

Original Paper

High-resolution reconstruction of 3D seismic data using generative adversarial networks with global grouped coordinate attention mechanism

Yun Long, Ji-Yuan Liu, Jian-Bin Zheng^{*}, Xin-Hao Zhang, Shuang Yan, Xiao-Kai Ma

School of Instrumentation and Electrical Engineering, Jilin University, Changchun, 130061, Jilin, China

ARTICLE INFO

Article history:

Received 29 December 2025

Received in revised form

12 March 2026

Accepted 12 May 2026

Available online 19 May 2026

Edited by Jie Hao

Keywords:

Three-dimensional seismic data reconstruction
Generative adversarial network
Global grouped coordinate attention
Natural seismic training set
Noise robustness

ABSTRACT

Due to limitations in exploration costs and surface conditions, actual seismic data are often sparsely sampled and exhibit irregular coverage, which weakens the effects of migration imaging and wavefield inversion, thereby reducing the accuracy and reliability of geological interpretation. Deep learning, particularly generative adversarial networks (GANs), has shown considerable promise in seismic data reconstruction. However, existing GAN-based methods often suffer from limitations such as restricted feature extraction, blurring and artifacts. Additionally, 2D networks struggle to capture the spatio-temporal correlations of 3D seismic data, leading to inter-layer inconsistencies and the loss of critical tectonic features, which ultimately limits overall reconstruction quality. To address complex missing-data patterns in seismic datasets, this study proposes a 3D generative adversarial network integrating a Global Grouped Coordinate Attention (GGCA-GAN). The proposed network employs group-based feature modeling and, by jointly training on natural images and real seismic data, achieves high-resolution 3D seismic reconstruction with moderate robustness to noise. This method further reduces computational overhead by enhancing global spatial feature extraction in both the generator and discriminator via grouped processing. In addition, to address the scarcity of training data under limited observation conditions, natural image datasets are incorporated into the network training process. Their structural features are used to construct cross-domain priors, which, when combined with measured seismic data, enhance the network's ability to generalize to complex seismic signals. The effectiveness of the proposed method is validated on publicly available datasets.

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

1. Introduction

Seismic wavefield data serve as the foundation for seismic inversion, offset imaging and interpretation of reservoir parameters. However, due to equipment and cost limitations, seismic data are often missing, which compromises the quality of data processing and the accuracy of inversion of the subsurface medium parameters, and reduces the reliability of geological formation interpretation (Abedi and Pardo, 2022; Goyes-Peñañiel et al., 2024; Kong et al., 2022). Therefore, high-resolution reconstruction of missing seismic data has become a critical task for improving

seismic data processing quality (Chen et al., 2019; Dou et al., 2023; Wang et al., 2024).

Current research on reconstruction methods for seismic data can be broadly classified into two main categories: traditional methods and deep learning-based methods, based on their implementation mechanisms. Traditional seismic data reconstruction algorithms primarily rely on the seismic wave propagation mechanism and digital signal processing techniques (Spitz, 1991; Ronen, 1987), which typically require the pre-setting of several physical constraints, including, but not limited to, seismic in-phase axial linear kinematics, the assumption of spatial

^{*} Corresponding author.

E-mail address: jbzheng22@mails.jlu.edu.cn (J.-B. Zheng).

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

continuity of the local wavefield and the sparsity characterization in the transform domain (Fomel, 2003; Porsani, 1999; Sacchi et al., 1998; Wang et al., 2010). These constraints are then applied to reconstruct the data through numerical methods. However, significant limitations exist in the traditional methodology: First, complex geological structures in actual exploration environments often lead to nonlinear motion characteristics in the seismic wavefield, resulting in substantial deviations between theoretical assumptions and actual characteristics of the data (Latif and Mousa, 2017; Liu et al., 2016). Second, the algorithm requires determining the optimal solution space through multi-parameter experiments, which considerably reduces computational efficiency (Huang et al., 2016; Innocent Oboué et al., 2021). Finally, the traditional algorithm depends on an iterative optimization strategy to achieve numerical convergence and its computational complexity increases exponentially when applied to large-scale seismic datasets (Abedi and Pardo, 2022; Saad et al., 2022a).

Advances in artificial intelligence technology have accelerated the rapid development of deep learning methods, which have become an important paradigm in modern machine learning, with their powerful nonlinear mapping and feature self-learning capabilities offering a new paradigm for technological innovation in geophysical exploration (Jiang et al., 2023; Mousavi et al., 2020; Wang et al., 2022). Compared to traditional seismic data reconstruction algorithms, deep learning methods offer four key advantages: First, their end-to-end learning mechanism eliminates the need for a priori physical assumptions, automatically extracting hidden features of the seismic wavefield through a data-driven approach, significantly improving the algorithm's adaptability to complex geological structures (Dou et al., 2023; Wang et al., 2020). Second, the fixed characteristics of the network parameters within a supervised learning framework effectively prevent model bias introduced by manual intervention, ensuring robustness and repeatability (Fang et al., 2021; Wang et al., 2018, 2020). Third, computational resource consumption is primarily concentrated in the model training phase, which substantially increases computational efficiency once training is complete. Fourth, recent advancements in unsupervised frameworks have enabled deep learning to handle complex, non-uniformly sampled field data without the need for labeled training sets. For instance, by integrating projection onto convex sets (POCS) with deep image prior (DIP) or attention-based architectures, deep learning can achieve high-fidelity reconstruction and denoising directly on off-the-grid (OTG) seismic data (Li et al., 2025b). This paradigm not only preserves weak signals better than conventional methods but also avoids the amplitude and phase distortions typically introduced by the binning process in tasks like simultaneous deblending and reconstruction (Li et al., 2025a). Therefore, incorporating deep learning algorithms into seismic data reconstruction will significantly enhance both the accuracy and efficiency of the process.

To address the issue of missing data and complex interference in seismic data acquisition, various model algorithms have been proposed for adaptive optimization in seismic waveform feature reconstruction (Saad et al., 2022b; Yu and Wu, 2022). Generative adversarial network (GAN), as a deep learning method, is widely applied in seismic data reconstruction (Chang et al., 2021; Creswell et al., 2018; Wei et al., 2021a, 2021b). Compared to traditional probabilistic generative models, GAN offers a significant advantage by eliminating the need for the Markov chain learning mechanism and performing gradient optimization solely through backpropagation. It also provides a flexible framework capable of integrating various loss functions, delivering robust algorithmic support for constructing unsupervised natural data generation networks. Mirza and Osindero pioneered conditional

generative adversarial networks, enhancing the expressive power of GANs by introducing generative conditions that guided the network to produce specific types of data (Mirza and Osindero, 2014). In contrast, Radford et al. combined deep convolutional GANs with convolutional neural networks (CNNs), improving the realism and naturalness of the generated data (Radford et al., 2015). To drive innovation, Pathak et al. introduced context encoder networks and used noise coding as deep convolutional GAN inputs for image restoration tasks with remarkable results (Pathak et al., 2016). Similarly, Isola et al. demonstrated the versatility of the GAN architecture in image processing and enhancement by utilizing deep convolutional GAN with conditional GAN for image processing (Isola et al., 2017).

Attention mechanisms have become a fundamental component of modern deep learning architectures, particularly in the field of computer vision. Although the traditional self-attention mechanism effectively captures dependencies between global features, its high computational cost and limited positional encoding capability hinder spatial localization. Meanwhile, channel attention is applied only in the channel dimension, neglecting spatial structure information (Li et al., 2024; Woo et al., 2018). In recent years, several variants of the attention mechanism have been proposed to address specific challenges in various scenarios. Global Grouped Coordinate Attention (GGCA) is a variant of the attention mechanism that builds upon the Coordinate Attention (CA) structure by introducing the concepts of "Grouped Modeling" and "Global Interaction." Its core idea enhances the model's ability to capture global dependencies while preserving spatial localization and enabling global interaction. GGCA is more computationally efficient, reducing the computational cost of large-scale data processing and enabling more efficient handling of complex tasks (Pathak et al., 2016; Xiao et al., 2025).

Moreover, beyond optimizing the model structure, the limited availability of seismic datasets is a crucial factor influencing the performance of deep learning methods. Although synthetic datasets can offer some training support, they are often constructed in a manner that fails to fully capture the complexity of real seismic records, limiting the network's generalization capabilities. To address the issue of limited seismic data available for practical exploration, Wu et al. proposed using "natural seismic data" as training labels for deep learning interpolation (Wu et al., 2023). Public natural image datasets offer valuable structural information owing to their large size, diverse categories, and rich structure. Following Liner and Monk, a training set of "natural seismic data" was created by converting natural images into seismic-like structural patterns using a frequency-wave number (f - k) domain filter (Liner, 1997; Monk, 2002). Compared to traditional synthetic seismic data, this type of data exhibits greater structural diversity, which improves the generalization performance and robustness of the network, thereby alleviating the constraints imposed by limited seismic data samples on model training.

In this paper, we propose a GAN model, the Global Grouped Coordinate Attention Generative Adversarial Network (GGCA-GAN), which incorporates the global grouped coordinate attention mechanism for high-resolution seismic data reconstruction. The method effectively captures long-range spatial correlations by expanding the receptive field and integrating multi-scale features through the global grouped attention module. Grouped feature processing reduces computational complexity while preserving feature diversity, thereby improving reconstruction efficiency. Noise was introduced to the missing data to verify the network's noise immunity. Additionally, natural images were used as the training set to enhance the network's information extraction and generalization, alleviating the limitations imposed by the scarcity of training samples.

2. Methodology

2.1. Seismic interpolation problem

Due to constraints in the acquisition environment and equipment, as well as the removal of waste guns and channels during processing, the observed seismic data are often missing or contaminated by random noise. These data suffer from spatial sampling inhomogeneity and exhibit features such as energy discontinuity, event fragmentation, and local structural ambiguity, increasing the uncertainty in subsequent offset imaging, velocity modeling, and petrophysical inversion. The noise-containing irregular seismic data can be expressed as follows:

$$y = M \odot (s + n) \quad (1)$$

where y represents the irregular seismic data contaminated by random noise, M is the sampling operator, having the same dimensions as the complete seismic data, with a value of 0 at missing data locations and 1 at sampled locations, $\odot$ denotes the element-wise multiplication operator, s represents the valid seismic signal, and n denotes the random noise.

Seismic data denoising and reconstruction can be formulated as an inverse problem. A deep neural network can establish a nonlinear mapping between noise-contaminated missing observations and clean, complete data through supervised or unsupervised learning strategies, enabling effective recovery of the complete signal. The optimization objective function for the supervised learning network is expressed as:

$$\min_{\theta} \|f(y; \theta) - s\|_2^2 \quad (2)$$

where f denotes the network architecture used in training, θ is the network parameter to be optimized, y denotes the noisy and incomplete input data used for training, and s represents the desired clean and complete output data, which is the target signal that the network aims to predict and is usually obtained from real data, serving as labeled data for supervised learning.

Existing seismic interpolation methods based on 2D slices often perform poorly in complex scenarios, such as large-scale or continuous missing data, and their slice-by-slice reconstruction results are prone to incoherent events in 3D stacked profiles. To address these issues, this study shifts the interpolation and reconstruction process of seismic data from the 2D to the 3D case, enabling more effective handling of large-scale missing data in complex scenarios.

2.2. Generative adversarial networks (GANs)

GANs are widely used in computer vision (CV) for image restoration, learning deep generative models through adversarial training. As a novel approach to training generative models, the GAN framework consists of two “adversarial” models: the generator (G), which captures the data distribution and the discriminator (D), which estimates the probability that the samples come from the training data. Both G and D are nonlinear mapping functions.

The core idea of GAN originates from the zero-sum game in game theory, where G and D optimize each other through adversarial training, ultimately converging to a Nash equilibrium and yielding improved model parameters. The objective function of GAN is defined as:

$$\min_{\mathcal{G}_{\theta_G}} \max_{\mathcal{D}_{\theta_D}} V(\mathcal{D}_{\theta_D}, \mathcal{G}_{\theta_G}) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log \mathcal{D}_{\theta_D}(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - \mathcal{D}_{\theta_D}(\mathcal{G}_{\theta_G}(z)))] \quad (3)$$

where $V(\mathcal{D}_{\theta_D}, \mathcal{G}_{\theta_G})$ is the value function with respect to $\mathcal{G}_{\theta_G}$ and $\mathcal{D}_{\theta_D}$, x is the input real data, z is the random variable following a Gaussian distribution, $p_z(z)$ is the probability distribution of z , $p_{\text{data}}(x)$ is the probability distribution of x , $z \sim p_z(z)$ indicate that the random variable subject to $p_z(z)$ sampling, $x \sim p_{\text{data}}(x)$ indicate that the real data is subject to $p_{\text{data}}(x)$ sampling, $\mathcal{D}_{\theta_D}(x)$ and $\mathcal{D}_{\theta_D}(\mathcal{G}_{\theta_G}(z))$ are the probabilities that x and $\mathcal{G}_{\theta_G}(z)$ are judged to be real data, respectively, $\mathcal{G}_{\theta_G}(z)$ is the data generated by z after the generator $\mathcal{G}_{\theta_G}$, and $\mathbb{E}_z(\cdot)$ is the expected value of “ $\cdot$ ” under z -sampling. The objective function aims to maximize the likelihood that the discriminator makes correct decisions, while training the generator to minimize $\log(1 - \mathcal{D}_{\theta_D}(\mathcal{G}_{\theta_G}(z)))$.

In the GAN framework, the local receptive field in the convolutional layer limits the model's ability to capture long-range dependencies. Standard convolution struggles to effectively integrate global information across the width and height dimensions in high-dimensional data, limiting the model's ability to capture features in complex visual tasks.

2.3. Global grouped coordinate attention mechanism (GGCA)

Coordinate attention decomposes spatial attention into directional representations along coordinate axes, enhancing spatial sensitivity while maintaining a lightweight design. GGCA partitions features into multiple channel groups, applies coordinate attention independently within each group, and achieves multi-scale responses through global fusion. This design reduces computational complexity, efficiently couples spatial and channel dimensions, and ensures both orientation awareness and global consistency, making it well suited for structural recognition in texture-dense data such as seismic images.

