Research Article

Xiaohu He, Sanfu Li*, Jianxiang Pei, Lin Hu, Fang Li, Yazheng Zhang, Zhongyu Fang, Min Zhang, Guochang Liu and Hanming Gu

# Enhancing the resolution of seismic post-stack data based on deep learning

https://doi.org/10.1515/geo-2025-0931
Received June 16, 2025; accepted November 27, 2025;
published online January 29, 2026

**Abstract:** Seismic interpretation is often restricted by the low resolution of field seismic data, which limits the characterization of subsurface geological structures and stratigraphic features. Improving the vertical resolution of seismic data has long been a challenge in seismic exploration. Recent studies have employed deep learning techniques to enhance seismic vertical resolution; however, limited attention has been paid to the role of loss function design in such enhancement. To address this gap, we propose a modified U-Net – based approach that employs a hybrid loss function to improve the vertical resolution of post-stack seismic data. Specifically, we first synthesized a substantial training dataset by convolving reflectivity models with wavelets of varying dominant frequencies. The prepared dataset was then used to train a modified U-Net architecture. The training process was guided by a loss function combining the Structural Similarity Index Measure and mean square error to optimize network parameters. We subsequently applied the trained network to a synthetic seismic image and two field datasets. The results confirm the efficacy of our proposed method in significantly improving the vertical resolution of seismic data while simultaneously attenuating random noise to a certain degree.

**Keywords:** deep learning; seismic resolution; post-stack data; interpretation; U-Net

***Corresponding author: Sanfu Li**, China Oilfield Services Limited, ZhanJiang, 524057, China, E-mail: lisf2@cosl.com.cn
**Xiaohu He,** State Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum (Beijing), Beijing, 102249, China; Hainan Branch of CNOOC Limited, Haikou, 570100, China; and CNOOC South China Sea Oil & Gas Energy Academician Workstation, Haikou, 570100, China, E-mail: hexh@cnooc.com.cn
**Jianxiang Pei, Lin Hu, Fang Li and Yazheng Zhang,** Hainan Branch of CNOOC Limited, Haikou, 570100, China; and CNOOC South China Sea Oil & Gas Energy Academician Workstation, Haikou, 570100, China, E-mail: peijx@cnooc.com.cn (J. Pei), hulin2@cnooc.com.cn (L. Hu), lifang9@cnooc.com.cn (F. Li), zhangyz@cnooc.com.cn (Y. Zhang)
**Zhongyu Fang and Min Zhang,** China Oilfield Services Limited, ZhanJiang, 524057, China, E-mail: fangzhy@cosl.com.cn (Z. Fang), zhangmin31@cosl.com.cn (M. Zhang)
**Guochang Liu,** State Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum (Beijing), Beijing, 102249, China, E-mail: guochang.liu@cup.edu.cn
**Hanming Gu,** School of Geophysics and Geomatics, China University of Geosciences, Wuhan, 430074, China, E-mail: hmgu@cug.edu.cn

## 1 Introduction

The geoscientists typically use seismic exploration to investigate subsurface structures and stratigraphic features in the search for energy resources, such as oil and gas. However, limitations in acquisition equipment, attenuation of high-frequency signals as they propagate through geological layers, and complex geological conditions result in insufficient seismic resolution. This hinders the accurate interpretation of subsurface geological structures and stratigraphy; consequently, enhancing seismic resolution is critical for improved subsurface characterization and has been pursued across various stages of seismic exploration, including acquisition, processing, and imaging. In acquisition, Naghizadeh et al. provide a way to test the effect of various geometries on both lateral and vertical resolution, which can guide the design of the acquisition geometry [1]. In processing, signal decomposition [2] and time–frequency methods [3] have been employed to increase vertical resolution, and Compensation techniques have also been introduced to mitigate attenuation effects and enhance resolution [4, 5]. Moreover, Sajid et al. implement the nonstationary differential resolution algorithm to improve seismic resolution and better preserve the relative amplitude [6]. In imaging, Perez and Marfurt correct for wavelet stretch in common-angle migration during imaging, improving lateral and vertical resolution of seismic image [7]. Despite these advancements, traditional methods still encounter significant challenges, such as difficulties in wavelet acquisition, uncertainties regarding stratigraphic $Q$ values, and severe noise disruptions. Additionally, the

complex parameters in conventional methods need to be carefully selected, and their performance is highly sensitive to parameter choices. This sensitivity can lead to unsatisfactory results if the parameters are not accurately tuned, making it difficult to achieve consistent performance across different datasets or geological settings. Because seismic data generally contain varying degrees of noise and noise cannot be completely separated from seismic signals, forcing traditional methods to repeatedly adjust parameters to balance resolution and signal-to-noise ratio of the processing results.

Recently, deep learning techniques have gained widespread application in computer vision and have also sparked significant interest in the geophysics community. This growing interest stems from the ability of deep learning models to effectively capture complex nonlinear relationships, which are often difficult to model with traditional methods. In particular, deep learning has proven to be highly effective in addressing a variety of geophysical challenges. For instance, it has been used for seismic data reconstruction [8], denoising [9], seismic inversion [10], fault detection [11–13], channel recognition [14, 15], and seismic facies classification [16, 17], etc. The successful implementation of deep learning methods within various geophysical domains demonstrates their significant potential to enhance both the accuracy and efficiency of seismic data processing. These methods often exhibit superior performance compared to conventional methods, reducing the reliance on labor-intensive manual tuning and automating the process of interpreting critical seismic features.

Numerous deep learning techniques have been introduced for seismic image resolution enhancement. In terms of network models, Li et al. train the convolutional neural network for simultaneous super-resolution and denoising of seismic image [18]. Generative networks, including generative adversarial networks (GANs), are also used to enhance seismic resolution by suppressing random noise and improving the dominant frequency [19, 20]. Additionally, diffusion models have been applied to the task of seismic resolution enhancement [21, 22], further demonstrating the potential of generative models in seismic data processing. Apart from the diversity of models, several studies have focused on dataset construction. Some authors have considered specific features of seismic field data [23, 24], while Zhang et al. develop a domain adaptation method aimed at bridging the gap between synthetic and real seismic data [25]. Jo et al. take into account the seismic attenuation by using time-variant wavelets [26]. To better support structural interpretation, additional prior knowledge [27–29] and edge information [30] have been incorporated into training process. Moreover, Kaur et al. employ deep learning techniques to estimate inverse Hessian matrices, thereby improving seismic image resolution [31]. Dong et al. use a deep learning framework to establish the mapping relationship from the sparse-shot image to the dense-shot image, thereby acquiring the high-resolution seismic images [32]. Overall, previous studies have improved deep learning–based seismic resolution enhancement from different perspectives, including network architectures, dataset preparation, and incorporation of geological constraints. However, relatively little attention has been given to the role of the loss function, which fundamentally determines the optimization direction of the network. Therefore, we address this gap by designing an effective loss function that balances structural preservation and detail recovery to improve seismic vertical resolution.

This paper presents a convolutional neural network (CNN) – based method for enhancing seismic image resolution, specifically employing a modified U-Net architecture. To optimize the model performance, we generate a diverse and comprehensive dataset that includes a large batch of noisy, low-resolution seismic images, along with their corresponding high-resolution counterparts. The dataset is carefully constructed to represent a variety of subsurface geological features, including different fault structures, folds, and stratigraphic variations, which are essential for capturing the complexity of real seismic data. A hybrid loss function, combining structural similarity (SSIM) and mean squared error (MSE), is adopted to guide the optimization process, thereby ensuring a balance between preserving structural fidelity and recovering fine details. To evaluate the efficacy of the proposed method, we eventually test the trained network using one synthetic seismic dataset and two real datasets. The performance of our trained network is thoroughly assessed by comparing the resolution-enhanced results with the original seismic images.

## 2 Methods

We address vertical-resolution enhancement of seismic images via supervised learning. First, we construct a large synthetic training set by convolving identical reflectivity models with two Ricker wavelets of different dominant frequencies (see Figure 1). The lower-frequency result serves as the training sample, and the higher-frequency counterpart serves as the supervision target. We then train the proposed architecture on this paired dataset using a composite loss that linearly combines Structural Similarity Index Measure (SSIM), which emphasizes global structural fidelity, with Mean Square Error (MSE), which constrains pixel-level details; the relative weights of the two terms are determined

empirically for optimal performance. The trained model is subsequently evaluated on independent synthetic data that are not involved in training. Finally, we provide additional implementation details of the network training. In the following subsections, each step of the workflow will be described in greater detail.

