ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

Multi-stage progressive network for seismic random noise suppression

Guanghui Li*, Huiwei Li, Shoufeng He, and Li Wang

Department of Electronic Information Engineering, College of Physics and Electronic Engineering, Shanxi University, Taiyuan, Shanxi, China

Abstract

Seismic data quality frequently deteriorates due to random noise contamination, substantially impeding subsequent processing and geological interpretation. While deep learning approaches have emerged as powerful tools for noise suppression, conventional single-stage architectures exhibit inherent limitations in handling complex seismic features while preserving subtle geological details. These challenges motivate the development of advanced multi-stage neural networks for seismic data enhancement. The proposed multi-stage progressive U-shaped convolutional network (MPU-Net) architecture addresses these limitations through supervised cross-stage attention mechanisms that maintain feature connectivity throughout the network. Building upon this foundation, group enhanced convolutional blocks (GEB)-MPU-Net introduces GEB to specifically counteract the progressive attenuation of shallow features in deep networks. This dual-stage enhancement strategy combines hierarchical feature preservation, adaptive information fusion, and stable gradient propagation. Comprehensive evaluation using both synthetic and field datasets demonstrates GEB-MPU-Net's superior performance compared to conventional time-frequency analysis methods and established networks, such as U-Net, residual dense network, residual dense block U-Net, and MPU-Net. The architecture consistently achieves enhanced reflection continuity, improved geological feature resolution, and robust noise suppression. These advancements provide more reliable input for seismic interpretation, better preservation of subtle stratigraphic features, and increased applicability to challenging field conditions.

Keywords: Noise suppression; Deep learning; Multi-stage networks; Seismic exploration

***Corresponding author:**
Guanghui Li
(ligh1986@sxu.edu.cn)

Citation: Li G, Li H, He S, Wang L. Multi-stage progressive network for seismic random noise suppressing. *J Seismic Explor.* 2025;34(1): 43-59. doi: 10.36922/JSE025240011

Received: June 10, 2025

1st revised: July 22, 2025

2nd revised: July 26, 2025

Accepted: August 6, 2025

Published online: August 14, 2025

Copyright: © 2025 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

The focus of seismic exploration has progressively transitioned to complex structural traps, deep-buried reservoirs, and unconventional hydrocarbon systems as conventional resources become increasingly depleted. This evolution demands seismic data of substantially improved quality. Nevertheless, field-acquired seismic records are invariably contaminated by ambient noise originating from diverse environmental and operational factors, significantly compromising both subsurface imaging resolution and geological interpretation accuracy. Consequently, noise suppression and signal-to-noise ratio (SNR) enhancement remain fundamental challenges in modern seismic data processing.

Traditional seismic denoising approaches primarily rely on mathematical transformations and filtering techniques, including Fourier transforms,[1] wavelet transforms,[2] Curvelet transforms,[3,4] Seislet transforms,[5] empirical mode decomposition,[6] variational mode decomposition,[7,8] low-rank approximation,[9] compressed sensing,[10] and dictionary learning.[11] Although these methods have proven effective for certain types of noise, they exhibit several inherent limitations. A primary challenge lies in their limited adaptability to handle diverse noise distributions, particularly in complex geological settings or unconventional reservoirs. Furthermore, their performance typically depends on manual parameter tuning, which may lead to suboptimal results when processing conditions vary.

Recent advances in artificial intelligence have established neural networks as powerful tools for seismic data denoising, offering substantial improvements over conventional approaches. In particular, deep convolutional neural networks[12,13] and their numerous enhanced algorithms demonstrate superior performance in both seismic noise suppression and signal preservation, exhibiting enhanced robustness to noise variability and improved generalization across diverse geological settings compared to traditional transform-based methods, such as the self-supervised framework,[14] the modular convolutional neural network that incorporates multi-scale attention mechanisms,[15] the singular value decomposition combined with deep learning,[16] the residual dense blocks integrated with time-frequency analysis,[17] and the advanced U-shaped convolutional network (U-Net) architectures through Atropos convolutions and dense connections.[18,19]

