GEOPHYSICS, VOL. 90, NO. 3 (MAY-JUNE 2025); P. 1–17, 25 FIGS., 1 TABLE.
10.1190/GEO2024-0502.1

From shallow to deep: Enhancing seismic resolution with weak supervision

Dawei Liu[1], Yijie He[2], Xiaokai Wang[2], Mauricio D. Sacchi[3], Wenchao Chen[2], Guanghong Du[4], and Mengbo Zhang[4]

ABSTRACT

High-resolution (HR) seismic data are crucial for accurately identifying subsurface geologic formations and reservoir properties. However, as widely observed in prestack and poststack data, seismic wave attenuation often leads to a gradual decrease in resolution with depth. This degradation leads to resolution differences between data from shallow and deep layers, characterized by high-frequency energy loss in deeper target layers. The latter could complicate the extraction of crucial deep-layer structure information, leading to increased uncertainty in reservoir modeling and significant economic losses due to our inability to image deep targets with sufficient resolution. Traditional methods with simplified assumptions about this nonlinear attenuation hinder modeling geologic complexity and variability. Deep-learning methods excel at capturing complex, nonlinear relationships but often rely heavily on scarce and costly paired labels for supervised learning. As a result, deep learning without requiring paired labels for seismic resolution enhancement has become a key research focus. This study explores a method to enhance 3D seismic data resolution using weakly supervised learning, leveraging the inherent similarity of sedimentary structures across depths. Specifically, our approach uses a relatively shallow, HR data window to transfer and extrapolate learned high-frequency information to a deeper window containing attenuated responses from deeper layers. Thereby, we achieve enhanced resolution in the target region. A tailored 3D convolutional neural network with a bidirectional cycle structure, custom-designed loss functions, and data preprocessing techniques specifically addresses the challenge of resolution differences caused by seismic wave attenuation. The effectiveness of our method is validated with synthetic and real 3D poststack migration data, demonstrating its robustness for regions near the target layer. Compared with conventional spectral whitening, our approach leverages intrinsic data characteristics more adaptively and robustly, making geologic structures more discernible.

INTRODUCTION

Exploration seismic data plays a pivotal role in uncovering crucial geologic features, such as stratigraphic structures, underlying strata, and fault information (Bagheri et al., 2017; Bagheri and Riahi, 2018; KhodAgholi and BAgheri, 2020). However, the seismic data acquired often undergo significant resolution degradation, particularly in the high-frequency range. This degradation primarily stems from absorption phenomena within geologic layers, leading to the dissipation of seismic wave energy, thus resulting in widened waveforms and diminished resolution. Given the indispensable nature of high-resolution (HR) seismic data sets for precise geologic interpretations, extensive research has been devoted to enhancing seismic data resolution.

Conventional approaches for enhancing seismic data resolution encompass deconvolution (Wiggins, 1978; Taylor et al., 1979; Levy and Fullagar, 1981; Sacchi, 1997; Velis, 2008; Gholami and Sacchi, 2013; Sui and Ma, 2020; Zhang et al., 2022), spectral blueing

Manuscript received by the Editor 17 July 2024; revised manuscript received 8 January 2025; published ahead of production 28 January 2025.
[1]Xi'an Jiaotong University, School of Information and Communications Engineering, Xi'an, China and University of Alberta, Department of Physics, Alberta, Canada. E-mail: 409791715@qq.com.
[2]Xi'an Jiaotong University, School of Information and Communications Engineering, Xi'an, China. E-mail: 13186150985@163.com; xkwang@xjtu.edu.cn; wencchen@xjtu.edu.cn (corresponding author).
[3]University of Alberta, Department of Physics, Alberta, Canada. E-mail: msacchi@ualberta.ca.
[4]National Engineering Laboratory of Low-Permeability Oil & Gas Exploration and Development, Xi'an, China and PetroChina Changqing Oilfield Company, Exploration and Development Research Institute, Xi'an, China. E-mail: dghong_cq@petrochina.com.cn; zmb_cq@petrochina.com.cn.

(Lancaster and Whitcombe, 2000; Kazemeini et al., 2010), and inverse Q filtering (Wang, 2006; Xue et al., 2019; Ke et al., 2023). These methods, which rely on predefined assumptions, have boosted seismic data resolution and are widely used within the industry. However, their effectiveness is constrained by the limitations of handcrafted priors, hindering further advancements in resolution enhancement. The absence of an active pursuit of data-driven prior knowledge derived from HR seismic data poses a significant challenge to overcome these limitations.

The advent of deep learning has revitalized data-driven approaches for enhancing seismic data resolution (Kaur et al., 2020; Yu and Ma, 2021; Liu et al., 2022, 2024; Wu et al., 2024; Zhang et al., 2024). Using an end-to-end training strategy, neural networks can autonomously extract prior knowledge from HR data, improving resolution more effectively and flexibly than conventional methods. Moreover, deep networks excel at capturing complex, nonlinear features, aligning well with the nonlinear attenuation process of seismic data. Notably, in the field of image superresolution, numerous classical networks (Dong et al., 2015; Kim et al., 2016a, 2016b; Mao et al., 2016; Lai et al., 2017; Ledig et al., 2017; Tong et al., 2017; Liang et al., 2021) have demonstrated remarkable success, underscoring the potential of deep learning for enhancing seismic data resolution. This rapid progress has spurred growing academic interest in deep-learning-based seismic data resolution enhancement over the past years. Choi et al. (2019) use a convolutional U-Net architecture to enhance seismic data vertical resolution, proving its efficacy in thin bed resolution. Chen et al. (2019) propose an iterative deep neural network for HR seismic inversion, enabling simultaneous wavelet and reflectivity estimation. Generative adversarial networks (GANs) are renowned for their ability to simulate data distributions. This capability is crucial for generating plausible HR details in image superresolution tasks. Inspired by this, Zhang et al. (2019) implement a GAN to enhance the resolution of seismic data, producing more subtle and continuous reflections than traditional methods. To address noisy seismic data, Li et al. (2021) develop a deep network that simultaneously enhances resolution and performs seismic denoising. Chen et al. (2021) design an optimization-inspired deep-learning inversion solver that accelerates the solution of blind HR inverse problems, enhancing network interpretability and demystifying the black-box nature of deep-learning models.

Research on deep-learning-based seismic resolution enhancement has focused on supervised deep learning, which relies on paired low-resolution (LR) and HR data. Specifically, these supervised approaches require HR labels that precisely correspond to LR data. However, in real-world scenarios, data pairs for training are not available. One could obtain HR labels via forward modeling. Still, this approach requires an accurate understanding of the geologic information in the study area.

Recent advances in weakly supervised learning (Jiang et al., 2021; Wei et al., 2021) offer promising solutions for signal enhancement tasks using unpaired LR-HR data. As a prominent deep-learning algorithm, GANs have demonstrated remarkable effectiveness in addressing weakly supervised learning problems (Ignatov et al., 2018; Li et al., 2018, 2020; Wang et al., 2020a, 2020b, 2021). Unlike supervised GANs, weakly supervised GANs overcome the need for paired labels, effectively tackling the scarcity of strictly matched samples in a real-world situation. Weakly supervised GANs typically incorporate an additional reverse generator, creating a Cycle-in-Cycle network architecture. This architecture has been successfully applied to image-to-image translation algorithms using weakly supervised learning, with notable examples such as CycleGANs (Zhu et al., 2017), DiscoGANs (Kim et al., 2017), and DualGANs (Yi et al., 2017). In particular, CinCGAN (Yuan et al., 2018) has demonstrated the feasibility of single image superresolution using the Cycle-in-Cycle architecture. The success of weakly supervised learning in image processing suggests the potential for enhancing the resolution of seismic data without the need for LR-HR pairs. A recent study by Liu et al. (2023) demonstrate the feasibility of weakly supervised learning for seismic data resolution enhancement. Their improved CinCGAN architecture effectively recovered high-frequency components from LR seismic data. However, this work requires the availability of unpaired HR data sets in proximity to the LR data set. This requirement often proves challenging, as it is more common to encounter existing data sets without adjacent HR equivalents. Moreover, due to the inherent 3D structure of poststack seismic data, using a 2D convolutional network architecture (Liu et al., 2023) is insufficient to comprehensively capture the complexity of the 3D subsurface layers. Using a 2D slice-wise approach to processing 3D data often results in discontinuities along the direction perpendicular to the slices. Alternatively, using 3D convolutional kernels allows more effective exploration of spatial structure relationships within seismic data than conventional 2D methods. However, the increased number of parameters in 3D networks, relative to 2D networks, poses challenges when solely imposing cycle consistency constraints on LR during network training. As a result, modifications are necessary to integrate additional cycle consistency constraints for HR, facilitating accurate convergence during network training.

Due to seismic absorption or attenuation, the resolution of seismic data decreases with increasing depth, as observed in pre- and poststack data. This causes resolution degradation between shallow- and deep-layer data, as shown in Figure 1. Shallower data typically exhibit higher resolution, providing valuable subsurface information. However, deep-layer data, often including target layers, suffer significant high-frequency loss, leading to reduced resolution. Acquiring HR seismic data incurs substantial costs and requires complex processing steps integrating extensive prior knowledge. Underusing these data sets diminishes their value and leads to significant economic losses.

Deep and shallow seismic data often reveal similar geologic features. The critical question then arises: how can we improve the resolution of deep-layer target data to match that of shallow-layer data by leveraging prior knowledge from the latter? To address this, consider a particular case. Seismic data recorded on sedimentary formations in the same survey area generally exhibit minimal structural changes from shallow to deep strata. Typically, these strata have undergone similar geologic processes, unlike surveys that contain structures, such as rocky mounds or volcanic intrusions in deeper layers. This similarity suggests that shallow and deep seismic responses have comparable distributions, with attenuation effects often exhibiting analogous patterns. In fact, the traditional approach determines attenuation factor Q using the spectral ratio method, which relies on shallow events to help correct for attenuation at deeper locations (Cheng and Margrave, 2012; Raji and Rietbrock, 2013). This approach underscores the potential of using shallow-layer seismic characteristics as proxies for the deeper, attenuated layers. Given this inherent similarity, it becomes possible to enhance the vertical resolution of deep-layer seismic responses

by learning from the distribution of shallow data. The presence of remarkably similar geologic structural features (e.g., Figure 1a) makes training a network to enhance resolution a natural option, as this similarity promotes strong generalization ability.

Accordingly, we introduce a 3D bidirectional CycleGAN architecture to extract HR features from shallow seismic data and transfer them to enhance deep, LR target data through weakly supervised learning. We develop a robust training workflow that includes carefully designed loss functions and comprehensive data preprocessing techniques by integrating the latest advancements in weakly supervised learning with the unique characteristics of 3D seismic data. These preprocessing steps include data patching and augmentation techniques, all contributing to stable training. Once trained with unpaired shallow HR data above the target layer, the network maps deep LR target seismic features to their shallow HR counterparts. Experiments with synthetic and field data validate our approach, demonstrating improved clarity and detail. In addition, they confirm that the unpaired training strategy alleviates the need for paired labeled data. Unlike traditional methods, this end-to-end, data-driven approach adaptively enhances resolution by learning from inherent data relationships rather than relying on fixed assumptions. This flexible exploitation of prior knowledge gives the method broader applicability across varying geologic settings, unlocking greater potential for HR imaging under diverse subsurface conditions. Furthermore, our experimental results confirm that networks imposing bidirectional cycle consistency constraints on LR and HR data, as opposed to the previously mentioned CinCGAN, significantly enhance training stability and reduce misleading artifacts.

METHODS

This study uses a 3D GAN combined with weakly supervised learning methods to enhance the resolution of deep seismic data $\boldsymbol{\mathcal{X}}$ using shallow data $\boldsymbol{\mathcal{Y}}$ from the same work area. The dimensions of 3D tensors $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$ are ordered by time t, crossline x, and inline i. The overall workflow is shown in Figure 2. The main idea is that deep data resolution is lower than shallow data in the same data set, and the enhanced deep data should exhibit features similar to shallow HR data. To achieve this, we adopt CycleGAN as our backbone and adapt it in three ways: we first introduce a specially designed 3D bidirectional CycleGAN architecture. Then, we formulate a novel loss function for weakly supervised learning, which is used to guide the training process and preserve useful seismic signal structures. Last, data preprocessing techniques, such as data augmentation, address the problem of insufficient high-quality 3D training data. We achieve a more stable training process and better results by improving the network structure and loss function. Our weakly supervised learning approach is distinct from traditional classifications by Zhou (2018). In our approach, all training data are inaccurately labeled, distinguishing it from strong supervision, which relies on accurate and precise labels.

The 3D bidirectional CycleGAN architecture

Bidirectional cycle consistency

The initial CycleGAN had constraints solely with unidirectional forward cycle consistency (LR→HR→LR) on LR input data (Liu et al., 2023). This often traps the network model in local minima and causes instability when handling complex mappings between LR and HR domains, resulting in HR outputs with misleading features. In our tasks, this presents challenges in training neural networks on data sets with intricate correspondences between shallow and deep data. To remedy this, we introduce bidirectional cycle consistency, involving forward (LR→HR→LR) and reverse (HR→LR→HR) cycles. This adjustment enhances training stability and allows the network to better understand nuanced connections across different domains.

