such as activations following convolutional layers. n represents the number of selected intermediate layers for comparison. Through the feature space loss, the student network module can learn the deep seismic structural information contained in the feature extraction module without prior knowledge, thereby improving the detail representation capability of the reconstruction.

2.3. Domain Adaptive Transfer stage

Due to significant differences between synthetic seismic data and actual seismic data in terms of wavelet morphology, frequency characteristics, and stratigraphic structure, directly applying a trained model to an actual work area may encounter several issues. For instance, discrepancies in wavelet morphology and geological structure between the actual work area's seismic data and the synthetic data can hinder the model's adaptation to new data. Alternatively, field data is often limited in volume, making it impractical to train complex deep learning models from scratch. Even if training is feasible, it demands substantial computational resources and time (Pan and Yang, 2010). To address this, DAKD-Net incorporates a domain adaptation mechanism to enable effective model transfer from the synthetic domain to the real domain.

Domain adaptation can both fine-tune models to accommodate variations between synthetic and real-world data and significantly reduce training time by transferring existing model weights (Tang et al., 2021). Furthermore, when actual field data is scarce, transfer

learning substantially reduces reliance on large labeled datasets. Therefore, after initially training the high-resolution seismic network DAKD-Net with synthetic data, this paper transfers the pre-trained network weights to the new task network. The lower-level features are frozen, while only parameters in the higher-level layers undergo fine-tuning. Updating a portion of the pre-trained model's parameters using a small amount of actual field data enables the model to better adapt to the characteristics of real-world seismic data. This strategy effectively mitigates the “domain shift” issue, allowing the model to retain its strong constrained learning capabilities from synthetic data while adapting to the properties of field seismic data, thereby enhancing its generalization ability and application scope.

The specific implementation process is as follows:

Loading and Freezing: First, load the pre-trained U-net student network (with architecture kept consistent with the source domain), freeze all parameters, then individually unfreeze the parameters of the final decoding layers dec2, dec1, out_conv, and the residual block to carry the real-domain increment, achieving “low-level freezing and high-level fine-tuning”.

$$\theta_{\text{real}} = \theta_{\text{syn}} + \Delta, \Delta = m_H \odot (\theta_{\text{real}} - \theta_{\text{syn}}) \quad (13)$$

where θ_{real} is the parameter set initialized from the real domain, used for the U-net after domain adaptation; θ_{syn} is the parameter set from the simulated domain, i.e., the pre-trained student network parameters; Δ is the incremental parameter set corresponding to the real domain relative to the synthetic domain, which only acts on the unfrozen high-level channels; m_H is a 0–1 mask with the same dimension as θ_{syn} , taking value 1 at the positions of the high-level channels to be unfrozen and 0 elsewhere; $\odot$ denotes the Hadamard element-wise multiplication operation.

Fine-tuning and Loss: Randomly split the real-world dataset into training and validation subsets, take low-frequency input, and perform forward computation to obtain output, i.e.,

$$\hat{X}_H^r = G_{\theta_{\text{real}}} (X_L^r) \quad (14)$$

where $\hat{X}_H^r$ is the output obtained by forward computation through the domain-adapted network; X_L^r is the low-frequency input.

Compute the composite loss function:

$$L_{DA} = \lambda \|\hat{X}_H^r - X_H^r\|_1 + (1 - \lambda)(1 - \text{SSIM}(\hat{X}_H^r, X_H^r)) \quad (15)$$

where L_{DA} is the total loss in the domain adaptation stage; λ is the weight of the pixel-level loss; X_H^r is the high-frequency target data in the real domain; $\|\cdot\|_1$ denotes the element-wise L_1 norm; $\text{SSIM}(\cdot, \cdot)$ is the structural similarity index. Only trainable parameters are updated during each training iteration, enabling an “optimization + differential fine-tuning” optimization strategy.

Evaluation and Saving: After each training epoch, save the model weights and evaluate them on the validation set. Compute the RSNR, SSIM, and PSNR values, where the PSNR calculation formula is:

$$\text{PSNR}(X_H^r, \hat{X}_H^r) = 10 \log_{10} \frac{R^2}{\frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (X_H^r(i, j) - \hat{X}_H^r(i, j))^2} \quad (16)$$

where PSNR is the peak signal-to-noise ratio; R is the maximum possible value of the data dynamic range (e.g., the upper limit after amplitude normalization); m, n are the dimensions of the 2D window; i, j are spatial indices.

2.4. High-resolution evaluation metrics

For evaluating the output results of high-resolution networks, this paper employs two quantitative metrics: peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM). These metrics comprehensively assess output quality from the perspectives of data accuracy and structural similarity, respectively.

PSNR is a widely used metric for assessing image reconstruction quality, calculated as the ratio of the signal's maximum possible power to the signal-to-noise power. It is commonly employed to evaluate the quality of high-resolution processed images, where higher values indicate smaller differences between the reconstructed image and the original, signifying superior image quality.

The specific formula is shown in Eq. (16).

SSIM measures the structural similarity between high-resolution predictions and high-resolution label data, rather than merely pixel-to-pixel differences. Compared to PSNR, SSIM values more closely align with human visual perception of image quality. An SSIM value closer to 1 indicates higher structural similarity between the high-resolution prediction and the high-resolution label data.

The specific formula is shown in Eq. (6).

3. Experiments and discussion

3.1. Network training and validation

Using the reflection coefficients from the Marmousi2 model, convolution is performed with both low-frequency and high-frequency wavelets. This process constructs low-resolution seismic data as input and high-resolution seismic data as labels. Subsequently, intermediate-frequency wavelets are randomly selected from both low-frequency and high-frequency wavelets for convolution, serving as high-resolution prior information. To enhance the network's generalization capability, the low-frequency wavelet was randomly selected from the Leake wavelet with a dominant frequency range of 20–30 Hz, the high-frequency wavelet was chosen as the Leake wavelet with a dominant frequency of 80 Hz, and the intermediate-frequency band wavelet was selected as the Leake wavelet with a frequency of 40 Hz. With a sampling time of 1 ms, 7000 sets of high-

resolution data of size 1000×300 were constructed along with corresponding low-resolution data and high-resolution prior information, as shown in Fig. 4. To enhance the network's robustness, Gaussian noise with varying variances was randomly added to the low-resolution synthetic data. Subsequently, all synthetic data were partitioned into training and validation datasets of size 128×128 by setting an appropriate sliding step size, as illustrated in Fig. 5. The training-to-validation ratio was set to 9:1.

Fig. 6(a)–(c) show the loss curves of teacher networks with different loss weights α and the corresponding PSNR and SSIM values. Fig. 6(d)–(f) show the high-resolution processing results corresponding to different loss weights. It can be seen that when the weight is 0.5, the recovery of bedding structure is not sufficient. When the weight is 0.9, although the structure is clear, some unnatural artifacts appear in the smooth region, leading to the decline of PSNR. When the weight is 0.7, it not only clearly describes the fault edge, but also maintains the continuity of the horizon. These results indicate that when $\alpha = 0.7$, the trade-off between SSIM loss and L_1 loss is most reasonable, with both PSNR and SSIM reaching high levels, and the best high-resolution processing results (see Table 1).

The training parameters for the feature extraction module are set as follows: loss weight = 0.7, learning rate = 0.00001, optimizer = Adam optimizer, number of iterations = 30, and batch size = 16. During the self-recovery phase, the parameters of the feature extraction module remain frozen, with the intermediate layer features and high-resolution output serving as guidance information for the guidance learning module. The loss function L_{Guided} for the student network is defined as follows, where L_{Pixel} weight $\alpha = 1.4$, L_{Soft} weight $\gamma = 0.1$, and L_{SSIM} weight $\kappa = 0.4$. The L_{Feature} weight needs to be dynamically adjusted. The L_{Feature} weight β is defined as:

$$\beta = \begin{cases} 1 - 0.2 \times \text{epoch} & \text{if epoch} < d \\ 0 & \text{if epoch} \geq d \end{cases} \quad (17)$$

where epoch represents the number of iterations in the student network's training process, and d is the threshold. When training the Marmousi2 model, $d = 5$. Similarly, the Adam optimizer is used with 30 iterations and a batch size of 16.

3.2. Synthetic seismic data

To thoroughly evaluate the effectiveness of the DAKD-Net, synthetic datasets generated from the Marmousi2 model are used to train and test both the proposed high-resolution network and a traditional fully convolutional network (FCN).

Fig. 7 shows the loss curve of DAKD-Net during the training phase and the high-resolution evaluation metrics (PSNR and SSIM curves) on the validation set. It can be seen that the loss curve of DAKD-Net decreases steadily, and the PSNR and SSIM metrics both reach high levels. Additionally, the loss values of the training set and validation set are close, and the improvement process of PSNR and SSIM is consistent, indicating that DAKD-Net can effectively avoid overfitting and has good generalization ability.