However, these single-stage architectures frequently exhibit suboptimal trade-offs between multi-scale representation and spatial precision when processing field seismic data, particularly in scenarios involving complex noise distributions, low SNR, and subtle geological features. Recent developments in seismic denoising have seen the emergence of multi-stage unsupervised and self-supervised deep learning approaches. For instance, the multi-stage progressive U-Net (MPU-Net),[20] the self-supervised multi-stage network,[21] and the two-step deep image prior model.[22] These methods significantly reduce reliance on annotated training data while demonstrating robust performance across varied noise conditions. Nevertheless, the lack of explicit supervisory signals presents inherent limitations, particularly in reliably differentiating subtle seismic reflections from background noise and maintaining stable performance under diverse geological settings.

Based on this, we present a novel group enhanced convolutional blocks (GEB) MPU-Net (GEB-MPU-Net) architecture for seismic data denoising, which innovatively integrates GEB[23] within the MPU-Net framework. This synthesis enhances feature representation while maintaining the structural advantages of progressive processing. GEB significantly improves multi-scale feature integration through its unique combination of residual learning and grouped feature extraction. By incorporating channel attention blocks (CABs), the architecture further enhances long-range feature propagation and representation, leading to more stable and effective seismic denoising. This design achieves comprehensive feature fusion through systematic GEB-CAB integration at each processing stage. The framework strengthens low-frequency feature representation through inter-channel correlation analysis while implementing signal enhancement mechanisms to maintain critical long-range dependencies. Importantly, this approach successfully resolves the persistent shallow information loss problem inherent in conventional MPU-Net architectures.

Compared to other multi-stage unsupervised and self-supervised deep learning approaches,[22-25] first, GEB-MPU-Net's tight integration with multi-stage supervised learning enables precise feature optimization through explicit signal-noise differentiation. This coupled framework systematically enhances discriminative feature extraction while preserving structural relationships across processing stages. Second, the GEB module incorporates a specialized channel attention mechanism designed to optimize shallow-deep feature integration and enhance low-frequency representation. This tailored architecture enables GEB-MPU-Net to surpass conventional unsupervised approaches employing standard attention modules, demonstrating superior denoising robustness and generalization capacity. The processing results of synthetic and field seismic data both demonstrate their superior performance and signal fidelity, particularly in complex and challenging environments. Overall, GEB-MPU-Net represents a dual advancement in seismic denoising methodology, introducing both architectural innovations and demonstrable improvements in signal preservation. The framework surpasses existing approaches through its integrated design, achieving superior noise suppression while maintaining critical geological features.

2. Algorithm principle

2.1. U-shaped convolutional network structure

U-shaped convolutional network's symmetric encoder-decoder framework utilizes cross-connection pathways to maintain its characteristic U-topology while enabling multi-scale feature integration. Figure 1 presents the standard implementation of this architecture. The first

half of the network focuses on feature extraction, while the latter half emphasizes up-sampling. Specifically, the encoding section employs a series of 3 × 3 convolutional layers, rectified linear unit (ReLU) activation functions, and 2 × 2 max pooling layers to extract features from the input image. With each down-sampling operation, the dimensions of the feature maps are halved, while the number of channels is doubled. The primary function of the decoding section is to utilize transposed convolutional layers to reconstruct high-resolution representations from the encoded low-resolution features. Skip connections between the contracting and expanding channels facilitate the fusion of low-resolution and high-resolution features, enabling a more effective capture of both local and global characteristics of the image.

2.2. MPU-Net structure

As shown in Figure 2, the network employs a multi-patch hierarchical decomposition strategy, where input seismic data undergoes non-overlapping patch segmentation. This preprocessing stage enables localized feature extraction while maintaining structural relationships across spatial domains. It implements the strategy across three progressive stages: initial coarse segmentation into