## 2.1 Data preparation

We approach the resolution enhancement task as a supervised learning problem. However, in practice, seismic data is typically low-resolution and heavily contaminated with noise, acquiring high-resolution field data without noise is challenging. To address this, a substantial amount of training samples and corresponding labels is generated for the supervised learning task, ensuring the accuracy and validity of the labels.

Using the subsurface structural modeling workflow proposed by Wu et al. [33], we generate a large number of reflectivity models that incorporate stratigraphic dips, folds, and faults. In constructing these seismic reflectivity models, we carefully account for the geological characteristics of multiple target areas, creating a diverse range of reflectivity models to compensate for the lack of sufficient real-world training data. Each model has a size of $128 \times 128$ and consists of 40–60 layers. Each model includes 5–8 folds, with vertical heights ranging from 3 to 8 and lengths and widths ranging from 30 to 50, so that the undulations of the reflectivity interfaces remain relatively gentle. For the fault structures, only normal faults are considered. Each model contains 2–5 faults, with strikes between 45° and 135°, dip angles between 60° and 90°, and displacements ranging from 2 to 10 pixels. These parameter ranges, defined in accordance with the geological features of the work area, ensure the validity and

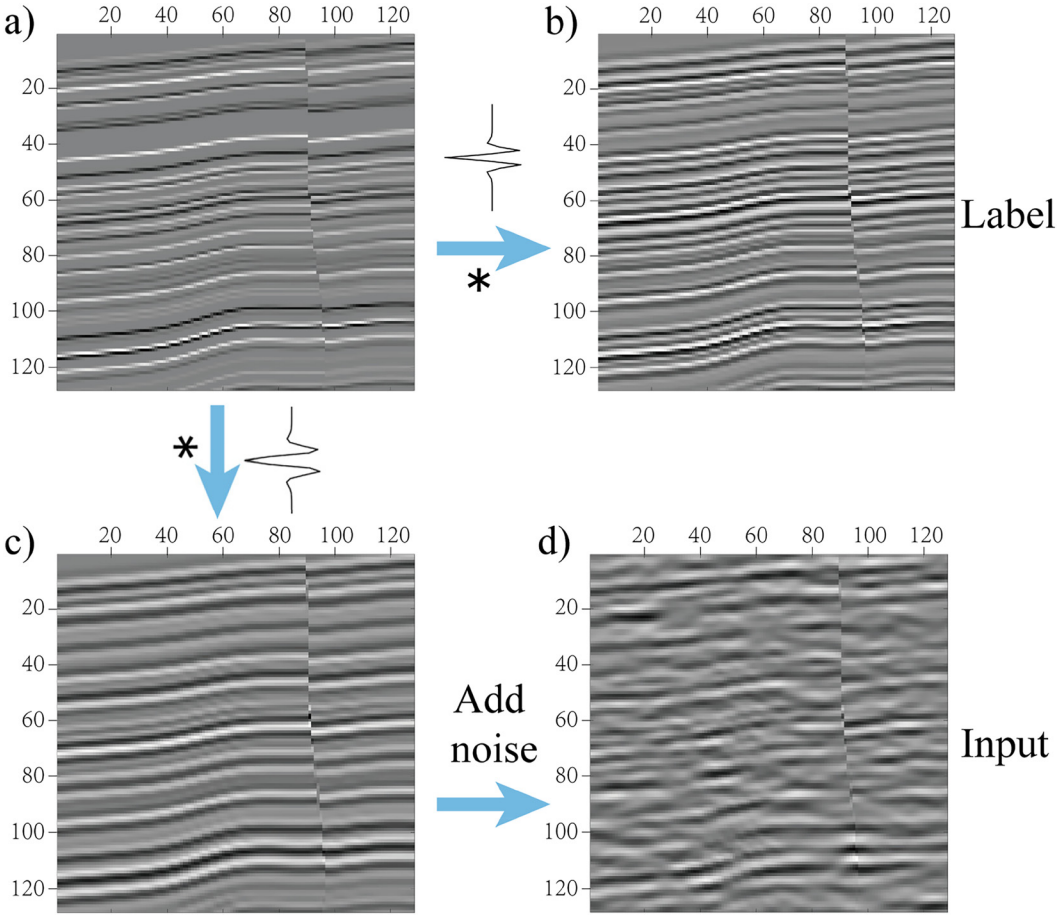


**Figure 1:** The process diagram of creating the training set. (a) Reflectivity model including stratigraphic dip, folds, and faults. (b) High-resolution seismic label generated by convolving the reflectivity with a higher-frequency Ricker wavelet (dominant frequency 30–40 % higher than the low-frequency case). (c) Low-resolution seismic image generated by convolving the reflectivity with a lower-frequency Ricker wavelet (dominant frequency 15–25 Hz). (d) Training input obtained by adding random noise to the low-resolution image. Asterisks in the figure indicate convolution operations.

diversity of the synthetic reflectivity models. Seismic training samples are generated by convolving the reflectivity model with a lower-frequency wavelet, while the corresponding high-resolution labels are obtained by applying a higher-frequency Ricker wavelet to the same reflectivity model. To ensure the diversity of the seismic training samples, the main frequency of the wavelet used for generating the training samples is randomly selected within the range of [15, 25] Hz. The main frequency of the wavelet used for synthesizing the high-resolution labels is increased by 30–40 % compared to that used for the training samples. Since noise is commonly present in real seismic data, we add colored random noise, making the training samples more representative of real seismic data. Figure 1 illustrates the workflow for generating the training dataset used for seismic resolution enhancement.

Based on the workflow diagram shown in Figure 1, we automatically generated 1,500 pairs of training data for training the resolution enhancement network. Figure 2 displays a portion of the dataset, where the first row represents low-resolution seismic training samples with noise, while the second row depicts corresponding noise-free high-resolution labels.

## 2.2 Network architecture

The network we used is a modified version of the U-Net architecture [34]. Due to its encoder-decoder architecture and the incorporation of skip connections, U-Net exhibits high versatility for various tasks within seismic exploration. This inherent flexibility allows for easy modifications and adaptations, making it a widely adopted deep learning framework in this domain. The network we employed has 4 down-sampling steps, and the channel number of the first convolutional layer is 64 (see Figure 3). The decision to use 4 down-sampling steps and to set the number of channels in the first convolutional layer to 64 was guided primarily by widely adopted practices in prior U-Net implementations for seismic and medical image processing tasks. In general, increasing the network depth allows the extraction of higher-level features and can improve performance. However, our input seismic images have a resolution of $128 \times 128$. With each down-sampling operation halving the spatial dimensions, four steps reduce the feature maps to $8 \times 8$. Further down-sampling would result in excessively small feature maps and loss of spatial detail. While we did not perform an exhaustive hyperparameter optimization, we conducted preliminary experiments with different numbers of down-sampling steps and initial channel sizes. The selected configuration achieved stable training and satisfactory results without excessive computational cost, and is consistent with common setups reported in related work.