Let the input feature map be $X \in \mathbb{R}^{B \times C \times H \times W \times D}$, where B, C, H, W , and D denote the batch size, number of channels, height, width, and temporal depth, respectively. The channel dimension is divided into G groups, each containing C/G channels, i.e.,

$$X = [X_1, X_2, \dots, X_G], X_g \in \mathbb{R}^{B \times (C/G) \times H \times W \times D} \quad (4)$$

For each subgroup X_g , global average pooling is performed along the height, width and temporal depth dimensions to extract directional descriptors:

$$\begin{aligned} F_g^h(h) &= \frac{1}{WD} \sum_{w=1}^W \sum_{d=1}^D X_g(:, :, h, w, d) \\ F_g^w(w) &= \frac{1}{HD} \sum_{h=1}^H \sum_{d=1}^D X_g(:, :, h, w, d) \\ F_g^d(d) &= \frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W X_g(:, :, h, w, d) \end{aligned} \quad (5)$$

where $F_g^h \in \mathbb{R}^{B \times (C/G) \times H \times 1 \times 1}$, $F_g^w \in \mathbb{R}^{B \times (C/G) \times 1 \times W \times 1}$ and $F_g^d \in \mathbb{R}^{B \times (C/G) \times 1 \times 1 \times D}$ denote the global responses along the height, width and temporal depth dimensions, respectively. Then, the features from the three dimensions are concatenated and fed into a shared, bottleneck convolutional layer:

$$f_g = \delta(\text{Conv}_{1 \times 1 \times 1}([F_g^h, F_g^w, F_g^d])) \quad (6)$$

where δ denotes a nonlinear activation function. Next, three independent $1 \times 1 \times 1$ convolutions are used to generate direction-dependent weight mappings:

$$\begin{aligned}
A_g^h &= \sigma(\text{Conv}_{1 \times 1 \times 1}^{(h)}(f_g)) \\
A_g^w &= \sigma(\text{Conv}_{1 \times 1 \times 1}^{(w)}(f_g)) \\
A_g^d &= \sigma(\text{Conv}_{1 \times 1 \times 1}^{(d)}(f_g))
\end{aligned} \quad (7)$$

where σ denotes the sigmoid activation function, which normalizes the weights. Finally, the in-group, position-aware output is obtained:

$$Y_g = X_g \times A_g^h \times A_g^w \times A_g^d \quad (8)$$

This process applies directional weighting to spatial locations and channel features, enabling the model to recognize target locations and structural contours. At the intergroup level, to prevent feature fragmentation due to group processing, GGCA introduces a global fusion operation:

$$Y = W_o[Y_1, Y_2, \dots, Y_G] \quad (9)$$

where W_o denotes a linear transformation or lightweight convolution operation for global integration and semantic propagation across feature groups, ensuring globally consistent feature representation.

GGCA integrates grouped feature modeling with coordinate attention to enhance global spatial awareness while improving computational efficiency. By partitioning feature channels and establishing coordinate dependencies, local and global feature fusion is achieved to capture tectonic boundaries and non-homogeneous features in seismic images, thereby improving the recognition performance of complex geological bodies. With a small number of parameters, high feature representation and strong spatial dependency capture, the method offers an efficient extension for GAN feature modeling and information aggregation.

2.4. Generative adversarial network architecture with global grouped coordinate attention mechanism (GGCA-GAN)

This study proposes the GGCA-GAN for high-resolution 3D seismic reconstruction. By integrating the GGCA module into the GAN, thereby improving the network's ability to capture global information and enhance feature representations from 3D seismic data. The GGCA module captures global context along three dimensions—height, width and temporal depth—and reweights channel features to produce direction-specific attention maps, alleviating the limitations of a single attention mechanism in modeling cross-dimensional dependencies. In addition, grouped modeling combined with cross-group fusion reduces computational complexity and runtime, while improving both reconstruction accuracy and efficiency. The GGCA-GAN architecture is shown in Fig. 1.

2.4.1. Generator model

This study develops a generator architecture that fuses residual convolutional units with GGCA. Through multi-scale feature extraction and cross-dimensional information interaction, the generator jointly refines local details while preserving 3D structural consistency. The generator architecture of GGCA-GAN is shown in Fig. 2.

As shown in the figure, the 3D seismic data are first processed by a 3D convolutional layer with a kernel size of $9 \times 9 \times 9$, followed by a LeakyReLU activation. The resulting feature maps are then fed into five 3D convolutional Blocks (Conv H-Blocks) with residual connections. Each Conv H-block contains two 3D convolutional layers with kernel size $3 \times 3 \times 3$, each followed by batch normalization, while a PReLU activation is inserted between the

two convolution layers. In addition, a GGCA module is inserted at the end of each convolutional block in the generator to enhance global context modeling in 3D space. A skip connection is applied to perform residual learning. The GGCA module is inserted after the residual feature extraction stage to enhance global contextual modeling across the three spatial dimensions. After the five Conv H-Blocks, the output of an additional convolutional layer followed by BN is element-wise summed with the earlier feature maps via skip connections. Subsequently, convolutional layers are used to increase the number of feature maps from 64 to 256. The final convolutional layer reconstructs the high-resolution seismic volume without additional activation to preserve amplitude fidelity. This design alleviates the vanishing gradient problem and enhances the network's ability to model complex seismic structures.

2.4.2. Discriminator model

In the adversarial learning framework, the discriminator determines whether an input sample is real or generated, thereby providing training signals for the generator. To improve adversarial training stability and increase sensitivity to seismic structural differences, we design a lightweight discriminator based on multi-layer convolutional feature extraction and global coordinate attention enhancement. The discriminator architecture is shown in Fig. 3.

The discriminator distinguishes real seismic volumes from generated samples. As shown in the figure, the input volume is first processed by an initial 3D convolutional layer followed by a LeakyReLU activation, and then passed through seven Conv L-Blocks for hierarchical feature extraction. Each Conv L-Block consists of one 3D convolutional layer with kernel size $3 \times 3 \times 3$, each followed by batch normalization and a LeakyReLU activation function with a negative slope of 0.2. To enhance the modeling of long-range dependencies and structural consistency, the GGCA module is inserted at the end of each convolutional block in the discriminator, thereby improving the discriminator's ability to distinguish tectonic structures. The number of feature maps gradually increases from 64 to 512 in a hierarchical manner ($64 \rightarrow 128 \rightarrow 256 \rightarrow 512$), enabling the network to capture increasingly abstract structural representations. Finally, the extracted features are flattened and passed through a fully connected layer with 1024 neurons followed by a LeakyReLU activation, then through a second fully connected layer with 1 neuron followed by a sigmoid activation function to produce the discriminator output. Furthermore, spectral normalization is applied to all convolutional layers to constrain the singular values of the weight matrices, stabilizing adversarial training and mitigating overfitting. After multi-layer feature extraction and global enhancement, the discriminator outputs a scalar in $[0, 1]$, representing the probability that the input is real.

2.4.3. Loss function design

This study addresses high-resolution 3D seismic data reconstruction. We construct a multi-objective composite loss that integrates content, adversarial and perceptual terms. A weighted combination strategy coordinates these objectives during training, enhancing high-frequency detail and texture representation while preserving overall structural consistency. The objective function for training the generator is defined as a multi-objective composite loss:

$$\mathcal{L}_G = \lambda_1 \mathcal{L}_{\text{Content}} + \lambda_2 \mathcal{L}_{\text{Adv}} + \lambda_3 \mathcal{L}_{\text{Perc}} \quad (10)$$

where $\mathcal{L}_{\text{Content}}$ denotes the content loss, $\mathcal{L}_{\text{Adv}}$ denotes the adversarial loss, $\mathcal{L}_{\text{Perc}}$ denotes the perceptual loss and $\lambda_1, \lambda_2, \lambda_3$ are the

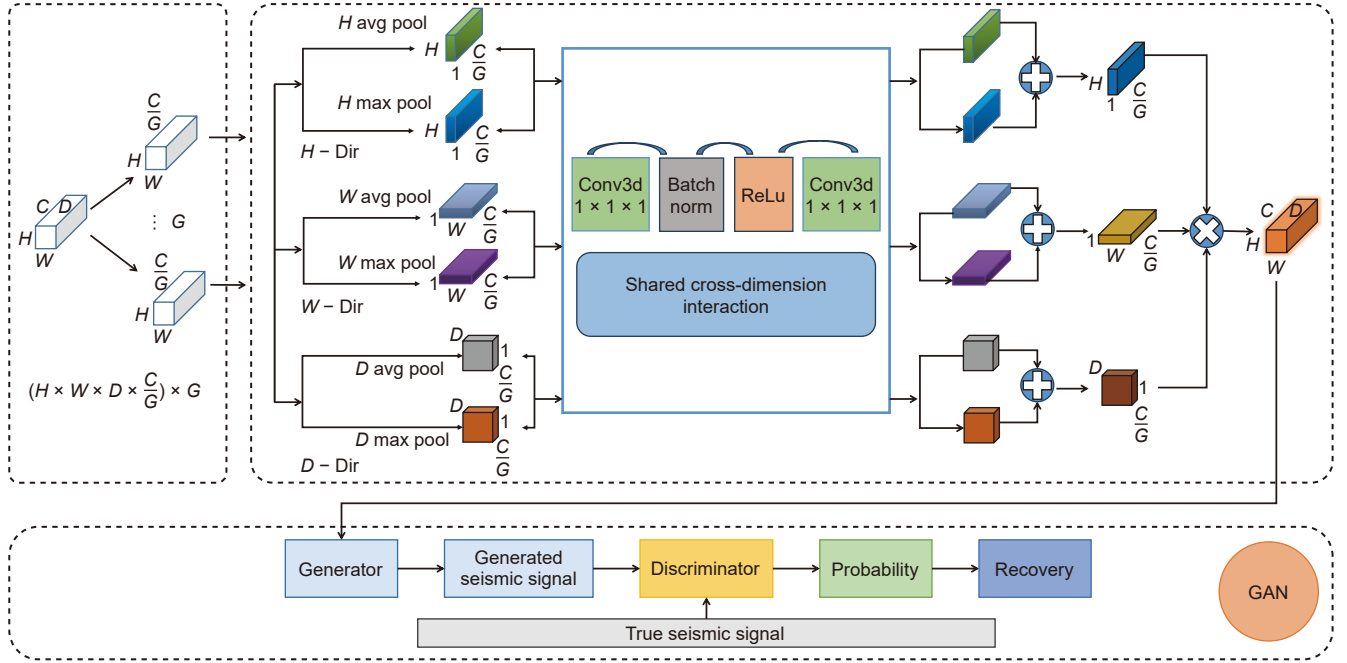


Fig. 1. Generative adversarial network architecture with global grouped coordinate attention mechanism.

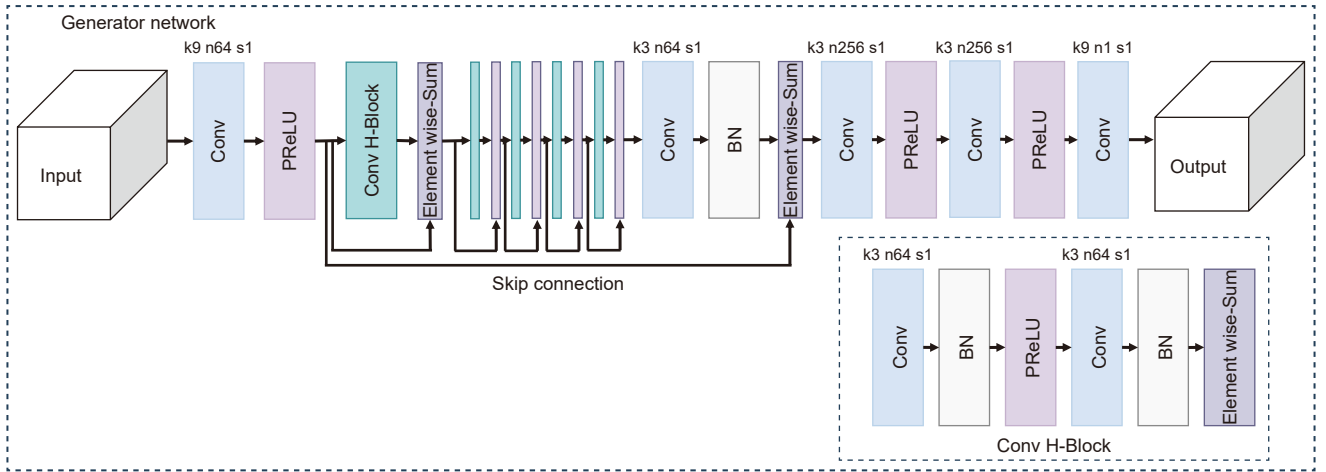


Fig. 2. Schematic diagram of the generator network structure. The generator (G) consists of five Conv H-Blocks and each convolutional layer is specified by the kernel size (k), number of feature maps (n) and stride (s).

corresponding weights to adjust the loss terms to the same order of magnitude. The adversarial loss is expressed as:

$$\mathcal{L}_{Adv} = \log(1 - D_{\theta_D}(G_{\theta_G}(I_{LR}))) \quad (11)$$

Most of the current reconstruction losses are based on L_1 or L_2 norm. Compared with the L_2 norm, the L_1 loss offers improved gradient stability, noise robustness and reduced sensitivity to outliers. Let the low-resolution input be I_{LR} , the ground-truth high-resolution volume be I_{HR} and the generator output be $\hat{I} = G_{\theta_G}(I_{LR})$. When the spatial dimensions and temporal depth are upsampled by a factor s (i.e., the output size is $sH \times sW \times sD$), the content loss is defined as:

$$\begin{aligned} \mathcal{L}_{Cont} &= \mathcal{L}_{L_1} \\ &= \frac{1}{sH sW sD} \sum_{i=1}^{sH} \sum_{j=1}^{sW} \sum_{k=1}^{sD} |I_{HR}(i, j, k) - \hat{I}(i, j, k)| \end{aligned} \quad (12)$$

To align with the 3D reconstruction task, the discriminator $D(\cdot)$ takes volume data as input. The discriminator loss and the generator adversarial loss are defined as follows:

$$\mathcal{L}_D = \log(1 - D_{\theta_D}(I_{HR})) + \log(D_{\theta_D}(G_{\theta_G}(I_{LR}))) \quad (13)$$

The parameters of G and D are updated iteratively. The objective of $D(\cdot)$ is to distinguish between real data I_{HR} and generated

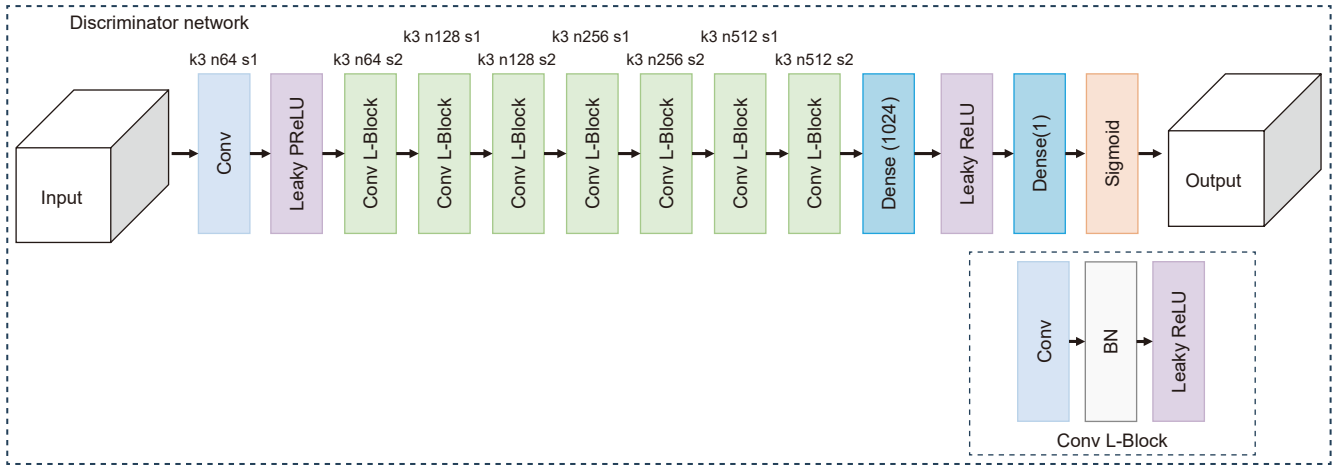


Fig. 3. Schematic diagram of the discriminator network architecture. The discriminator (D) consists of seven Conv L-Blocks and each convolutional layer is specified by the kernel size (k), number of feature maps (n) and stride (s).

data $\hat{I} = G_{\theta_G}(I_{LR})$, while the objective of $G(\cdot)$ is to maximize the probability that the generated samples are classified as real.