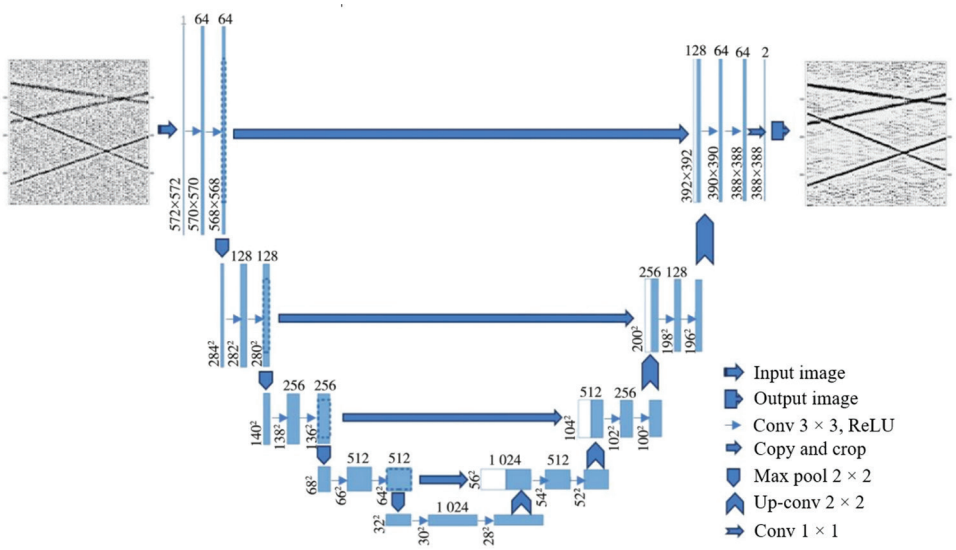

Figure 1. Structure of a U-shaped convolutional network
Abbreviations: conv: Convolution; ReLU: Rectified linear unit

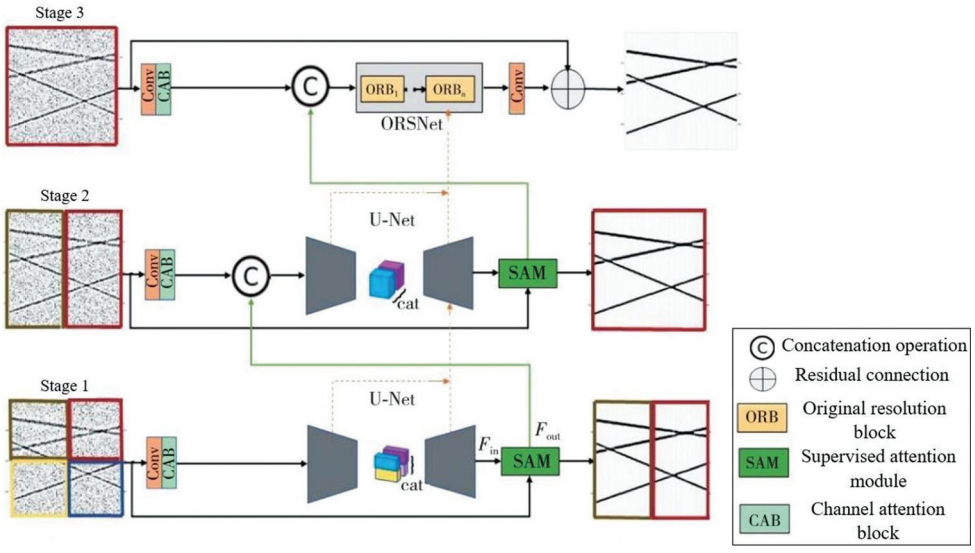

Figure 2. Multi-stage progressive U-Net structure
Abbreviations: cat: Concatenate; Conv: Convolution; ORSNet: Original resolution subnetwork; U-Net: U-shaped convolutional network

four patches, intermediate division into two patches, and final processing of the full-resolution input. The initial processing stages implement an enhanced U-Net framework, where each encoder-decoder level incorporates dual CABs. This design enables efficient multi-scale feature extraction while maintaining spatial relationships across different resolution levels. The attention mechanisms selectively emphasize informative channels, optimizing feature representation throughout the network hierarchy. Conventional transposed convolution operations in U-Net decoders frequently generate checkerboard artifacts due to uneven kernel overlap patterns. These artifacts manifest as spurious seismic events in processed records, particularly adjacent to genuine reflections, potentially compromising interpretation accuracy. This phenomenon motivates the development of more stable up-sampling alternatives for seismic data restoration. To mitigate this limitation, MPU-Net modifies the conventional U-Net architecture by implementing bilinear interpolation for up-sampling, coupled with subsequent convolutional layers for spatial feature restoration. This adaptation reduces artifacts while maintaining resolution fidelity during the decoding phase. The original resolution subnetwork (ORSNet) consists of three original resolution blocks, which are connected to the input in the third stage.