## 2.3 Loss functions

During training, the loss function plays a crucial role. It measures the discrepancy between the model's predictions and the ground truth, and guides the optimization of network parameters by minimizing this difference. Selecting an appropriate loss function for a specific task has a substantial impact on the training process and the final performance. A well-designed loss function can effectively guide the model toward faster convergence and lead to a better

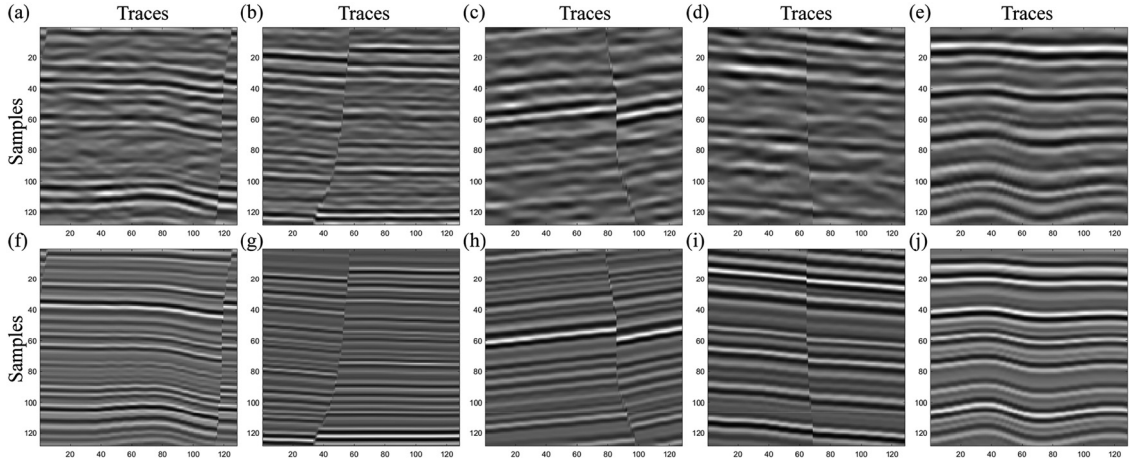


**Figure 2:** A subset of our training set, where the first row represents low-resolution seismic training samples with noise, while the second row depicts corresponding noise-free high-resolution labels.

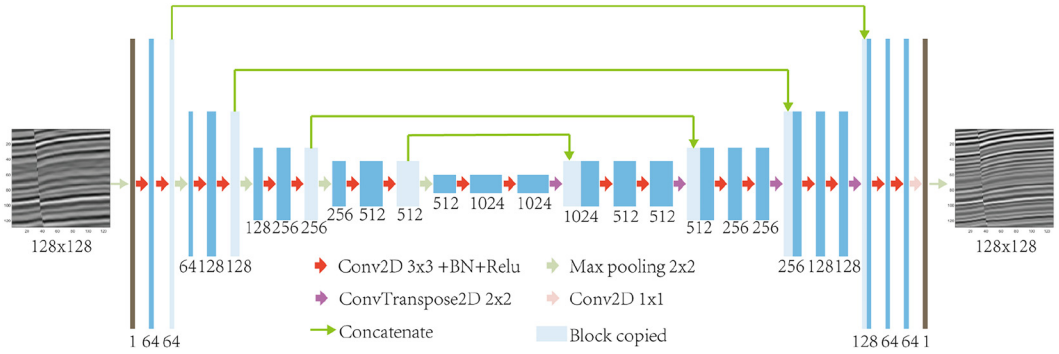


**Figure 3:** The network used for seismic resolution enhancement.

solution [35, 36]. In the task of enhancing seismic resolution, to better preserve the structural characteristics of seismic images and restore stratigraphic details, we optimized the network using a weighted combination of the Structural Similarity Index Measure (SSIM) and Mean Squared Error (MSE). MSE emphasizes the high-frequency components of the image, while SSIM captures the texture and structural information of the image and provides favorable optimization properties. By combining these two complementary metrics, the proposed loss function improves both pixel-level accuracy and structural fidelity. The details of the loss function are described below.

### 2.3.1 Structural Similarity Index Measure (SSIM)

Pixel-level image evaluation metrics are unable to preserve image structure [37]. SSIM, on the other hand, accounts for texture structure, aligning more closely with the human visual system (HVS) and is widely used in image quality assessment [38]. SSIM is primarily composed of three components [38]: luminance $l\left(\widehat{\mathbf{Y}},\mathbf{Y}\right)$, contrast $c\left(\widehat{\mathbf{Y}},\mathbf{Y}\right)$, and structure $s\left(\widehat{\mathbf{Y}},\mathbf{Y}\right)$, which are defined as follows:

$$l\left(\widehat{\mathbf{Y}},\mathbf{Y}\right)=\frac{2\mu_{\widehat{\mathbf{Y}}}\mu_{\mathbf{Y}}+C_1}{\mu_{\widehat{\mathbf{Y}}}^{2}+\mu_{\mathbf{Y}}^{2}+C_1}, \tag{1}$$

$$c\left(\widehat{\mathbf{Y}},\mathbf{Y}\right)=\frac{2\sigma_{\widehat{\mathbf{Y}}}\sigma_{\mathbf{Y}}+C_2}{\sigma_{\widehat{\mathbf{Y}}}^{2}+\sigma_{\mathbf{Y}}^{2}+C_2}, \tag{2}$$

$$s\left(\widehat{\mathbf{Y}},\mathbf{Y}\right)=\frac{\sigma_{\widehat{\mathbf{Y}}\mathbf{Y}}+C_3}{\sigma_{\widehat{\mathbf{Y}}}\sigma_{\mathbf{Y}}+C_3}, \tag{3}$$

where $\widehat{\mathbf{Y}}$ and $\mathbf{Y}$ denote the predicted seismic image and label respectively. $\mu_{\mathbf{Y}}$ correspond to the average value of $\mathbf{Y}$, while the standard deviation of $\mathbf{Y}$ is represented as $\sigma_{\mathbf{Y}}$. $\sigma_{\widehat{\mathbf{Y}}\mathbf{Y}}$ signifies the covariance of $\widehat{\mathbf{Y}}$ and $\mathbf{Y}$. Luminance is perceived by humans as the average intensity of a region; therefore, the mean pixel value serves as a direct and appropriate estimator of this perceptual attribute. Contrast is captured by the standard deviation because it reflects the magnitude of intensity variations around the mean. Structure is derived by subtracting the mean and normalizing by the standard deviation, producing a unit-vector signal that preserves only the spatial patterns. The correlation between these normalized signals thus quantifies structural similarity independently of luminance and contrast [38]. Small constants $C_1$, $C_2$, and $C_3$ are introduced to mitigate numerical instability when the denominator approaches zero. The general formula of SSIM is

$$\mathrm{SSIM}\left(\widehat{\mathbf{Y}},\mathbf{Y}\right)=\left[l\left(\widehat{\mathbf{Y}},\mathbf{Y}\right)\right]^{\alpha}\cdot\left[c\left(\widehat{\mathbf{Y}},\mathbf{Y}\right)\right]^{\beta}\cdot\left[s\left(\widehat{\mathbf{Y}},\mathbf{Y}\right)\right]^{\gamma}, \tag{4}$$

luminance, contrast, and structure are relatively independent; specifically, a change in one component does not affect the evaluation of the others. Moreover, their corresponding indices $\alpha$, $\beta$, and $\gamma$ are all positive numbers [38]. They are assigned different weights according to the importance of the three components. For simplicity of the expression, the parameters are usually set as $\alpha=\beta=\gamma=1$, so that SSIM is also derived as

$$\mathrm{SSIM}\left(\widehat{\mathbf{Y}},\mathbf{Y}\right)=\frac{\left(2\mu_{\widehat{\mathbf{Y}}}\mu_{\mathbf{Y}}+c_1\right)\left(2\sigma_{\widehat{\mathbf{Y}}\mathbf{Y}}+C_2\right)}{\left(\mu_{\widehat{\mathbf{Y}}}^{2}+\mu_{\mathbf{Y}}^{2}+c_1\right)\left(\sigma_{\widehat{\mathbf{Y}}}^{2}+\sigma_{\mathbf{Y}}^{2}+C_2\right)}. \tag{5}$$

It is not difficult to conclude that the upper bound of SSIM is 1 when the two images are identical. When we use SSIM to measure the error between the generated seismic image and the ground truth image, it is defined as:

$$\mathscr{L}_{\mathrm{SSIM}}\left(\widehat{\mathbf{Y}},\mathbf{Y}\right)=1-\mathrm{SSIM}\left(\widehat{\mathbf{Y}},\mathbf{Y}\right). \tag{6}$$

#### 2.3.2 Mean Square Error (MSE)

The Mean Square Error (MSE) quantifies the difference between predicted and ground-truth pixel values by squaring the differences and averaging them across all pixels. It is commonly used across various fields [39, 40]. Another widely used loss function is the Mean Absolute Error (MAE), which computes the average of the absolute differences between predictions and ground truth. The key difference between the two lies in their sensitivity to large errors: MSE penalizes larger errors more heavily due to the squaring operation. As a result, MSE tends to emphasize high-frequency details and sharp variations in the image, while MAE may lead to smoother results and is less sensitive to outliers. In our task of enhancing seismic image resolution, preserving high-frequency details is critical. Therefore, we chose MSE over MAE, as it better reduces over-smoothing in the high-frequency regions during training. MSE is expressed as:

$$\mathscr{L}_{\mathrm{MSE}}=\frac{1}{N}\sum_{i=1}^{n}\left(\widehat{y}_i-y_i\right)^2. \tag{7}$$

Where $\widehat{y}_i$ represents the amplitude values at each point in the predicted image, $y_i$ denotes the corresponding amplitude values in the ground truth, and $N$ is the total number of pixels present within the seismic image.