As shown in Figure 2, the dual-cycle process uses two generators (the forward generator G_{LR2HR} and the reverse generator G_{HR2LR}) and two discriminators (the LR discriminator D_{LR} and the HR discriminator D_{HR}). The generator G_{LR2HR} maps LR data to HR data, whereas G_{HR2LR} maps HR data to LR data. The discriminators D_{LR} and D_{HR} determine whether the input data are genuine LR or HR data, respectively.

Specifically, the forward cycle ($\boldsymbol{\mathcal{X}} \rightarrow \boldsymbol{\mathcal{X}}'_{\mathrm{HR}} \rightarrow \boldsymbol{\mathcal{X}}''_{\mathrm{LR}} \approx \boldsymbol{\mathcal{X}}$) starts with the generator G_{LR2HR}, which learns to map deep LR seismic data to HR data with a similar distribution to shallow HR data. The input is deep LR seismic data $\boldsymbol{\mathcal{X}}$, and the learning target is the shallow HR seismic data $\boldsymbol{\mathcal{Y}}$. The output of G_{LR2HR} is the HR result $\boldsymbol{\mathcal{X}}'_{\mathrm{HR}}$ of the deep data. Following this, the reverse generator G_{HR2LR} simulates the inverse process of G_{LR2HR}, essentially mapping HR seismic data back to the original LR data domain. Accordingly, $\boldsymbol{\mathcal{X}}'_{\mathrm{HR}}$ is input into G_{HR2LR}, resulting in an approximate reconstruction $\boldsymbol{\mathcal{X}}''_{\mathrm{LR}}$ of the original LR data $\boldsymbol{\mathcal{X}}$. By applying consistency constraints to $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{X}}'$ in the loss function, the reversibility of G_{LR2HR} can be uniquely ensured. In addition, the HR discriminator D_{HR} is used to evaluate and refine the performance of G_{LR2HR}. It distinguishes

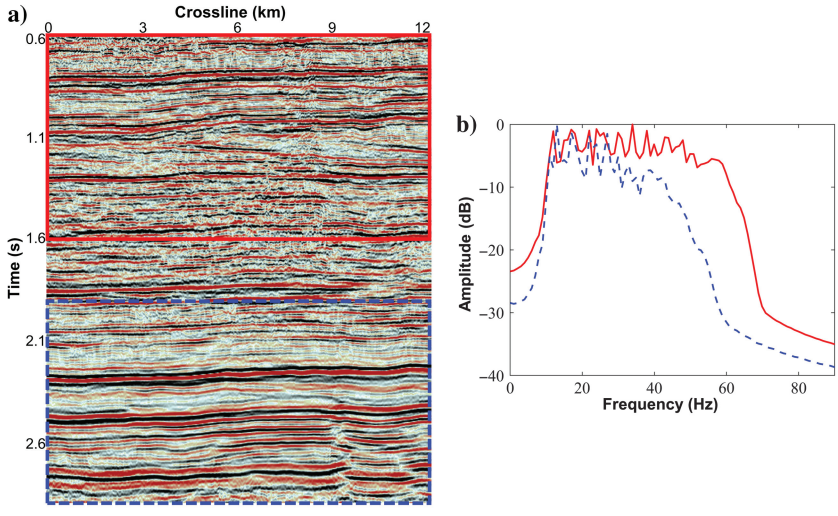


Figure 1. Resolution comparison between shallow-layer and deep-layer seismic data. (a) Inline section from real data. The shallow portion (0.6 to 1.6 s) exhibits relatively high resolution. In contrast, a deep portion (1.9 to 2.9 s) demonstrates lower resolution due to seismic wave attenuation. (b) Amplitude spectrum comparison between the shallow HR data (solid red) and deep LR data (dashed blue).

between the HR data $\mathcal{X}'_{HR}$ generated by G_{LR2HR} and the real target HR data $\mathcal{Y}$, estimating the probability that the generated HR data matches the target HR data, thereby guiding the training of the generator and improving robustness. This adversarial training process forces G_{LR2HR} to synthesize more realistic HR data, ultimately enabling it to generate HR data closely related to the target HR data, achieving a meaningful resolution enhancement of the LR seismic data. The reverse cycle ($\mathcal{Y} \rightarrow \mathcal{Y}'_{LR} \rightarrow \mathcal{Y}''_{HR} \approx \mathcal{Y}$) mirrors the forward cycle, swapping shallow and deep data, generators, and discriminators, constituting the reverse cycle process. By introducing additional constraints, the dual-cycle process significantly enhances training stability, reduces artifacts, and improves the overall quality of the generated HR seismic data.

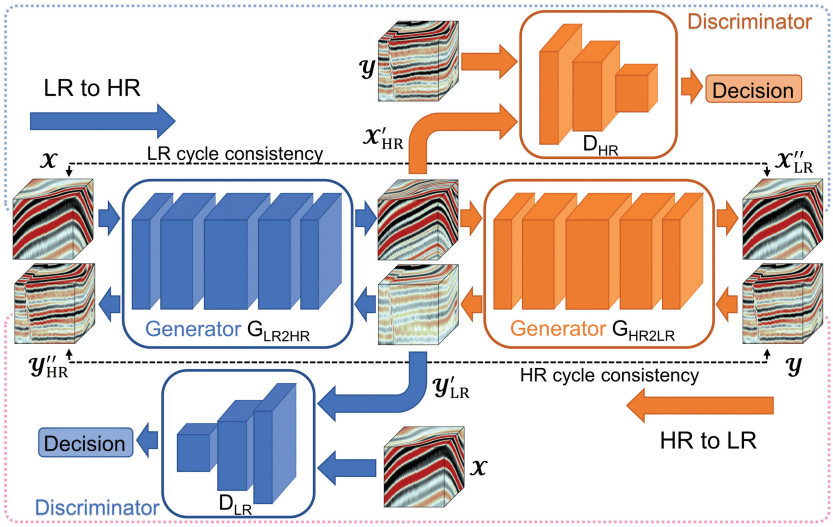


Figure 2. Workflow of bidirectional CycleGAN for enhancing seismic data resolution through weakly supervised learning. The notation $\mathcal{X}$ represents LR data from the deep layer. The HR version generated by the generator G_{LR2HR} is $\mathcal{X}'_{HR}$, and the reversed LR version generated by the second generator G_{HR2LR} is $\mathcal{X}''_{LR}$. Similarly, $\mathcal{Y}$ represents raw HR data from the shallow layer. The LR and HR versions obtained by G_{HR2LR} and G_{LR2HR} are $\mathcal{Y}'_{LR}$ and $\mathcal{Y}''_{HR}$, respectively. This workflow illustrates the dual-cycle process of transforming and regenerating between HR and LR seismic data to enhance resolution.

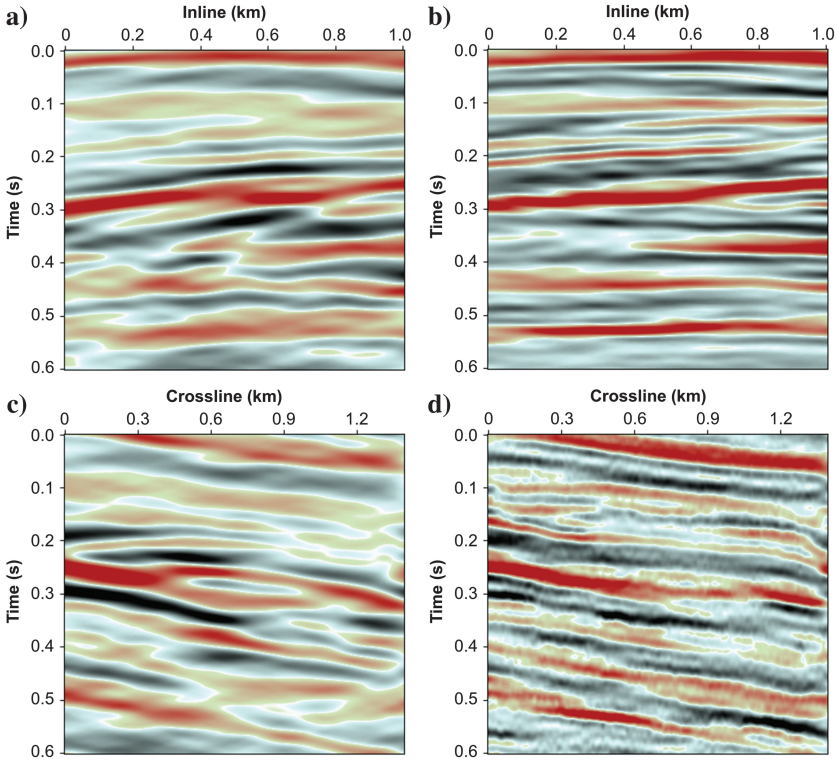


Figure 3. Illustration of discontinuities caused by 2D networks: (a) original crossline section. The 2D network is trained in this direction. (b) Crossline section results based on 2D networks, preserving reflection continuity, (c) original inline section, and (d) inline section results based on 2D networks showing significant discontinuities.

Generator architecture

Poststack seismic data are typically 3D, whereas most convolutional neural networks (CNNs) used in image processing are 2D. Applying these conventional networks requires slicing the 3D seismic data into a series of inline or crossline sections at the training and testing stages, which faces significant limitations. The 2D networks, constrained to learning within these sections, need help to effectively capture the 3D spatial structure of seismic data. Consequently, this often leads to discontinuities when investigating dimensions other than the trained ones. As shown in Figure 3, training with 2D CNNs on crossline sections of nearby HR 3D data and outputting results in crossline sections maintains high reflection continuity. However, it exhibits significant lateral discontinuity in the inline section, complicating stratigraphic tracking and fault identification. This challenge stems from the inherent limitations of 2D networks, which are insufficient to capture the 3D spatial coherency.

This challenge is particularly evident in our tasks. For shallow data layers where the signal-to-noise ratio (S/N) is relatively high, the output typically retains this high S/N and preserves the continuity of the original valuable signals. However, for deeper layers where the S/N is lower, outputting data by individual sections may lead to pronounced discontinuities on the other dimension, exacerbated by significant differences between adjacent sections.

Intuitively, 3D CNNs are better suited to capture the features of 3D seismic data, thereby delivering superior results. Accordingly, we design a 3D CNN architecture to extract the inherent characteristics of 3D seismic data. Unlike 2D networks that process single 2D sections individually, our approach processes 3D data blocks, better preserving spatial structure information. The architecture of our 3D generator is shown in Figure 4. It begins with an initial convolutional layer, followed by 16 consecutive residual blocks (He et al., 2016). Each residual block comprises two convolutional layers and a residual connection, with an overarching

skip connection encompassing the entire block group. The hidden representations generated by the residual blocks are further refined through additional convolutional layers and finalized with a tanh activation function. Key features of our generator include the absence of pooling layers to prevent the loss of important information, the use of residual blocks to enhance deep network performance, and the exclusion of batch normalization (BN) layers. Removing BN layers, especially in GAN training, helps mitigate artifacts and improve generalization, enhancing training stability and consistency (Wang et al., 2020b).

Discriminator architecture

The discriminator's architecture, as shown in Figure 5, starts with a 3D convolutional layer. This is followed by a series of five convolutional blocks: the first two blocks use 3D convolutions for feature extraction, whereas the next three switch to 2D convolutions, aiming to streamline dimensionality and reduce computational demands. Each block integrates a convolutional layer, a BN layer, and a Leaky ReLU activation function (Maas et al., 2013) with $\alpha = 0.2$, accelerating convergence and facilitating stable training. The processed feature maps then pass through a sigmoid activation function, which outputs probabilities from zero to one.

Distinctively, this discriminator uses a PatchGAN architecture (Isola et al., 2017), replacing the traditional fully connected output layer with a convolutional layer. This adaptation allows the discriminator to focus on localized features, which is particularly beneficial for enhancing HR details in seismic data. This strategic modification also reduces the model parameter count, promoting quicker convergence and yielding higher resolution and precision outputs.

Furthermore, the generator and discriminator architectures are mirrored in the forward and reverse cycles but their parameters are updated independently. This symmetry ensures a unified learning process across both translating directions, eliminating the need for separate adjustments of training hyperparameters. Such a setup boosts overall model efficiency and efficacy, ensuring stable reversibility performance regardless of resolution enhancement directionality.

Loss functions

Our proposed loss function comprises two components: generator loss and discriminator loss. The generator loss focuses on accurately reconstructing seismic images. In contrast, the discriminator loss ensures that the generated images are indistinguishable from real seismic data. These components work together to optimize the network performance, enhancing the reliability of seismic resolution enhancement processing.

Generator loss

The forward and reverse generators share the same network architecture and use the same loss function. The generator loss L_{Gen} comprises four components: adversarial loss L_{GAN}, cycle consistency loss L_{cyc}, identity loss L_{idt}, and total variation loss L_{TV}, as follows:

$$L_{\mathrm{Gen}} = w_1 L_{\mathrm{GAN}} + w_2 L_{\mathrm{cyc}} + w_3 L_{\mathrm{idt}} + w_4 L_{\mathrm{TV}}, \quad (1)$$

where w_1, w_2, w_3, and w_4 represent the trade-off weights assigned to each loss function component. Ideally, the individual contributions of these four loss functions to the overall loss should be cooperative, collectively enhancing seismic resolution. However, determining the best option is problematic. In our examples, these weights are initially determined through experiments on synthetic data to select appropriate values. They are then applied to field data. The robustness of our hyperparameter settings eliminates laborious parameter selection.