Perceptual loss measures the discrepancy between the generated result and the ground-truth data in a high-level semantic feature space. Since VGG-16 is a 2D network, we adopt a slice-wise feature alignment strategy along the temporal depth D . Specifically, the data volume is represented as $I^{(k)}$ at the k -th temporal slice. The ground-truth slice $I_{HR}^{(k)}$ and the generated slice $\hat{I}^{(k)}$ are fed into a pretrained VGG-16, the parameters of the pretrained VGG-16 network remain fixed during training, features are extracted from selected layers and compared, then the loss is averaged over all reconstructed slices. The perceptual loss is defined as:

$$\mathcal{L}_{\text{Perc}} = \frac{1}{sD} \sum_{k=1}^{sD} \sum_{l=1}^L \frac{1}{H_l W_l C_l} \left\| \phi_l(I_{HR}^{(k)}) - \phi_l(\hat{I}^{(k)}) \right\|_1 \quad (14)$$

where $\phi_l(\cdot)$ denotes the feature map extracted from the l -th selected layer of the pretrained VGG-16 network and L represents the total number of selected feature maps. H_l , W_l and C_l denote the height, width and number of channels of the corresponding feature map, respectively. sD denotes the reconstructed depth dimension after upsampling.

2.5. Training enhancement of natural image priors

In 3D seismic data reconstruction tasks, the actual acquired seismic data are often affected by factors such as limited observation conditions, incomplete sampling and noise interference. The number of high-quality 3D seismic data samples used as training data is limited and the statistical diversity in seismic lines, profiles and time directions is insufficient, making it challenging to adequately train deep networks. Natural image (NI) datasets are typically large in scale and rich in categories, providing additional statistical prior information to the model, thereby helping to enhance the network's generalization ability in structural reconstruction and detail representation. Therefore, this study introduces natural image data as auxiliary training data in the training process of GGCA-GAN.

Because natural images differ significantly from seismic reflection characteristics in spectral structure, they are difficult to directly use for training seismic data reconstruction models. Therefore, this study preprocesses natural images in the f - k domain to make their statistical properties more closely resemble seismic data at the spectral level. First, the natural image (data volume) is transformed into the f - k domain. The low-frequency energy of natural images primarily reflects large-scale structural information, while high-frequency components correspond to local detail features, which closely resemble the spectral characteristics of seismic data to some extent. Second, a bandpass filter is applied to the f - k spectrum to suppress spectral components unrelated to seismic reflection structure and retain effective frequency bands that better match the frequency distribution of seismic signals, thereby achieving spectral constraint and structural statistical matching of the natural image. In this study, the frequency band of the filter is selected according to the typical frequency range of seismic reflection data, which is widely used in conventional seismic processing workflows. The mathematical expression of the bandpass filter is as follows:

$$Q(f, k) = \begin{cases} 1, & \text{if } \left| \frac{f}{k} \right| \geq v \\ 0, & \text{otherwise} \end{cases} \quad (15)$$

where v is the velocity threshold, determined by the minimum apparent velocity of the effective seismic wave, thereby controlling the spectral range of the f - k filter. In the dataset generation process, v is empirically set within the range of 800 to 3500 m/s to mimic the apparent velocity distribution of effective seismic events. The filtered spectrum is then inversely transformed back to the time domain to obtain NI samples with seismic pattern structural characteristics.

To adapt to the training requirements of 3D networks and expand the training sample size, a set of preprocessed 2D natural images based on f - k domain statistical alignment preprocessing is obtained. These 2D slices are then stacked along the third dimension in a fixed stacking order to construct a NI data volume.

This data volume serves as a 3D extension of structural statistical priors, enhancing data diversity and spatial structure representation capabilities during the training phase. The process of constructing the training dataset is shown in Fig. 4.

Finally, the constructed pseudo-seismic volumes, along with real 3D seismic samples, is used for network training to improve the model's robust reconstruction performance for 3D structural continuity, local details and complex reflection morphologies and to mitigate the risk of overfitting due to insufficient training samples.

3. Experimental program

3.1. Assessment of indicators

To quantitatively assess the reconstruction quality, three commonly used image-quality metrics—Peak Signal-to-Noise Ratio (PSNR), Signal-to-Noise Ratio (SNR) and Structural Similarity Index Metric (SSIM)—were employed in the experiments.

PSNR quantifies the discrepancy between the reconstructed image and the ground-truth image and is widely used to evaluate image super-resolution performance. PSNR can be calculated as follows:

$$\text{PSNR} = 10 \left(\log_{10} \left(\frac{\text{MAX}_I^2}{\text{MSE}} \right) \right) \quad (16)$$

where MAX_I denotes the maximum pixel value of the high-resolution, noise-free reference seismic image and MSE is the mean squared error between the image reconstructed by the generator and the reference image.

SNR quantifies the relative strength of the useful information in a signal (e.g. an image, sound, or other data) with respect to the background noise. In image processing, SNR characterizes the ratio between the true image content (signal) and random errors or perturbations (noise). In image processing, SNR can be calculated as follows:

$$\text{SNR} = 10 \log_{10} \left(\frac{\|I\|_2^2}{\|I - \hat{I}\|_2^2} \right) \quad (17)$$

where I denotes the pixel value of the original image and $\hat{I}$ denotes the pixel value of the reconstructed image.

SSIM is also employed to evaluate the similarity between the reconstructed image and the ground-truth image in terms of brightness, contrast and structure. SSIM can be defined as follows:

$$\begin{aligned} L(X, Y) &= \frac{2u_X u_Y + C_1}{u_X^2 + u_Y^2 + C_1} \\ C(X, Y) &= \frac{2\sigma_X \sigma_Y + C_2}{\sigma_X^2 + \sigma_Y^2 + C_2} \\ S(X, Y) &= \frac{\sigma_{XY} + C_3}{\sigma_X \sigma_Y + C_3} \end{aligned} \quad (18)$$

$$\text{SSIM}(X, Y) = L(X, Y) \times C(X, Y) \times S(X, Y)$$

where X and Y denote the reconstructed image and the corresponding ground-truth image, respectively, u_X and u_Y denote their mean values, σ_X and σ_Y denote their standard deviations and σ_{XY} denotes the covariance between X and Y . Here, C_1 , C_2 and C_3 are small positive constants introduced to prevent division by zero. In the experiments, we set $C_1 = 10^{-4}$, $C_2 = 9 \times 10^{-4}$, $C_3 = C_2/2$. $L(X, Y)$, $C(X, Y)$ and $S(X, Y)$ denote the luminance, contrast and structural similarity between X and Y , respectively.

3.2. Training set

The training dataset comprises both field seismic data and pseudo-seismic data generated from natural images through f - k filtering. The field datasets include the Kerry-3D dataset and the F3-3D dataset, which inherently contain realistic acquisition and processing noise.

To enrich structural variability and alleviate the limitation of labeled seismic samples, natural images are introduced as supplementary training data. Compared with field seismic data, natural images do not contain acquisition-related noise and typically exhibit continuous structural patterns with clear energy distributions, characteristics that resemble simulated seismic responses. To adapt them to the seismic domain, the proposed NI transformation strategy is applied to transform natural images into pseudo-seismic sections. These sections are subsequently stacked to construct seismic volumes, which serve as simulated training samples.

All seismic volumes, including both field and pseudo-seismic datasets, are partitioned into overlapping 3D patches of size $128 \times 128 \times 128$ using a sliding-window strategy. Among the

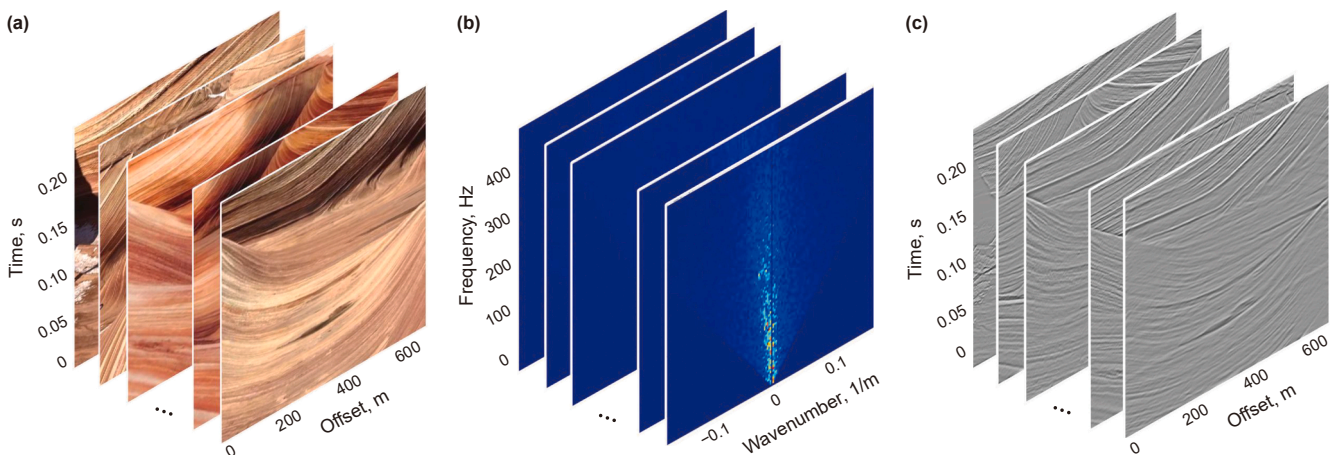


Fig. 4. Workflow for converting 2D natural images into natural-image-derived volumes via f - k spectral alignment. (a) Original natural images, (b) f - k spectra after spectral alignment, (c) resulting pseudo-seismic volume.

extracted patches, 70% are randomly selected for training, 15% for validation and the remaining 15% for testing.

Regarding noise modeling, synthetic noise is not uniformly injected into all samples. The field seismic datasets inherently provide realistic noise conditions, enabling the network to learn reconstruction features under practical acquisition and processing environments. To further enhance robustness and mitigate the discrepancy between pseudo-seismic and field data, Gaussian noise is introduced to 20% of the samples derived from the transformed natural images. This strategy improves stability under noisy conditions while preventing the model from overfitting to clean simulated structures.

The Adam optimization algorithm is used, with a learning rate set to 1×10^{-5} . The batch size is set to 2 due to the high memory consumption of 3D convolutional operations. The model is trained for 500,000 steps on a server equipped with three RTX 4090 GPUs and the model with the lowest loss on the validation datasets is selected as the final test model.

3.3. Training convergence analysis

To further illustrate the convergence behavior of GGCA-GAN and clarify the checkpoint selection criterion, the training and validation loss curves are presented in Fig. 5. The curves are obtained using the mixed training dataset described above.

Fig. 5(a) shows the adversarial loss curves of the generator and discriminator during training. After initial fluctuations in the early stages of training, both losses gradually stabilize, indicating that the adversarial training process reaches a dynamic equilibrium.

Fig. 5(b) presents the L_1 loss ($\mathcal{L}_{Cont}$) for the validation datasets. As training progressed, the validation loss followed a similar downward trend without noticeable divergence. This behavior indicates stable convergence and suggests that no significant overfitting occurs. The optimal checkpoint is selected at the 4.973×10^5 training step, where the validation loss reaches its minimum. It is therefore used for all subsequent quantitative and qualitative evaluations.

This analysis demonstrates that the proposed model exhibits stable convergence behavior, with no evident signs of overfitting or underfitting observed throughout the training process.

3.4. Kerry-3D

To validate the effectiveness of GGCA-GAN for 3D seismic data reconstruction, experiments were conducted on the Kerry-3D

dataset. Two baseline models were selected for comparison: a 2D U-Net and a GCA-GAN variant without the grouped feature modeling module. The 2D U-Net is effective at capturing local features in seismic data, whereas GCA-GAN captures long-range spatial dependencies via global coordinate attention, making it suitable for global-context modeling. All models were trained under the same training settings as the proposed model.

3.4.1. Discrete missing

To validate the efficacy of GGCA-GAN in reconstructing high-resolution seismic data, three-dimensional reconstruction experiments were conducted using the Kerry-3D dataset. The original dataset served as the baseline, as shown in Fig. 6.

The experiment increased the complexity of the discrete missing scenario by raising the missing rate from 30% to 50% and 80%. The reconstruction results and corresponding residuals obtained by the three methods under different missing rates were visualized using 3D slices, as presented in Figs. 7–9. Table 1 presents a quantitative analysis using three evaluation metrics: PSNR, SNR and SSIM.

Under a 30% discrete missing rate, 2D U-Net, GCA-GAN and GGCA-GAN all produced reconstructions close to the ground truth,

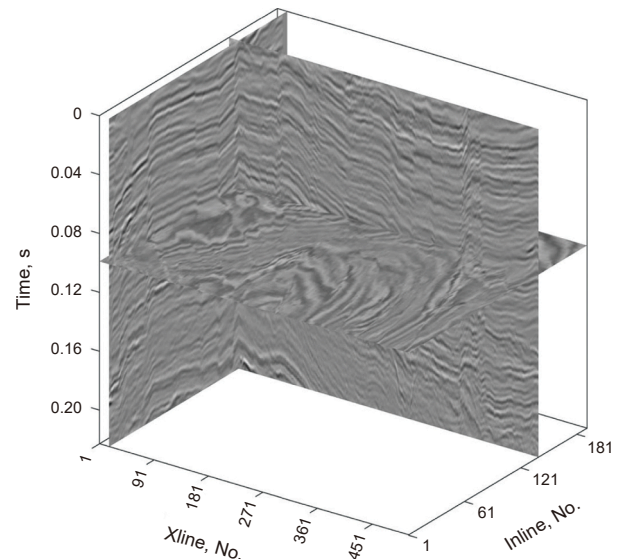
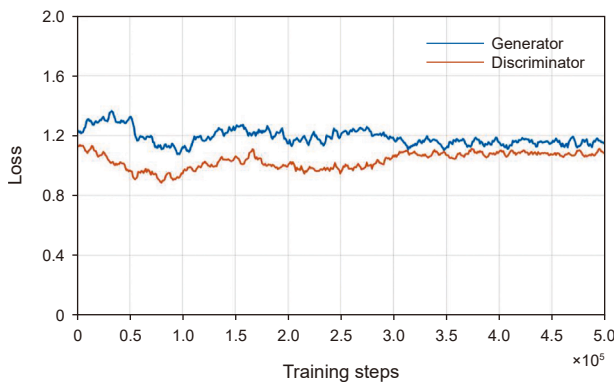
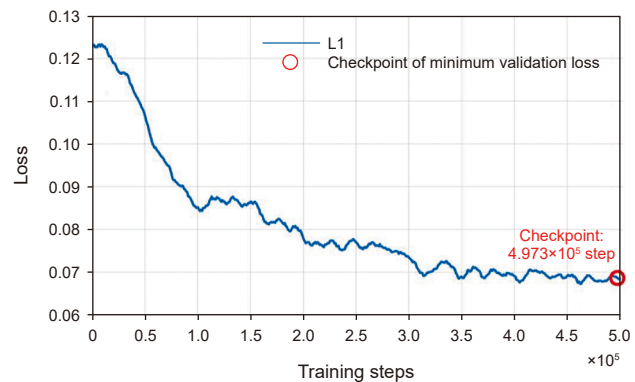


Fig. 6. Kerry-3D test data.