Fig. 8 shows the high-resolution prediction results of DAKD-Net for the Marmousi2 model data. Fig. 8(a) is a low-resolution seismic section with a dominant frequency of 20 Hz, where the phase axis contours are blurred and details are unclear. Fig. 8(b) is a true high-resolution seismic section with a dominant frequency of 80 Hz, where structural details are more distinct. Fig. 8(c) shows the prediction results of DAKD-Net. It can be observed that DAKD-Net can refine the blurred details in the low-resolution profile while preserving the overall structure of the seismic profile, significantly enhancing the precision of the phase axis. In particular, complex

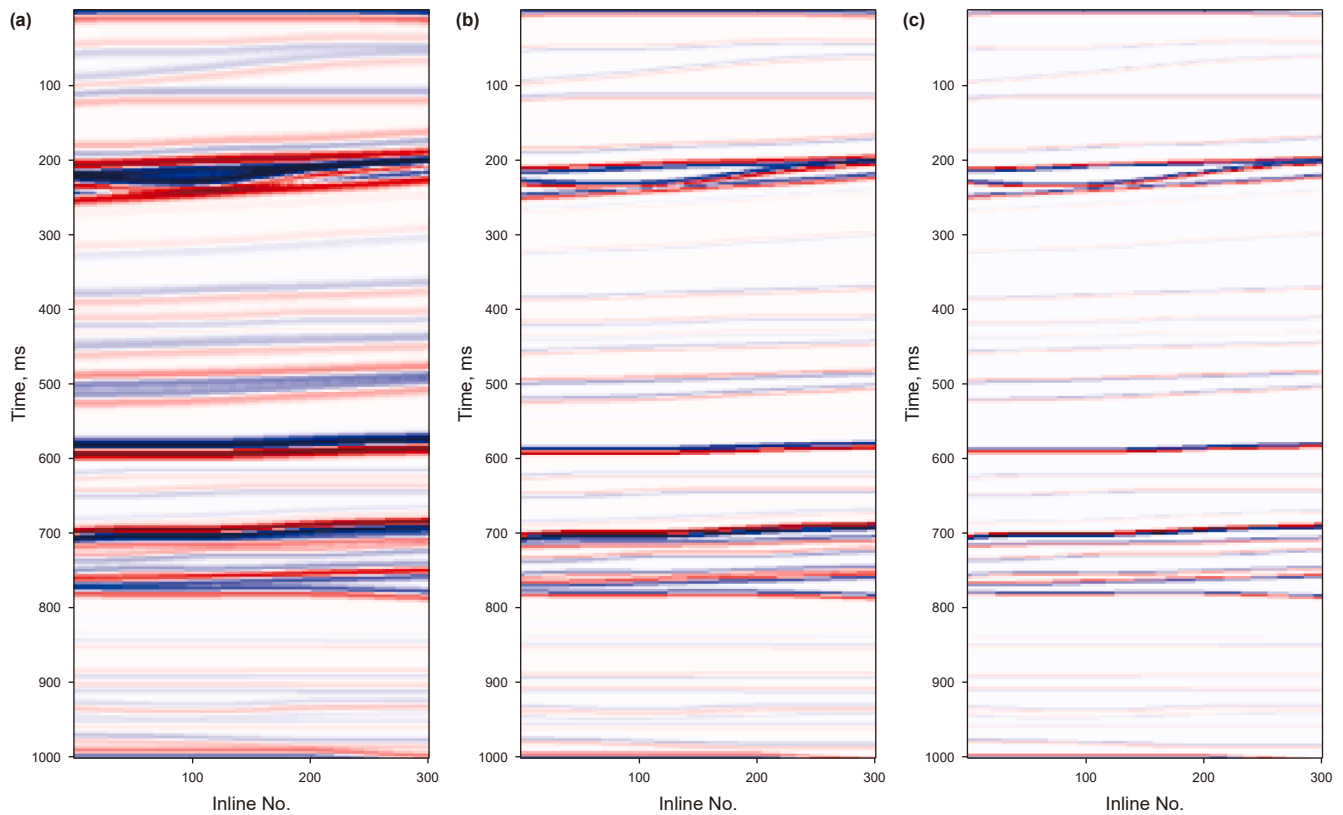


Fig. 4. Synthetic data of size 1000×300 . (a) Low-resolution synthetic data; (b) high-resolution prior data; (c) high-resolution synthetic data.

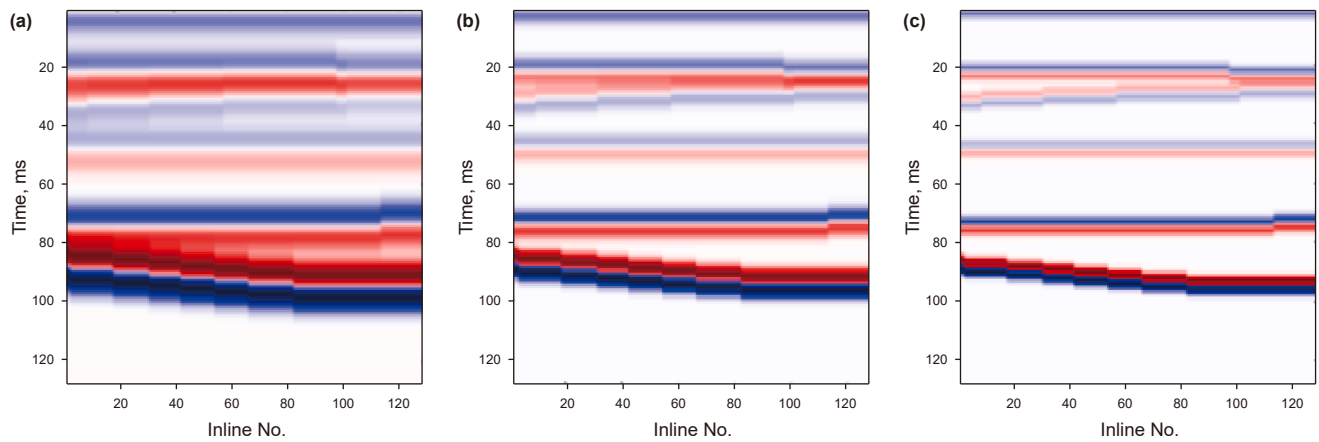


Fig. 5. Synthetic dataset. (a) Low-resolution synthetic data of size 128×128 ; (b) high-resolution prior data of size 128×128 ; (c) high-resolution synthetic data of size 128×128 .

geological structures become clearer in the prediction results, with a noticeable improvement in the precision of the phase axis compared to the low-resolution data, and an overall increase in resolution. The results also exhibit a high degree of similarity to the true high-resolution data in Fig. 8(b), indicating that DAKD-Net can effectively enhance the resolution of seismic data.

To more intuitively demonstrate the spectral extension effect of the student network, a single trace is randomly selected from the prediction results for comparison in both the time and frequency domains. Fig. 9 compares the normalized amplitude spectra of the high-resolution prediction results from DAKD-Net with those of actual high-resolution data and low-resolution data. It can be observed that the normalized amplitude spectra of the student

network's prediction results exhibit higher dominant frequencies and bandwidths compared to the original low-resolution seismic data, and are more closely aligned with the actual high-resolution seismic data. This indicates that the student network can successfully recover the high-frequency components in the seismic data, with the amplitude spectrum improved by approximately two octaves relative to the original data.

To highlight the advantages of DAKD-Net, we constructed a traditional FCN for high-resolution prediction on synthetic data using the same dataset and training process. FCN is a classic deep learning architecture commonly used for image reconstruction tasks, featuring strong feature extraction and recovery capabilities. It was employed as a control experiment against DAKD-Net. Both