The adversarial loss attempts that the outputs of G_{LR2HR} and G_{HR2LR} effectively deceive their corresponding discriminators, D_{HR} and D_{LR}. Specifically, the smaller the adversarial loss, the more convincingly D_{HR} recognizes $\boldsymbol{\mathcal{X}}_{\mathrm{HR}}$ as genuine HR data. This loss function gradually drives the generated data toward the desired resolution. We use the negative log-likelihood (Goodfellow et al., 2020) as the adversarial loss, defined as

$$L_{\mathrm{GAN}} = -\frac{1}{N}\sum_{i=1}^{N}(\log(D_{\mathrm{HR}}(G_{\mathrm{LR2HR}}(\boldsymbol{\mathcal{X}}_i))) + \log(D_{\mathrm{LR}}(G_{\mathrm{HR2LR}}(\boldsymbol{\mathcal{Y}}_i)))). \quad (2)$$

Here, $\{\boldsymbol{\mathcal{X}}_i, \boldsymbol{\mathcal{Y}}_i\}_{i=1}^{N}$ represents the data set used during a training epoch, where N is the total number of samples processed.

The cycle consistency loss is designed to ensure the reversibility of the forward and reverse generators, thereby maintaining data consistency. It helps to address the nonuniqueness problem of seismic resolution enhancement and improve the realism of the generated data. The cycle consistency loss is defined as follows:

$$L_{\mathrm{cyc}} = \frac{1}{N}\left(\sum_{i=1}^{N}\|G_{\mathrm{HR2LR}}(G_{\mathrm{LR2HR}}(\boldsymbol{\mathcal{X}}_i)) - \boldsymbol{\mathcal{X}}_i\|_1 + \sum_{i=1}^{N}\|G_{\mathrm{LR2HR}}(G_{\mathrm{HR2LR}}(\boldsymbol{\mathcal{Y}}_i)) - \boldsymbol{\mathcal{Y}}_i\|_1\right). \quad (3)$$

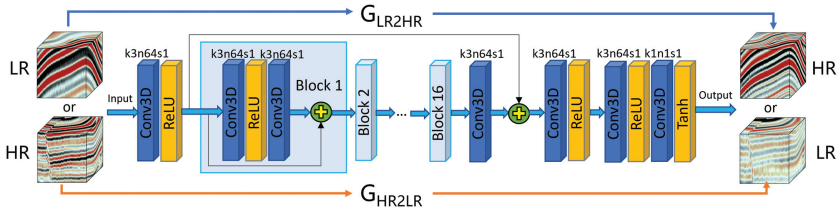


Figure 4. Generator architecture.

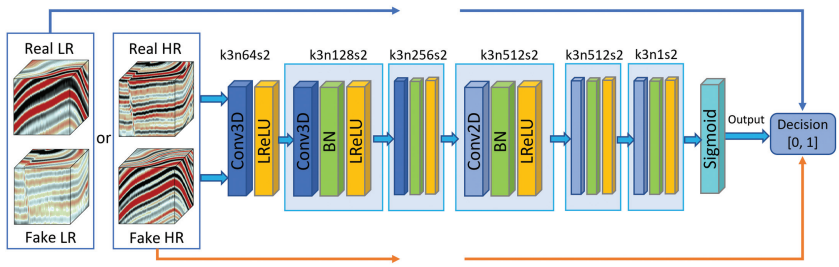


Figure 5. Discriminator architecture.

Identity loss is commonly used in image transformation models to preserve key image characteristics, such as color consistency, texture details, and structural integrity (Zhu et al., 2017). For seismic resolution enhancement, we modify the input of identity loss to retain key seismic characteristics and produce more accurate and realistic results. Specifically, the HR data $\mathcal{Y}$ is degraded through low-pass filtering to obtain the corresponding degraded data $\bar{\mathcal{Y}}_i$, which is then fed into the forward generator to produce HR data $G_{\mathrm{LR2HR}}(\bar{\mathcal{Y}}_i)$. By maintaining the identity between $\mathcal{Y}$ and $G_{\mathrm{LR2HR}}(\bar{\mathcal{Y}}_i)$, the forward generator G_{LR2HR} is constrained. Similarly, the HR data $\mathcal{Y}$ is fed into the reverse generator to obtain LR data $G_{\mathrm{HR2LR}}(\mathcal{Y}_i)$, and by maintaining the identity between $\bar{\mathcal{Y}}$ and $G_{\mathrm{LR2HR}}(\mathcal{Y}_i)$, the reverse generator G_{HR2LR} is constrained to ensure stable resolution improvement.

Notably, this degradation process is also crucial to help the forward generator learn an effective mapping from LR to HR data. However, this mapping might differ from the actual LR-HR mapping since real LR data are not just a low-pass filtered version of HR data. Despite this, the degradation process helps the generator capture the key elements necessary for generating HR data. In addition, since the generator is subject to multiple constraints, its final output is collaboratively improved by all constraints, contributing jointly to successful resolution enhancement.

Recent research indicates that using a combined approach of L_1 loss and multiscale structural similarity (MS-SSIM) loss $L_{\mathrm{MS\text{-}SSIM}}$ can yield better image restoration results (Zhao et al., 2017). Following this, the definition of the identity loss used in this study is

$$L_{\mathrm{idt}}=\frac{1}{N}\sum_{i=1}^{N}(L_1(G_{\mathrm{LR2HR}}(\bar{\mathcal{Y}}_i),\mathcal{Y}_i)+L_{\mathrm{MS\text{-}SSIM}}(G_{\mathrm{LR2HR}}(\bar{\mathcal{Y}}_i),\mathcal{Y}_i) +L_1(G_{\mathrm{HR2LR}}(\mathcal{Y}_i),\bar{\mathcal{Y}}_i)+L_{\mathrm{MS\text{-}SSIM}}(G_{\mathrm{HR2LR}}(\mathcal{Y}_i),\bar{\mathcal{Y}}_i)), \quad (4)$$

where the L_1 loss measures the difference between two seismic cubes. Given two 3D seismic data $\mathcal{U}$ and $\mathcal{V}$, the L_1 loss is calculated as

$$L_1(\mathcal{U},\mathcal{V}) = \|\mathcal{U}-\mathcal{V}\|_1. \quad (5)$$

The MS-SSIM loss is an image quality assessment method sensitive to local structural variations (Wang et al., 2003). The term "multiscale" refers to applying a low-pass filter sequentially followed by a twofold downsampling of the image. Given 2D seismic sections $\mathbf{U}$ and $\mathbf{V}$, the formula for MS-SSIM is determined over all M scales as follows:

$$\mathrm{MS-SSIM}(\mathbf{U},\mathbf{V})=[l_M(\mathbf{U},\mathbf{V})]^{\alpha_M}\cdot\prod_{j=1}^{M}[c_j(\mathbf{U},\mathbf{V})]^{\beta_j}[s_j(\mathbf{U},\mathbf{V})]^{\gamma_j}, \quad (6)$$

where $l_j(\mathbf{U},\mathbf{V})$, $c_j(\mathbf{U},\mathbf{V})$, and $s_j(\mathbf{U},\mathbf{V})$ represent the similarities of $\mathbf{U}$ and $\mathbf{V}$ in luminance, contrast, and structure at scale j, respectively. Their corresponding weights are α_j, β_j, and γ_j. The respective definitions are

$$l(\mathbf{U},\mathbf{V}) = \frac{2\mu_{\mathbf{U}}\mu_{\mathbf{V}}+c_1}{\mu_{\mathbf{U}}^2+\mu_{\mathbf{V}}^2+c_1}, \quad (7a)$$

$$c(\mathbf{U},\mathbf{V}) = \frac{2\sigma_{\mathbf{UV}}+c_2}{\sigma_{\mathbf{U}}^2+\sigma_{\mathbf{V}}^2+c_2}, \quad (7b)$$

$$s(\mathbf{U},\mathbf{V}) = \frac{\sigma_{\mathbf{UV}}+c_3}{\sigma_{\mathbf{U}}\sigma_{\mathbf{V}}+c_3}, \quad (7c)$$

where μ and σ indicate the mean and standard deviation, respectively. Constants c_1, c_2, and c_3 help stabilize calculations and prevent division errors. Since the calculation is based on 2D data, we extract 2D seismic sections with time and larger spatial dimensions (either inline or crossline) from 3D data to accommodate this loss. A larger spatial dimension optimizes our ability to capture intricate subsurface geologic features and enhances measurement accuracy. Specifically, when the crossline length is larger than the inline, we apply the following formula:

$$L_{\mathrm{MS-SSIM}}(\mathcal{U},\mathcal{V}) = 1-\mathrm{MS-SSIM}(\mathbf{U}_{:,:,i},\mathbf{V}_{:,:,i}), \quad (8)$$

where $\mathbf{U}_{:,:,i}$ and $\mathbf{V}_{:,:,i}$ represent randomly selected inline 2D sections from the 3D data sets $\mathcal{U}$ and $\mathcal{V}$, respectively, with i indicating a randomly chosen inline index. In contrast, when the inline length exceeds the crossline length, the analysis shifts to crossline sections $\mathbf{U}_{:,x,:}$ and $\mathbf{V}_{:,x,:}$, using a randomly chosen crossline index x. This random selection strategy provides a comprehensive and accurate evaluation of seismic data tailored specifically to seismic dimensional characteristics.

Furthermore, we introduce total variation loss, which enhances spatial smoothness by mitigating excessively high gradients during generator training. For the output data $\mathcal{D}$ from either the forward or reverse generators, its definition is

$$L_{\mathrm{TV}} = \frac{1}{N}\sum_{i=1}^{N}(\|\nabla_t\mathcal{D}\|_1+\|\nabla_x\mathcal{D}\|_1+\|\nabla_i\mathcal{D}\|_1), \quad (9)$$

where ∇_t, ∇_x, and ∇_i represent the gradients of $\mathcal{D}$ across the three dimensions, respectively. This formulation minimizes noise and artifacts in the generated seismic data by penalizing these abrupt changes across different orientations.

Discriminator loss

Building on the fundamental principles of GAN, we use discriminators D_{HR} and D_{LR} to facilitate adversarial training for the generators G_{LR2HR} and G_{HR2LR}. The discriminators are designed to output a score of one when assessing real data as authentic and zero when evaluating generator-generated data. Discriminator loss quantifies the effectiveness of the discriminators in distinguishing genuine data from synthetic outputs. This loss is determined by the negative log-likelihood as

$$L_{\mathrm{Dis}} = -\frac{1}{N}\sum_{i=1}^{N}(\log D_{\mathrm{HR}}(\mathcal{Y}_i)+\log(1-D_{\mathrm{HR}}(G_{\mathrm{LR2HR}}(\mathcal{X}_i))) +\log D_{\mathrm{LR}}(\mathcal{X}_i)+\log(1-D_{\mathrm{LR}}(G_{\mathrm{HR2LR}}(\mathcal{Y}_i)))). \quad (10)$$

By minimizing this loss, the discriminators effectively distinguish between real and generated data, enhancing the generator's

translation capabilities and ensuring the authenticity of the outputs against real-world data.

Training data set construction and preprocessing

When using 3D CNNs for seismic resolution enhancement, more network parameters than 2D networks lead to a surge in graphics processing unit (GPU) memory consumption. This requires a systematic approach to partitioning and aggregating 3D data. In addition, the parameter increase necessitates a larger data set to prevent overfitting, which is addressed through data set augmentation. Furthermore, uneven energy distribution across reflection layers can cause training instability. Amplitude balancing is crucial to mitigate the problem mentioned previously.

Data partitioning and aggregating

To enhance 3D seismic data resolution, we partition the original data set into fixed-size blocks with a 50% overlap. This overlap value is based on empirical experience to ensure seamless integration with affordable computational costs during processing. Each block is independently processed and reassembled to approximate the original data dimensions. This approach, while effective, often introduces discontinuities at the block boundaries, posing significant challenges to preserving data integrity across the data set.

To mitigate these discontinuities, we implement an effective windowing aggregation strategy. Before reassembly, each block undergoes processing through a triangular window function tailored to smooth transitions between blocks. This window function, shown in Figure 6, ensures that the sum of the weights at each point equals one, thus preserving the original data's amplitude and integrity. Our subsequent experiments confirm that this technique significantly improves the overall quality of seismic data by effectively minimizing boundary artifacts, resulting in a more precise and consistent representation of geologic features across the data set.

Data augmentation

Due to budget constraints, some exploration data sets are small scale, exacerbating the challenge of insufficient 3D training data. Taking training data $\mathcal{D}$ with a 3D structure as an example, we use three data augmentation strategies to address this issue.

1) Flipping inline and crossline spatial dimensions: $\mathcal{D}^{\text{aug}}_{t,x,i} = \mathcal{D}_{t,x',i'}$, where $x' = N_{\text{x}} - x$ and $i' = N_{\text{i}} - i$. Here, N_{i} and N_{x} represent the lengths of data in inline and crossline dimensions, respectively. This technique mirrors the data along these axes, enhancing data set diversity.
2) Swapping inline and crossline spatial dimensions: $\mathcal{D}^{\text{aug}}_{t,x,i} = \mathcal{D}_{t,i,x}$. Interchanging these axes enhances spatial feature understanding, crucial for comprehensive model training.
3) Inverting data values: $\mathcal{D}^{\text{aug}} = -\mathcal{D}$. This technique negates the data values, adding a layer of complexity by introducing contrasts, which can be particularly useful in feature identification.