#### 2.3.3 Weighted SSIM–MSE loss

To enhance the perceptual quality of seismic images and emphasize detailed stratigraphic features, we employ a loss function that combines SSIM and MSE, defined as follows:

$$\mathscr{L}_{\mathrm{Mix}}\left(\widehat{\mathbf{Y}},\mathbf{Y}\right)=\alpha\cdot\mathscr{L}_{\mathrm{SSIM}}\left(\widehat{\mathbf{Y}},\mathbf{Y}\right)+(1-\alpha)\cdot\mathscr{L}_{\mathrm{MSE}}\left(\widehat{\mathbf{Y}},\mathbf{Y}\right). \tag{8}$$

Among them, $\alpha$ determines the relative importance of the two components. When $\alpha = 0$ or $\alpha = 1$, they represent two special cases: $\alpha = 0$ corresponds to using only the MSE loss function, and $\alpha = 1$ corresponds to using only the SSIM loss function. To better leverage the strengths of SSIM and MSE, we experimented with different combination weights to determine the final loss function. Visual examples comparing the results obtained with different loss functions are presented in Figure 4. Figure 4a shows a low-resolution synthetic seismic image, while Figure 4b is the corresponding high-resolution label for comparison. Figure 4c–h display the network predictions trained with combination weights of $\alpha = 0$, 0.2, 0.4, 0.6, 0.8, and 1, respectively. It can be observed that when $\alpha = 0$, i.e., the network is trained using only MSE, the predicted discontinuities near faults are clearer, as indicated by the green circles in Figure 4c. In contrast, when $\alpha = 1$, i.e., the network is trained using only SSIM, the predicted discontinuities are relatively blurred, as shown in the green circles in Figure 4h. Moreover, the SSIM-trained network exhibits some loss of boundary information. However, the overall reflection features are better preserved in the SSIM predictions, as indicated by the blue arrows in Figure 4h. These qualitative observations suggest that SSIM and MSE each have their own advantages. For intermediate weights ($\alpha = 0.2$, 0.4, 0.6, 0.8), the visual differences in the predictions are less clear, making qualitative assessment difficult. Therefore, we further conducted a quantitative analysis of the results. In addition to SSIM and MSE, we used signal-to-noise ratio (SNR) and peak signal-to-noise ratio (PSNR) as evaluation criteria. The higher values of SSIM, SNR, and PSNR indicate better performance, while a lower MSE value signifies better results. Table 1 presents the evaluation metrics for different weight parameters, where bold indicates that the network trained with the corresponding loss function performs better than those trained with other loss functions. It shows that the network performs better when the weight coefficient $\alpha$ of the combined loss function is set to 0.2. When $\alpha = 0.2$, all evaluation metrics except for SSIM show the best performance. The network trained with only the SSIM loss function (i.e., when $\alpha = 1$) yields the best SSIM score. This outcome is expected: a network trained using the SSIM loss is directly optimized to minimize SSIM, so it naturally achieves higher SSIM scores compared to networks trained with other loss functions. When combined, these loss functions complement each other: SSIM helps preserve the structural aspects of the image, while MSE ensures pixel-level accuracy, resulting in more robust and visually coherent outputs in image processing tasks.

### 2.4 Training details

By applying geometric transformations, including horizontal and vertical flips, for data augmentation, we ultimately obtained 4,500 pairs of resolution-enhanced training data. Before training, we preprocess the data by performing normalization through the following formula [15]:

$$\mathbf{X}^{*}=\frac{\mathbf{X}-\mu}{\sigma}. \tag{9}$$

Here, $\mathbf{X}^{*}$ represents the normalized seismic volume of $\mathbf{X}$. $\mu$ correspond to the average value of the input seismic volume,

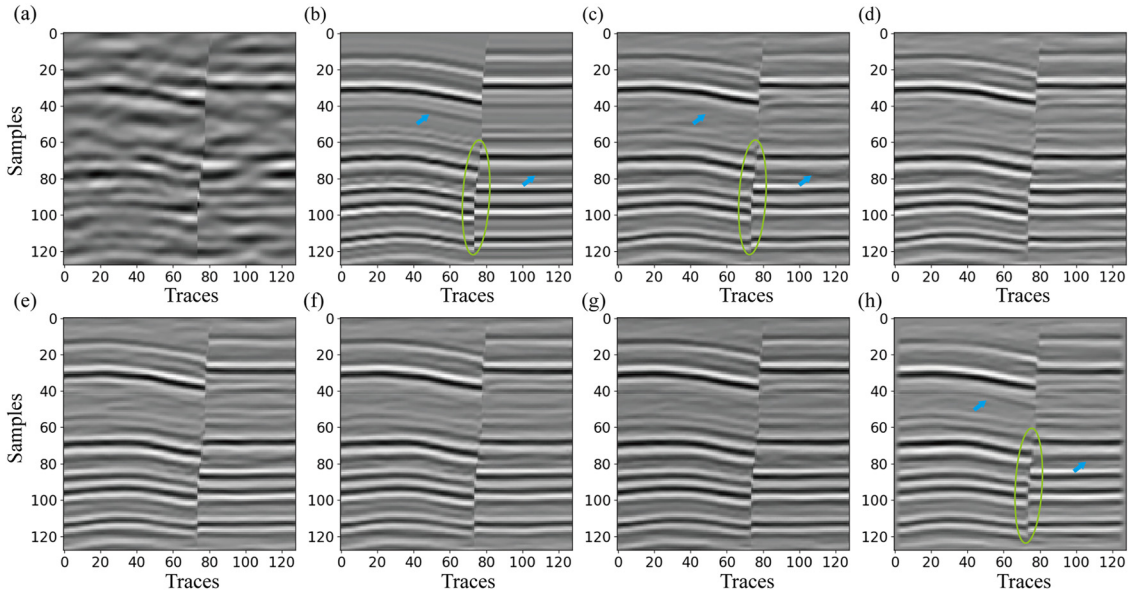


**Figure 4:** Comparison of low-resolution and high-resolution seismic images, and network predictions trained with different weighted combination losses. (a) Low-resolution synthetic seismic image. (b) Corresponding high-resolution label. (c–h) Network predictions trained with the weighted combination loss using $\alpha = 0$, 0.2, 0.4, 0.6, 0.8, and 1, respectively. Here, $\alpha = 0$ indicates training with MSE only, and $\alpha = 1$ indicates training with SSIM only.

**Table 1:** The values of evaluation criteria on the test data for the hybrid loss function with different weights. Each row represents different training loss functions, and each column represents different values of evaluation metrics. The best evaluation results are highlighted in bold, where smaller values for $\mathscr{L}_{\text{MSE}}$ are preferable, and larger values for SSIM, SNR, and PSNR are better.

| Quality metric | Input | Training loss function | | | | | |
|---|---|---|---|---|---|---|---|
| | | $\alpha = 0$ | $\alpha = 0.2$ | $\alpha = 0.4$ | $\alpha = 0.6$ | $\alpha = 0.8$ | $\alpha = 1$ |
| 100* $\mathscr{L}_{\text{MSE}}$ ↓ | 61.653 | 8.531 | **7.131** | 9.078 | 9.122 | 14.725 | 9.571 |
| SSIM↑ | 0.370 | 0.801 | 0.830 | 0.800 | 0.807 | 0.774 | **0.855** |
| SNR↑ | 3.458 | 12.048 | **12.826** | 11.778 | 11.757 | 9.677 | 11.548 |
| PSNR↑ | 2.100 | 10.690 | **11.468** | 10.420 | 10.399 | 8.320 | 10.190 |

*100 × LMSE.

while the standard deviation of the input seismic volume is represented as $\sigma$. During the training phase, we first initialize the network's weights and biases randomly. The network was implemented in Python using the PyTorch framework. We then use the Adam algorithm to optimize the network and update its parameters [41]. The training was conducted for 100 epochs. We evaluated the performance of learning rates $10^{-5}$, $5 \times 10^{-5}$, and $10^{-4}$, and the learning rate is subsequently reduced according to the decay of the validation loss. As evidenced by the loss curve trends and the corresponding test predictions, the best performance was achieved with a learning rate of $10^{-4}$. The other two learning rates were relatively small, which resulted in slower convergence, and by the end of 100 epochs, they had not yet reached a desirable convergence state. The batch size of the network is set to 10.