Figure 3 illustrates the original resolution block module's composition, featuring eight CABs integrated with a final convolutional layer. The CAB consists of four 3×3 convolutional layers, one global average pooling layer, and three activation functions designed to enhance the representational capability of valuable features. Since ORSNet does not perform down-sampling, it retains high-resolution spatial details. As indicated by the dashed lines in Figure 2, the three stages are not independent; rather, a supervised attention module (SAM) is incorporated between each pair of stages to weigh the significant features. These features are then closely cascaded through a cross-stage feature fusion (CSFF) process. While all other convolutional layers in the MPU-Net architecture are 3×3, the convolutional layers in the SAM and CSFF are 1×1.

Cross-stage feature fusion mechanisms are introduced between the U-Net of the first and second stages, as well as between the U-Net of the second stage and the ORSNet of the third stage, as illustrated separately in Figure 4.

The architecture processes encoder and decoder outputs through parallel 1×1 convolutional layers for dimensional refinement and feature conditioning. These optimized feature maps subsequently undergo cross-level fusion, creating an enriched representation for stage-transition processing. This dual-path approach maintains feature integrity while enabling information exchange across network depths. The CSFF mechanism enables systematic integration of multi-scale features throughout the network hierarchy. This architecture provides three key advantages: (i) Preservation of critical information across processing stages, (ii) enhanced model robustness through diversified feature representation, and (iii) flexible network optimization via adjustable stage connectivity. The improved inter-stage information flow additionally facilitates architectural diagnostics and stage-number optimization during network development.

A SAM, illustrated in Figure 5, is integrated at the end of each encoder-decoder sub-network during the first two stages.

The SAM module processes incoming features through a 1×1 convolutional operation to produce residual representations. The network utilizes the learned residual representations to systematically attenuate noise components in the input seismic data. The processed seismic data undergoes feature optimization through a 1×1 convolutional layer with sigmoidal activation.

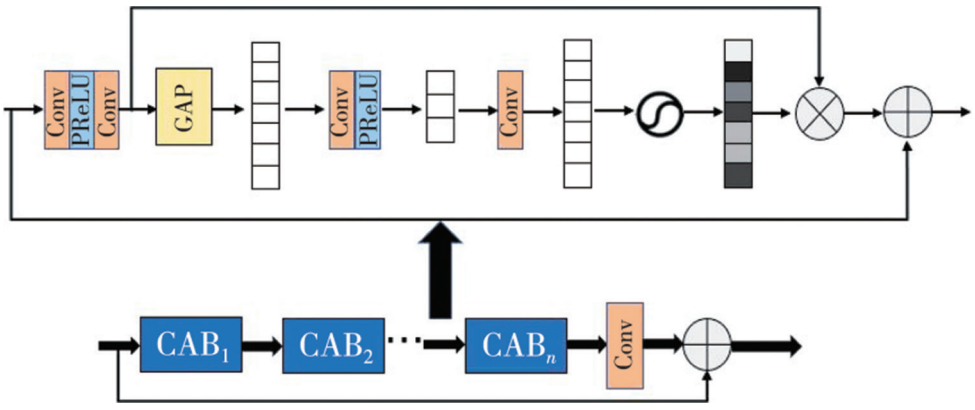

Figure 3. Structure of the original resolution block
Abbreviations: CAB: Channel attention block; Conv: Convolution; GAP: Global average pooling; PReLU: Parametric rectified linear unit