(a) Adversarial loss curves



(b) L1 loss curves

Fig. 5. Training convergence curves of GGCA-GAN. (a) Adversarial losses of the generator and discriminator during training. (b) L_1 loss ($\mathcal{L}_{Cont}$) on the validation datasets. The marked point indicates the selected checkpoint (4.973×10^5 step) corresponding to the minimum validation loss.

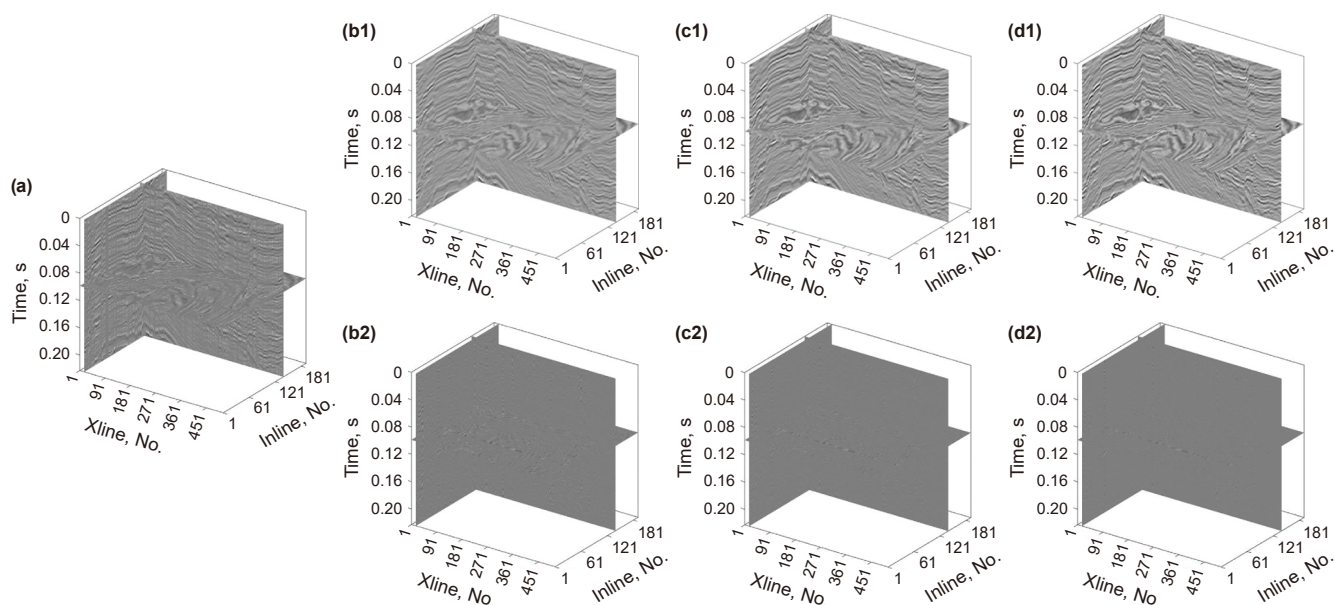


Fig. 7. Reconstruction performance of 2D U-Net, GCA-GAN and GGCA-GAN on the Kerry-3D dataset under 30% discrete missing rate. (a) Instance with 30% discrete missing data, (b)–(d) display the reconstruction results of 2D U-Net, GCA-GAN and GGCA-GAN, along with their corresponding residuals.

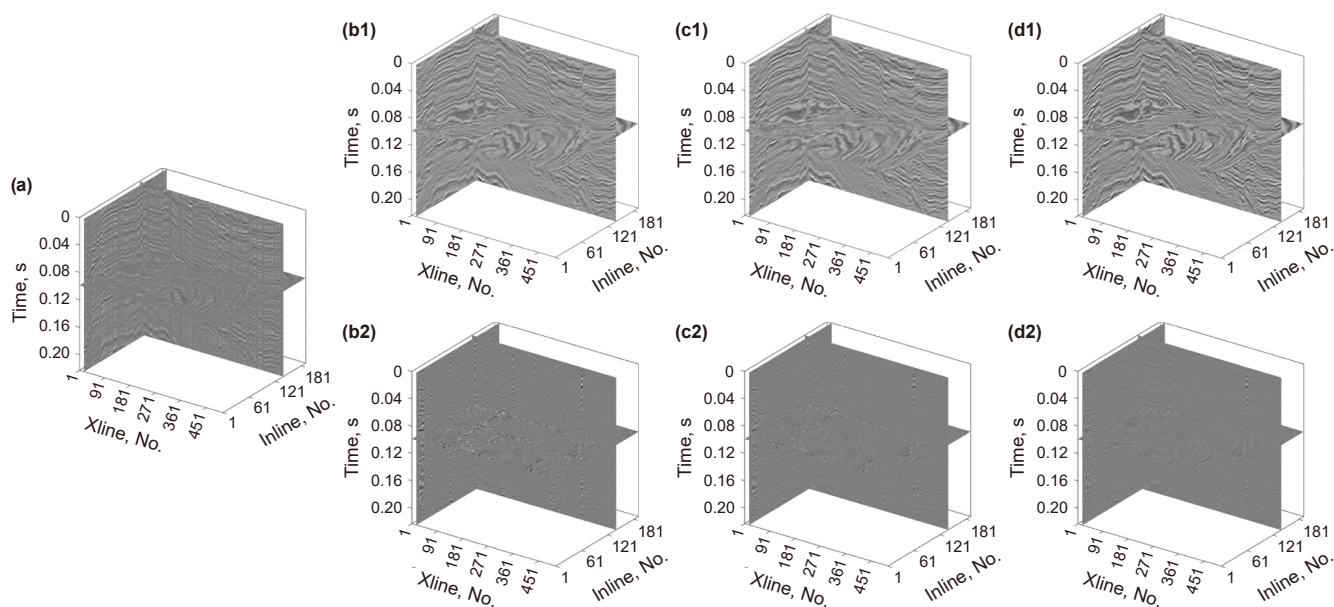


Fig. 8. Reconstruction performance of 2D U-Net, GCA-GAN and GGCA-GAN on the Kerry-3D dataset under 50% discrete missing rate. (a) Instance with 50% discrete missing data, (b)–(d) display the reconstruction results of 2D U-Net, GCA-GAN and GGCA-GAN, along with their corresponding residuals.

with GAN-based models better preserving fine details and structural continuity. When the missing rate increased to 50%, 2D U-Net exhibited a marked degradation in structural consistency, whereas GCA-GAN and GGCA-GAN remained relatively stable, yielding better detail recovery and structural consistency in the interpolated results.

When 80% of the seismic traces were removed under discrete missing, the reconstruction problem became severely undersampled and highly ill-posed, as only limited observations remained to constrain lateral continuity, high-frequency reflectors and amplitude variations. Consequently, under such extremely sparse conditions, the performance of all methods deteriorated substantially. However, GAN-based models degraded more gradually than 2D U-Net, and GGCA-GAN

achieved the highest overall fidelity. Qualitatively, GGCA-GAN produced the smoothest data volumes with the most coherent structures and the fewer artifacts, indicating the strongest regularization and improved robustness under extreme data loss.

3.4.2. Continuous missing

The spatial continuity of missing regions makes the reconstruction task more complex. It requires not only recovering the missing values but also maintaining structural coherence and spatial consistency. We further evaluated the reconstruction performance of the proposed method under continuous missing patterns. Data volumes with 30% and 50% continuous missing rates were reconstructed, as shown in Figs. 10 and 11.

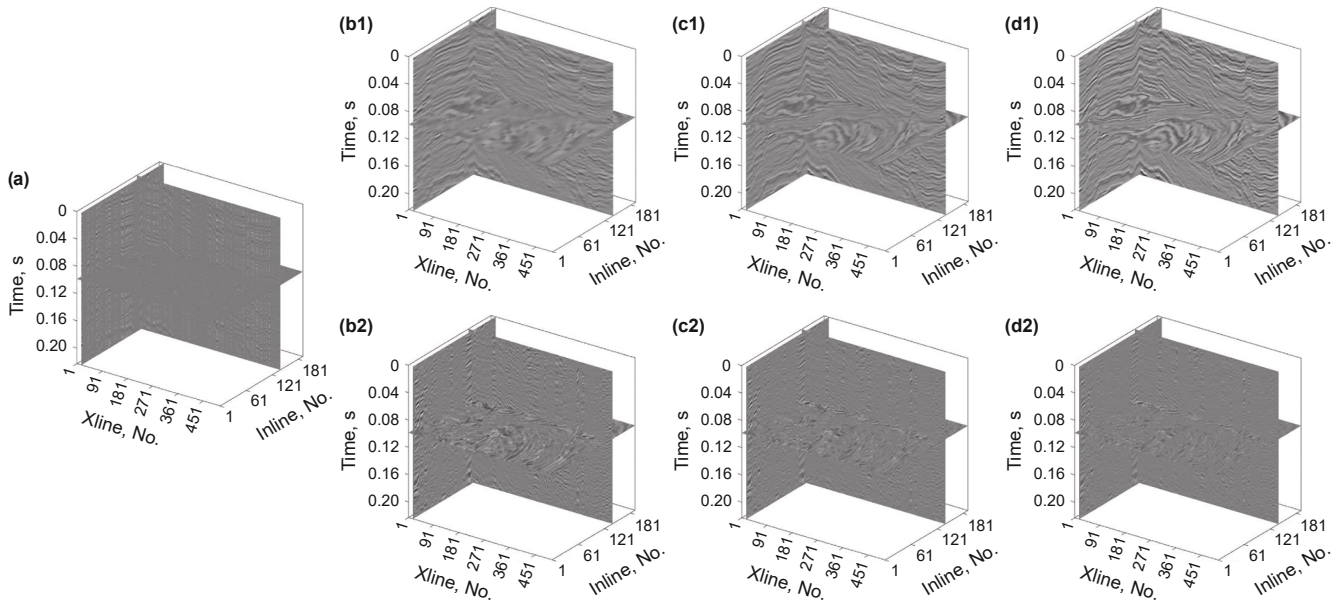


Fig. 9. Reconstruction performance of 2D U-Net, GCA-GAN and GGCA-GAN on the Kerry-3D dataset under 80% discrete missing rate. (a) Instance with 80% discrete missing data, (b)–(d) display the reconstruction results of 2D U-Net, GCA-GAN and GGCA-GAN, along with their corresponding residuals.

Table 1

Comparison of the reconstruction metrics of different reconstruction methods for the Kerry-3D dataset under varying discrete missing rates.

Missing rate, %	Method	PSNR, dB	SNR, dB	SSIM
30	2D U-Net	28.46	23.00	0.835
	GCA-GAN	31.76	26.29	0.926
	GGCA-GAN	34.51	29.04	0.962
50	2D U-Net	26.00	20.53	0.730
	GCA-GAN	28.92	23.45	0.856
	GGCA-GAN	32.40	26.94	0.938
80	2D U-Net	21.32	19.85	0.568
	GCA-GAN	27.14	20.78	0.751
	GGCA-GAN	28.21	22.54	0.804

Qualitative analysis of the reconstructions at a 30% missing rate indicates that all three methods closely match the reference data. However, as the missing rate increases to 50%, the continuous missing scenario becomes markedly more challenging and 2D U-Net exhibits pronounced structural distortions. In contrast, both GCA-GAN and GGCA-GAN remain comparatively stable, with GGCA-GAN providing the most consistent structural recovery.

Table 2 further confirms that GAN-based methods substantially outperform conventional methods under continuous missing conditions. The severe data loss at 50% continuous missing makes it difficult for all methods to preserve high-frequency content and lateral continuity, rendering the reconstruction more error-prone. Despite these challenges, GAN-based models—particularly GGCA-

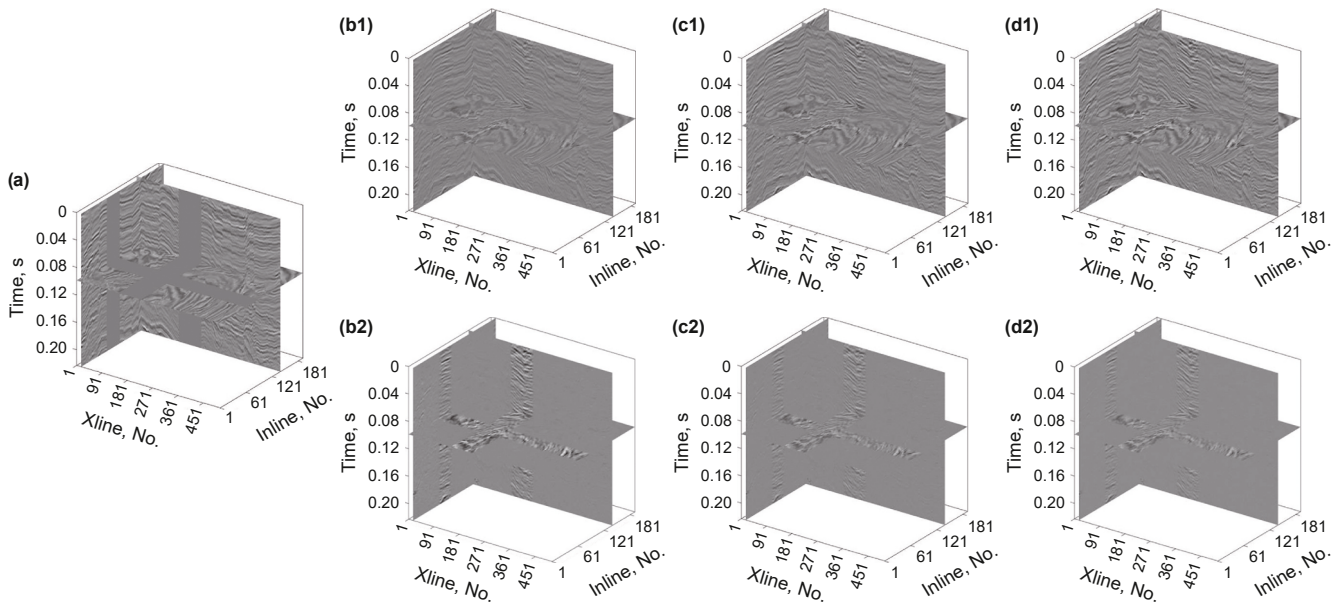


Fig. 10. Reconstruction performance of 2D U-Net, GCA-GAN and GGCA-GAN on the Kerry-3D dataset under 30% continuous missing rate. (a) Instance with 30% continuous missing data, (b)–(d) display the reconstruction results of 2D U-Net, GCA-GAN and GGCA-GAN, along with their corresponding residuals.