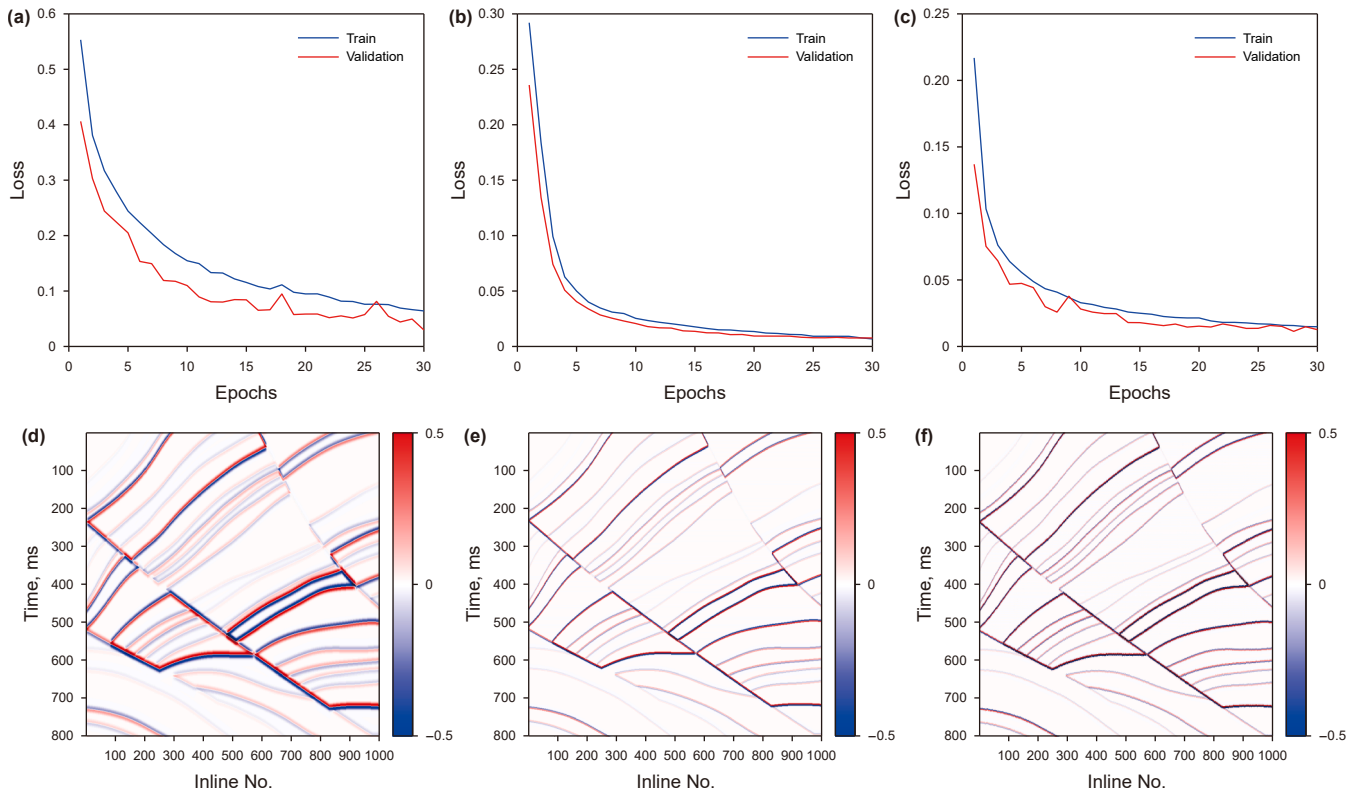


Fig. 6. Teachers' network loss curves and high-resolution results corresponding to different loss weights. (a) 0.5 loss weight; (b) 0.7 loss weight; (c) 0.9 loss weight; (d) 0.5 weighted high resolution result; (e) 0.7 weighted high resolution result; (f) 0.9 weighted high resolution result.

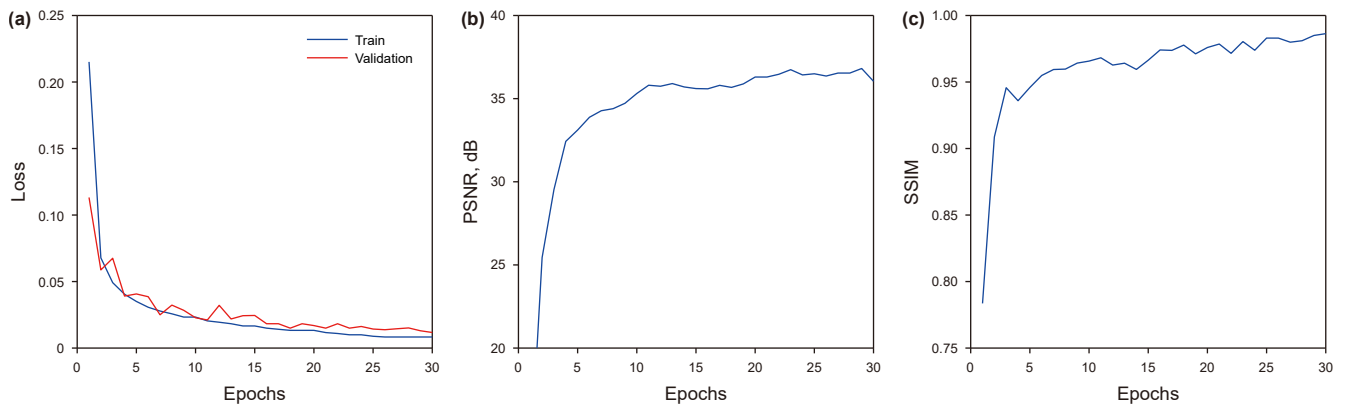


Fig. 7. Loss functions and evaluation metrics of DAKD-Net. (a) Loss function of DAKD-Net on the training and validation sets; (b) PSNR values of DAKD-Net on the validation sets; (c) SSIM values of DAKD-Net on the validation sets.

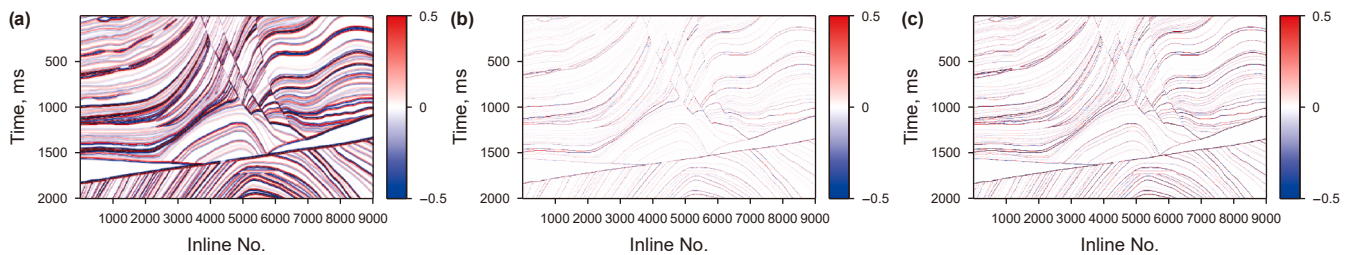


Fig. 8. Prediction results of DAKD-Net on Marmousi2 model data. (a) Original low-resolution seismic section; (b) ground-truth high-resolution seismic section; (c) high-resolution prediction result from DAKD-Net.

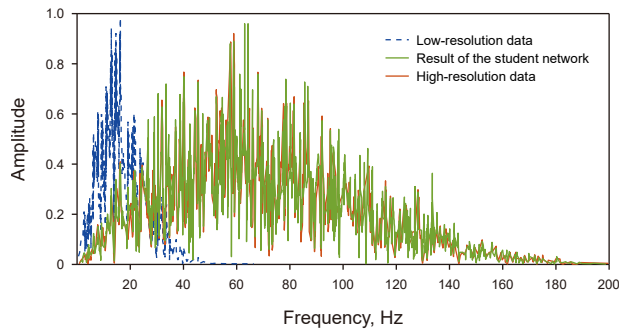


Fig. 9. Comparison of normalized amplitude spectra for single-trace data from the Marmousi2 model.

networks utilized 7000 sample sets of 1000×300 images, randomly cropped into 128×128 training patches. The optimizer was Adam with a learning rate of 1×10^{-4} , batch size of 16, and 30 iterations, consistent with the experimental settings in Fig. 11. DAKD-Net employs the proposed teacher-student network combined with domain adaptation, while FCN consists of a standard image reconstruction fully convolutional architecture with a 5-layer encoder/decoder and three 3×3 convolutional layers.

Regarding parameter scale and efficiency, PyTorch 2.0 statistics reveal that DAKD-Net contains approximately 6.84 million trainable parameters, compared to only 1.29 million for FCN. Despite having more channels and a deeper structure, DAKD-Net achieves a single training iteration (forward + backward) for a 128×128 patch in approximately 0.193 s on CPU, compared to 0.118 s for

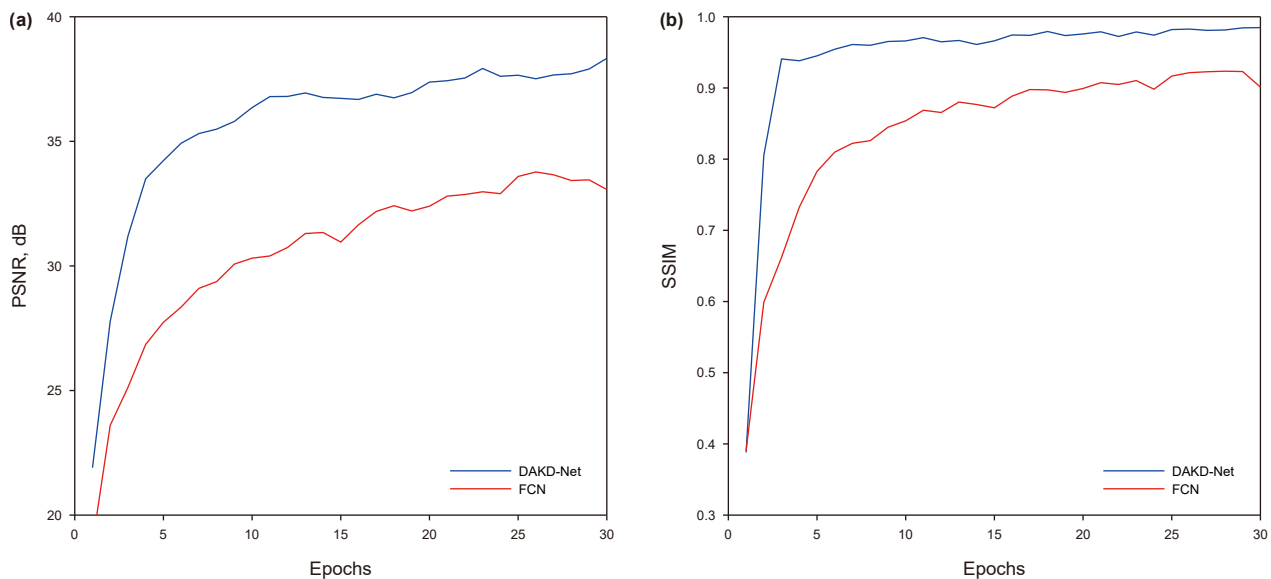


Fig. 10. Evaluation metrics of FCN and DAKD-Net. (a) PSNR values of FCN and DAKD-Net on the validation set; (b) SSIM values of FCN and DAKD-Net on the validation set.