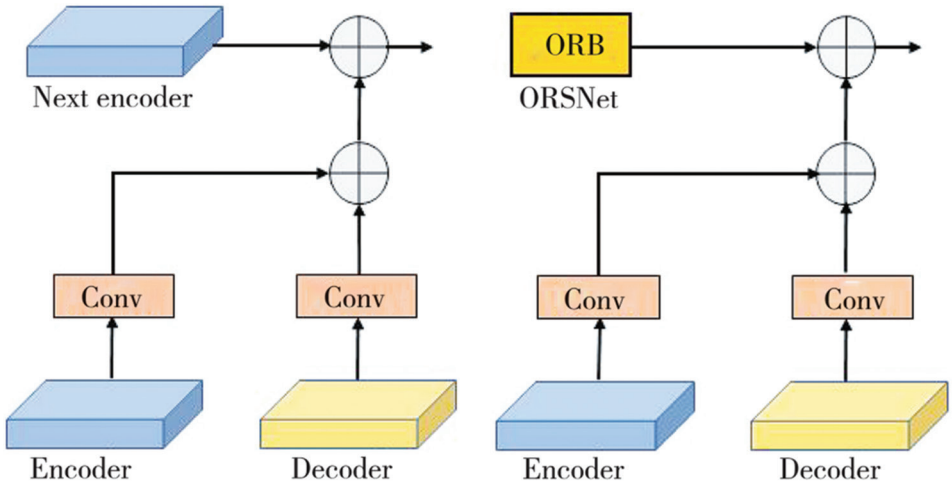

Figure 4. Structure of CSFF. (A) CSFF between stage 1 and 2; (B) CSFF between stage 2 and 3
Abbreviations: Conv: Convolution; CSFF: Cross-stage feature fusion; ORB: Original resolution block; ORSNet: Original resolution subnetwork

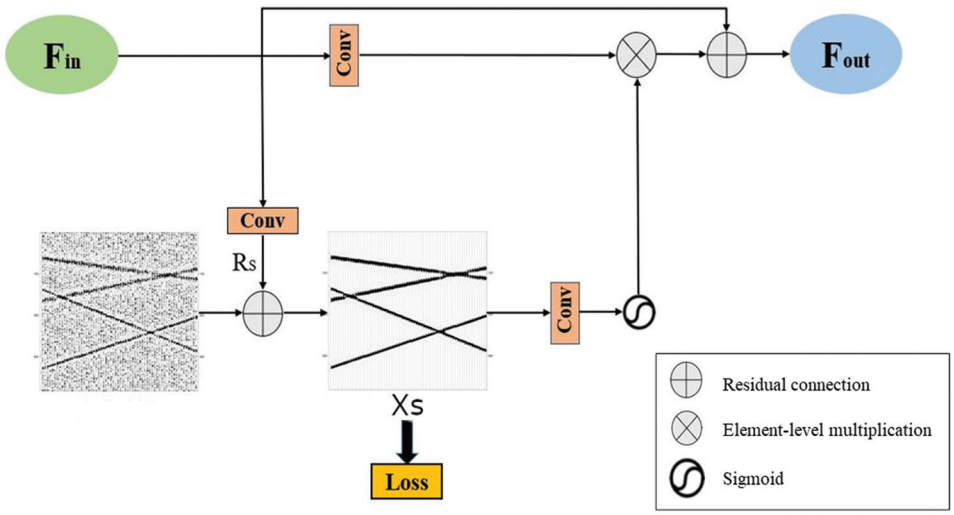

Figure 5. Structure of a supervised attention module
Abbreviations: Conv: Convolution; F_{in}: Input feature; F_{out}: Output feature; Rs: Residual representation; Xs: Input seismic feature

This operation performs channel-wise feature rescaling, suppression of non-informative components, and selective propagation of geophysically valid features. This gating mechanism ensures only the most salient signal characteristics propagate through the network hierarchy, substantially improving the denoising efficacy while maintaining geological plausibility. The three stages are tightly integrated through the SAM and CSFF, leveraging feature information extracted from each stage to achieve improved learning outcomes.

2.3. GEBs module structure

Figure 6 illustrates the structure of the GEB. The block comprises two components: the known extraction portion (GConv1), which encompasses one-fifth of the feature channels from the current convolutional layer, and the remaining portion (GConv2), which includes four-fifths of the feature channels from the same layer. In each GEB, GConv1 serves as a convolutional layer with 16 input channels, 16 output channels, and a filter size of 3 × 3, whereas GConv2 has an input and output channel count of 60, also utilizing a 3 × 3 filter size. The remaining portion is utilized as the input for the subsequent convolutional layer in the main network, facilitating the extraction of additional deep features.