Here, $\mathcal{D}^{\text{aug}}$ represents the augmented data, where each augmentation is randomly performed with equal probability. These strategies have significantly expanded the training sample size, enabling the network to thoroughly explore seismic features from multiple perspectives. Different data augmentation techniques can be used separately for input data and learning targets. This offers greater flexibility than traditional supervised learning methods with paired training samples.

EXAMPLES

This section starts by showcasing the viability of our unpaired training, illustrated by a 3D synthetic data example. Subsequently, we assess the efficacy of our proposed network using field data sets. Furthermore, we validate our methodology by comparing with a conventional time-varying spectrum whitening method (Bian and Zhang, 1986).

Synthetic data experiment

Synthetic data construction

We use the open-source Synthoseis model (Merrifield et al., 2022) to construct synthetic seismic data. This model, featuring a 3D structure and diverse subsurface scenarios, is particularly suitable for evaluating our 3D approach. It generates P-wave velocity and density models, from which the acoustic impedance is derived.

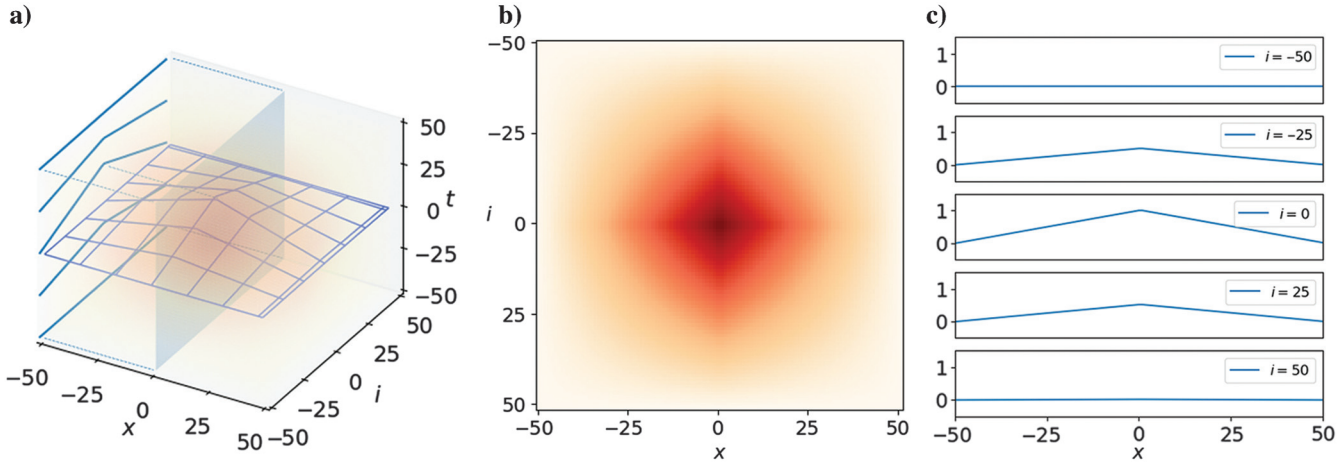


Figure 6. Illustration of the 3D windowing function applied to data blocks of size $100 \times 100 \times 100$: (a) Weight distributions across the three dimensions, (b) a 2D cross section at $t = 0$, and (c) variations at different x values with $i = 0$.

Reflection coefficients are then calculated based on impedance contrasts between different layers. Finally, they are convolved with seismic wavelets of varying central frequencies to produce seismic data at different resolutions. We use a Ricker wavelet (Hosken, 1988; Gholamy and Kreinovich, 2014) for the convolution.

The synthetic seismic data dimensions are $600 \times 400 \times 400$ (time, crossline, and inline, respectively), with a time sampling interval of 1 ms. The spatial sampling intervals are 10 m in the crossline and inline directions. To briefly mirror the shallow and deep segments of actual seismic data, the synthetic data are divided into two blocks, each of size $300 \times 400 \times 400$ points. The upper block is convolved with a 45 Hz Ricker wavelet to simulate the shallow part. In contrast, the lower block uses a 30 Hz wavelet for the deep region. Figure 7 shows the resulting synthetic seismic data. Figure 8 shows the normalized multitrace amplitude spectra for both sections. From this, we can clearly distinguish their differences in frequency. Before training, the data undergo min-max normalization to ensure stable network performance. Using uniform distribution sampling, shallow and deep synthetic seismic data are segmented into blocks of $80 \times 80 \times 11$ (or $80 \times 11 \times 80$), serving as input and label data for the network. This segmentation strategy uses thin, rectangular prisms with a longer temporal axis, encompassing multiple wavelet lengths to enhance analytical precision while optimizing GPU memory utilization. Following data augmentation, the prepared data are inputted into the network for training, as shown in Figure 9.

Training parameter setting

The Adam optimizer (Kingma and Ba, 2014) is used for model training, configured with parameters $\beta_1 = 0.5$, $\beta_2 = 0.999$, and $\varepsilon = 10^{-8}$. Training samples are randomly extracted from the data set constructed in the previous section through uniform distribution sampling. The settings for equation 1 specify $w_1 = 0.05$, $w_2 = 25$, $w_3 = 5$, and $w_4 = 2$. The initial learning rate of the network is set at 0.0001 and is halved every 10,000 epochs. The input data are processed in batches of two, culminating in comprehensive 20,000 training epochs.

Experimental results

After finalizing network training, the deep-layer LR data are tested. There are typically two critical evaluation metrics for seismic resolution enhancement: extending the effective signal bandwidth to enhance high-frequency features while retaining low-frequency information. Following this rule, our results are evaluated by displaying inline, crossline, and time-slice results and comparing their normalized multitrace average amplitude spectra. The efficacy and superiority of the network method are also confirmed by comparing it with spectral whitening techniques. We also perform ablation experiments to verify the validity of each model component.

Figure 10 shows results at inline 200, showcasing the original LR and corresponding HR data of the deep layer in Figure 10a and 10b, respectively. The outcomes of applying spectral whitening and network-based resolution enhancement to the LR data are shown in Figure 10c and 10d, respectively. These comparisons highlight a significant enhancement in the deep-layer resolution when processed by the network, notably enhancing the ability to recognize critical geologic structures. Fault structures are remarkably well preserved, demonstrating high-fidelity structural maintenance during resolution enhancement. Notably, within the green-boxed area, the network method sharply contrasts with spectral whitening, as the latter blurs seismic events and obscures structural details. In contrast, the network not only preserves the integrity of structural information but also clearly presents subtle geologic features, demonstrating its superior resolution enhancement capabilities over traditional methods.

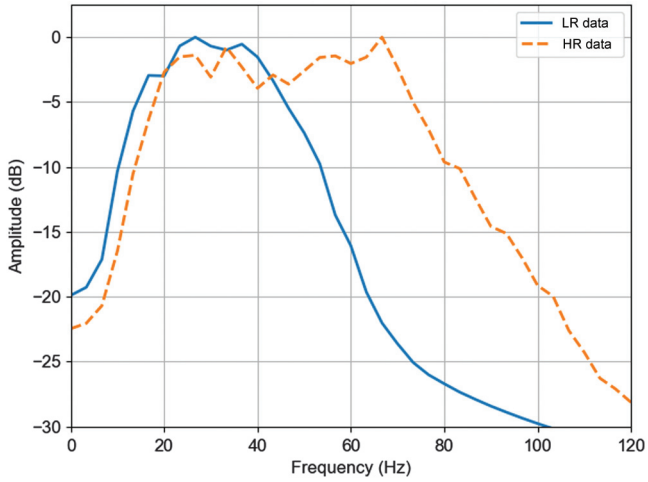


Figure 8. Normalized multitrace amplitude spectrum for HR and LR data.

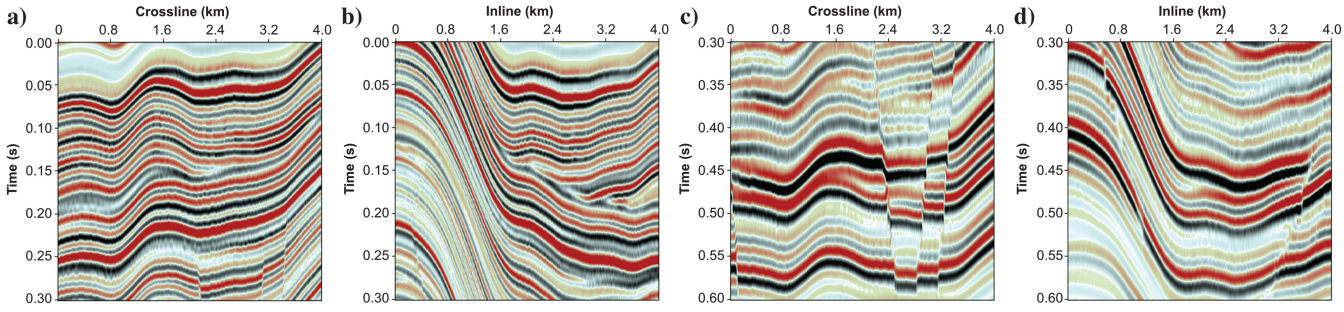


Figure 7. Examples of synthetic sections. (a) Shallow-layer HR inline section, (b) shallow-layer HR crossline section, (c) deep-layer LR inline section, and (d) deep-layer LR crossline section.

Figure 11 shows result at crossline 200. The side-by-side comparison shown in Figure 11c (results from spectral whitening) and Figure 11d (results from our method) highlights our method's performance. Our approach enhances resolution while effectively preserving stratigraphic details, clearly surpassing spectral whitening. The green box highlights regions where our method exhibits more distinct stratigraphic sequences, making target layers pronounced and easily identifiable. In addition, the yellow arrows emphasize that our technique improves the continuity of events over spectral whitening, facilitating more accurate horizon tracking.

Figure 12 shows the time-slice results at time = 450 ms. The comparison between spectral whitening (Figure 12c) and our HR network method (Figure 12d) underscores our effectiveness. Our approach not only boosts the resolution of deep-layer data but also preserves geologic details with greater fidelity than spectral whitening. Notably, our method demonstrates its superior capability within the green-boxed area by capturing nuanced seismic structures more accurately, maintaining structural integrity, and aligning closely with the HR reference (Figure 12b).

The normalized multitrace amplitude spectra are shown in Figure 13 to further compare the spectral whitening and our network method. The original LR data are represented by a solid blue line. At the same time, the corresponding HR data are shown with a dashed yellow line. We can observe significant bandwidth differences. Results from the spectral whitening method are depicted as a green dotted line, and the red dash-dotted line indicates outcomes from the network method. Both methods effectively enhance high-frequency energy and expand bandwidth. However, the network method maintains greater fidelity to the original low-frequency characteristics, particularly below 20 Hz. In contrast, spectral whitening is designed to artificially boost these frequencies without accurately preserving their original content.

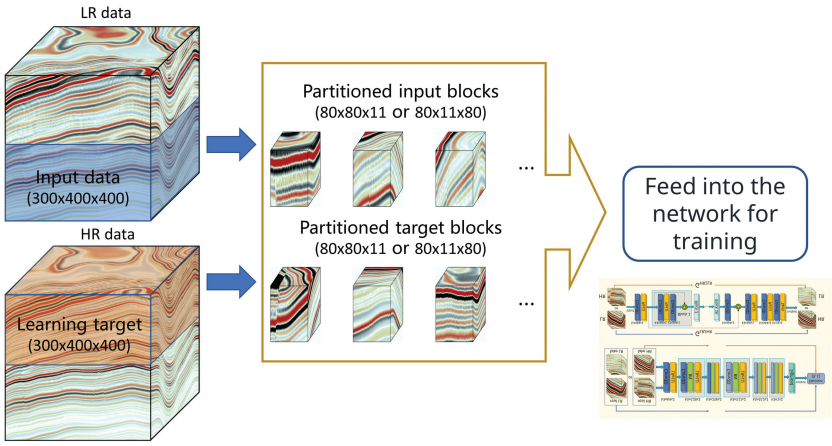


Figure 9. Schematic diagram of synthetic seismic data preparation.

Ablation experiments

To quantitatively evaluate the proposed method, this section presents an ablation study. We systematically eliminate each loss component and data augmentation during the training process while keeping the network architecture and training parameters constant. This retraining approach allows us to assess the impact of these strategies by measuring the similarity between network outputs and the corresponding HR data. In addition, we benchmark these results against those from spectral whitening using the same HR data. The Pearson correlation coefficient (PCC) and S/N serve as evaluation metrics, where higher values indicate a closer resemblance between the network outputs and HR data.

As shown in Table 1, data augmentation significantly enhances network performance, effectively mitigating severe overfitting caused by 3D training scarcity. In addition, identity loss L_{idt} plays a second important role. It strategically uses degraded data to finely guide the optimization of network weights, ensuring the precision of the training process. The exclusion of any loss components notably impacts network performance, with specific declines in PCC and S/N, underscoring their individual importance. Compared with conventional spectral whitening, the network method delivers HR outputs closer to the intended HR data, confirming its advanced efficacy.