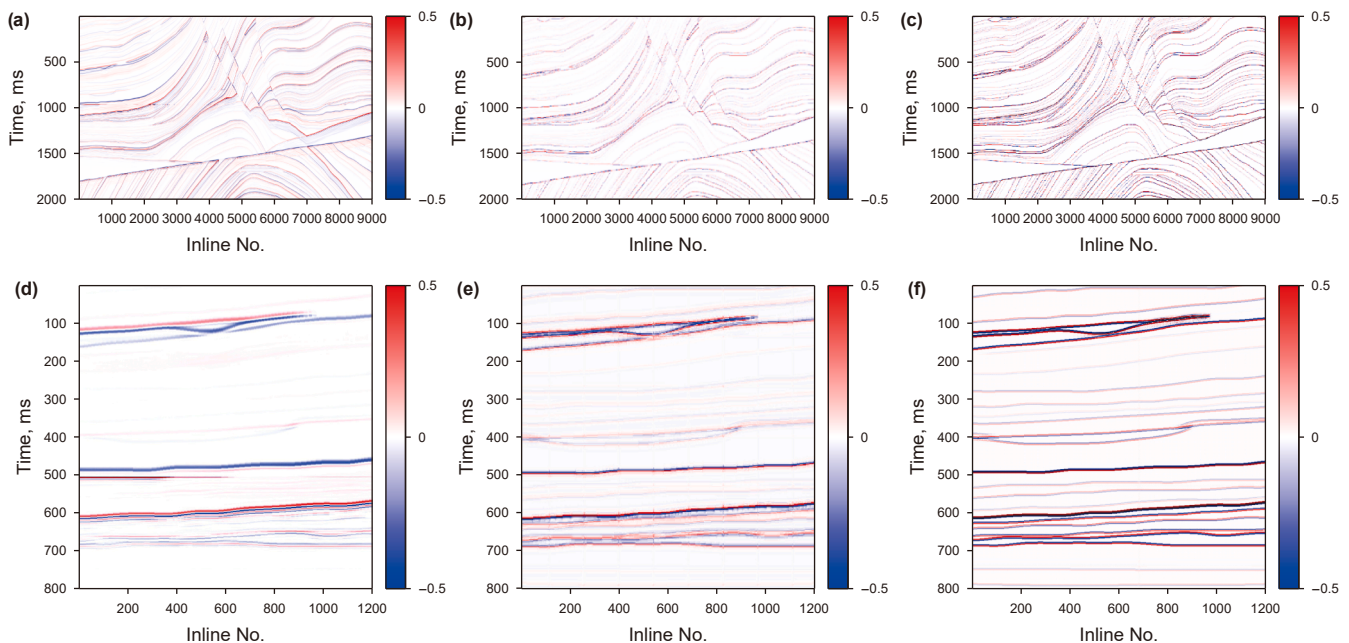


Fig. 11. Comparison of de-convolution, FCN, and DAKD-Net prediction results. (a) De-convolution results; (b) FCN results; (c) DAKD-Net prediction results; (d) local region of de-convolution results; (e) local region of FCN prediction results; (f) local region of DAKD-Net prediction results.

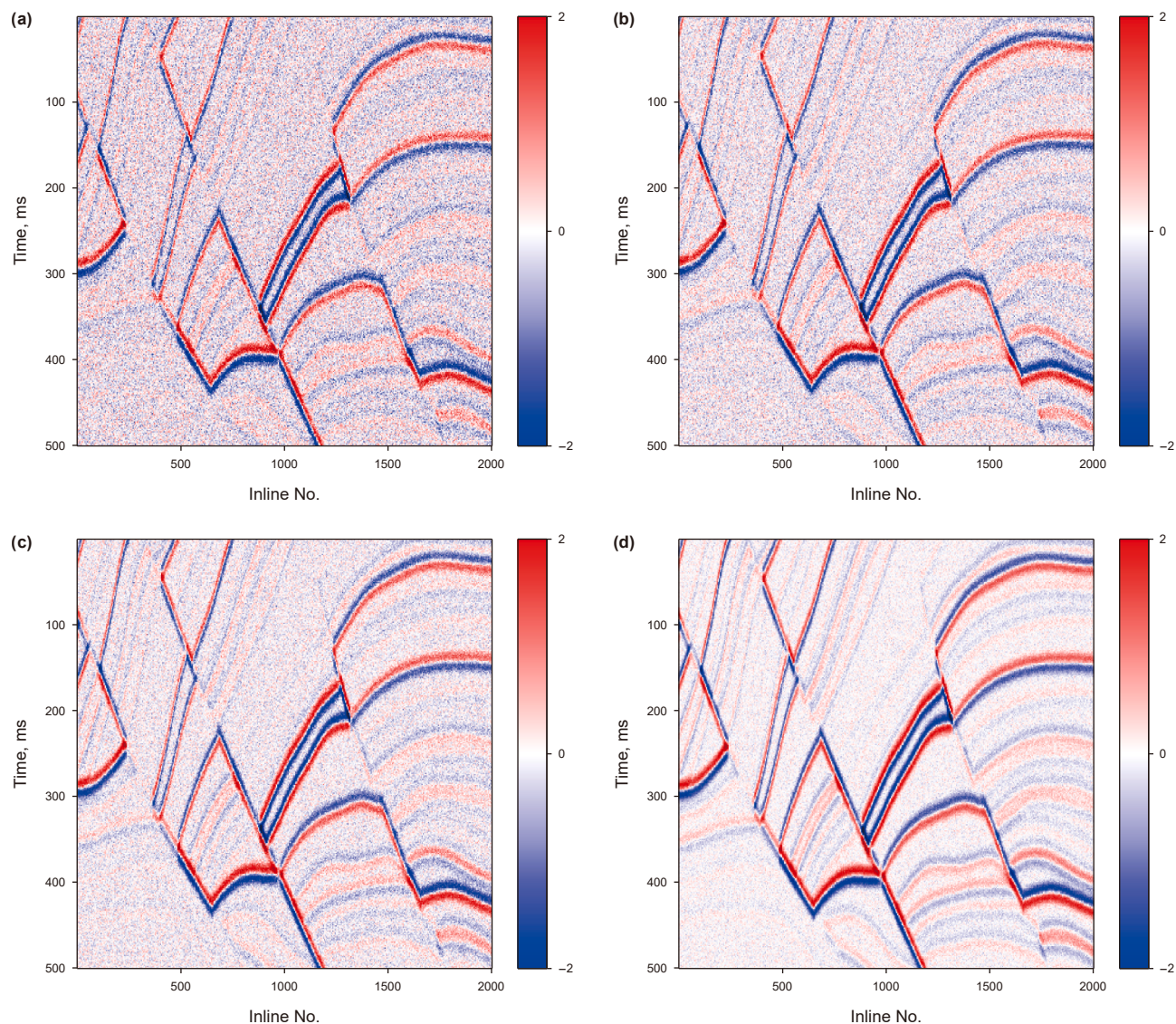


Fig. 12. Synthetic data with different SNR. (a) SNR = 1 dB; (b) SNR = 3 dB; (c) SNR = 5 dB; (d) SNR = 10 dB.

FCN. Pure inference forward times are 9.6 ms and 6.1 ms respectively, representing a factor-of-1.6 difference without introducing an order-of-magnitude increase in computational cost. As shown in the PSNR/SSIM curves in Fig. 10, despite having more parameters, DAKD-Net significantly enhances detail fitting capabilities in both real and synthetic domains: its PSNR and SSIM consistently outperform FCN, with smoother curves indicating superior capture of high-frequency information. In contrast, traditional FCN outputs appear relatively smooth but lack sufficient detail representation. In summary, DAKD-Net achieves higher quantitative metrics and clearer seismic structure representation while maintaining acceptable training/inference efficiency.

Fig. 11 displays high-resolution prediction results from the three methods and provides a local magnification of the seismic profile, further highlighting their differences. It can be observed that deconvolution enhances longitudinal resolution to some extent at the main reflection interface but struggles to fully leverage inter-channel structural continuity. Improvements in thin-layer boundary and phase axis continuity remain limited, and its noise resistance is weaker than that of deep learning methods. In contrast, while FCN can restore relatively clear primary structures in seismic profile predictions, it exhibits certain

shortcomings in recovering high-frequency detail information. Interlayer information appears somewhat blurred, with spurious frequencies particularly evident between 280 ms and 300 ms. In contrast, DAKD-Net delivers superior prediction performance. It not only effectively restores interlayer information but also maintains better phase axis continuity and clarity of high-frequency reflection interfaces. This results in seismograms that appear richer in detail overall, with particularly notable restoration effects in complex structural regions.