To enhance the expressiveness of low-frequency features, the GEB module employs a fusion mechanism between every two adjacent GConv2 layers. Specifically,

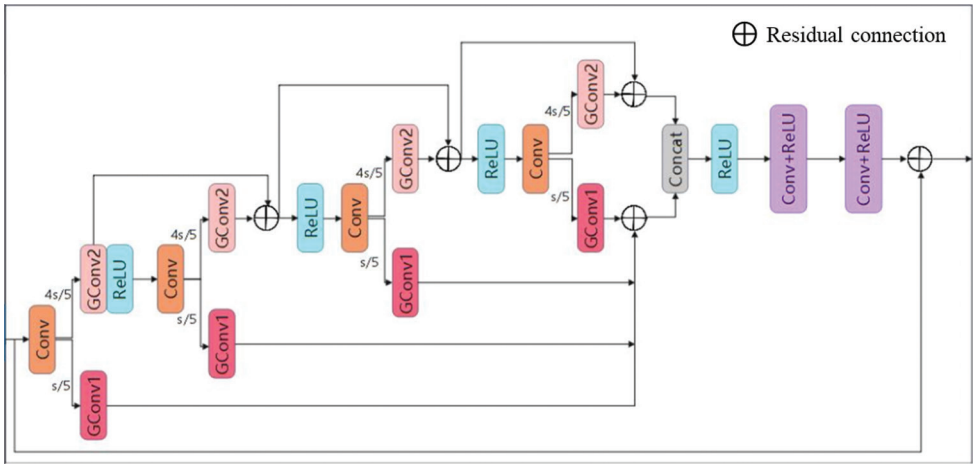

Figure 6. Structure of the group enhanced convolutional blocks
Abbreviations: Concat: Concatenation operation; Conv: Convolution; GConv1: Known extraction portion; GConv2: Remaining portion; ReLU: Rectified linear unit

for each pair of consecutive GConv2 layers within the GEB, their outputs are combined using a residual connection (Figure 6). This process enables the network to effectively aggregate deep neighborhood information across layers, thereby capturing broader contextual dependencies and improving the representation of low-frequency features. The network architecture establishes progressive feature interdependence, where each subsequent extraction module builds upon transformed representations from preceding stages. It provides two key advantages, including cascading information refinement through the network depth and complementary broad-context supplementation to deep features. The resulting multi-scale integration enhances denoising performance by simultaneously preserving both local details and global seismic characteristics. The feature processing within the GEB can be categorized into the following four steps[21]:

(i) Step 1. To enhance the features of neighboring layers and improve the accuracy of deep features across various channels, the features from two adjacent GConv2 layers are fused using a residual learning strategy. This combined information is subsequently fed into the next convolutional layer. While the outputs of the later GConv1 layers are derived from the preceding GConv2 layers, the output of the first GConv1 is derived from one-fifth of the output channels of the initial convolutional layer. Specifically, linear features are transformed into non-linear ones through the connection of the upper GConv2 to a ReLU activation function. These non-linear characteristics then act as a convolutional layer that learns additional low-frequency features; GConv1 receives the output data from the final one-fifth of the channels.

(ii) Step 2. By employing the residual learning technique, the features extracted from all GConv1 layers are integrated to enhance the connections among the various extraction segments.

(iii) Step 3. To obtain additional complementary features, the outputs from the final GConv1 and GConv2 are integrated along the channel dimension using a concatenation operation, as indicated by "Concat" in Figure 6.

(iv) Step 4. To address the limitation of shallow feature memory capacity across the network, we employ the concept of signal augmentation to preserve long-distance features. This approach entails superimposing shallow features, obtained through the residual learning technique, onto the deep features acquired, thereby enhancing the significance of the shallow features.

2.4. GEB-MPU-Net structure

The progressive depth of MPU-Net may compromise shallow feature retention, potentially limiting its contribution to final representations. To mitigate this limitation, GEBs are systematically integrated following each CAB module across three critical processing stages, as detailed in Figure 7. This architecture optimizes shallow feature integration while effectively attenuating noise through enhanced low-frequency representation. It strengthens inter-channel correlations, thereby improving feature discrimination across multiple scales without compromising signal integrity.