# 3 Results

To demonstrate the performance of the proposed method, one synthetic dataset and two real offshore datasets from different work areas were fed into the trained network. The detailed network architecture is described in Section 2.2. The training dataset was prepared following the procedure outlined in Section 2.1, and the training was terminated after 100 epochs. For validation, we selected the network parameters obtained at the 80th epoch. The same parameters were consistently applied across both the synthetic dataset and the two field datasets. Since the network input size is fixed at $128 \times 128$, whereas the actual seismic sections are of varying dimensions, a sliding-window strategy was adopted to divide the data into smaller patches, which were then fed into the network. The enhanced patches were subsequently

merged to reconstruct the full section. Before entering the network, each patch was normalized using the same preprocessing procedure described in Section 2.4. In the following section, the application results on both the synthetic and real datasets are analyzed and discussed.

## 3.1 Application of synthetic data

We first validate the effectiveness of our proposed method on synthetic data. Figure 5 shows a pair of synthetic seismic data derived from a common reflectivity model. This reflectivity model incorporates complex geological features, including varying stratigraphy, multiple folds, and four faults with distinct displacement parameters. Figure 5a is the noisy, low-resolution seismic image, while Figure 5b is the corresponding high-resolution label, both with dimensions of 512 [trace] × 256 [sample]. It is worth noting that this synthetic dataset is used exclusively for testing purposes and was not involved in network training. Figure 5c shows the results generated by the proposed method. Compared to the original noisy, low-resolution seismic image, the processed result demonstrates a significant improvement in vertical resolution, particularly evident in areas with prominent edge features. Moreover, the processed result indicates that the trained network possesses inherent denoising capabilities. Stratigraphic features that were previously obscured by noise are now clearly visible.

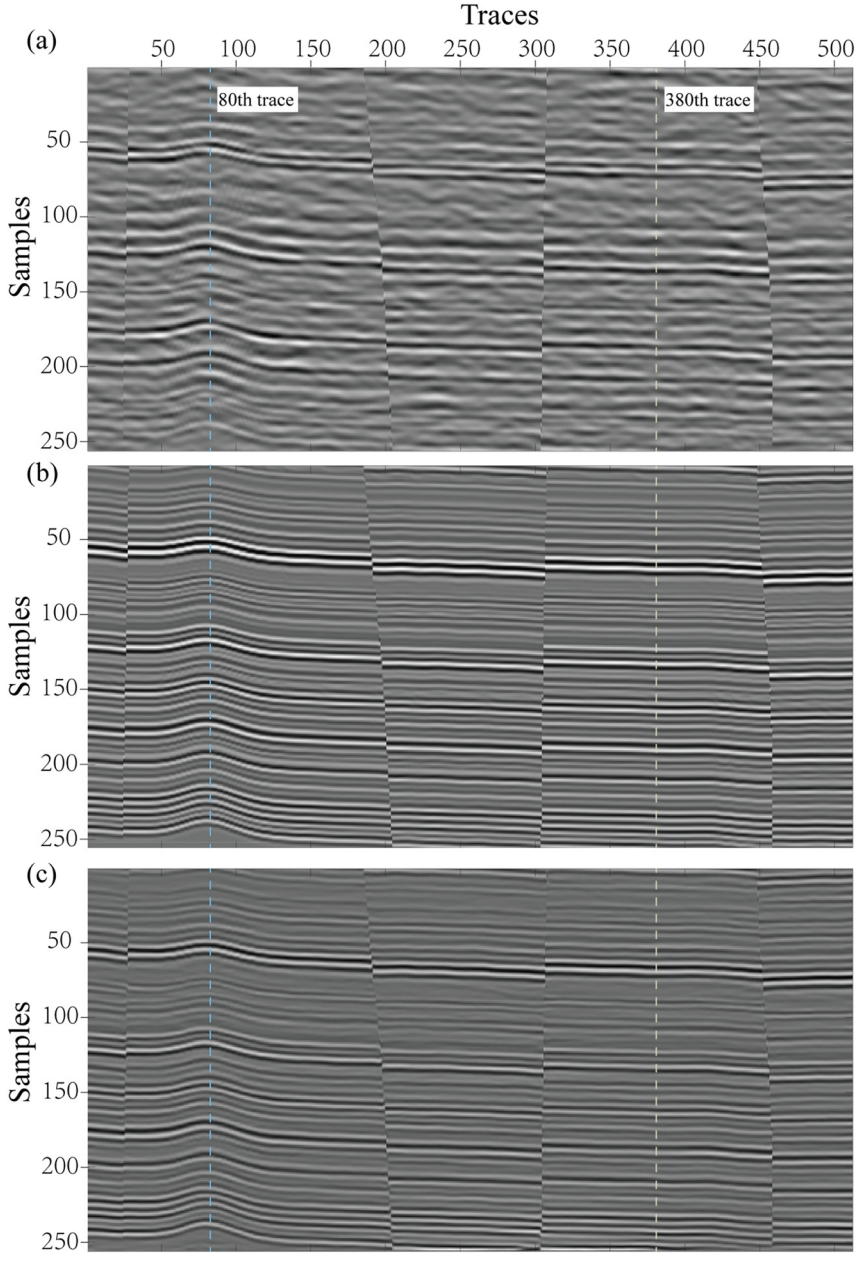


**Figure 5:** Synthetic test data, including (a) seismic image characterized by low resolution and noise; (b) associated high-resolution label; (c) the high-resolution processing result.

To further evaluate the amplitude characteristics before and after processing, we compare individual seismic traces from the high-resolution output with those from the original and label. Specifically, the 80th and 380th traces, marked by the blue and yellow dashed lines in Figure 5, are randomly selected for comparison. The corresponding results are presented in Figure 6a and b. In these figures, the red line represents the seismic trace extracted from the original low-resolution seismic image (see Figure 5a), the green line indicates the seismic trace extracted from ground truth (see Figure 5b), and the cyan line represents the single seismic trace from the predicted high-resolution result (see Figure 5c). Compared to the original low-resolution seismic trace, the processed trace retains a similar waveform trend. However, it is evident that the processed trace has restored much of the high-frequency information, with waveform characteristics closely matching the ground truth.

In addition, we analyze the spectral characteristics of the seismic images, as shown in Figure 7. The amplitude values across different frequency bands are averaged over all seismic traces in the image. The amplitude spectra of the original data, ground truth, and processed output are represented by the red, green, and cyan curves, respectively. The results demonstrate a broadened frequency bandwidth along with an enhanced dominant frequency in the processed image, which are crucial for improving vertical resolution and facilitating the identification of finer geological details.