To evaluate the robustness of the DAKD-Net method under different signal-to-noise ratio (SNR) conditions, Gaussian noise with varying variances was randomly added to the Marmousi2 model data to generate noisy datasets with SNR levels of 1, 3, 5, and 10 dB. DAKD-Net was then used to perform high-resolution predictions on the noisy synthetic data. Figs. 12 and 13 show the noisy data and the corresponding high-resolution prediction results under different SNR conditions. It can be observed that DAKD-Net maintains good robustness under various noise levels, effectively suppressing random noise while achieving high-resolution reconstruction. Overall, the higher the SNR, the better the noise suppression performance. From the comparison, it is evident that as the SNR decreases, the degree of noise interference

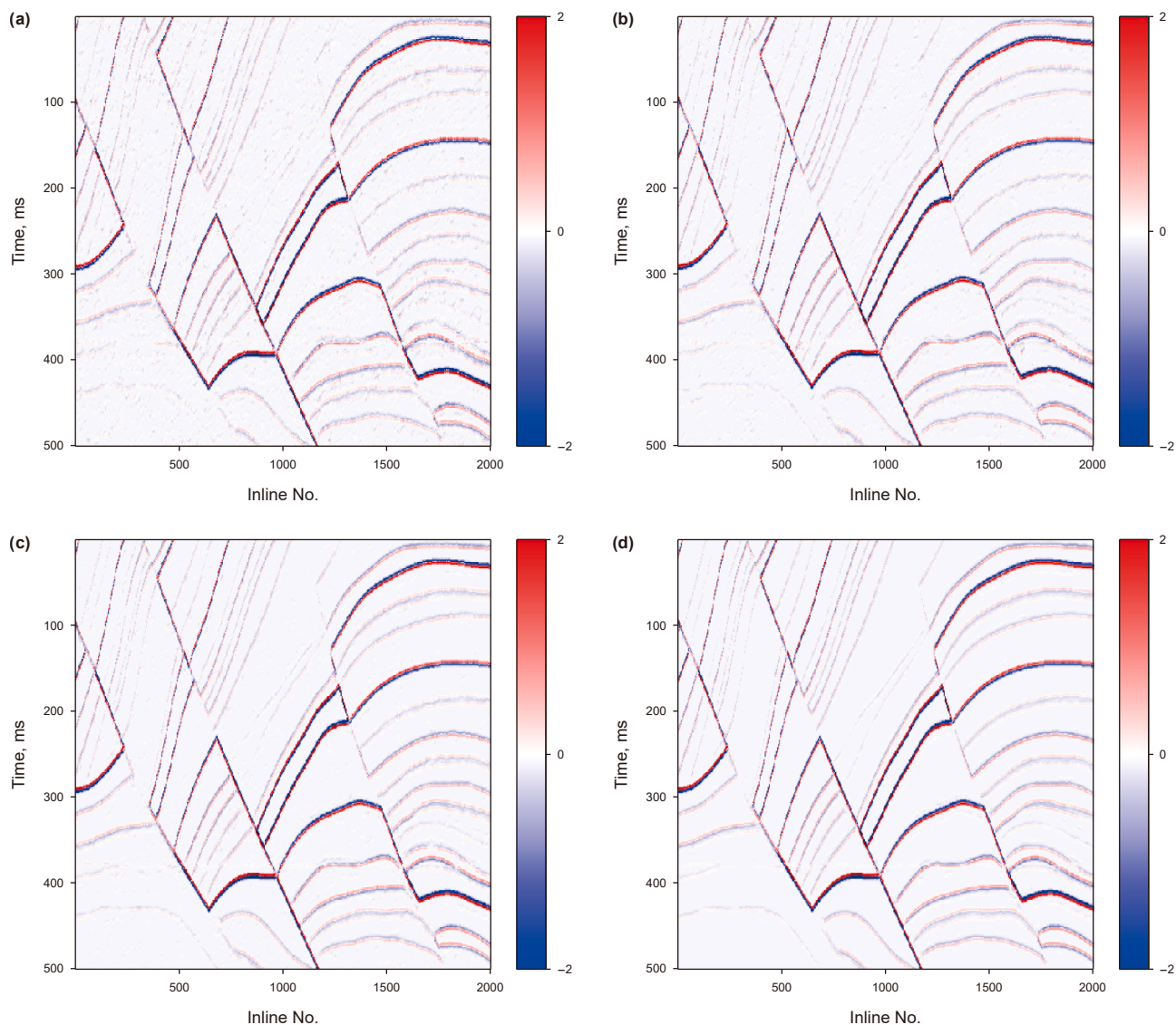


Fig. 13. High-resolution prediction results under different SNR conditions. (a) SNR = 1 dB; (b) SNR = 3 dB; (c) SNR = 5 dB; (d) SNR = 10 dB.

increases, and the effectiveness of the high-resolution processing is reduced. In particular, when the SNR drops below 3 dB, the influence of noise becomes significantly more pronounced, making it difficult for the high-resolution network to achieve the expected reconstruction accuracy. Under such conditions, it is recommended to apply noise suppression processing before performing high-resolution reconstruction.

To rigorously evaluate the generalization ability and physical validity of DAKD-Net, we performed a cross-dataset validation on the SEG Salt Model, selecting two representative profiles (Inline 100 and Inline 340) for demonstration. Unlike the stratigraphy-dominated Marmousi2 model used for training, the Salt model possesses distinct geological characteristics (as shown in Fig. 14(d)), including large homogeneous salt bodies and strong velocity contrasts, representing a significant domain shift.

To verify the generalization, we constructed a test set containing low-frequency inputs and corresponding high-frequency labels. During the testing phase, we fed the low-frequency data (25 Hz) as the input into the pre-trained network (Fig. 14(a) and

(d)). The high-frequency labels (80 Hz) were used exclusively for subsequent comparison and result evaluation (Fig. 14(b) and (e)) and were not involved in network inference or parameter updates. In the low-frequency input, salt dome boundaries are blurred, and thin sedimentary layers are indistinguishable. In contrast, the DAKD-Net predictions (Fig. 14(c) and (f)) significantly improve the vertical resolution, recovering sharp salt boundaries and continuous events that closely match the high-frequency ground truth. Furthermore, the DAKD-Net results maintain a clean and artifact-free background within the homogeneous regions of the salt dome interior, while effectively sharpening structural edges. This confirms that the network has learned the fundamental physical mapping of spectral extension rather than overfitting to specific textures, thereby demonstrating that the network possesses robust generalization capability and physical reliability.

3.3. Real seismic data

The trained model is tested using real seismic data. The actual seismic data is from the Tarim Basin in Xinjiang, specifically from

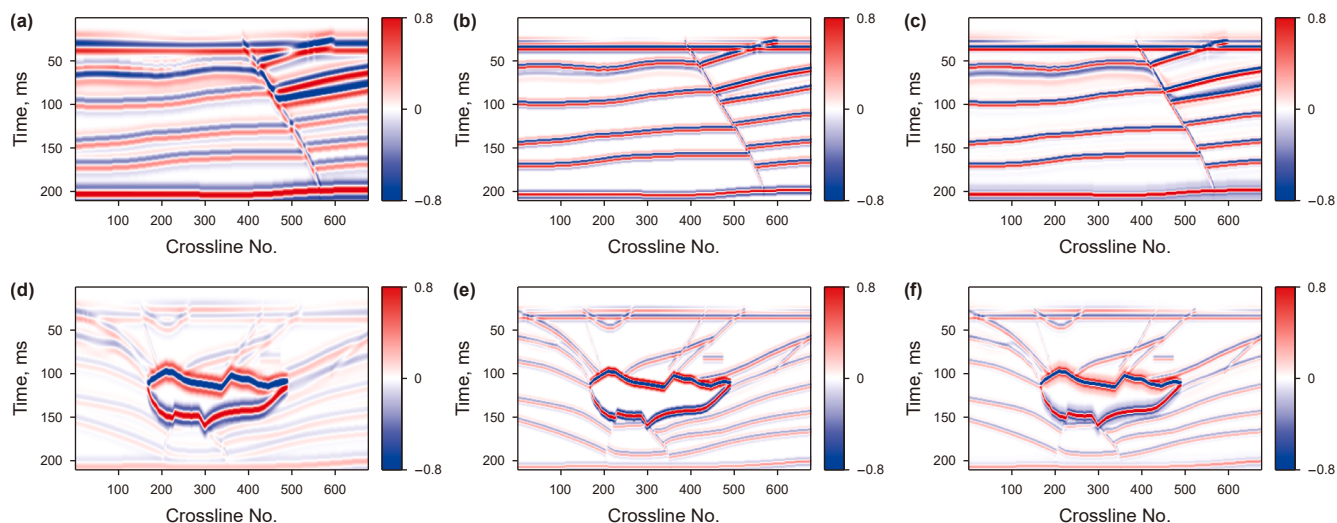


Fig. 14. Cross-dataset generalization results on the SEG Salt Model. Top row: Inline 100; bottom row: Inline 340. (a), (d) Low-resolution inputs (25 Hz). (b), (e) high-resolution ground truth labels (80 Hz). (c), (f) super-resolution results predicted by DAKD-Net.

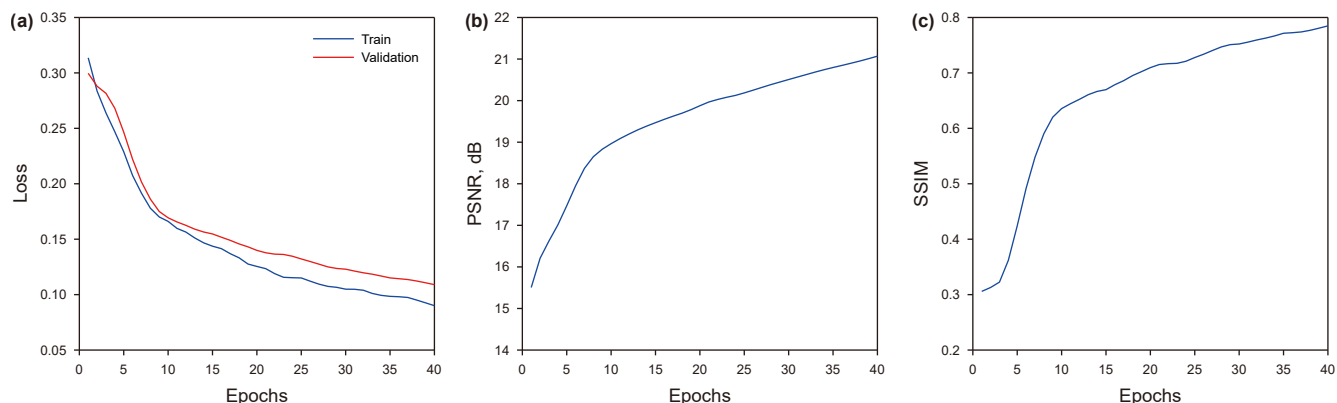


Fig. 15. Loss functions and evaluation metrics for DAKD-Net during domain adaptation. (a) Loss curves for training and validation sets; (b) PSNR curve; (c) SSIM curve.

the Tazhong Oil Field. The study area is located in the southwest part of the Akekule uplift, a third-level structural unit on the northeastern sag of the Tazhong Depression. The research target layer is the Middle Triassic oil group, with sand body thickness varying between 0 and 60 m, exhibiting strong geological heterogeneity and significant lateral variation. The actual seismic data has a size of Inline 401 $\times$ Xline 401 and a sampling rate of 2 ms.