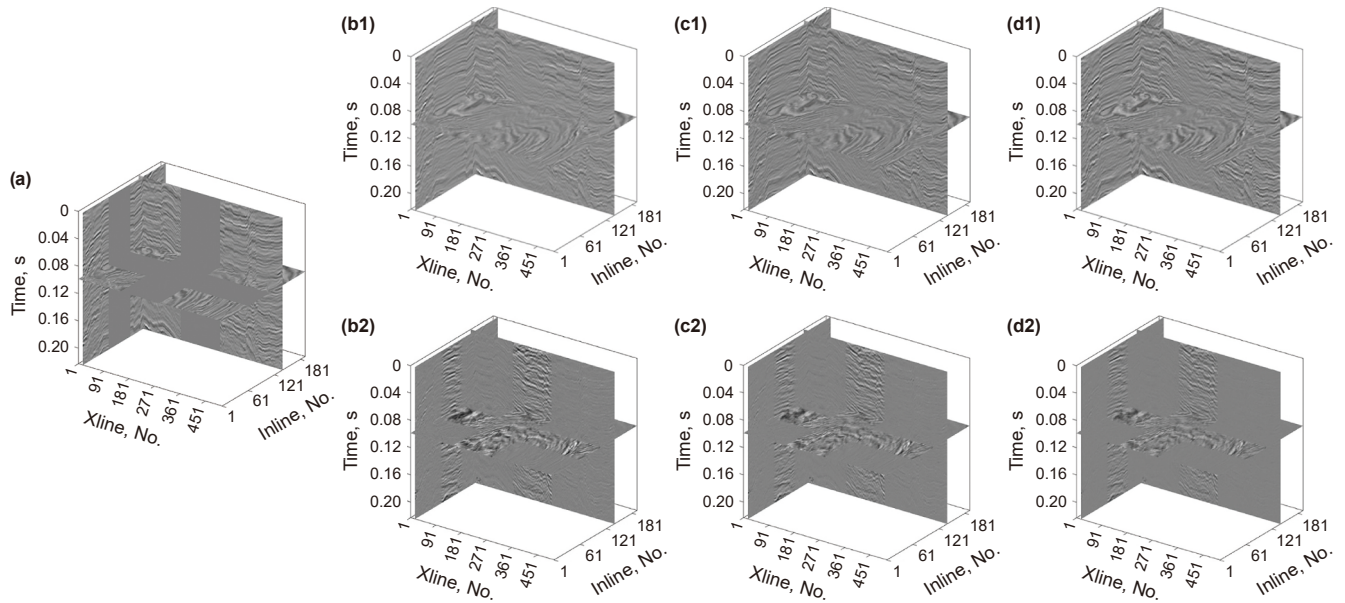


Fig. 11. Reconstruction performance of 2D U-Net, GCA-GAN and GGCA-GAN on the Kerry-3D dataset under 50% continuous missing rate. (a) Instance with 50% continuous missing data, (b)–(d) display the reconstruction results of 2D U-Net, GCA-GAN and GGCA-GAN, along with their corresponding residuals.

Table 2

Comparison of the reconstruction metrics of different reconstruction methods for the Kerry-3D dataset under varying continuous missing rates.

Missing rate, %	Method	PSNR, dB	SNR, dB	SSIM
30	2D U-Net	25.45	19.99	0.731
	GCA-GAN	26.41	20.95	0.746
	GGCA-GAN	27.38	21.92	0.813
50	2D U-Net	23.02	17.56	0.591
	GCA-GAN	24.15	18.68	0.655
	GGCA-GAN	25.86	20.33	0.742

GAN—are more robust, with GGCA-GAN achieving better detail recovery and more coherent structural consistency than GCA-GAN. This suggests that incorporating global coordinate attention helps mitigate the adverse effects of missing data, enabling GGCA-GAN to better control residuals and maintain smoother reconstructions under extreme sparsity.

3.5. F3-3D

3.5.1. Different missing patterns

To further evaluate the performance of GGCA-GAN and verify its effectiveness, experiments were conducted on the F3-3D test data, with comparisons against conventional 2D U-Net and GCA-GAN. We reconstructed the pre-stack gathers into volumes for qualitative analysis and an example from the F3-3D test data is shown in Fig. 12.

Experiments were conducted under different missing patterns, with missing rates varied across settings. The reconstruction results of the three methods were visualized in 3D, as shown in Figs. 13–17.

In the 30% discrete missing scenario shown in Fig. 13, all methods are able to reconstruct the data reasonably well under both discrete missing and continuous missing conditions. When comparing GCA-GAN and GGCA-GAN, the latter exhibits superior detail recovery. As shown in Figs. 14 and 15, the residual maps reveal pronounced discontinuities in the U-Net reconstructions, whereas GCA-GAN and GGCA-GAN maintain stable reconstruction

performance even under high discrete missing rates. Reconstruction under continuous missing conditions is more challenging, as shown in Figs. 16 and 17. In particular, at a 50% missing rate, U-Net begins to exhibit noticeable distortions, whereas GAN-based networks still achieve reliable results, with GGCA-GAN recovering more fine-scale details. These results indicate that GGCA-GAN can maintain effective reconstruction performance even under high missing rates.

Fig. 18 shows a waveform comparison extracted from the same seismic trace in the reference data and in the reconstructions obtained by 2D U-Net, GCA-GAN and GGCA-GAN. Compared with the reference data, 2D U-Net shows noticeable deviations, indicating limited ability to preserve 3D structural consistency. In contrast, GCA-GAN and GGCA-GAN produce waveforms that closely match the reference, with the major reflection events exhibiting well-preserved phase and amplitude variations. Meanwhile, GGCA-GAN achieves better waveform consistency with the reference data in high-amplitude regions.

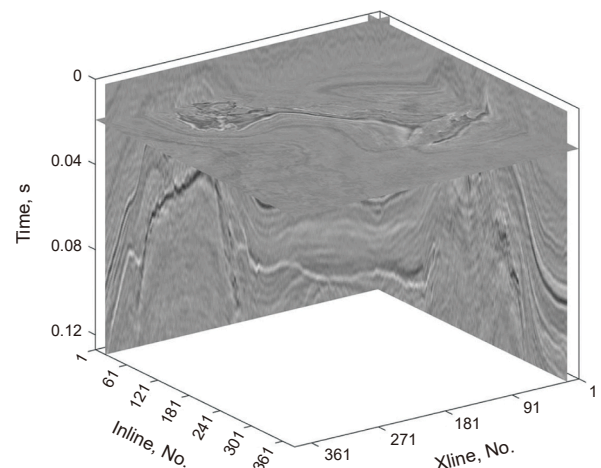


Fig. 12. F3-3D test data.

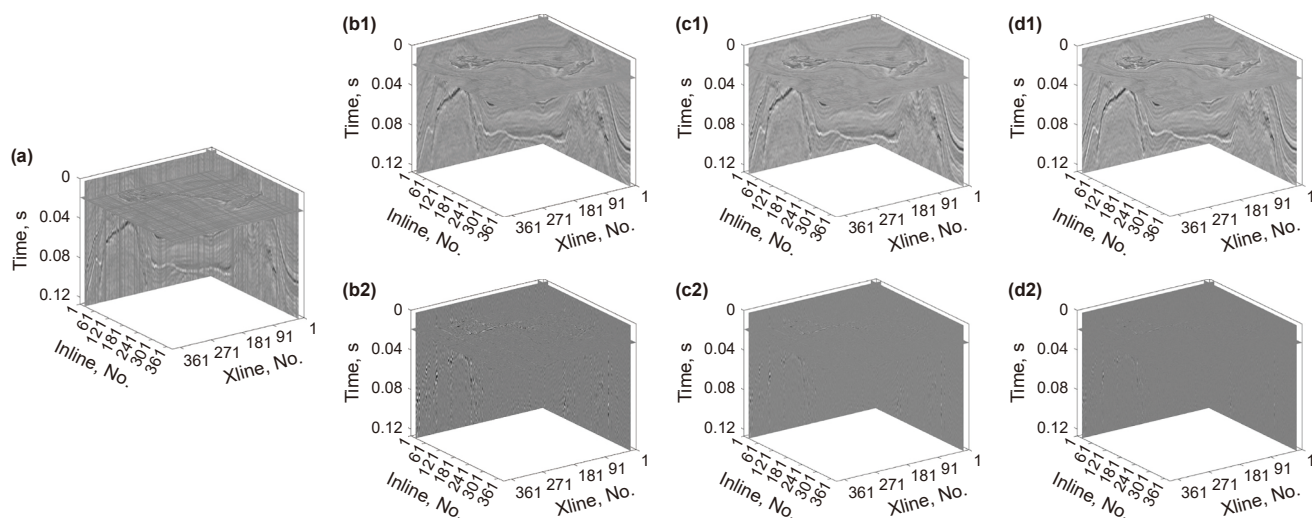


Fig. 13. Reconstruction performance of 2D U-Net, GCA-GAN and GGCA-GAN on the F3-3D dataset under 30% discrete missing rate. (a) Instance with 30% discrete missing data, (b)–(d) display the reconstruction results of 2D U-Net, GCA-GAN and GGCA-GAN, along with their corresponding residuals.

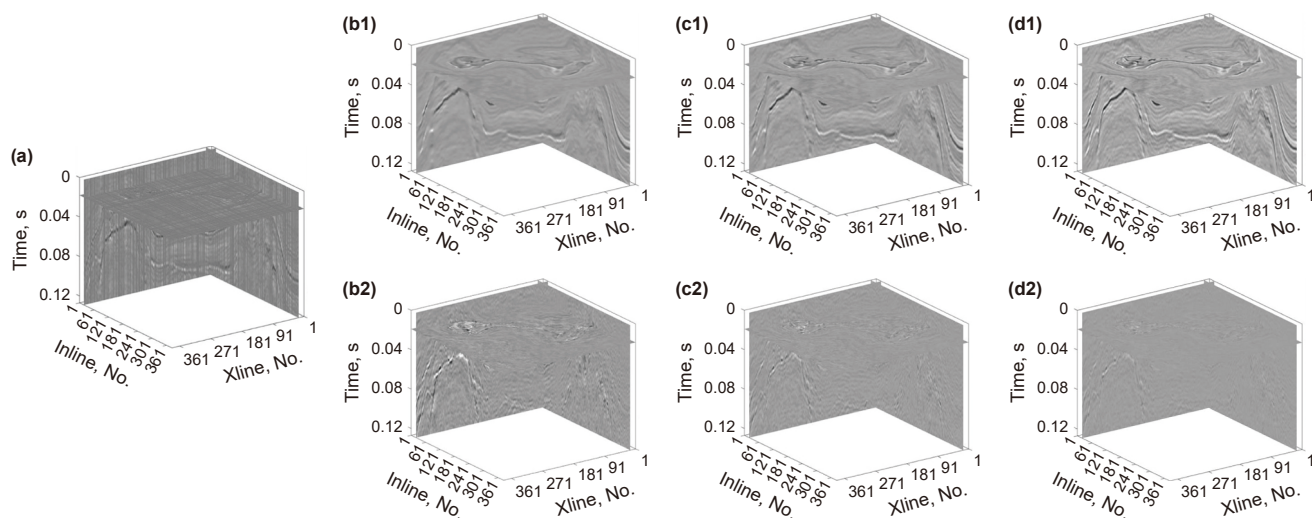


Fig. 14. Reconstruction performance of 2D U-Net, GCA-GAN and GGCA-GAN on the F3-3D dataset under 50% discrete missing rate. (a) Instance with 50% discrete missing data, (b)–(d) display the reconstruction results of 2D U-Net, GCA-GAN and GGCA-GAN, along with their corresponding residuals.

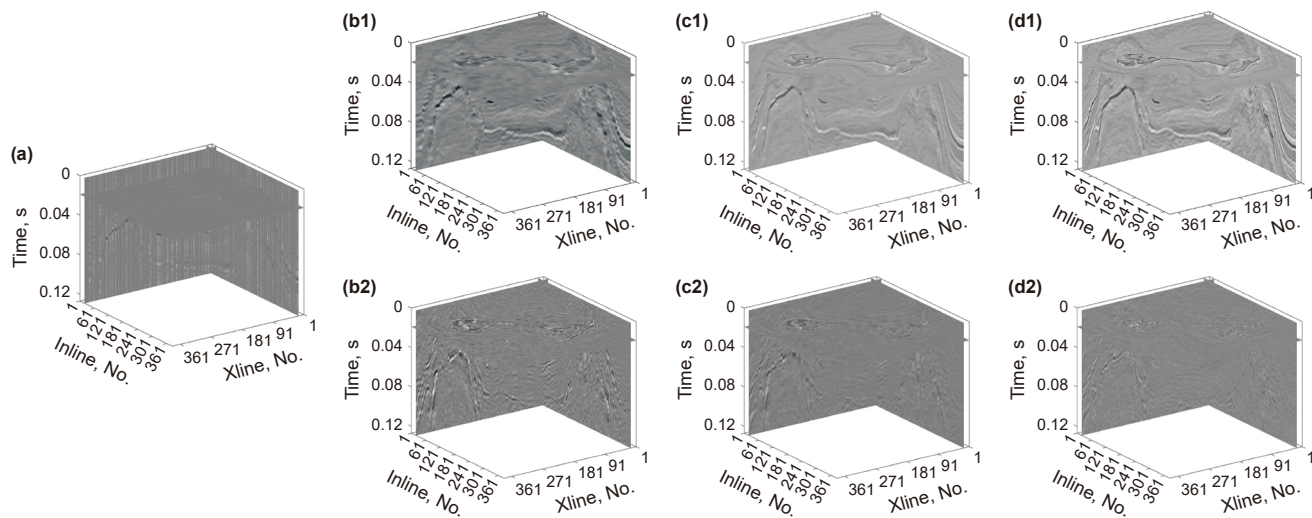


Fig. 15. Reconstruction performance of 2D U-Net, GCA-GAN and GGCA-GAN on the F3-3D dataset under 80% discrete missing rate. (a) Instance with 80% discrete missing data, (b)–(d) display the reconstruction results of 2D U-Net, GCA-GAN and GGCA-GAN, along with their corresponding residuals.