An ablation experiment was conducted on the network architecture to justify the choice of architecture and demonstrate its optimality. Figure 7 depicts Strategy 1's organizational framework, in which the GEB module is placed after the CAB modules in the three stages of

MPU-Net. Figure 8 illustrates Strategy 2's organizational structure. The GEB module is placed respectively after the encoder-decoder modules of the first and second stages, as well as the ORSNet + Conv of the third stage. Figure 9 shows the construction of Strategy 3, in which the CSFF module is followed by the GEB module.

The synthetic records with varying noise levels were denoised using the three different GEB-MPU-Net network structures. Figures 10 and 11 display the denoising and residual results, respectively. Figure 10 reveals characteristic waveform distortion at the intersection of Strategy 2 and Strategy 3's in-phase axes, as highlighted by the red annotation box. The residual data presented in Figure 11 indicate superior amplitude retention in Strategy 1, as evidenced by reduced phase-coherent artifacts compared to Strategies 2 and 3. Table 1 presents the peak SNR and mean squared error (MSE) of the denoising outcomes. The results demonstrate Strategy 1's superior performance in both noise suppression and amplitude preservation across varying noise conditions,

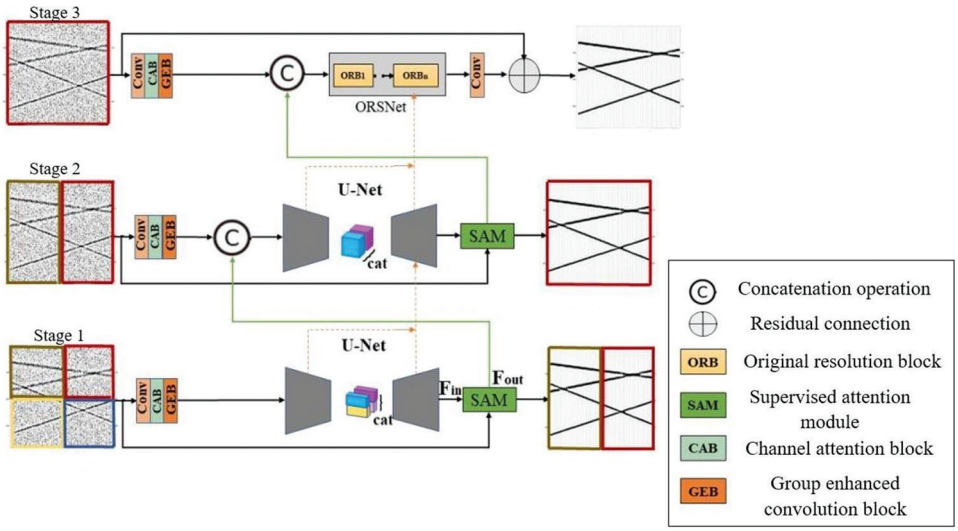

Figure 7. Structure of group enhanced convolution block, multi-stage progressive U-shaped convolutional network of Strategy 1
Abbreviations: cat: Concatenate; Conv: Convolution; U-Net: U-shaped convolutional network