Since the Marmousi2 model is a simulated seismic data model, it can be effectively used to validate and train high-resolution reconstruction methods. However, there are certain differences between its seismic wavelets and frequency characteristics and those of actual field data. These differences cause the prediction results to fall short of expectations when a network trained using the Marmousi2 model is directly tested on real seismic data. To address this issue, a transfer learning approach is introduced. The specific steps are as follows: First, a seismic profile of size time 500 $\times$ Xline 401 in the Inline direction is selected from the actual seismic data. This profile is divided into several 128 $\times$ 128 sized data blocks, from which 20 blocks are randomly selected as low-resolution input data. Then, using the spectral extension method, high-resolution data is constructed from these selected low-resolution data, providing high-precision label data for subsequent high-resolution predictions. A total of 20 pairs of high-

resolution and low-resolution seismic data are obtained, which are used as the training dataset for transfer learning fine-tuning. The pre-trained DAKD-Net is fine-tuned using these training datasets by freezing the parameters of the first few layers of the network. Only the last two convolutional layers and residual layers are trained, with their weights updated to enhance the network's generalization ability. A seismic profile that was not involved in the transfer learning process is selected for testing to verify the effectiveness and applicability of this method on actual field data.

Fig. 15 demonstrates that during domain-adaptive training, the loss function of DAKD-Net steadily decreases with iterations while the training/validation curves closely align, indicating that high-level parameter transfer does not introduce significant overfitting. Concurrently, PSNR and SSIM steadily improve over 40 epochs, ultimately stabilizing near 21 dB and 0.82 respectively. This demonstrates that moderate fine-tuning of higher-level parameters explicitly enhances real-domain reconstruction quality, achieving stronger generalization and structural preservation capabilities.

Fig. 16 presents the high-resolution prediction results of various methods applied to actual seismic data. This includes the original seismic profile (Fig. 16(a)), the predictions from the DAKD-Net method before fine-tuning (Fig. 16(b)), the predictions from the

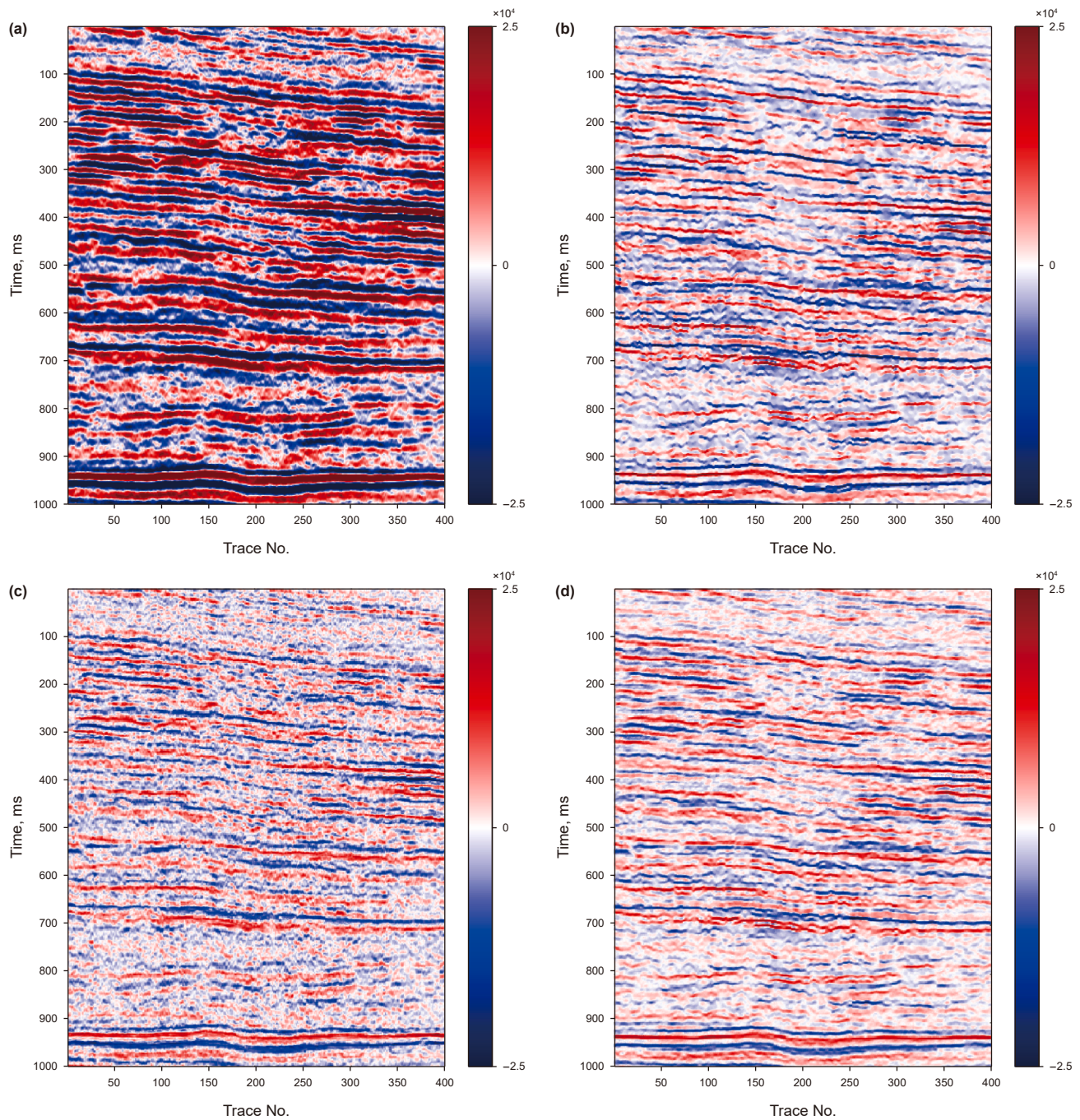


Fig. 16. Comparison of real seismic profile prediction results. (a) Original seismic profile; (b) DAKD-Net prediction result without transfer learning; (c) FCN prediction result; (d) DAKD-Net prediction result after transfer learning.

FCN method (Fig. 16(c)), and the predictions from the DAKD-Net method after fine-tuning (Fig. 16(d)).

It can be seen that when the network trained with the Marmousi2 synthetic data is directly applied to high-resolution prediction of actual seismic data, the prediction results show significant deviations. The high-resolution prediction results after fine-tuning via transfer learning are clearly superior to those before fine-tuning. The predicted profile shows more continuous reflection axes, clearer structural features, and better consistency with the actual high-resolution data. When comparing with the

FCN method's prediction results, it is observed that the FCN method improves resolution to some extent. However, the continuity of the reflection axes is not as good as that of the high-resolution seismic data network. Some stratigraphic boundaries exhibit blurring and distortion, especially in complex structural areas, where the stability of the prediction results is weaker.

A comparison is made between the DAKD-Net method and traditional spectral whitening techniques. Spectral whitening is one of the commonly used methods for high-resolution seismic processing, which aims to flatten the amplitude spectrum within

the effective frequency range without altering the phase of the signal, thereby compensating for high-frequency components. The results are illustrated in Fig. 17. The spectral whitening method processes the data on a single trace basis in the frequency domain, neglecting inter-trace relationships. As a result, the entire profile after spectral whitening exhibits noticeable noise, leading to a lack of clarity. In contrast, the DAKD-Net method effectively suppresses noise while preserving more structural details, thereby avoiding the issue of overall enhanced noise that can arise from frequency compensation.

Further comparison of the spectral characteristics between the original seismic data and the high-resolution prediction results is shown in Fig. 18. As illustrated in the figure, the effective frequency band of the original data is mainly concentrated between 20 Hz and 40 Hz, with energy gradually attenuating as frequency increases. After high-resolution processing using DAKD-Net, the data's frequency bandwidth is effectively broadened, and the energy in the high-frequency range is significantly enhanced. The frequency range is expanded to 20–80 Hz. The dominant frequency is increased, indicating that the network has successfully compensated for the high-frequency components of the seismic data. Moreover, in the low-frequency range, the spectral curve of the predicted data closely matches that of the original data. This demonstrates that while enhancing high-frequency information, DAKD-Net does not introduce additional low-frequency distortion or energy loss.