Field data experiment

Data set preparation

This section demonstrates the effectiveness of our proposed method using a 3D real seismic data set. The spatial dimensions consist of 400

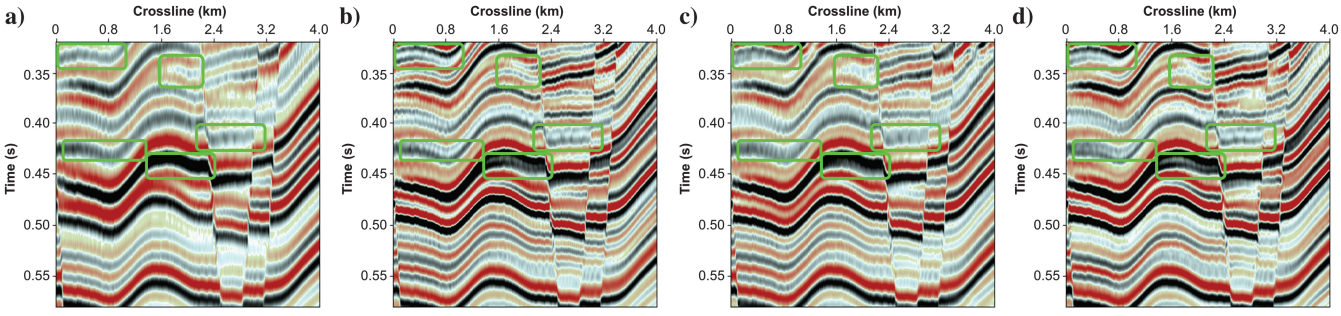


Figure 10. Inline results (inline = 200) of synthetic seismic data: (a) original deep-layer LR inline section using a 30 Hz wavelet; (b) corresponding HR inline section for comparison, using a 45 Hz wavelet; (c) HR inline result achieved through spectral whitening; and (d) HR inline result obtained using our method. The green boxes highlight areas where our method significantly improves resolution, enabling better recognition of subtle geologic events compared with spectral whitening.

inlines and 600 crosslines, with a bin size of 40 m × 20 m, covering an area of 16 km × 12 km. The time sampling is 2 ms. Figure 1 shows an inline section of the poststack 3D seismic data collected from Northern China. The shallow portion is selected across 0.6–1.6 s, whereas the deep portion is chosen from a relatively high S/N region spanning 1.9–2.9 s. After training, the network is applied to the deep data to enhance resolution. A crossline example is shown in Figure 14a, and the normalized multitrace amplitude spectra of the upper and lower parts are shown in Figure 14b. The results indicate a significant resolution difference between the shallow and deep layers, similar to the inline section. Both layers exhibit stratified sedimentary formations with comparable geologic characteristics, making them suitable for resolution enhancement using a weakly supervised network.

Training parameter setting

The Adam optimizer is used for training with parameters set as $\beta_1 = 0.5$, $\beta_2 = 0.999$, and $\varepsilon = 10^{-8}$. Training samples are randomly cropped from areas selected previously using uniform distribution sampling. The parameters in equation 1 are set as $w_1 = 0.05$, $w_2 = 25$, $w_3 = 5$, and $w_4 = 2$. The initial learning rate of our network is 0.0001, which decreases to 0.5 times the initial learning rate for every 10,000 sample. The batch size of the input data are two, and the network is trained for 20,000 epochs. By keeping the parameters the same as the previous synthetic data, we can assess the transferability of our method to other data sets. This ensures network performance remains consistent and allows for a seamless transition to different training environments, avoiding time-consuming parameter selection.

Experimental results

We evaluate the proposed method using crossline profiles, inline profiles, time slices, spectral analysis, and well-log tie. A comparative study with the widely implemented traditional time-variant spectral whitening technique is used to validate the accuracy and effectiveness of our method.

Figure 15 shows the real data results for inline = 8 km. Figure 15a shows the original deep LR data distorted by moderate noise, causing event discontinuities. Figure 15b shows the enhanced

Table 1. Comparison of ablation experiment results for synthetic seismic data.

Evaluation method	PCC	S/N (dB)
Our full approach	0.9707	12.377
Spectral whitening	0.9382	9.216
Without L_{GAN}	0.9199	8.117
Without L_{cyc}	0.9307	8.620
Without L_{idt}	0.9105	7.668
Without L_{TV}	0.9461	9.622
Without data augmentation	0.8391	5.246

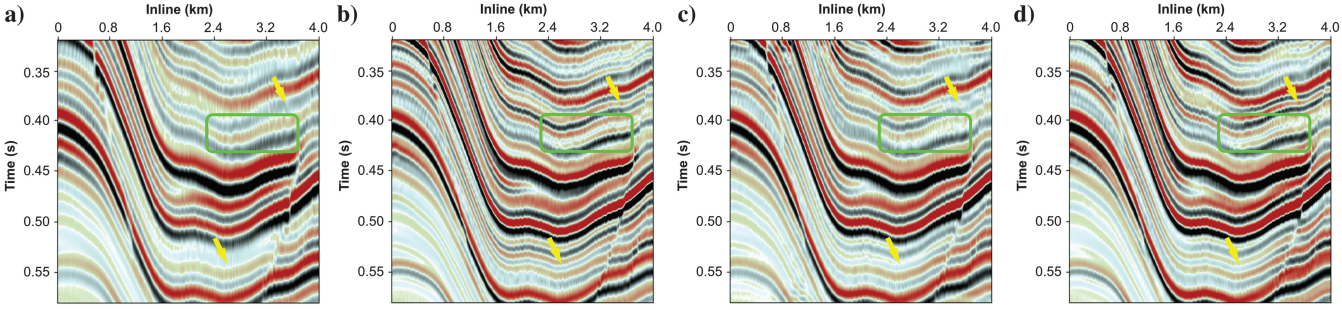


Figure 11. Crossline results (crossline = 200) of synthetic seismic data: (a) original deep-layer LR crossline section using a 30 Hz wavelet; (b) corresponding HR crossline section for comparison, using a 45 Hz wavelet; (c) HR crossline result achieved through spectral whitening; and (d) HR crossline result obtained using our method. The green boxes highlight regions where our method exhibits clearer stratigraphic sequences. At the same time, the yellow arrows point to areas where our approach better enhances event continuity than spectral whitening.

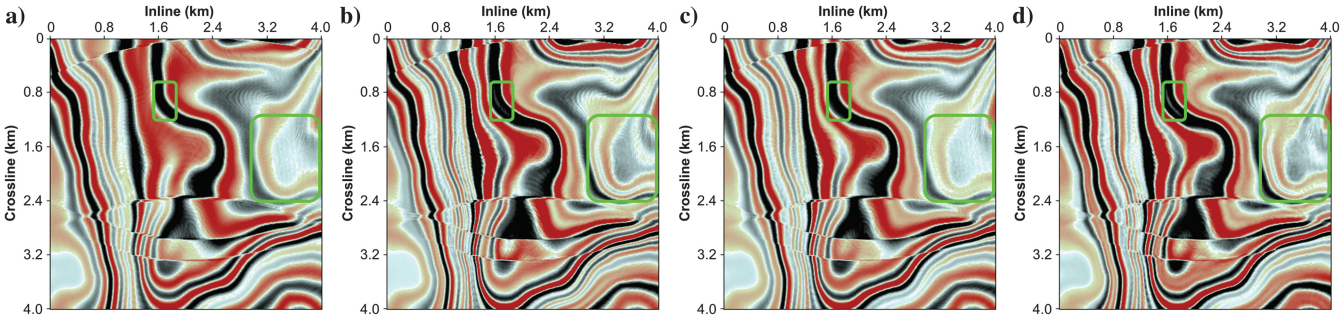


Figure 12. Time-slice results (time = 450 ms) of synthetic seismic data: (a) original deep-layer LR time slice using a 30 Hz wavelet; (b) corresponding HR time slice for comparison, using a 45 Hz wavelet; (c) HR result achieved through spectral whitening; and (d) HR result obtained using our method. The green boxes highlight that our method better captures nuanced seismic structures.

result after spectral whitening. In contrast, Figure 15c shows the result from the network. The results indicate a significant enhancement in the resolution of the original seismic data with more detailed reflections and finer geologic structures. Moreover, we observe substantial structural similarity between the two methods, confirming the reliability of our method.

However, notable differences exist between our method and spectral whitening. Although spectral whitening enhances resolution, it inevitably introduces high-frequency noise, reducing S/N. In contrast, our method effectively restores thin events with good continuity, enhancing seismic resolution without sacrificing data fidelity. The green boxes further underscore our method's capability to substantially boost the resolution of weak events, demonstrating its robustness in revealing subtle geologic formations. This improvement is attributed to the carefully designed loss function and the bidirectional cycle structure that identifies the intrinsic features of seismic data. Specifically, despite the low S/N of the shallow HR data being designated as the target for learning in the forward cycle, high S/N LR deep data are also targeted in the reverse cycle for training the generator. This strategic dual targeting enables the generator to enhance resolution consistently across varying input noise levels, boosting network robustness. The high S/N outcomes in HR data illustrate that this strategy successfully leverages the benefits of shallow and deep data layers. Accordingly, weakly supervised learning mirrors visual features and accurately characterizes the seismic wavefield, leading to superior resolution enhancement.

Figure 16 shows the crossline section results at 6 km. Figure 16a shows the original section, characterized by relatively broad events due to a lack of high-frequency energy. Compared with the original profile, both the spectral whitening technique in Figure 16b and our method in Figure 16c show significant resolution improvements and effectively restore valuable high-frequency information. However, spectral whitening struggles to maintain spatial continuity and energy consistency, resulting in blurred and noisy events. In contrast, our method significantly enhances lateral continuity, improving spatial energy consistency. Even under noisy conditions, the network maintains superior enhancement performance, demonstrating its strong resilience to noise, as highlighted by the yellow boxes. This enhancement is crucial for reliable horizon tracking and detailed stratigraphic interpretations.

Further examination of two time slices at 2284 and 2504 ms in Figures 17 and 18 emphasizes the superior resolution recovery capabilities of our method. After processing, deep time slices show a more pronounced improvement than with spectral whitening. As delineated by the yellow rectangles, our method delineates geologic structures well. It also reveals details that spectral whitening fails to capture. This enhanced delineation accurately depicts spatial relationships between different strata with increased precision, making geologic interfaces more discernible. We also compute a similarity attribute map to support this further. As shown in Figure 19, our method highlights more detailed spatial variations while preserving the integrity of the original data, thereby enabling the identification of subtle structural changes. Notably, our method yields a higher S/N than the spectral whitening method, as indicated by the yellow elliptical circles.

Figure 20 shows a comparative analysis of normalized multitrace amplitude spectra. The dashed blue, solid red, and dotted green lines represent the original LR data, the HR results from spectral whitening, and our network. An amplitude spectrum analysis shows that the proposed method and spectral whitening effectively enhance high-frequency components, achieving comparable bandwidth extensions. They align closely with the original LR data, preserving intrinsic features while improving resolution. Notably, both approaches preserve the low-frequency range (0–15 Hz) with

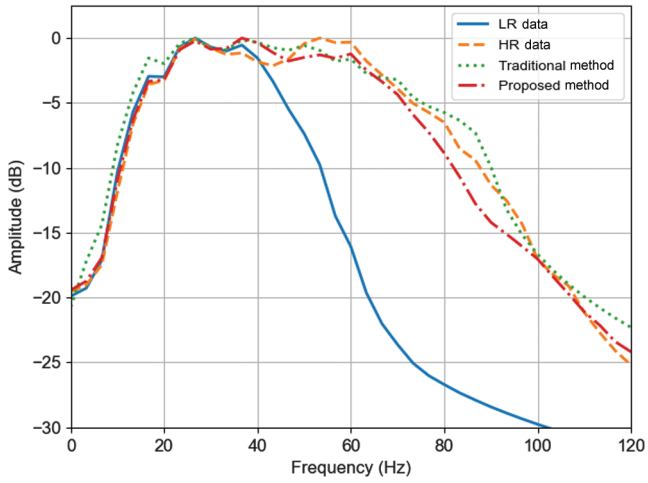


Figure 13. Normalized multitrace amplitude spectrum comparison for synthetic data. Our method preserves low-frequency energy more accurately than spectral whitening.