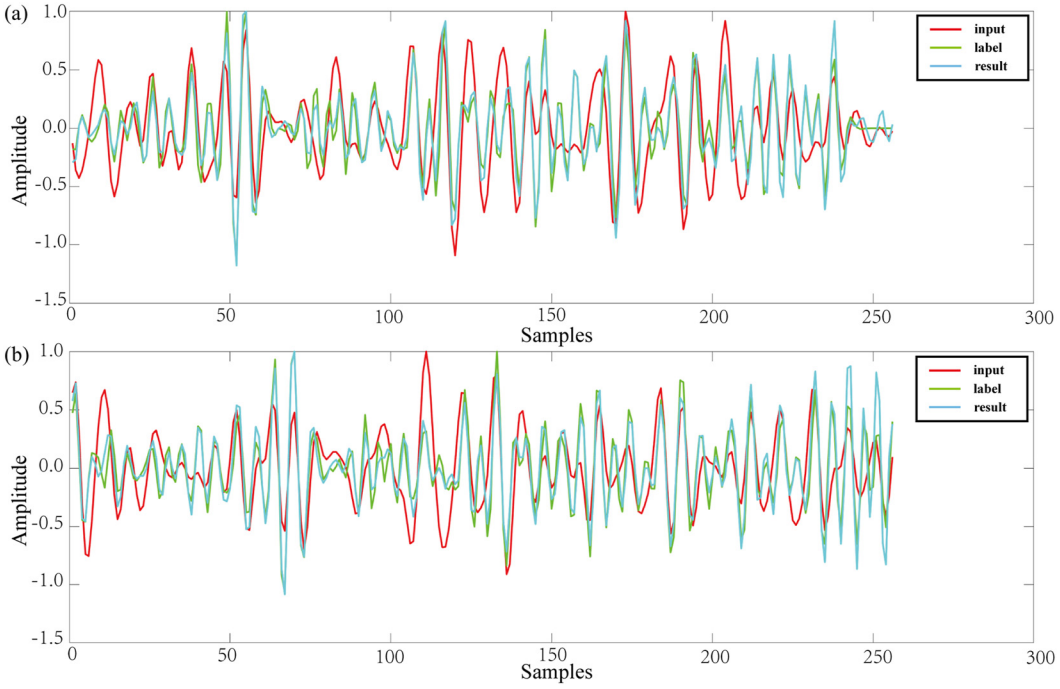


**Figure 6:** Seismic trace analysis of the (a) 80th trace (indicated by the blue dashed line in Figure 5) and (b) 380th trace (indicated by the yellow dashed line in Figure 5). Shown are the traces from the input seismic image prior to processing (red), the true label (green), and the processed high-resolution result (cyan).

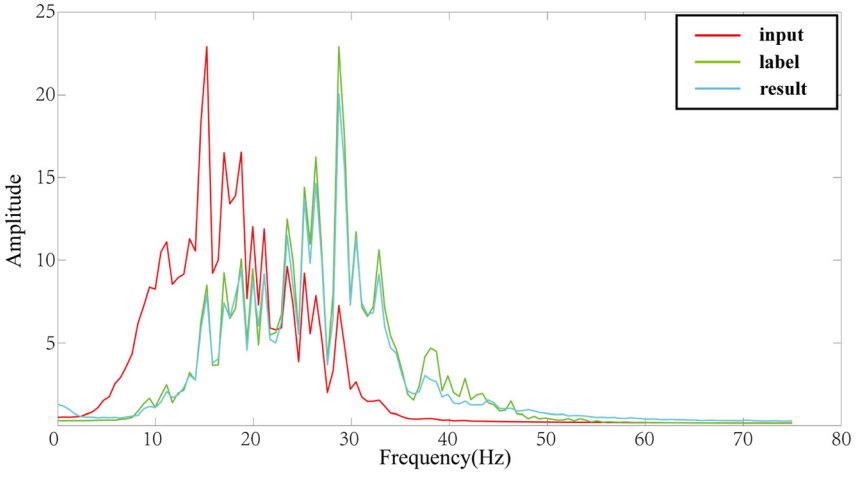


**Figure 7:** Spectrum analysis of the original seismic image from Figure 5a (red), the ground truth from Figure 5b (green), and the output high-resolution results from Figure 5c (cyan). The results show broader frequency bandwidth and enhanced dominant frequency, leading to improved vertical resolution and clearer identification of fine geological details.

## 3.2 Application of actual data

Having demonstrated the resolution enhancement capabilities of the proposed method on synthetic datasets in the previous section, we then evaluate the trained network's performance on two field datasets obtained from different work areas.

The first actual seismic dataset is from the 2D profiles of the Poseidon-3D Marine Seismic Survey. Figure 8a–c display three original profiles from the Poseidon dataset. Figure 8d–f show the corresponding results obtained by applying a spectral shaping method [42] to enhance resolution, while Figure 8g–i present the high-resolution profiles produced by our trained network. In both cases, an enhancement of

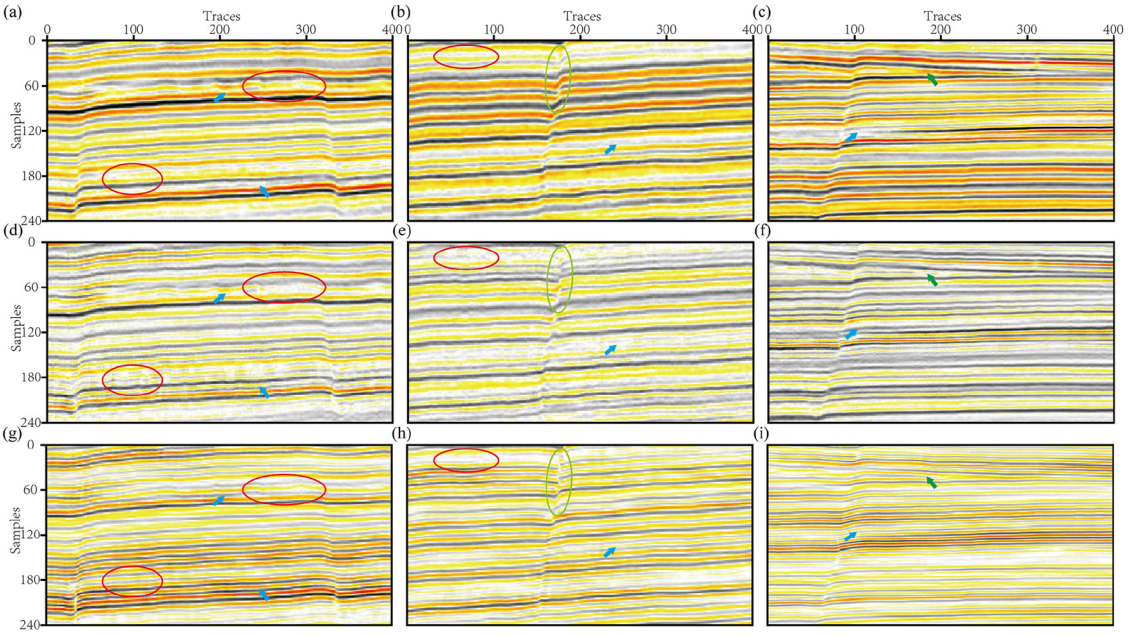


**Figure 8:** 2D seismic profiles from the Poseidon-3D dataset. (a)–(c) Original profiles, (d)–(f) results processed using the spectral shaping method, (g)–(i) enhanced profiles obtained with our proposed approach.

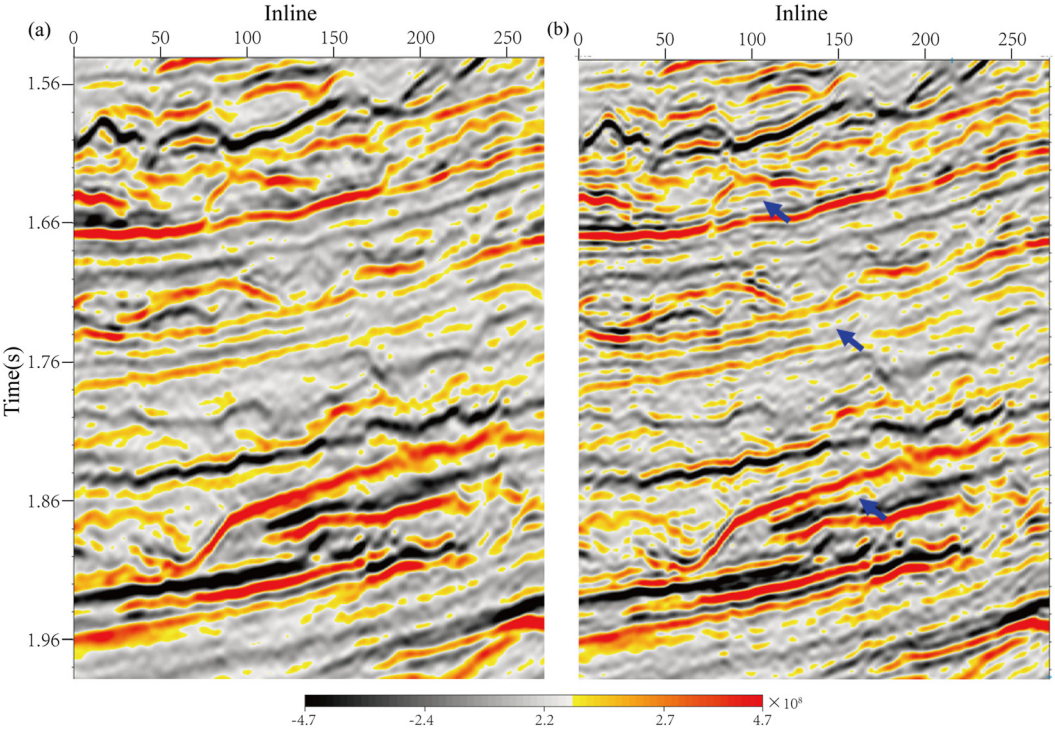


**Figure 9:** A field seismic section and its enhanced result. (a) Field seismic section from Data2; (b) corresponding enhanced result.