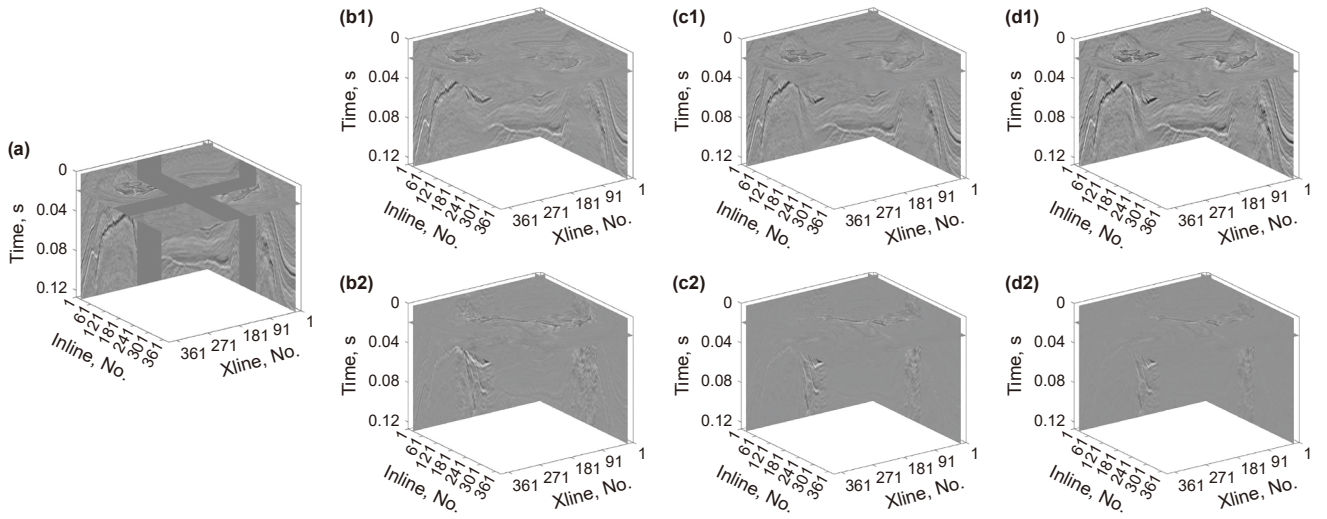


Fig. 16. Reconstruction performance of 2D U-Net, GCA-GAN and GGCA-GAN on the F3-3D dataset under 30% continuous missing rate. (a) Instance with 30% continuous missing data, (b)–(d) display the reconstruction results of 2D U-Net, GCA-GAN and GGCA-GAN, along with their corresponding residuals.

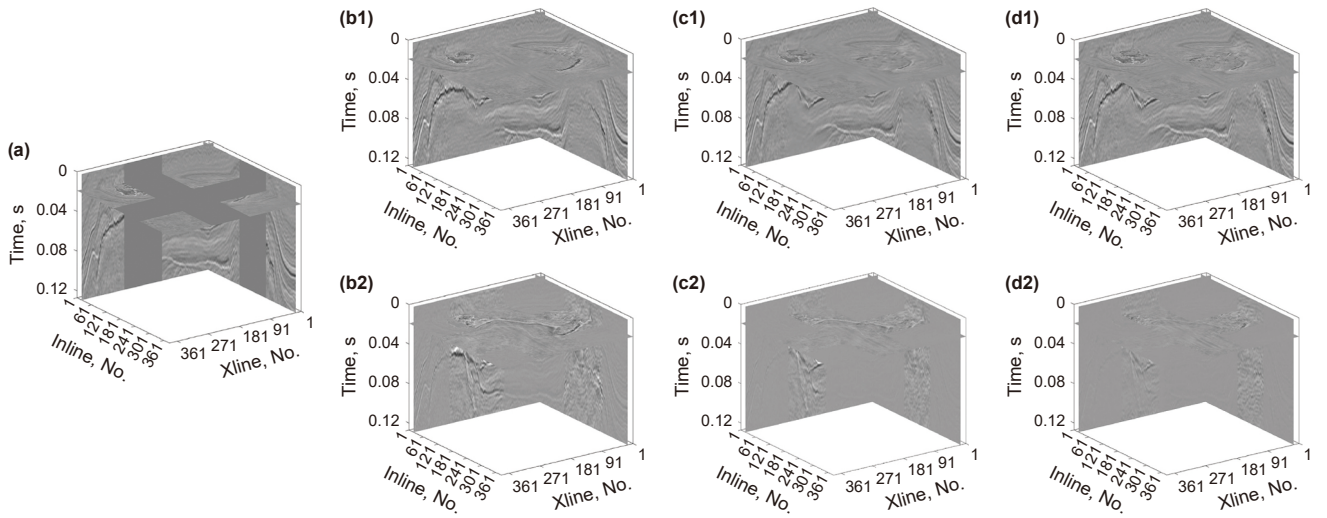


Fig. 17. Reconstruction performance of 2D U-Net, GCA-GAN and GGCA-GAN on the F3-3D dataset under 50% continuous missing rate. (a) Instance with 50% continuous missing data, (b)–(d) display the reconstruction results of 2D U-Net, GCA-GAN and GGCA-GAN, along with their corresponding residuals.

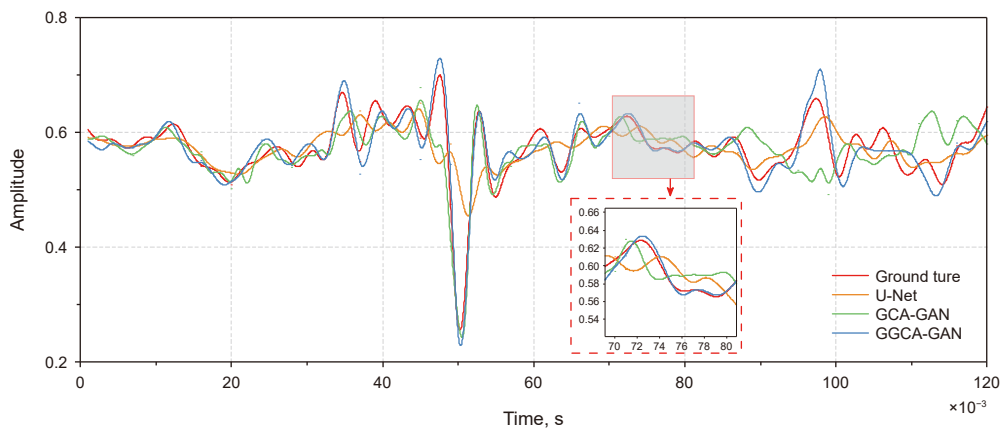


Fig. 18. Comparison of waveforms between the reference seismic record (red curve) and that reconstructed by 2D U-Net (orange curve), GCA-GAN (green curve) and GGCA-GAN (blue curve).

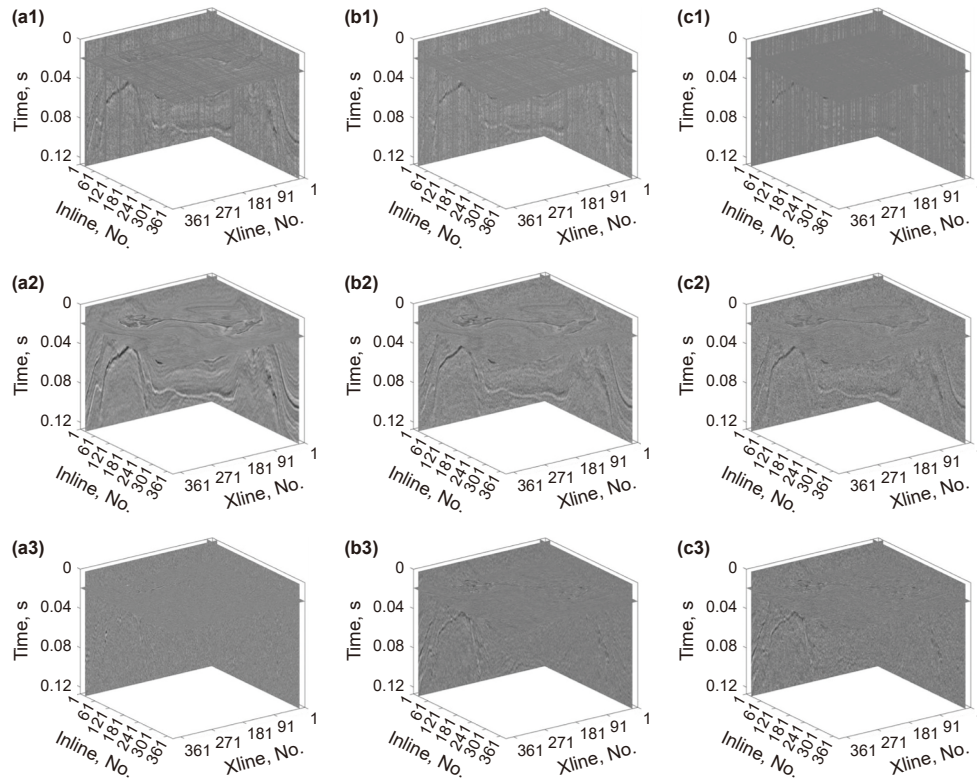


Fig. 19. Comparison of GGCA-GAN reconstruction results on the noise-contaminated F3-3D dataset under different discrete missing rates. (a1)–(a3) show the 30% discrete missing noisy input, along with the corresponding reconstruction result and residual map. (b1)–(b3) present the 50% discrete missing noisy input, along with the corresponding reconstruction result and residual map. (c1)–(c3) display the 80% discrete missing noisy input, along with the corresponding reconstruction result and residual map.

3.5.2. Noise robustness test

Next, to evaluate the noise robustness of GGCA-GAN under noisy conditions, we applied the model to F3-3D dataset under discrete missing patterns at different missing rates. Although field seismic data inherently contain noise, the actual noise level and characteristics are unknown and difficult to control. Therefore, synthetic noise with predefined amplitude was added to the data to construct a controlled evaluation environment. Gaussian noise was introduced to datasets with missing rates of 30%, 50% and 80%, resulting in SNR of 19.12 dB, 18.55 dB and 17.91 dB, respectively. The results, shown in Fig. 19, demonstrate that the network effectively recovers the main reflection structures while substantially suppressing noise, maintaining reliable reconstruction quality under noisy inputs.

Table 3 summarizes the reconstruction performance of GGCA-GAN on discrete missing data with noise. As the missing rate increases to 50% and 80%, reconstruction performance decreases, as fewer observations are available and noise further corrupts the inputs. In particular, at an 80% missing rate, the reconstruction becomes severely underdetermined and the added noise further exacerbates the ill-posedness. Nevertheless, GGCA-GAN maintains relatively high reconstruction quality, indicating strong noise robustness and the ability to recover underlying structures with minimal artifacts. Overall, these quantitative evaluation metrics highlight the noise-

suppression capability of GGCA-GAN under challenging conditions, including high data sparsity and noise contamination.

3.5.3. Contribution of natural images to training

To further evaluate the impact of training set composition on network performance, Table 4 compares reconstruction results on F3-3D dataset with a 30% discrete missing rate using two training datasets: a hybrid dataset combining natural images and measured data and a dataset using only field data. The hybrid dataset substantially outperforms training on the seismic dataset alone, achieving higher PSNR, SNR and SSIM.

These findings indicate that incorporating natural images improves generalization, enhancing the recovery of fine details and the preservation of structural consistency. The hybrid dataset compensates for the limited size and diversity of the seismic training data, suggesting that natural images provide useful variability for alleviating 3D data scarcity. Overall, this strategy improves accuracy and robustness, leading to better reconstruction performance under missing-data conditions.

3.6. SEG-C3

3.6.1. Discrete missing

To further evaluate the generalization capability of the proposed GGCA-GAN, experiments were conducted on the SEG-C3

Table 3

GGCA-GAN reconstruction metrics for the F3-3D dataset under varying discrete missing rates with noise.

Missing rate, %	PSNR, dB	SNR, dB	SSIM
30	28.87	24.16	0.839
50	26.85	23.06	0.805
80	24.38	22.15	0.738

Table 4

GGCA-GAN reconstruction metrics for the F3-3D dataset under 30% discrete missing data with different training datasets.

Training set	PSNR, dB	SNR, dB	SSIM
Hybrid dataset	31.94	27.15	0.879
Seismic dataset	31.07	26.29	0.840

simulated dataset, including reconstruction tasks under both discrete missing and continuous missing scenarios. In each case, the performance of GGCA-GAN was compared with that of 2D U-Net in terms of reconstruction accuracy. Then, the noise robustness of the network was systematically evaluated. As illustrated in Fig. 20, an example from the SEG-C3 test dataset, which was not involved in the training process, is presented.

To evaluate the reconstruction performance of GGCA-GAN on the SEG-C3 dataset under varying missing rates, experiments were first conducted for discrete missing scenarios. By

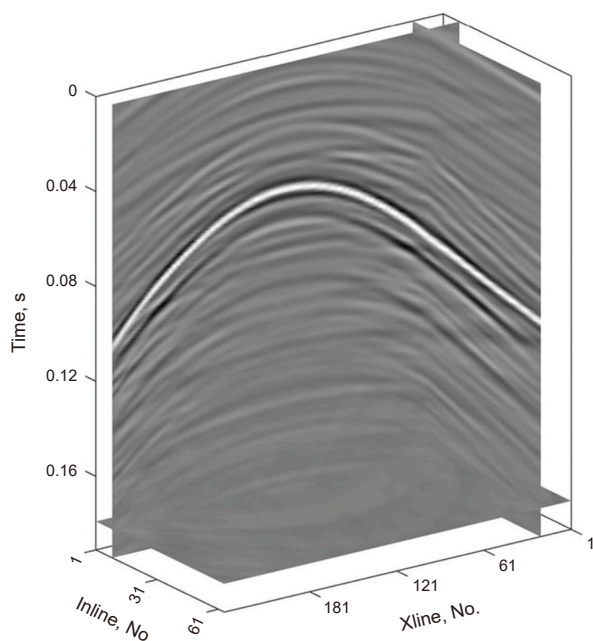


Fig. 20. SEG-C3 test data.

randomly removing 30%, 50% and 80% of the data samples, reconstruction was performed using both GGCA-GAN and 2D U-Net. The reconstruction results of the two methods, along with their corresponding residual maps, are presented in Figs. 21–23.

As shown in Figs. 21–23, GGCA-GAN consistently outperforms 2D U-Net across all missing rates. The reconstruction performance gap becomes more pronounced as the missing rate increases. At a missing rate of 80%, 2D U-Net exhibits severe structural distortion, whereas GGCA-GAN better preserves structural continuity and spatial coherence. These results indicate that GGCA-GAN demonstrates superior robustness and reconstruction fidelity compared with 2D U-Net.

3.6.2. Continuous missing

We further evaluated the reconstruction performance of GGCA-GAN on the SEG-C3 dataset under different continuous missing rates, aiming to simulate more challenging and realistic data loss scenarios. Specifically, 30% and 50% of the data were removed from spatially continuous regions, representing large-scale missing areas commonly encountered in seismic data acquisition.