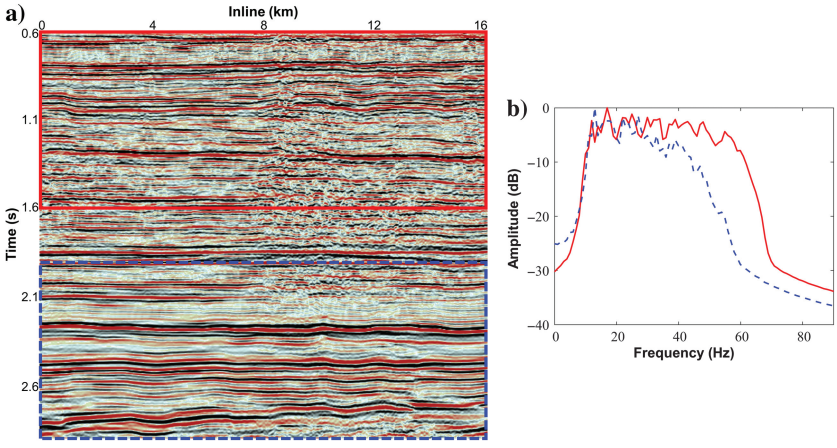


Figure 14. Resolution comparison between shallow-layer and deep-layer seismic data. (a) Crossline section from real data at 6 km. The red box highlights the shallow portion, which exhibits relatively high resolution and serves as the network input. In contrast, the blue box marks the deep portion, demonstrating lower resolution due to seismic wave attenuation, and is selected as the training target for weakly supervised learning. (b) Normalized amplitude spectrum comparison between the shallow HR data (solid red) and the deep LR data (dashed blue).

high fidelity, maintaining the fundamental characteristics of original seismic data. However, our method demonstrates superior alignment over the entire LR effective bandwidth, suggesting enhanced preservation of midfrequency (15–35 Hz) details. In contrast, spectral whitening diverges from the original spectra beyond 20 Hz. It fails to follow the original frequency transitions around 35 Hz, losing signal integrity in the midfrequency range. By keeping low-frequency information while enhancing high frequencies, our approach provides a more balanced and reliable results. Moreover, our approach preserves data integrity and structural information well.

To evaluate the preceding results quantitatively, we compare the seismic data with well-logging data from the borehole marked by the red circle in Figures 17a and 18a. We extract statistical wavelets from nine nearby traces of each data set within a 2.0 to 2.3 s time window, targeting the layer of interest. Figure 21 shows the synthetic traces derived from well-log data alongside the closest seismic traces to the borehole location. A close match between the original seismic and synthetic data, shown in Figure 21a, demonstrates a reliable well-log tie, further confirmed by a normalized crosscorrelation of 0.65. Spectral whitening enhances resolution, as shown in Figure 21b but achieves a crosscorrelation of only 0.51, indicating limited accuracy and some misalignment with the well-log data. In contrast, our method in Figure 21c yields a crosscorrelation of 0.61, showing a closer alignment. This quantitative evidence, combined with qualitative enhancements, confirms that our approach significantly improves seismic resolution.

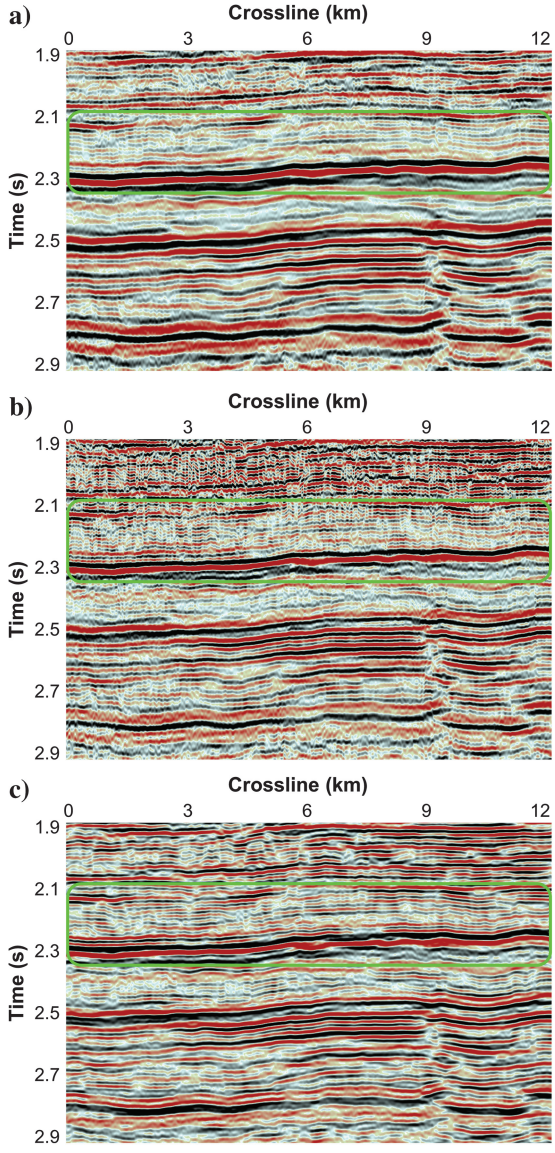


Figure 15. Inline section results at 8 km. (a) Deep LR seismic data, (b) the HR result obtained by spectral whitening, and (c) the HR result obtained by the proposed method. The green boxes highlight that our method significantly enhances the resolution of weak events, demonstrating its robustness in revealing subtle geologic formations.

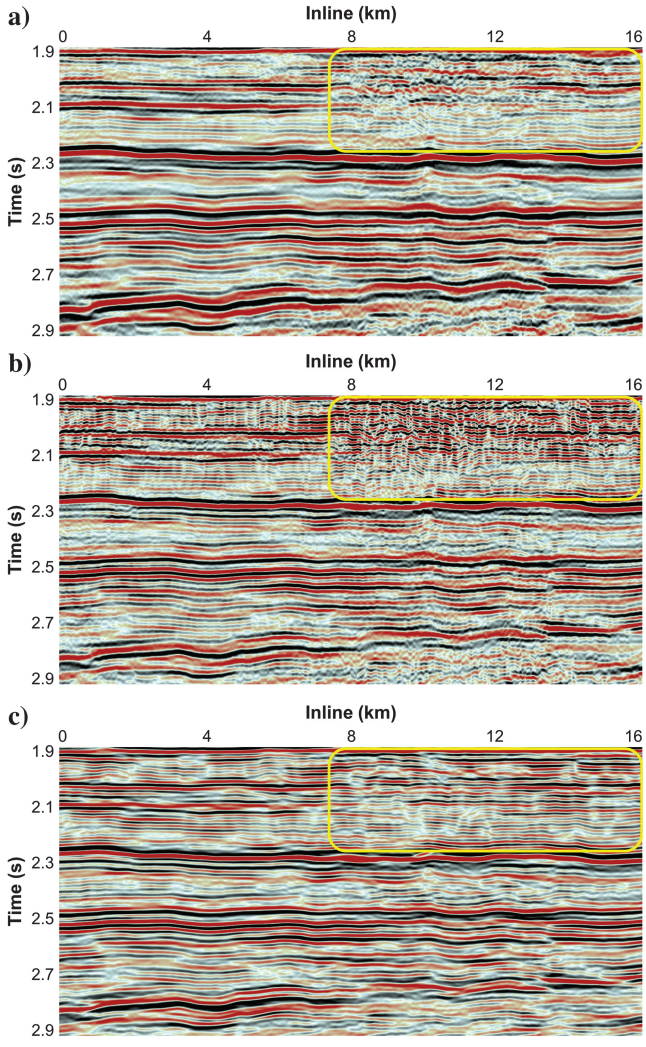


Figure 16. Crossline section results at 6 km. (a) Deep LR seismic data, (b) the HR result obtained by spectral whitening, and (c) the HR result obtained by our method. The yellow boxes highlight that our method significantly enhances the resolution under noisy conditions, indicating its high resilience to noise.

DISCUSSION

Resolution enhancement for mismatched structures

Our method has demonstrated effectiveness in enhancing the resolution of deep-layer seismic data for horizontal sedimentary formations with strong robustness to noise. However, to further assess its robustness, evaluating its performance on more complex geologic structures is important. As shown in Figures 22 and 23, the target layer, situated between 1.8 and 2.1 s, exhibits distinct features that are not present in the shallow data. Consequently, directly selecting nonoverlapping shallow-layer data as the learning target, although improving resolution, would introduce fake geologic patterns due to feature mismatch, thus limiting the reliability of the results. Through trial and error tests, we discover that the input data encompassing a portion of the learning target regions can ensure consistency and minimize the risk of introducing artificial patterns. This overlapping training data set division yields satisfactory outcomes, interpreted as data augmentation. As shown in Figures 24 and 25, the resolution within the 1.8–2.1 s interval is significantly enhanced, recovering clearer events compared with spectral whitening. Furthermore, the geologic features below 2.1 s are preserved, underscoring the reliability of our approach. Notably, while the results are only marginally superior to those obtained through spectral whitening for sedimentary layers, they nonetheless demonstrate the

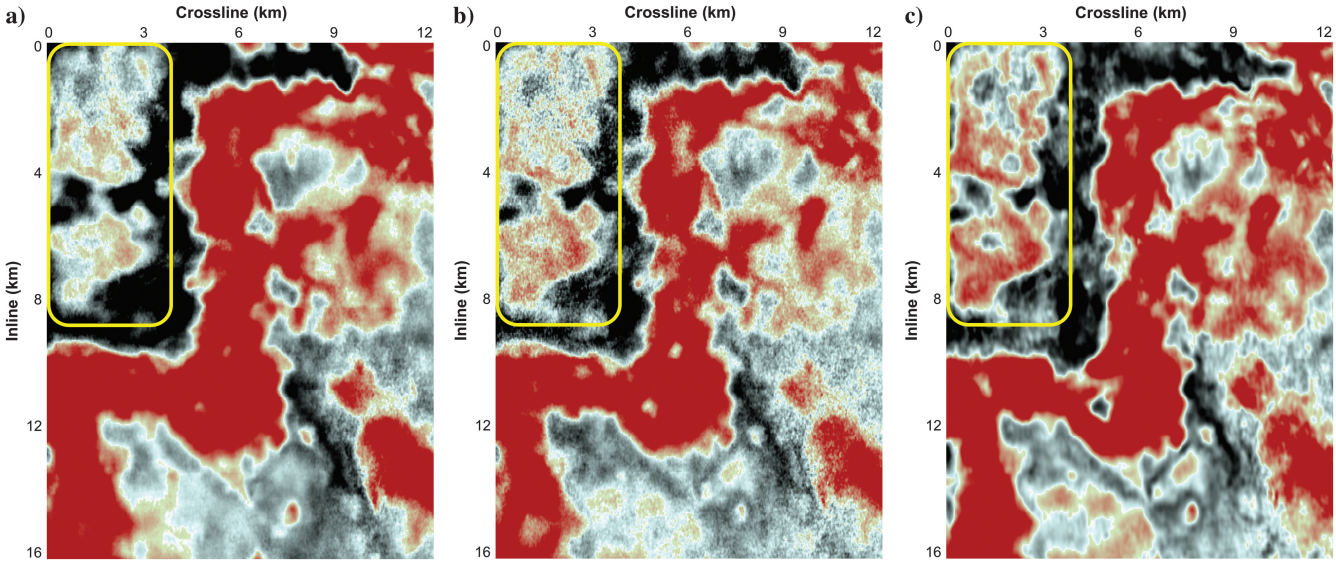


Figure 17. Time-slice results at 2284 ms. (a) Original LR seismic data, (b) the HR result obtained by spectral whitening, and (c) the HR result obtained by our method. As delineated by the yellow rectangles, our method delineates clearer geologic structures than spectral whitening.

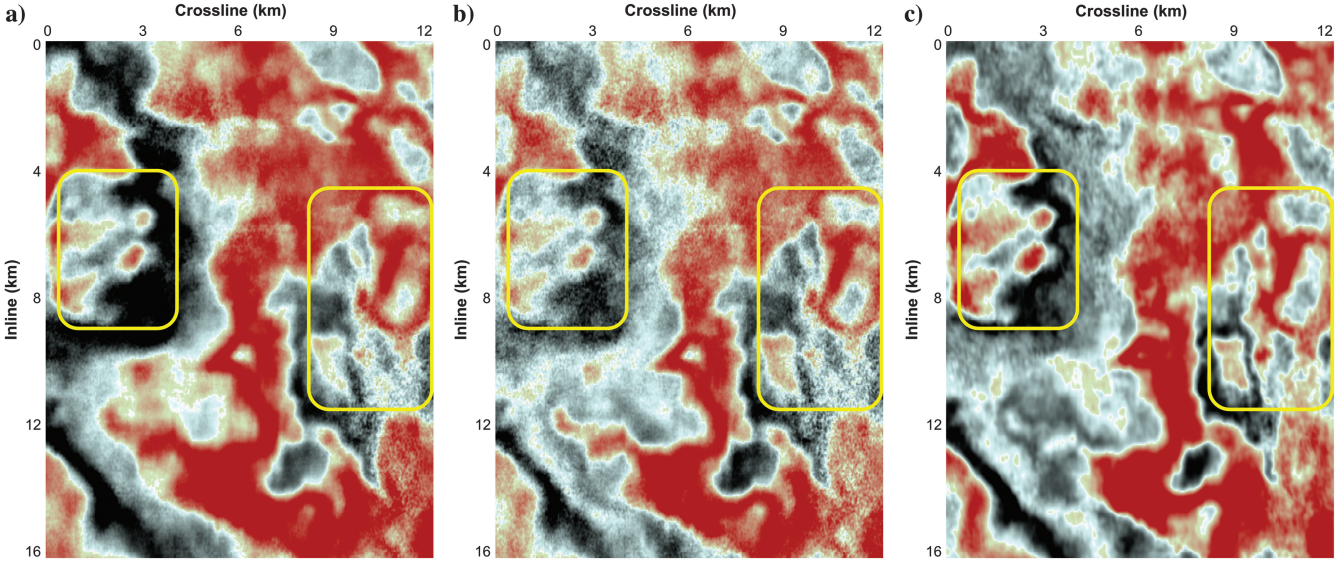


Figure 18. Time-slice results at 2504 ms. (a) Original LR seismic data, (b) the HR result obtained by spectral whitening, and (c) the HR result obtained by the proposed method. As the yellow rectangles indicate, our method uncovers more nuanced geologic structures.