seismic data resolution is clearly observed. In particular, in the results processed by our method, seismic responses from thin layers become more distinct, as indicated by the blue arrows, which greatly facilitates the detection and characterization of such layers. Compared with the spectral shaping method, the improvement achieved by our approach appears more evident. Notably, fault discontinuities also become more evident after processing with the proposed method. For example, in Figure 8b, faults are manifested only as subtle warping features with limited signs of horizon dislocation, which complicates reliable interpretation. By contrast, the processed result in Figure 8h reveals distinct fault displacement across the section, especially between 0 and 70 samples (green circle), where the fault breakpoints appear more prominent and easier to identify. These improvements appear less pronounced when using the spectral shaping method. This demonstrates that the proposed approach significantly enhances fault

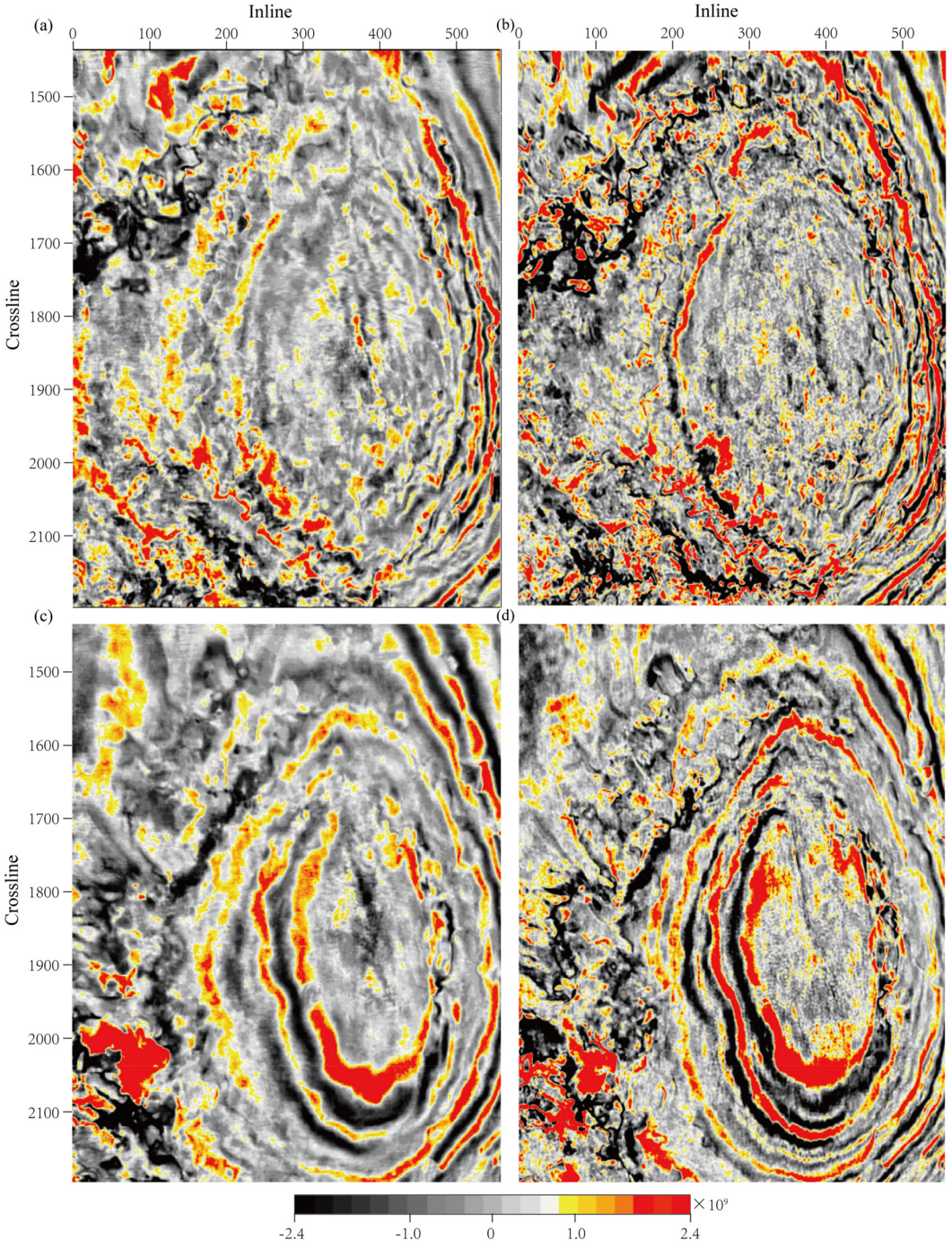


**Figure 10:** Original seismic time slices and corresponding resolution-enhanced results. (a) Original seismic time slice at 1,600 ms. (b) Resolution-enhanced result corresponding to (a). (c) Original seismic time slice at 1,800 ms. (d) Resolution-enhanced result corresponding to (c).

resolution in shallow intervals while simultaneously improving the visibility of thin layers across the section, thereby aiding the recognition of subtle structural features. Furthermore, the resolution of pinch-out formations, highlighted by the green arrows in Figure 8f and i, is markedly improved, enabling more accurate localization and tracking of stratigraphic variations. In addition, our method exhibits stronger robustness compared to the spectral shaping method. For instance, in the regions marked by red circles, the enhanced seismic images generated by our approach preserve textural features and stratigraphic continuity more effectively.

The second actual seismic data is from the South China Sea region, which we refer to as Data2. Figure 9a shows a two-dimensional profile of this data. The corresponding high-resolution seismic image through the trained network is shown in Figure 9b. The enhanced seismic images show a significant improvement in resolution, revealing many stratigraphic features that were previously hard to detect, as highlighted by the blue arrows.

We also demonstrate the resolution enhancement effect on the time slice. Figure 10a and c display the original time slices at 1,600 ms and 1800 ms, respectively, while Figure 10b and d present their corresponding resolution-enhanced results. It is evident that the enhanced time slices exhibit much clearer seismic detail responses, with a significant improvement in lateral resolution.

The enhanced vertical resolution allows for finer delineation of thin beds and subtle stratigraphic changes that were previously indistinguishable in the original data. These improvements help to reveal previously hidden geological features, such as minor faults, unconformities, and layer boundaries, which can now be more clearly identified. The enhanced resolution facilitates better temporal or depth accuracy in the depiction of seismic events, which is crucial for precise subsurface interpretation, especially in complex geological areas. Overall, the resolution enhancement significantly improves the clarity of seismic imaging, providing more detailed and reliable information for geological analysis.

# 4 Conclusions

This paper proposes a deep learning-based method for seismic resolution enhancement. To address the limitations of pixel-level losses, we design a hybrid loss function that combines the Structural Similarity Index Measure with the Mean Squared Error, allowing the network to preserve high-frequency details while improving overall seismic image quality. Experiments on both synthetic and field data demonstrate that our method effectively enhances resolution, recovers signals from thin layers, and provides additional denoising capabilities, resulting in clearer and more accurate representations of subsurface geological features. Notably, although the proposed network is trained on synthetic data, it achieves promising results on real seismic datasets. Nevertheless, its applicability to data from other regions requires further validation. Incorporating transfer learning with a limited amount of real seismic data may further enhance the network's generalization capability in practical applications. In future work, we will incorporate more realistic wavelets to enhance applicability to real seismic scenarios.