Finally, the DAKD-Net method was employed for high-resolution processing of three-dimensional seismic data, validated against well log data to analyze the thin layer characterization capability of the seismic data and the well-seismic consistency. Fig. 19 illustrates the frequency enhancement effect on vertical seismic profiles. The red solid line represents the reflection coefficients calculated from well data, convolved with synthetic seismic records generated from different dominant frequency Ricker wavelets. Fig. 19(a) shows the vertical seismic profile extracted from the original three-dimensional data volume, compared with the synthetic seismic record generated using a dominant frequency of 25 Hz Ricker wavelet. Fig. 19(b) displays the

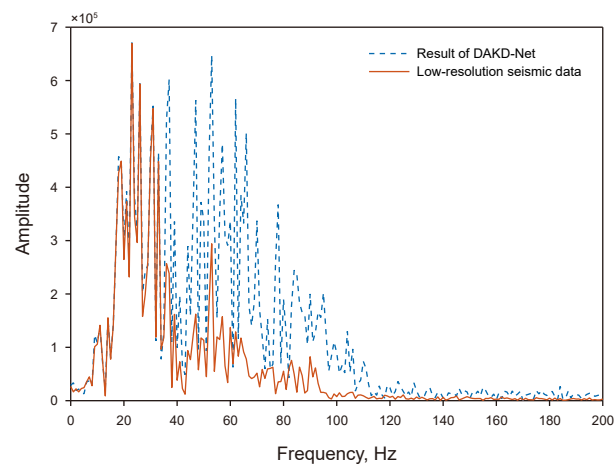


Fig. 18. Spectral comparison between the original seismic data and the high-resolution predictions obtained from DAKD-Net.

seismic profile extracted from the high-resolution processed three-dimensional data volume, overlaid with the synthetic seismic record generated using a dominant frequency of 65 Hz Ricker wavelet. It is evident that the synthetic records correspond well with the seismic profiles, indicating that the high-resolution seismic data post-frequency enhancement can accurately reflect subsurface geological information. The original seismic profile exhibits lower resolution, limiting its capability to characterize thin layers; for instance, two thin layers at a depth of 3040 ms are difficult to identify in the original profile, whereas these layers are clearly visible in the high-resolution profile, demonstrating the method's effectiveness in enhancing thin layer recognition. The high-resolution processed profiles not only exhibit higher vertical resolution, making details clearer, but also show significant improvement in the continuity of the same-phase axes laterally, resulting in clearer inter-layer boundaries and a more coherent overall structure. The reflection waveforms of the synthetic

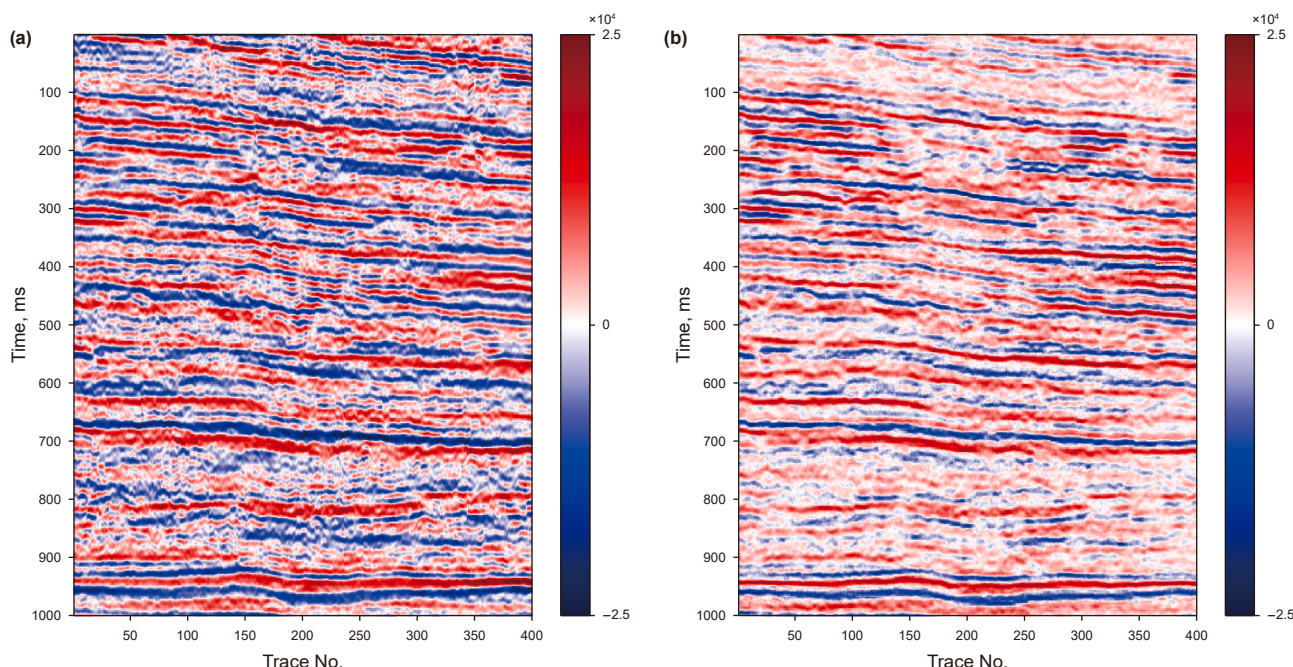


Fig. 17. Comparison between DAKD-Net and the spectral whitening method. (a) Results from the spectral whitening method; (b) predictions obtained using DAKD-Net.

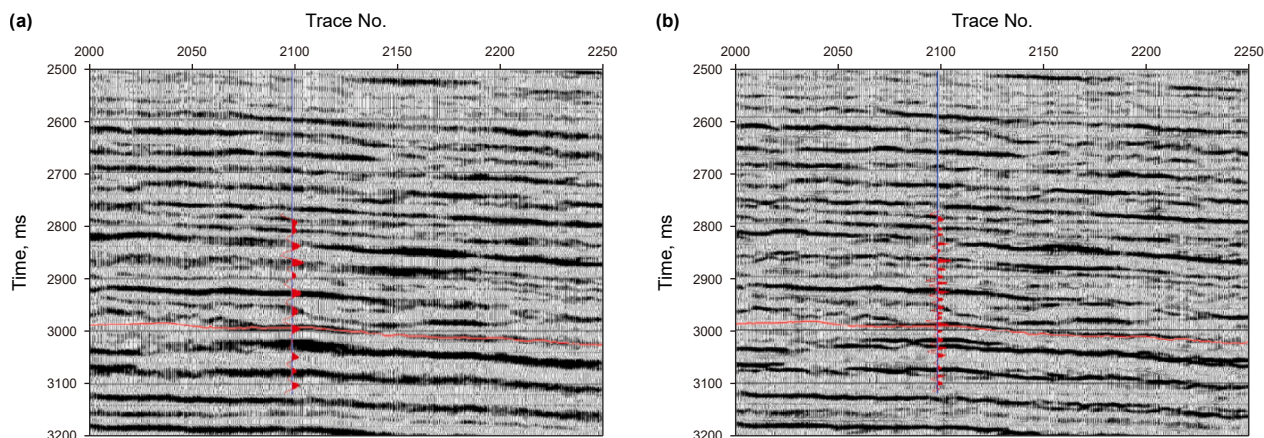


Fig. 19. Composite record and seismic profile overlay map. (a) Overlay display of original seismic profiles and 25 Hz Ricker wavelet composite records; (b) high-resolution processed seismic profiles and 65 Hz Ricker wavelet composite records are overlaid and displayed.

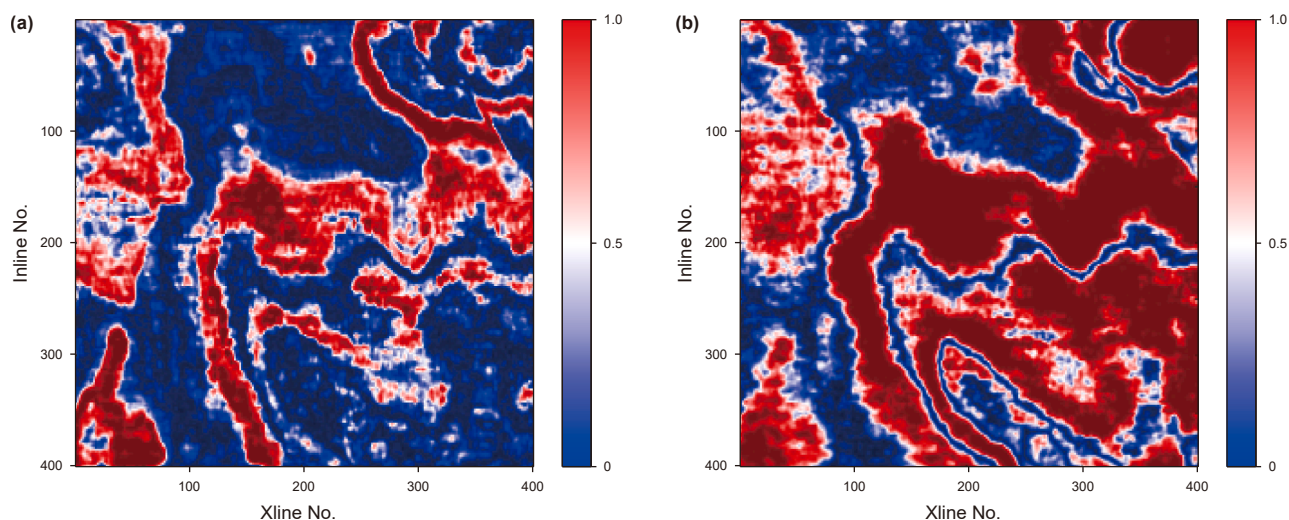


Fig. 20. RMS amplitude horizontal slice comparison. (a) RMS amplitude horizontal slice before high-resolution processing; (b) RMS amplitude horizontal slice after high-resolution processing.

seismic records closely match the processed seismic profiles. These analyses confirm that the DAKD-Net method not only enhances the resolution of seismic data but also preserves the authenticity of the seismic signals, demonstrating a high degree of correlation with well log data. This further validates the effectiveness and reliability of the proposed method.