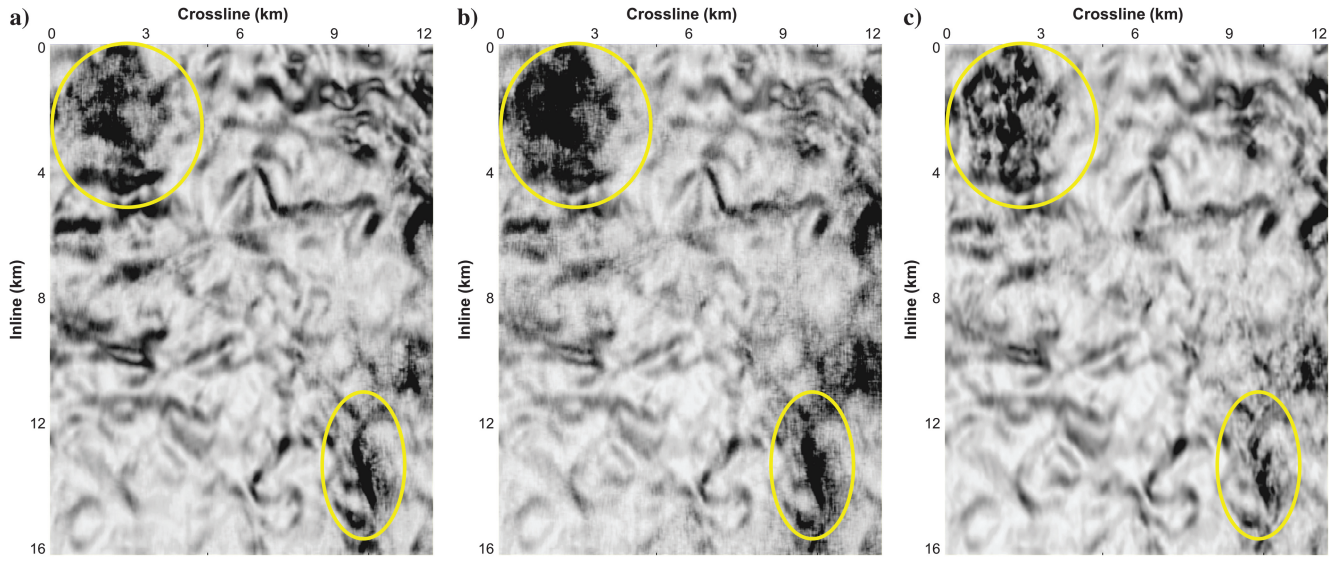


Figure 19. Similarity attribute map at 2284 ms. (a) Original data, (b) after resolution enhancement by spectral whitening, and (c) after resolution enhancement by our network. The yellow circles highlight areas where our method produces a higher S/N and reveals finer spatial variations, demonstrating improved resolution and interpretability.

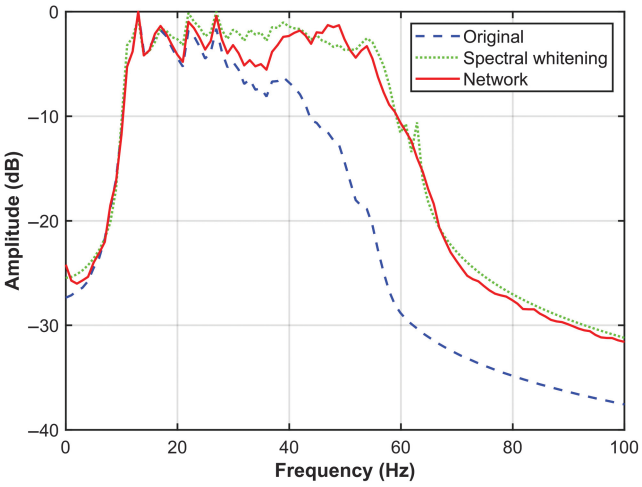


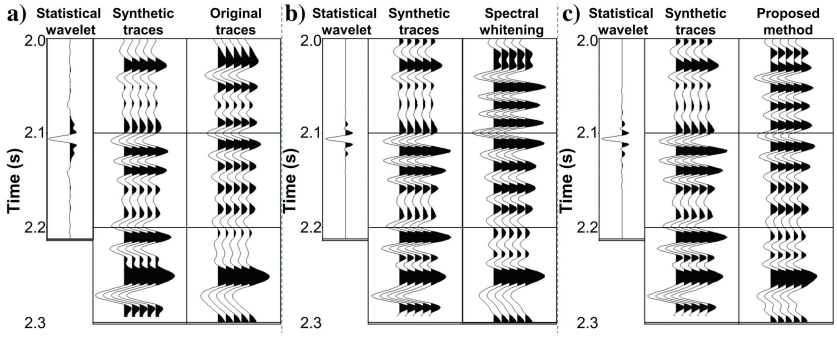
Figure 20. Comparison of field data results on normalized multi-trace amplitude spectrum.

potential of our method in tackling complex scenarios. In other words, the underlying geologic processes that shape shallow and deep subsurface structures share common characteristics, allowing fundamental properties learned from shallow layers to be effectively applied to deeper layers, even in complex structures. In future research, we aim to leverage more advanced deep-learning tools to address such intricate challenges with mismatched structures.

Perspectives on transfer learning for real data applications

Transfer learning offers a promising approach to further improve the performance of our method on real data. However, its practical application requires careful simulation of synthetic data. The success of transfer learning largely depends on the similarity between synthetic and real data; if the synthetic data fails to capture essential characteristics of actual geologic formations or seismic responses, the model may not generalize well to real-world scenarios. Therefore, reducing domain shift and creating synthetic data that accurately reflects the diversity and complexity of real seismic conditions, such as faults, stratigraphic variations, and intrusions, is crucial for effective model transfer.

Figure 21. Seismic-well tie analysis with statistical wavelets from proximal traces. Each panel displays a statistical wavelet, synthetic traces, and the extracted seismic traces near the borehole location from (a) original seismic data, (b) results by spectral whitening, and (c) results by our proposed method.



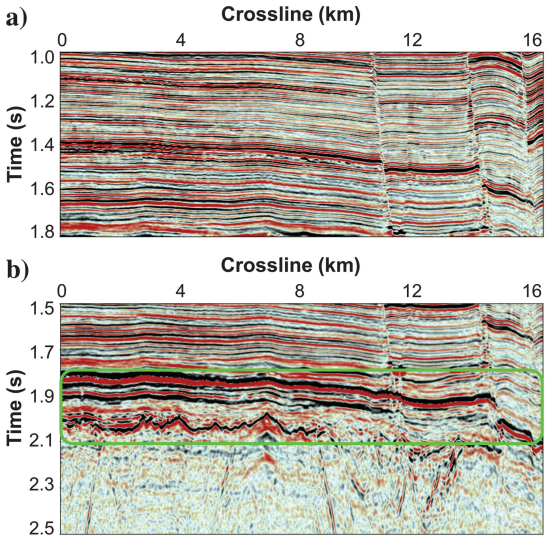


Figure 22. Resolution comparison for inline sections at 8 km in real data with structures. (a) Original shallow HR data and (b) original deep LR data. The green box highlights the target layer.

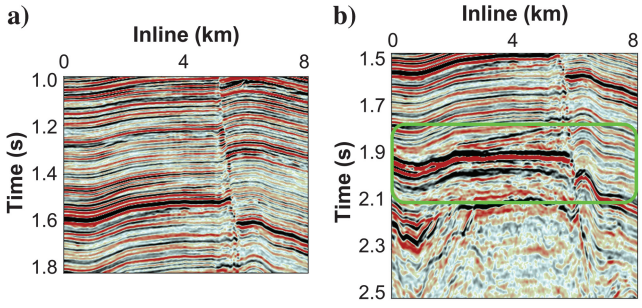


Figure 23. Resolution comparison for crossline sections at 4 km in real data with structures. (a) Original shallow HR data and (b) original deep LR data. The green box highlights the target layer.

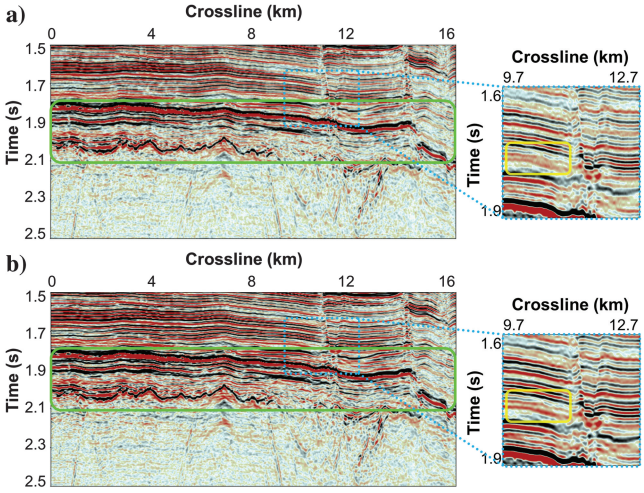


Figure 24. Inline section results at 8 km. (a) The HR result obtained by spectral whitening and (b) the HR result obtained by our method. Our method recovers clearer events than spectral whitening, especially in the target layer highlighted by the green box. The yellow boxes highlight areas where our method achieves enhanced resolution, allowing clearer differentiation between seismic sequences.

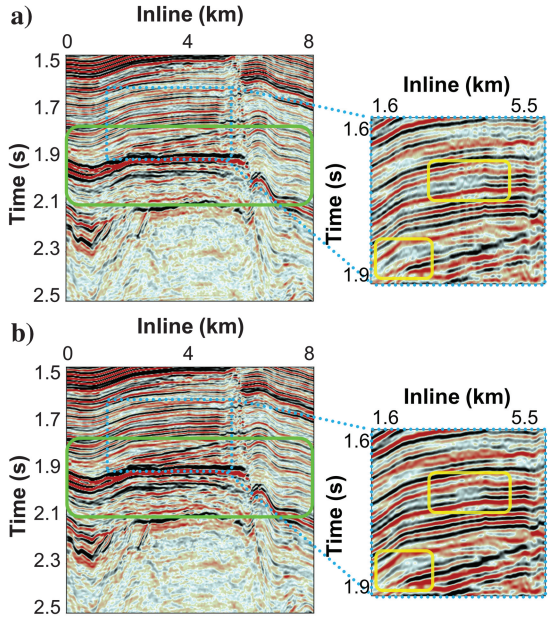


Figure 25. Crossline section results at 4 km. (a) The HR result obtained by spectral whitening and (b) the HR result obtained by our method. Our method improves event continuity and resolution, especially within the target layer highlighted by the green box. The yellow boxes highlight areas where our method achieves better S/N.

In addition, balancing the knowledge from the pretrained model with the adaptation needed for field data can require significant resources for fine-tuning, data augmentation, and model adjustment. These requirements can make transfer learning costly and time intensive, particularly for large-scale seismic data sets with high variability. Recently, pretrained foundation models (Sheng et al., 2023) have shown promise in alleviating the need for extensive initial training, which could be a valuable direction for future research.

Moreover, real seismic data always contain noise, artifacts, and distortions not typically present in synthetic data sets. Models trained on idealized, clean synthetic data may need help with these real-world scenarios. Since replicating identical, realistic noise is challenging, a practical and potentially lightweight strategy might involve fine-tuning the model by artificially deforming shallow, HR regions to match deep-layer data. This approach could help adapt the model to geologic complexities inherent in real data without additional training.

CONCLUSION

This study successfully demonstrates a new approach to enhancing 3D seismic data resolution through weakly supervised learning, circumventing the limitations posed by the absence of paired labels. We leverage a CycleGAN architecture to bridge HR and LR data within the same data set. The network successfully recovers high-frequency information for deeper layer data by exploiting shallow data, surpassing traditional spectral whitening in resolution improvement and noise reduction. These advances are crucial for accurately identifying subsurface structures and aiding oil exploration. The current scope of our research focuses primarily on stratified sedimentary formations. Future initiatives will extend this methodology to more complex geologic structures and explore its application to prestack data to

improve resolution for tasks, such as amplitude variation with offset inversion and prestack seismic migration.

ACKNOWLEDGMENTS

This work was supported in part by the National Key Research and Development Program of China under grant no. 2021YFA0716903, in part by the National Natural Science Foundation of China under grant no. 42374135, and in part by the Fundamental Research Funds for the Central Universities under grant no. xzy012023073. The contribution of the second author is supported by the industrial sponsors of the Signal Analysis and Imaging Group (SAIG) consortium at the University of Alberta.

DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be obtained by contacting the corresponding author.

REFERENCES

Bagheri, M., and M. A. Riahi, 2018, Using a novel method for random noise reduction of seismic records: Iranian Journal of Oil and Gas Science and Technology, **7**, 65–72, doi: 10.22050/ijogst.2018.75178.1381.

Bagheri, M., M. A. Riahi, and H. Hashemi, 2017, Denoising and improving the quality of seismic data using combination of DBM filter and fx deconvolution: Arabian Journal of Geosciences, **10**, 1–8, doi: 10.1007/s12517-017-3224-5.