**Acknowledgments:** The authors are sincerely grateful to the editor and the two anonymous reviewers for their valuable comments and constructive suggestions, which have significantly improved the quality of this manuscript. The authors also gratefully acknowledge the support of the CNOOC South China Sea Oil & Gas Energy Academician Workstation. This research was supported by the specific research fund of The Innovation Platform for Academicians of Hainan Province (grant No. YSPTZX202302).

**Author contributions:** Conceptualization, X.H. and S.L.; methodology, S.L. and J.P.; software, J.P. and F.L.; validation, L.H., Y.Z. and Z.F.; formal analysis, M.Z. and H.G.; investigation, G.L.; writing – original draft preparation, X.H.; all authors have read and agreed to the published version of the manuscript.

**Conflicts of interest:** The authors declare no conflicts of interest.

**Funding statement:** This work is supported by the specific research fund of The Innovation Platform for Academicians of Hainan Province (grant No. YSPTZX202302).

# References

1. Naghizadeh M, Vermeulen P, Crook A, Birce A, Ross S, Stanton A, et al. EcoSeis: a novel acquisition method for optimizing seismic resolution while minimizing environmental footprint. The Leading Edge 2023; 42:61–8.
2. Wang C, Zong Z, Yin X, Li K. Seismic resolution enhancement with variational modal-based fast-matching pursuit decomposition. Interpretation 2024;12:T77–86.
3. Niloofar A, Amin RK, Abolghasem KR, Mehrdad SM. Seismic resolution enhancement using scale transform in the time-frequency domain. Geophysics 2018;83:V305–14.
4. Wang Y. Inverse Q-filter for seismic resolution enhancement. Geophysics 2006;71:V51–60.
5. Li C, Liu G, Wang Z, Wu Q, Shi L. Multichannel joint resolution enhancement for nonstationary prestack data with adaptive structure regularization. Geophysics 2023;88:V361–70.

6. Sajid M, Ghazali AR. Nonstationary differential resolution: an algorithm to improve seismic resolution. Geophysics 2018;83:V149–56.
7. Perez G, Marfurt KJ. Improving lateral and vertical resolution of seismic images by correcting for wavelet stretch in common-angle migration. Geophysics 2007;72:C95–104.
8. Yoon D, Yeeh Z, Byun J. Seismic data reconstruction using deep bidirectional long short-term memory with skip connections. IEEE Geosci Remote Sens Lett 2021;18:1298–302.
9. Yu S, Ma J, Wang W. Deep learning for denoising. Geophysics 2019;84: V333–50.
10. Yang FS, Ma JW. Deep-learning inversion: a next-generation seismic velocity model building method. Geophysics 2019;84:R583–99.
11. Jing JK, Yan Z, Zhang Z, Gu HM, Han BK. Fault detection using a convolutional neural network trained with point-spread function-convolution-based samples. Geophysics 2023;88:IM1–14.
12. Wu XM, Liang LM, Shi YZ, Fomel S. FaultSeg3D: using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation. Geophysics 2019;84:IM35–45.
13. Zhang Z, Yan Z, Jing J, Gu H, Li H. Generating paired seismic training data with cycle-consistent adversarial networks. Remote Sens 2023;15:265.
14. Pham N, Fomel S, Dunlap D. Automatic channel detection using deep learning. Interpretation 2019;7:SE43–50.
15. Gao H, Wu XM, Liu GF. ChannelSeg3D: channel simulation and deep learning for channel interpretation in 3D seismic images. Geophysics 2021;86:IM73–83.
16. Mustafa A, AlRegib G. Active learning with deep autoencoders for seismic facies interpretation. Geophysics 2023;88:IM77–86.
17. Liu XY, Li B, Li JY, Chen XH, Li QC, Chen YK. Semi-supervised deep autoencoder for seismic facies classification. Geophys Prospect 2021; 69:1295–315.
18. Li J, Wu X, Hu Z. Deep learning for simultaneous seismic image super-resolution and denoising. IEEE Trans Geosci Remote Sens 2021;60:1–11.
19. Yu W, Dong S, Lu S, Dong X. MAE-GAN: a novel strategy for simultaneous super-resolution reconstruction and denoising of post-stack seismic profile. arXiv preprint arXiv:240519767 2024.
20. Lin L, Zhong Z, Cai C, Li C, Zhang H. SeisGAN: improving seismic image resolution and reducing random noise using a generative adversarial network. Math Geosci 2024;56:723–49.
21. Xiao Y, Li K, Dou Y, Li W, Yang Z, Zhu X. Diffusion models for multidimensional seismic noise attenuation and super-resolution. Geophysics 2024;89:V479–92.
22. Zhang H-R, Liu Y, Sun Y-H, Chen G. SeisResoDiff: seismic resolution enhancement based on a diffusion model. Pet Sci 2024;21:3166–88.
23. Choi Y, Jo Y, Seol SJ, Byun J, Kim Y. Deep learning spectral enhancement considering features of seismic field data. Geophysics 2021;86: V389–408.
24. Gao Y, Zhao D, Li T, Li G, Guo S. Deep learning vertical resolution enhancement considering features of seismic data. IEEE Trans Geosci Remote Sens 2023;61:1–13.
25. Zhang H, Alkhalifah T, Liu Y, Birnie C, Di X. Improving the generalization of deep neural networks in seismic resolution enhancement. IEEE Geosci Remote Sens Lett 2022;20:1–5.
26. Jo Y, Choi Y, Seol SJ, Byun J. Machine learning-based vertical resolution enhancement considering the seismic attenuation. J Petrol Sci Eng 2022;208:109657.
27. Zhou R, Zhou C, Wang Y, Yao X, Hu G, Yu F. Deep learning with fault prior for 3-D seismic data super-resolution. IEEE Trans Geosci Remote Sens 2023;61:1–16.
28. Hamida A, Alfarraj M, Al-Shuhail AA, Zummo SA. Facies-guided seismic image super-resolution. IEEE Trans Geosci Remote Sens 2023;61:1–13.
29. Gao Y, Zhang J, Li H, Li G. Incorporating structural constraint into the machine learning high-resolution seismic reconstruction. IEEE Trans Geosci Remote Sens 2022;60:1–12.
30. Min F, Wang L, Pan S, Song G. D 2 UNet: dual decoder U-Net for seismic image super-resolution reconstruction. IEEE Trans Geosci Remote Sens 2023;61:1–13.
31. Kaur H, Pham N, Fomel S. Improving the resolution of migrated images by approximating the inverse Hessian using deep learning. Geophysics 2020;85:WA173–83.
32. Dong X, Lu S, Lin J, Zhang S, Ren K, Cheng M. Can deep learning compensate for sparse shots in the imaging domain? A potential alternative for reducing the acquisition cost of seismic data. Geophysics 2024;89:V119–37.
33. Wu X, Geng Z, Shi Y, Pham N, Fomel S, Caumon G. Building realistic structure models to train convolutional neural networks for seismic structural interpretation. Geophysics 2020;85: WA27–39.
34. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In: Proceedings of 18th international conference on medical image computing and computer-assisted intervention –MICCAI 2015, October 5–9, 2015. Cham: Springer; 2015, vol III:234–41 pp.
35. Wang Q, Ma Y, Zhao K, Tian Y. A comprehensive survey of loss functions in machine learning. Annals of Data Science 2022;9:187–212.
36. Lecun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521: 436–44.
37. Zhao H, Gallo O, Frosio I, Kautz J. Loss functions for image restoration with neural networks. IEEE Transactions on Computational Imaging 2017;3:47–57.
38. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 2004;13:600–12.
39. Terven J, Cordova-Esparza DM, Romero-González J-A, Ramírez-Pedraza A, Chávez-Urbiola EA. A comprehensive survey of loss functions and metrics in deep learning. Artif Intell Rev 2025;58:195.
40. Elharrouss O, Mahmood Y, Bechqito Y, Serhani M, Badidi E, Riffi J, et al. Loss functions in deep learning: a comprehensive review. arXivarXiv: 2504.04242v1 2025.
41. Kingma DP, Ba J. Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980 2014.
42. Xue J, Cai C, Gu H, Luo H. Seismic resolution enhancement by spectral shaping using shaping-regularized inversion. IEEE Geosci Remote Sens Lett 2021;19:1–5.