The reconstruction results and residual maps for continuous missing are shown in Figs. 24 and 25. GGCA-GAN outperforms 2D U-Net in qualitative visualization evaluation. 2D U-Net performs poorly in these continuous missing regions, exhibiting pronounced amplitude artifacts and structural distortions. In contrast, GGCA-GAN effectively preserves the structural continuity of seismic reflections, even in regions with substantial data loss and provides more accurate and stable reconstruction results than 2D U-Net. These findings demonstrate that GGCA-GAN is well suited for handling continuous missing cases commonly encountered in seismic exploration.

It should be noted that the SEG-C3 dataset was not included in the training stage. The stable reconstruction performance under both discrete and continuous missing scenarios therefore indicates that GGCA-GAN generalizes well to unseen seismic datasets.

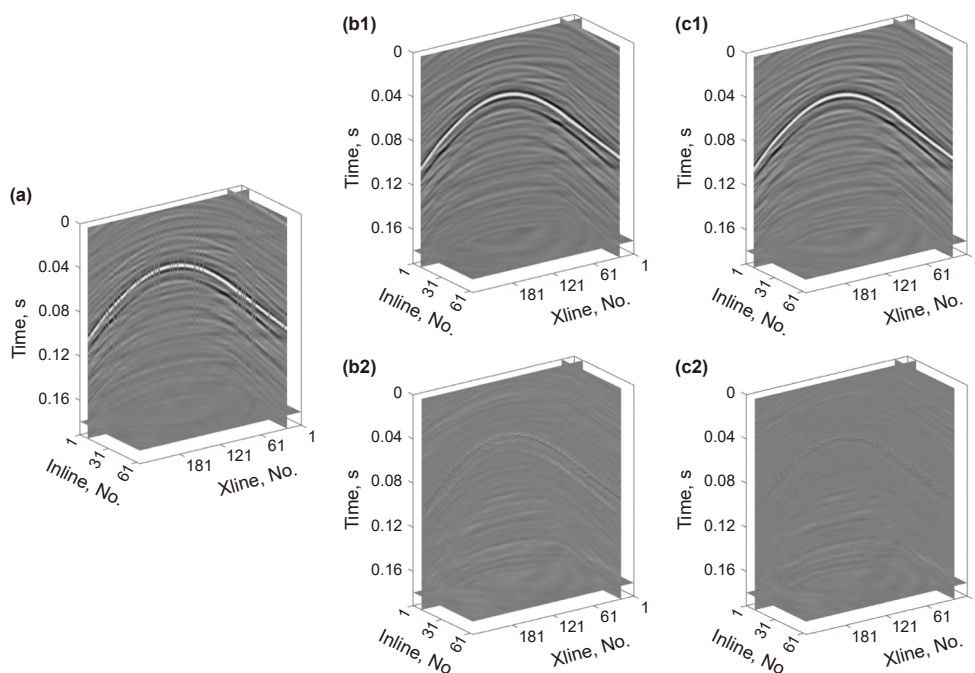


Fig. 21. Reconstruction performance of 2D U-Net and GGCA-GAN on the SEG-C3 dataset under 30% discrete missing rate. (a) Instance with 30% discrete missing data, (b)–(c) display the reconstruction results of 2D U-Net and GGCA-GAN, along with their corresponding residuals.

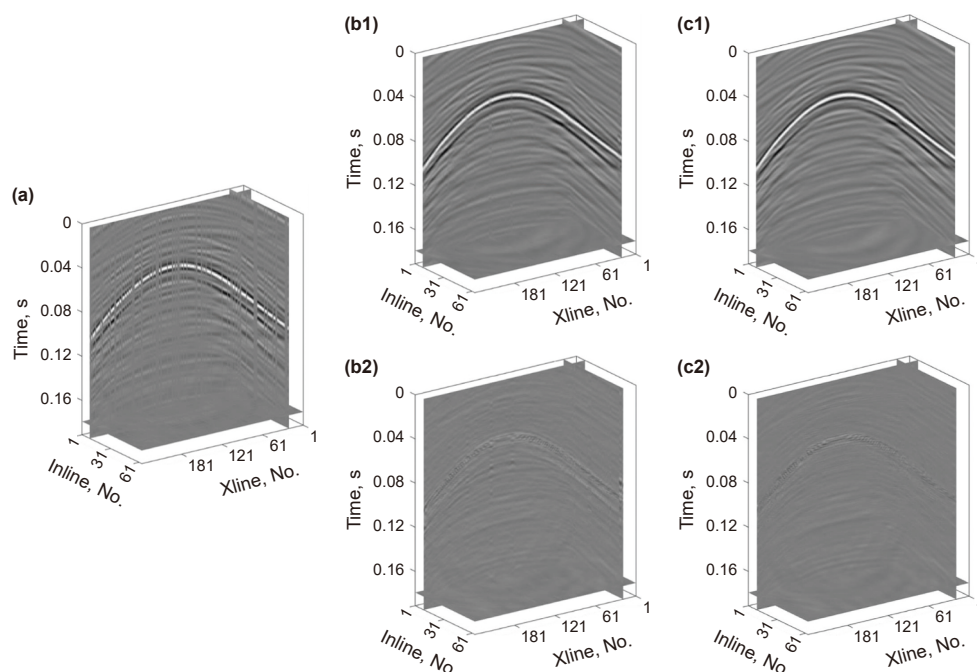


Fig. 22. Reconstruction performance of 2D U-Net and GGCA-GAN on the SEG-C3 dataset under 50% discrete missing rate. (a) Instance with 50% discrete missing data, (b)–(c) display the reconstruction results of 2D U-Net and GGCA-GAN, along with their corresponding residuals.

3.6.3. Noise robustness test

The introduction of noise during seismic data acquisition significantly affects the data quality. To assess the noise robustness of GGCA-GAN, Gaussian noise was added to discrete missing data with missing rates of 30%, 50% and 80%, with the corresponding SNR of 18.52 dB, 17.91 dB and 16.97 dB, respectively, thus simulating more realistic seismic data acquisition

conditions. Reconstruction was then performed using GGCA-GAN on the noisy datasets at each missing rate. Fig. 26 shows the noisy inputs, corresponding reconstruction results and residual maps. Meanwhile, to quantitatively analyze the noise robustness, Table 5 respectively summarizes the SNR, PSNR and SSIM of the above reconstruction results processed by the proposed model.

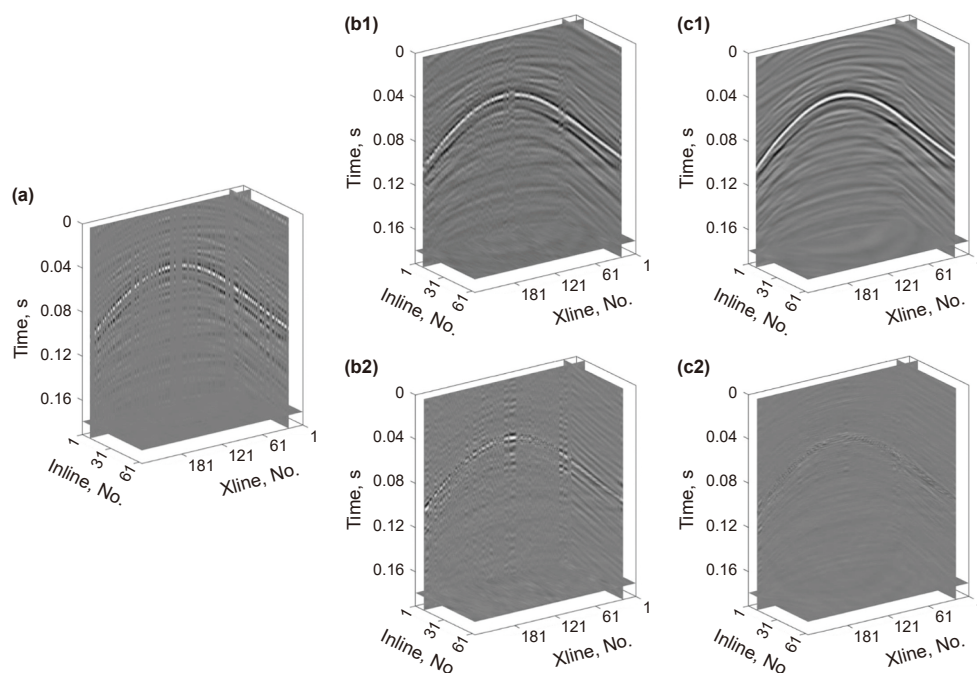


Fig. 23. Reconstruction performance of 2D U-Net and GGCA-GAN on the SEG-C3 dataset under 80% discrete missing rate. (a) Instance with 80% discrete missing data, (b)–(c) display the reconstruction results of 2D U-Net and GGCA-GAN, along with their corresponding residuals.

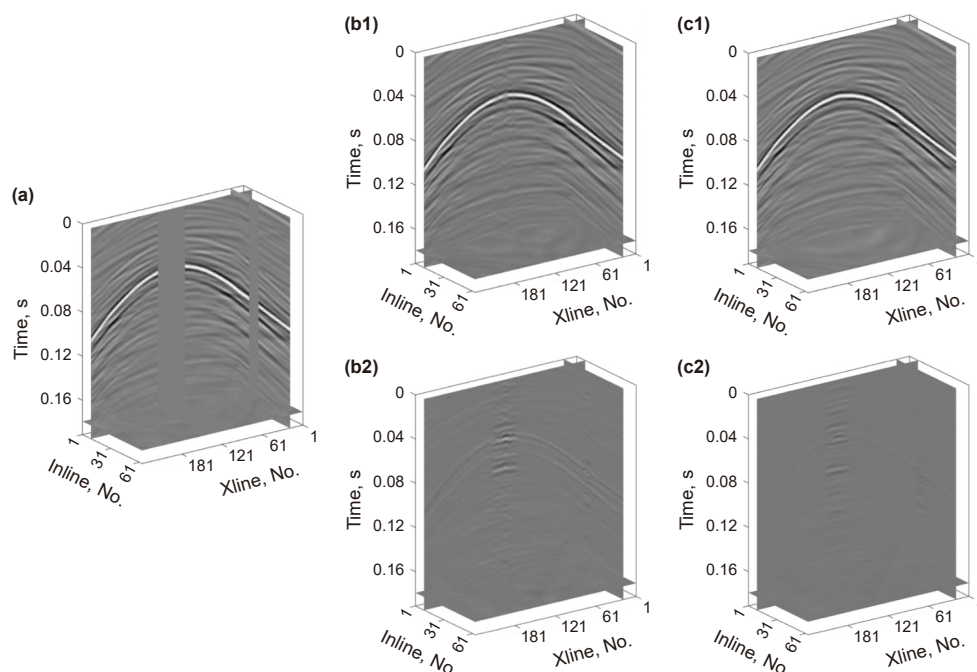


Fig. 24. Reconstruction performance of 2D U-Net and GGCA-GAN on the SEG-C3 dataset under 30% continuous missing rate. (a) Instance with 30% continuous missing data, (b)–(c) display the reconstruction results of 2D U-Net and GGCA-GAN, along with their corresponding residuals.

From the plots and table, as the missing rate increases, reconstruction performance naturally decreases. Moreover, the addition of noise further deteriorates the quality of the input data, making the reconstruction task more challenging. Especially when the missing rate reaches 80%, the data becomes highly sparse, significantly degrading reconstruction quality and the presence of noise

further exacerbates the reconstruction difficulty. Despite these challenges, GGCA-GAN achieves reliable seismic reconstruction under noisy conditions. Even with Gaussian disturbances and high missing rates, the key structural features are effectively restored, confirming the robustness of the proposed framework in noisy environments.

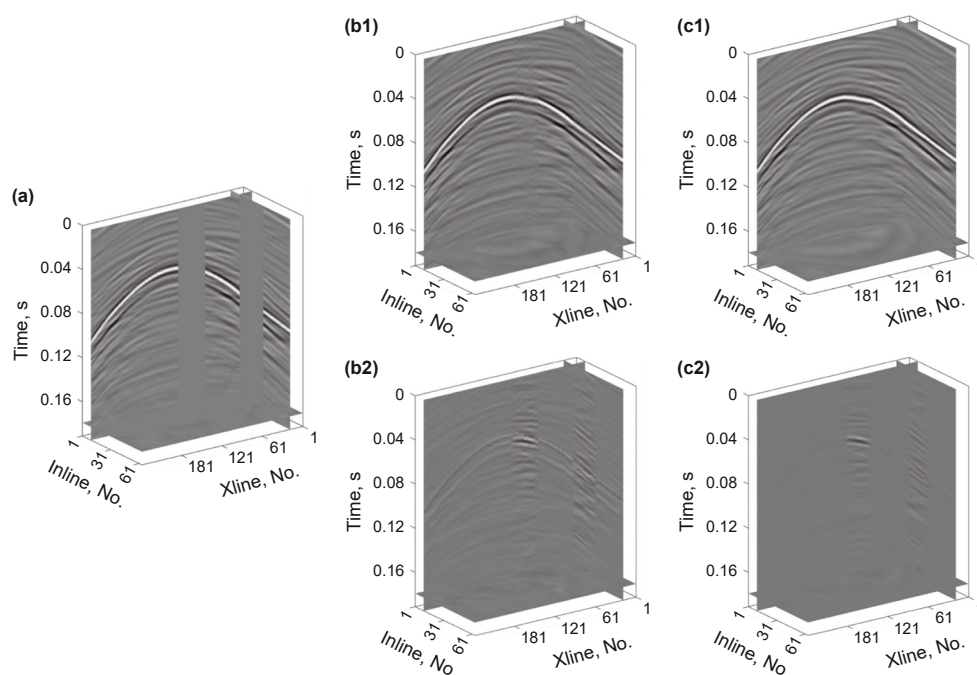


Fig. 25. Reconstruction performance of 2D U-Net and GGCA-GAN on the SEG-C3 dataset under 50% continuous missing rate. (a) Instance with 50% continuous missing data, (b)–(c) display the reconstruction results of 2D U-Net and GGCA-GAN, along with their corresponding residuals.