Bian, G., and L. Zhang, 1986, Spectral whitening of seismic data: Geophysical Prospecting for Petroleum, **2**, 26–33.

Chen, D., J. Gao, Y. Hou, and Z. Gao, 2019, High resolution inversion of seismic wavelet and reflectivity using iterative deep neural networks: 89th Annual International Meeting, SEG, Expanded Abstracts, 2538–2542, doi: 10.1190/segam2019-3216178.1.

Chen, H., J. Gao, X. Jiang, Z. Gao, and W. Zhang, 2021, Optimization-inspired deep learning high-resolution inversion for seismic data: Geophysics, **86**, no. 3, R265–R276, doi: 10.1190/geo2020-0034.1.

Cheng, P., and G. F. Margrave, 2012, Q estimation by a match-filter method: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: 10.1190/segam2012-1257.1.

Choi, Y., S. J. Seol, J. Byun, and Y. Kim, 2019, Vertical resolution enhancement of seismic data with convolutional U-Net: 89th Annual International Meeting, SEG, Expanded Abstracts, 2388–2392, doi: 10.1190/segam2019-3216042.1.

Dong, C., C. C. Loy, K. He, and X. Tang, 2015, Image super-resolution using deep convolutional networks: IEEE Transactions on Pattern Analysis and Machine Intelligence, **38**, 295–307, doi: 10.1109/TPAMI.2015.2439281.

Gholami, A., and M. D. Sacchi, 2013, Fast 3D blind seismic deconvolution via constrained total variation and GCV: SIAM Journal on Imaging Sciences, **6**, 2350–2369, doi: 10.1137/130905009.

Gholamy, A., and V. Kreinovich, 2014, Why Ricker wavelets are successful in processing seismic data: Towards a theoretical explanation: IEEE Symposium on Computational Intelligence for Engineering Solutions, 11–16.

Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, 2020, Generative adversarial networks: Communications of the ACM, **63**, 139–144, doi: 10.1145/3422622.

He, K., X. Zhang, S. Ren, and J. Sun, 2016, Deep residual learning for image recognition: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 770–778.

Hosken, J., 1988, Ricker wavelets in their various guises: First Break, **6**, 24–33, doi: 10.3997/1365-2397.1988002.

Ignatov, A., N. Kobyshev, R. Timofte, K. Vanhoey, and L. Van Gool, 2018, WESPE: Weakly supervised photo enhancer for digital cameras: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 691–700.

Isola, P., J.-Y. Zhu, T. Zhou, and A. A. Efros, 2017, Image-to-image translation with conditional adversarial networks: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1125–1134.

Jiang, Y., X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, and Z. Wang, 2021, EnlightenGAN: Deep light enhancement without paired supervision: IEEE Transactions on Image Processing, **30**, 2340–2349, doi: 10.1109/TIP.2021.3051462.

Kaur, H., N. Pham, and S. Fomel, 2020, Improving the resolution of migrated images by approximating the inverse Hessian using deep learning: Geophysics, **85**, no. 4, WA173–WA183, doi: 10.1190/geo2019-0315.1.

Kazemeini, S. H., C. Yang, C. Juhlin, S. Fomel, and C. Cosma, 2010, Enhancing seismic data resolution using the prestack blueing technique: An example from the Ketzin CO2 injection site, Germany: Geophysics, **75**, no. 6, V101–V110, doi: 10.1190/1.3483900.

Ke, X., Y. Shi, X. Fu, L. Song, H. Jing, J. Yang, and Z. Zhang, 2023, The nth power Fourier spectrum analysis for the generalized seismic wavelets: IEEE Transactions on Geoscience and Remote Sensing, **61**, 1–10, doi: 10.1109/TGRS.2023.3243184.

KhodAgholi, M., and M. BAgheri, 2020, Seismic data random noise attenuation using LLSP smoothing: Bollettino di Geofisica Teorica ed Applicata, **61**, 131–144, doi: 10.4430/bgta0297.

Kim, J., J. K. Lee, and K. M. Lee, 2016a, Accurate image super-resolution using very deep convolutional networks: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1646–1654.

Kim, J., J. K. Lee, and K. M. Lee, 2016b, Deeply-recursive convolutional network for image super-resolution: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1637–1645.

Kim, T., M. Cha, H. Kim, J. K. Lee, and J. Kim, 2017, Learning to discover cross-domain relations with generative adversarial networks: International Conference on Machine Learning, 1857–1865.

Kingma, D. P., and J. Ba, 2014, Adam: A method for stochastic optimization: arXiv preprint, doi: 10.48550/arXiv.1412.6980.

Lai, W.-S., J.-B. Huang, N. Ahuja, and M.-H. Yang, 2017, Deep Laplacian pyramid networks for fast and accurate super-resolution: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 624–632.

Lancaster, S., and D. Whitcombe, 2000, Fast-track 'coloured' inversion: 70th Annual International Meeting, SEG, Expanded Abstracts, 1572–1575, doi: 10.1190/1.1815711.

Ledig, C., L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, 2017, Photo-realistic single image super-resolution using a generative adversarial network: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 4681–4690.

Levy, S., and P. K. Fullagar, 1981, Reconstruction of a sparse spike train from a portion of its spectrum and application to high-resolution deconvolution: Geophysics, **46**, 1235–1243, doi: 10.1190/1.1441261.

Li, C., J. Guo, and C. Guo, 2018, Emerging from water: Underwater image color correction based on weakly supervised color transfer: IEEE Signal Processing Letters, **25**, 323–327, doi: 10.1109/LSP.2018.2792050.

Li, J., X. Wu, and Z. Hu, 2021, Deep learning for simultaneous seismic image super-resolution and denoising: IEEE Transactions on Geoscience and Remote Sensing, **60**, 1–11, doi: 10.1109/TGRS.2021.3057857.

Li, Y., H. Wang, and X. Dong, 2020, The denoising of desert seismic data based on Cycle-GAN with unpaired data training: IEEE Geoscience and Remote Sensing Letters, **18**, 2016–2020, doi: 10.1109/LGRS.2020.3011130.

Liang, J., J. Cao, G. Sun, K. Zhang, L. Van Gool, and R. Timofte, 2021, SwinIR: Image restoration using Swin transformer: Proceedings of the IEEE/CVF International Conference on Computer Vision, 1833–1844.

Liu, D., W. Niu, X. Wang, M. D. Sacchi, W. Chen, and C. Wang, 2023, Improving vertical resolution of vintage seismic data by a weakly supervised method based on cycle generative adversarial network: Geophysics, **88**, no. 6, V445–V458, doi: 10.1190/geo2023-0006.1.

Liu, D., Q. Wang, N. You, M. D. Sacchi, and W. Chen, 2024, Filling the gap: Enhancing borehole imaging with tensor neural network: Geophysics, **90**, no. 3, 1–64, doi: 10.1190/geo2024-0242.1.

Liu, D., X. Wang, X. Yang, H. Mao, M. D. Sacchi, and W. Chen, 2022, Accelerating seismic scattered noise attenuation in offset-vector tile domain: Application of deep learning: Geophysics, **87**, no. 5, V505–V519, doi: 10.1190/geo2021-0654.1.

Maas, A. L., A. Y. Hannun, and A. Y. Ng, 2013, Rectifier nonlinearities improve neural network acoustic models: Proceedings of the 30th International Conference on Machine Learning.

Mao, X., C. Shen, and Y.-B. Yang, 2016, Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections: Advances in Neural Information Processing Systems, 2802–2810.

Merrifield, T. P., D. P. Griffith, S. A. Zamanian, S. Gesbert, S. Sen, J. D. L. T. Guzman, R. D. Potter, and H. Kuehl, 2022, Synthetic seismic data for training deep learning networks: Interpretation, **10**, no. 3, SE31–SE39, doi: 10.1190/INT-2021-0193.1.

Raji, W., and A. Rietbrock, 2013, Attenuation (1/Q) estimation in reflection seismic records: Journal of Geophysics and Engineering, **10**, 045012, doi: 10.1088/1742-2132/10/4/045012.

Sacchi, M. D., 1997, Reweighting strategies in seismic deconvolution: Geophysical Journal International, **129**, 651–656, doi: 10.1111/j.1365-246X.1997.tb04500.x.

Sheng, H., X. Wu, X. Si, J. Li, S. Zhang, and X. Duan, 2023, Seismic foundation model (SFM): A new generation deep learning model in geophysics: arXiv preprint, doi: 10.48550/arXiv.2309.02791.

Sui, Y., and J. Ma, 2020, Blind sparse-spike deconvolution with thin layers and structure: Geophysics, **85**, no. 6, V481–V496, doi: 10.1190/geo2019-0423.1.
Taylor, H. L., S. C. Banks, and J. F. McCoy, 1979, Deconvolution with the L1 norm: Geophysics, **44**, 39–52, doi: 10.1190/1.1440921.
Tong, T., G. Li, X. Liu, and Q. Gao, 2017, Image super-resolution using dense skip connections: Proceedings of the IEEE International Conference on Computer Vision, 4799–4807.
Velis, D. R., 2008, Stochastic sparse-spike deconvolution: Geophysics, **73**, no. 1, R1–R9, doi: 10.1190/1.2790584.
Wang, H., Y. Li, and X. Dong, 2020a, Generative adversarial network for desert seismic data denoising: IEEE Transactions on Geoscience and Remote Sensing, **59**, 7062–7075, doi: 10.1109/TGRS.2020.3030692.
Wang, X., L. Xie, C. Dong, and Y. Shan, 2021, Real-ESRGAN: Training real-world blind super-resolution with pure synthetic data: Proceedings of the IEEE/CVF International Conference on Computer Vision, 1905–1914.
Wang, Y., 2006, Inverse Q-filter for seismic resolution enhancement: Geophysics, **71**, no. 3, V51–V60, doi: 10.1190/1.2192912.
Wang, Z., J. Chen, and S. C. Hoi, 2020b, Deep learning for image super-resolution: A survey: IEEE Transactions on Pattern Analysis and Machine Intelligence, **43**, 3365–3387, doi: 10.1109/TPAMI.2020.2982166.
Wang, Z., E. P. Simoncelli, and A. C. Bovik, 2003, Multiscale structural similarity for image quality assessment: The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, IEEE, 1398–1402.
Wei, Y., Z. Zhang, Y. Wang, M. Xu, Y. Yang, S. Yan, and M. Wang, 2021, DerainCycleGAN: Rain attentive CycleGAN for single image deraining and rainmaking: IEEE Transactions on Image Processing, **30**, 4788–4801, doi: 10.1109/TIP.2021.3074804.
Wiggins, R. A., 1978, Minimum entropy deconvolution: Geoexploration, **16**, 21–35, doi: 10.1016/0016-7142(78)90005-4.
Wu, Y., S. Pan, H. Lan, Y. Chen, J. Badal, and Z. Qin, 2024, Unsupervised-learning stable inverse Q filtering for seismic resolution enhancement: IEEE Transactions on Geoscience and Remote Sensing, **62**, 1–12, doi: 10.1109/TGRS.2024.3458870.
Xue, Y.-J., J.-X. Cao, and X.-J. Wang, 2019, Inverse Q filtering via synchrosqueezed wavelet transform: Geophysics, **84**, no. 2, V121–V132, doi: 10.1190/geo2018-0177.1.
Yi, Z., H. Zhang, P. Tan, and M. Gong, 2017, DualGAN: Unsupervised dual learning for image-to-image translation: Proceedings of the IEEE International Conference on Computer Vision, 2849–2857.
Yu, S., and J. Ma, 2021, Deep learning for geophysics: Current and future trends: Reviews of Geophysics, **59**, e2021RG000742, doi: 10.1029/2021RG000742.
Yuan, Y., S. Liu, J. Zhang, Y. Zhang, C. Dong, and L. Lin, 2018, Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 701–710.
Zhang, H., W. Wang, X. Wang, W. Chen, Y. Zhou, C. Wang, and Z. Zhao, 2019, An implementation of the seismic resolution enhancing network based on GAN: 89th Annual International Meeting, SEG, Expanded Abstracts, 2598–2602, doi: 10.1190/segam2019-3216229.1.
Zhang, W., M. Ravasi, J. Gao, and Y. Shi, 2024, Deep-unrolling architecture for imagedomain least-squares migration: Geophysics, **89**, no. 3, S215–S234, doi: 10.1190/geo2023-0428.1.
Zhang, Y., H. Zhou, Y. Wang, M. Zhang, B. Feng, and W. Wu, 2022, A novel multichannel seismic deconvolution method via structure-oriented regularization: IEEE Transactions on Geoscience and Remote Sensing, **60**, 1–10, doi: 10.1109/TGRS.2022.3141113.
Zhao, H., O. Gallo, I. Frosio, and J. Kautz, 2017, Loss functions for image restoration with neural networks: IEEE Transactions on Computational Imaging, **3**, 47–57, doi: 10.1109/TCI.2016.2644865.
Zhou, Z.-H., 2018, A brief introduction to weakly supervised learning: National Science Review, **5**, 44–53, doi: 10.1093/nsr/nwx106.
Zhu, J.-Y., T. Park, P. Isola, and A. A. Efros, 2017, Unpaired image-to-image translation using cycle-consistent adversarial networks: Proceedings of the IEEE International Conference on Computer Vision, 2223–2232.

Biographies and photographs of the authors are not available.