Fig. 20 compares the changes in root mean square (RMS) amplitude horizontal slices of the 3D seismic data before and after spectral extension. It can be observed that the RMS amplitude of the original data before spectral extension exhibits relatively blurred characteristics, with local details lost. In particular, in areas with complex structures, the amplitude variations are smoother, making boundary identification more difficult. After spectral extension, the RMS amplitude distribution becomes clearer. Local energy is enhanced, making thin-layer features and structural boundaries more prominent. This allows for a more accurate depiction of detailed variations in the seismic data, as well as reservoir boundaries and structural features. These results demonstrate that the spectral extension method significantly improves seismic data resolution and enhances the ability to interpret subsurface geological features.

4. Conclusion

This article presents a deep learning method for enhancing the high-resolution of seismic data, which integrates knowledge distillation and domain adaptation mechanisms—DAKD-Net. The DAKD-Net approach is not limited to single-channel seismic data; it is capable of extracting structural information from inter-channel seismic data, demonstrating higher robustness and efficiency compared to traditional spectral extrapolation methods. By incorporating physical priors through knowledge distillation, DAKD-Net enhances the structural interpretation capabilities relative to conventional fully convolutional networks (FCNs). Additionally, the implementation of a domain adaptation strategy utilizes real high-resolution spectral extrapolation results as supervision, effectively merging physical models with deep features. Experimental results on synthetic datasets and actual field data indicate that the DAKD-Net method significantly improves the clarity of seismic profile details and the continuity of the iso-chrones, which holds substantial significance for fine reservoir characterization and structural interpretation. Future work will consider the introduction of physics-constrained inversion

mechanisms and a broader range of pre-trained models to further enhance the robustness and adaptability of the network.

CRedit authorship contribution statement

Han-Peng Cai: Methodology, Conceptualization. **Hao-Nan Zhang:** Writing – original draft, Data curation. **Li-Yu Zhang:** Methodology, Conceptualization. **Suo Cheng:** Validation, Supervision.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (NSFC) under Grant No. 42474168 & No. 42130812, and R & D Center for Ultra-Deep Complex Reservoir Exploration and Development, CNPC and Engineering Research Center for Ultra-deep Complex Reservoir Exploration and Development, Xinjiang Uygur Autonomous Region (Grant No. RDCUD2024-CX01-03). We are also grateful to the related company for generously providing the field data.

References

- Chen, C., Zhou, X., 2000. Improving seismic data resolution using wavelet spectrum whitening. *Oil Geophys. Prospect.* 35 (6), 703–709 (in Chinese).
- Chen, Z., Wang, J., 2015. Research on sparse spike reflection coefficient spectrum inversion method based on compressed sensing. *Geophys. Prospect. Pet.* 54 (4), 459–466. <https://doi.org/10.3969/j.issn.1000-1441.2015.04.014> (in Chinese).
- Debye, H.W.J., Van Riel, P., 1990. Lp-norm deconvolution. *Geophys. Prospect.* 38 (4), 381–403. <https://doi.org/10.1111/j.1365-2478.1990.tb01852.x>.
- Dong, C., Loy, C.C., He, K., et al., 2014. Learning a deep convolutional network for image super-resolution. In: *Computer Vision – ECCV 2014: 13Th European Conference*, pp. 184–199. https://doi.org/10.1007/978-3-319-10593-2_13.
- Dragomireskiy, K., Zosso, D., 2013. Variational mode decomposition. *IEEE Trans. Signal Process.* 62 (3), 531–544. <https://doi.org/10.1109/TSP.2013.2288675>.
- Gao, H., Wu, X., Liu, G., 2021. ChannelSeg3D: Channel simulation and deep learning for channel interpretation in 3D seismic images. *Geophysics* 86 (4), IM73–IM83. <https://doi.org/10.1190/geo2020-0572.1>.
- Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., et al., 2014. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* 2, 2672–2680.
- Hale, D., 1981. An inverse Q-filter. *Stanford Explor. Proj. Rep.* 26, 231–243.
- Hinton, G., Vinyals, O., Dean, J., 2015. Distilling the knowledge in a neural network. *arXiv, preprint arXiv:1503.02531*. <https://doi.org/10.48550/arXiv.1503.02531>.
- Lan, N.-Y., Zhang, F.-C., Sang, K.-H., et al., 2023. Simultaneous denoising and resolution enhancement of seismic data based on elastic convolution dictionary learning. *Pet. Sci.* 20 (4), 2127–2140. <https://doi.org/10.1016/j.petsci.2023.03.011>.
- Ledig, C., Theis, L., Huszar, F., et al., 2017. Photo-realistic single image super-resolution using a generative adversarial network. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4681–4690. <https://doi.org/10.1109/CVPR.2017.19>.
- Li, J., Wu, X., Hu, Z., 2021. Deep learning for simultaneous seismic image super-resolution and denoising. *IEEE Trans. Geosci. Rem. Sens.* 60, 1–11. <https://doi.org/10.1109/TGRS.2021.3057857>.
- Li, Z., Chen, X., Li, J., et al., 2023. Time-frequency domain complex value U-Net-based spectral enhancement. In: *84th EAGE Annual Conference & Exhibition*, vol. 1. European Association of Geoscientists & Engineers, pp. 1–5. <https://doi.org/10.3997/2214-4609.202310246>.
- Lu, W., Mu, S., 1996. Neural network wavelet deconvolution. *Geophys. Prospect. Pet.* 31 (A02), 107–111 (in Chinese).
- Oppenheim, A.V., 1965. *Superposition in a Class of Nonlinear Systems*. MIT Research Laboratory of Electronics.
- Pan, S., Yang, Q., 2010. A survey on transfer learning. *IEEE Trans. Knowl. Discov. Data Eng.* 22 (10), 1345–1359. <https://doi.org/10.1109/TKDE.2009.191>.
- Robinson, E.A., 1967. Predictive decomposition of time series with application to seismic exploration. *Geophysics* 32 (3), 418–484. <https://doi.org/10.1190/1.1439873>.
- Sun, Y., Huang, J., Yu, W., et al., 2021. End-to-end high-resolution seismic processing method based on U-Net network. *Prog. Geophys.* 36 (3), 1297–1305. <https://doi.org/10.6038/pg2021EE0349> (in Chinese).
- Tang, J., Meng, T., Han, S., et al., 2021. Fault detection method of seismic data based on multi-resolution U-Net network. *Geophys. Prospect. Pet.* 56 (3), 436–445. <https://doi.org/10.3969/j.issn.1000-1441.2021.03.012> (in Chinese).
- Tian, J., Song, W., Yang, F., 2009. Enhancing the resolution of seismic data based on the generalized S-transform. *Pet. Sci.* 6, 153–157. <https://doi.org/10.1007/s12182-009-0024-x>.
- Wang, C., Jiang, J., Zhong, Z., et al., 2022. Propagating facial prior knowledge for multitask learning in face super-resolution. *IEEE Trans. Circ. Syst. Video Technol.* 32 (11), 7317–7331. <https://doi.org/10.1109/TCSVT.2022.3183578>.
- Wang, Y., 2006. Inverse Q-filter for seismic resolution enhancement. *Geophysics* 71 (3), V51–V60. <https://doi.org/10.1190/1.2192912>.
- Wu, X., Liang, L., Shi, Y., et al., 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. <https://doi.org/10.1190/geo2018-0646.1>.
- Yang, P.J., Yin, X.Y., Zhang, G.Z., 2007. Seismic signal time-frequency analysis and attributes extraction based on HHT. *Prog. Geophys.* 22 (5), 1585–1590 (in Chinese).
- Yu, S., Ma, J., Wang, W., 2019. Deep learning for denoising. *Geophysics* 84 (6), V333–V350. <https://doi.org/10.1190/geo2018-0668.1>.
- Yuan, Z., Huang, H., Jiang, Y., et al., 2019. An enhanced fault-detection method based on adaptive spectral decomposition and super-resolution deep learning. *Interpretation* 7 (3), T713–T725. <https://doi.org/10.1190/INT-2018-0243.1>.
- Yuan, Z., Jiang, Y., Huang, H., et al., 2021. Improving seismic resolution by a sequential convolutional neural network. In: *82nd EAGE Annual Conference & Exhibition 2021*, pp. 1–5. <https://doi.org/10.3997/2214-4609.202112836>.
- Zhang, H.-R., Liu, Y., Sun, Y.-H., et al., 2024. SeisResoDiff: Seismic resolution enhancement based on a diffusion model. *Pet. Sci.* 21 (5), 3166–3188. <https://doi.org/10.1016/j.petsci.2024.03.016>.
- Zhao, B., Yu, S., Nie, X., et al., 1996. Spectral modeling deconvolution and its application. *Geophys. Prospect. Pet.* (1), 101–116 (in Chinese).
- Zhou, W.H., Shi, M., Zhu, D.M., et al., 2021. Seismic data super-resolution method based on residual attention network. *Comput. Sci.* 48 (8), 24–31 (in Chinese).