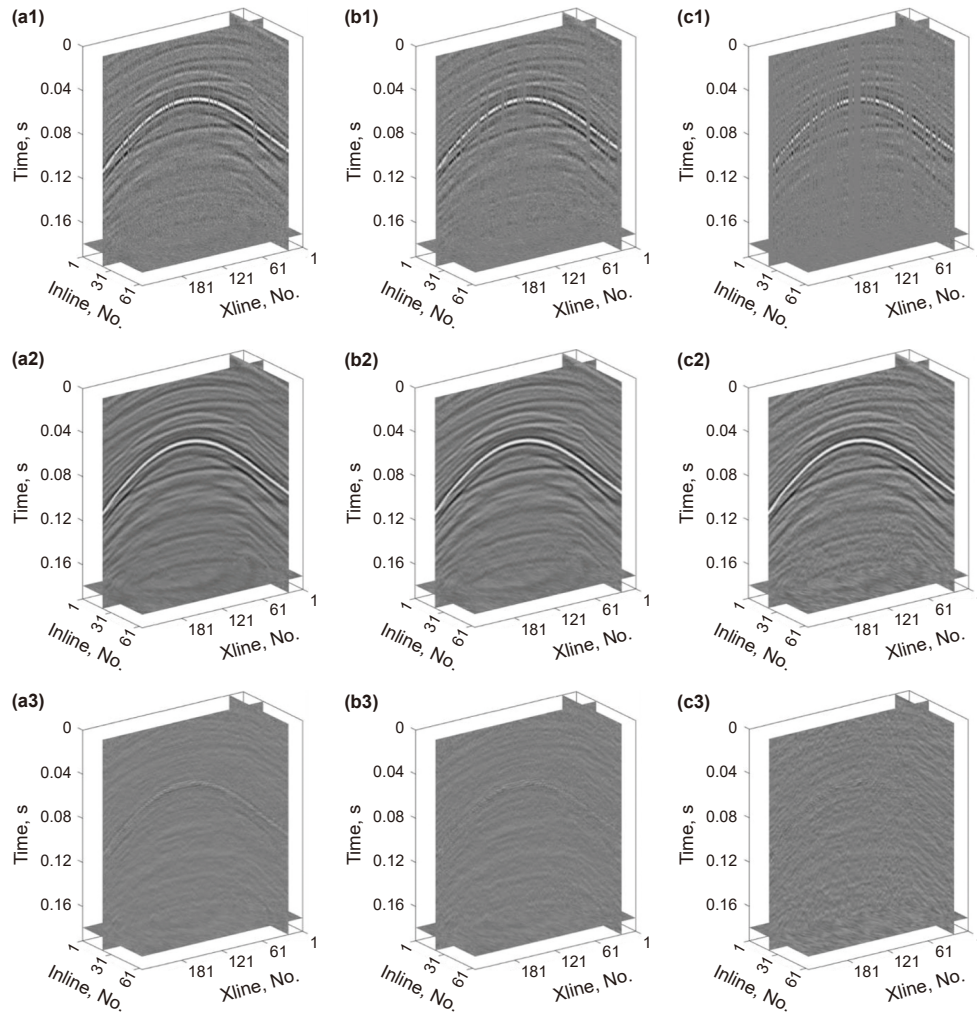


Fig. 26. Comparison of GGCA-GAN reconstruction results on the noise-contaminated SEG-C3 dataset under different discrete missing rates. **(a1)–(a3)** show the 30% discrete missing noisy input, along with the corresponding reconstruction result and residual map. **(b1)–(b3)** present the 50% discrete missing noisy input, along with the corresponding reconstruction result and residual map. **(c1)–(c3)** display the 80% discrete missing noisy input, along with the corresponding reconstruction result and residual map.

Table 5

GGCA-GAN reconstruction metrics for the SEG-C3 dataset under varying discrete missing rates with noise.

Missing rate, %	PSNR, dB	SNR, dB	SSIM
30	31.99	25.99	0.884
50	31.53	25.54	0.861
80	29.10	23.00	0.765

4. Conclusion and discussion

This paper proposes a Generative Adversarial Network based on the Global Grouped Coordinate Attention mechanism (GGCA-GAN) for high-resolution 3D seismic data reconstruction. The GGCA mechanism significantly enhances reconstruction performance in complex seismic scenarios by improving the global feature extraction capabilities of both the generator and discriminator. Furthermore, we construct a hybrid training dataset combining natural images and real seismic data, effectively addressing the scarcity of seismic data samples and further improving the network's generalization ability and reconstruction accuracy.

Experimental results validate the effectiveness of GGCA-GAN under various discrete missing and continuous missing

conditions, including high missing rates and noisy environments. Compared to traditional methods like U-Net, this model demonstrates clear advantages in both discrete missing and continuous missing scenarios, achieving significant improvements in PSNR, SNR and SSIM metrics. Notably, GGCA-GAN effectively suppresses noise and maintains high reconstruction quality under noisy conditions, fully demonstrating its noise robustness. The incorporation of natural images greatly enhances the network's generalization ability in complex seismic structures. The hybrid training approach alleviates the issue of insufficient seismic data samples, thereby improving reconstruction accuracy.

Although GGCA-GAN achieves significant progress in 3D seismic data reconstruction, the existing framework remains heavily reliant on data-driven learning mechanisms, lacking explicit integration of physical constraints (e.g., seismic wave propagation). Under conditions of extremely high data missing rates, complex geological structures, or significant variations in observation conditions, the model may face challenges related to insufficient physical consistency and limited generalization ability. To address this challenge, future research can incorporate physics-informed deep learning methods into the network architecture. By integrating prior knowledge from seismic wave propagation

models, reconstruction accuracy and robustness can be further enhanced in complex scenarios with high data loss rates. This will provide new theoretical foundations and practical guidance for applying deep learning to high-resolution seismic signal reconstruction.

CRediT authorship contribution statement

Yun Long: Project administration, Funding acquisition, Conceptualization. **Ji-Yuan Liu:** Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis. **Jian-Bin Zheng:** Writing – review & editing, Supervision. **Xin-Hao Zhang:** Data curation. **Shuang Yan:** Data curation. **Xiao-Kai Ma:** Software, Methodology, Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported in part by the General Program of the National Natural Science Foundation of China under Grant 42574187, and in part by the National Key Research and Development Program of China under Grant 2023YFC2811104.

References

- Abedi, M.M., Pardo, D., 2022. A multidirectional deep neural network for self-supervised reconstruction of seismic data. *IEEE Trans. Geosci. Rem. Sens.* 60, 1–9. <https://doi.org/10.1109/TGRS.2022.3227212>.
- Chang, D., Yang, W., Yong, X., Zhang, G., Wang, W., Li, H., Wang, Y., 2021. Seismic data interpolation using dual-domain conditional generative adversarial networks. *IEEE Geosci. Remote Sens. Lett.* 18, 1856–1860. <https://doi.org/10.1109/LGRS.2020.3008478>.
- Chen, Y., Chen, X., Wang, Y., Zu, S., 2019. The interpolation of sparse geophysical data. *Surv. Geophys.* 40, 73–105. <https://doi.org/10.1007/s10712-018-9501-3>.
- Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., Bharath, A.A., 2018. Generative adversarial networks: an overview. *IEEE Signal Process. Mag.* 35, 53–65. <https://doi.org/10.1109/MSP.2017.2765202>.
- Dou, Y., Li, K., Duan, H., Li, T., Dong, L., Huang, Z., 2023. MDA GAN: adversarial-learning-based 3-d seismic data interpolation and reconstruction for complex missing. *IEEE Trans. Geosci. Rem. Sens.* 61, 1–14. <https://doi.org/10.1109/TGRS.2023.3249476>.
- Fang, W., Fu, L., Liu, S., Li, H., 2021. Dealiased seismic data interpolation using a deep-learning-based prediction-error filter. *Geophysics* 86, V317–V328. <https://doi.org/10.1190/geo2020-0487.1>.
- Fomel, S., 2003. Seismic reflection data interpolation with differential offset and shot continuation. *Geophysics* 68, 733–744. <https://doi.org/10.1190/1.1567243>.
- Goyes-Pañafiel, P., Suárez-Rodríguez, L., Correa, C.V., Arguello, H., 2024. Gan supervised seismic data reconstruction: an enhanced learning for improved generalization. *IEEE Trans. Geosci. Rem. Sens.* 62, 1–10. <https://doi.org/10.1109/TGRS.2024.3434474>.
- Huang, W., Wang, R., Chen, Y., Li, H., Gan, S., 2016. Damped multichannel singular spectrum analysis for 3d random noise attenuation. *Geophysics* 81, V261–V270. <https://doi.org/10.1190/geo2015-0264.1>.
- Innocent Oboué, Y.A.S., Chen, W., Wang, H., Chen, Y., 2021. Robust damped rank-reduction method for simultaneous denoising and reconstruction of 5d seismic data. *Geophysics* 86, V71–V89. <https://doi.org/10.1190/geo2020-0032.1>.
- Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A., 2017. Image-to-image translation with conditional adversarial networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5967–5976. <https://doi.org/10.1109/CVPR.2017.632>.
- Jiang, P., Deng, F., Wang, X., Shuai, P., Luo, W., Tang, Y., 2023. Seismic first break picking through swin transformer feature extraction. *IEEE Geosci. Remote Sens. Lett.* 20, 1–5. <https://doi.org/10.1109/LGRS.2023.3248233>.
- Kong, F., Picetti, F., Lipari, V., Bestagini, P., Tang, X., Tubaro, S., 2022. Deep prior-based unsupervised reconstruction of irregularly sampled seismic data. *IEEE Geosci. Remote Sens. Lett.* 19, 1–5. <https://doi.org/10.1109/LGRS.2020.3044455>.
- Latif, A., Mousa, W.A., 2017. An efficient undersampled high-resolution radon transform for exploration seismic data processing. *IEEE Trans. Geosci. Rem. Sens.* 55, 1010–1024. <https://doi.org/10.1109/TGRS.2016.2618848>.
- Li, C., Liu, G., Wang, Z., Li, Z., Fomel, S., Chen, Y., 2025a. Simultaneous off-the-grid deblending and data reconstruction via unsupervised deep learning. *IEEE Trans. Geosci. Rem. Sens.* 63, 1–11. <https://doi.org/10.1109/TGRS.2025.3548644>.
- Li, C., Saad, O.M., Chen, Y., 2025b. Unsupervised deep learning for off-the-grid seismic reconstruction and denoising. *Geophysics* 90, V241–V254. <https://doi.org/10.1190/geo2024-0189.1>.
- Li, Z., Ren, M., Li, Q., Jiang, F., Sun, Z., 2024. Improving transferability of adversarial samples via critical region-oriented feature-level attack. *IEEE Trans. Inf. Forensics Secur.* 19, 6650–6664. <https://doi.org/10.1109/TIFS.2024.3404857>.
- Liner, C.L., 1997. On the history and culture of geophysics, and science in general. *Lead. Edge* 16, 1059–1060. <https://doi.org/10.1190/1.1437729>.
- Liu, W., Cao, S., Chen, Y., 2016. Seismic time-frequency analysis via empirical wavelet transform. *IEEE Geosci. Remote Sens. Lett.* 13, 28–32. <https://doi.org/10.1109/LGRS.2015.2493198>.
- Mirza, M., Osindero, S., 2014. Conditional generative adversarial nets. *ArXiv abs/1411.1784*. <https://doi.org/10.48550/arXiv.1411.1784>.
- Monk, D., 2002. Lena: a seismic model. *Lead. Edge* 21, 438–444. <https://doi.org/10.1190/1.1481249>.
- Mousavi, S.M., Ellsworth, W.L., Zhu, W., Chuang, L.Y., Beroza, G.C., 2020. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nat. Commun.* 11, 3952. <https://doi.org/10.1038/s41467-020-17591-w>.
- Pathak, D., Krähenbühl, P., Donahue, J., Darrell, T., Efros, A.A., 2016. Context encoders: feature learning by inpainting. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2536–2544. <https://doi.org/10.1109/CVPR.2016.278>.
- Porsani, M.J., 1999. Seismic trace interpolation using half-step prediction filters. *Geophysics* 64, 1461–1467. <https://doi.org/10.1190/1.1444650>.
- Radford, A., Metz, L., Chintala, S., 2015. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. <https://doi.org/10.48550/arXiv.1511.06434>.
- Ronen, J., 1987. Wave-equation trace interpolation. *Geophysics* 52, 973–984. <https://doi.org/10.1190/1.1442366>.
- Saad, O.M., Fomel, S., Abma, R., Chen, Y., 2022a. Unsupervised deep learning for 3d interpolation of highly incomplete data. *Geophysics* 88, WA189–WA200. <https://doi.org/10.1190/geo2022-0232.1>.
- Saad, O.M., Fomel, S., Abma, R., Chen, Y., 2022b. Unsupervised deep learning for 3d interpolation of highly incomplete data. *Geophysics* 88, WA189–WA200. <https://doi.org/10.1190/geo2022-0232.1>.
- Sacchi, M., Ulrych, T., Walker, C., 1998. Interpolation and extrapolation using a high-resolution discrete fourier transform. *IEEE Trans. Signal Process.* 46, 31–38. <https://doi.org/10.1109/78.651165>.
- Spitz, S., 1991. Seismic trace interpolation in the f-x domain. *Geophysics* 56, 785–794. <https://doi.org/10.1190/1.1443096>.
- Wang, B., Zhang, N., Lu, W., Wang, J., 2018. Deep-learning-based seismic data interpolation: a preliminary result. *Geophysics* 84, V11–V20. <https://doi.org/10.1190/geo2017-0495.1>.
- Wang, H., Zhang, J., Zhao, Z., Zhang, C., Long, L., Yang, Z., Geng, W., 2022. Automatic velocity picking using a multi-information fusion deep semantic segmentation network. *IEEE Trans. Geosci. Rem. Sens.* 60, 1–10. <https://doi.org/10.1109/TGRS.2022.3188669>.
- Wang, J., Ng, M., Perz, M., 2010. Seismic data interpolation by greedy local radon transform. *Geophysics* 75, WB225–WB234. <https://doi.org/10.1190/1.3484195>.
- Wang, S., Deng, F., Jiang, P., Gong, X., Wei, X., Wang, Y., 2024. Seisfusion: constrained diffusion model with input guidance for 3-d seismic data interpolation and reconstruction. *IEEE Trans. Geosci. Rem. Sens.* 62, 1–15. <https://doi.org/10.1109/TGRS.2024.3462414>.
- Wang, Y., Wang, B., Tu, N., Geng, J., 2020. Seismic trace interpolation for irregularly spatial sampled data using convolutional autoencoder. *Geophysics* 85, V119–V130. <https://doi.org/10.1190/geo2018-0699.1>.
- Wei, Q., Li, X., Song, M., 2021a. De-aliased seismic data interpolation using conditional wasserstein generative adversarial networks. *Comput. Geosci.* 154, 104801. <https://doi.org/10.1016/j.cageo.2021.104801>.
- Wei, Q., Li, X., Song, M., 2021b. Reconstruction of irregular missing seismic data using conditional generative adversarial networks. *Geophysics* 86, V471–V488. <https://doi.org/10.1190/geo2020-0644.1>.
- Woo, S., Park, J., Lee, J.Y., Kweon, I.S., 2018. Cbam: convolutional block attention module. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (Eds.), *Computer Vision – ECCV 2018*. Springer International Publishing, Cham, pp. 3–19. https://doi.org/10.1007/978-3-030-01234-2_1.
- Wu, G., Liu, Y., Liu, C., Zheng, Z., Cui, Y., 2023. Seismic data interpolation using deeply supervised u-net++ with natural seismic training sets. *Geophys. Prospect.* 71, 227–244. <https://doi.org/10.1111/1365-2478.13307>.
- Xiao, J., Zhang, J., Li, Z., 2025. Global grouped coordinate attention for transformer in pedestrian trajectory prediction. *Signal Image Video Process.* 19, 714. <https://doi.org/10.1007/s11760-025-04299-x>.
- Yu, J., Wu, B., 2022. Attention and hybrid loss guided deep learning for consecutively missing seismic data reconstruction. *IEEE Trans. Geosci. Rem. Sens.* 60, 1–8. <https://doi.org/10.1109/TGRS.2021.3068279>.