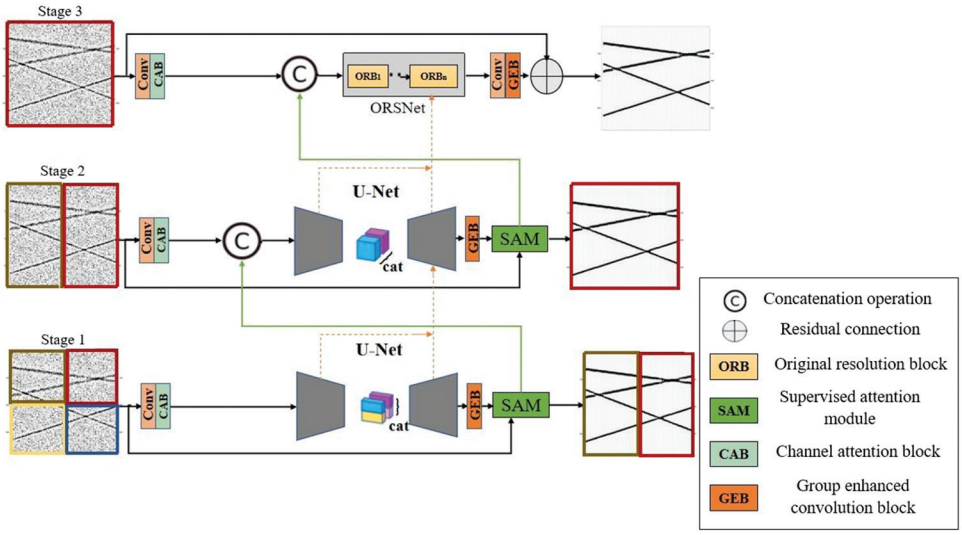

Figure 8. Structure of group enhanced convolution block, multi-stage progressive U-shaped convolutional network of Strategy 2
Abbreviations: cat: Concatenate; Conv: Convolution; U-Net: U-shaped convolutional network

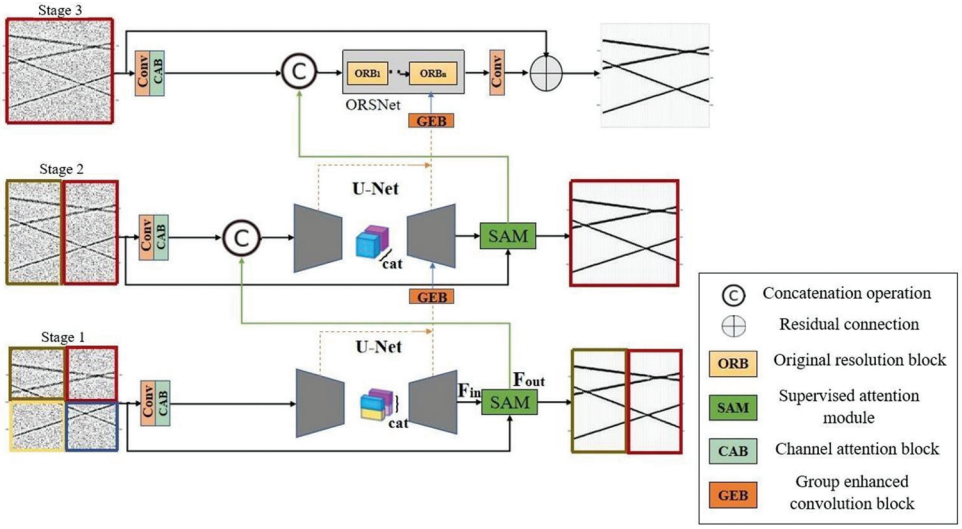

Figure 9. Structure of group enhanced convolution block, multi-stage progressive U-shaped convolutional network of Strategy 3
Abbreviations: cat: Concatenate; Conv: Convolution; U-Net: U-shaped convolutional network

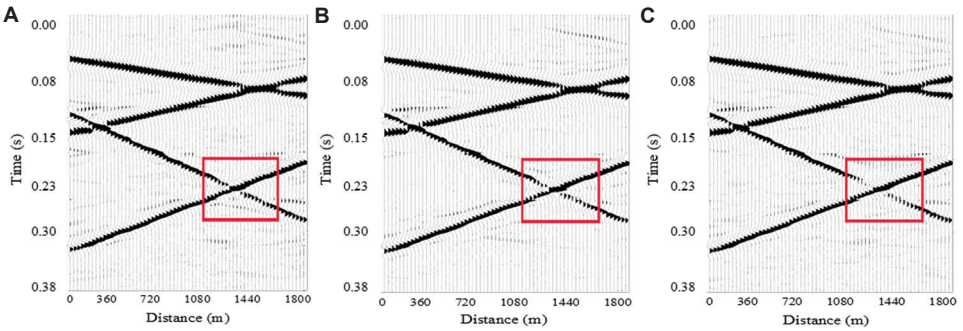

Figure 10. Denoising results of the three strategies. (A) Strategy 1. (B) Strategy 2. (C) Strategy 3

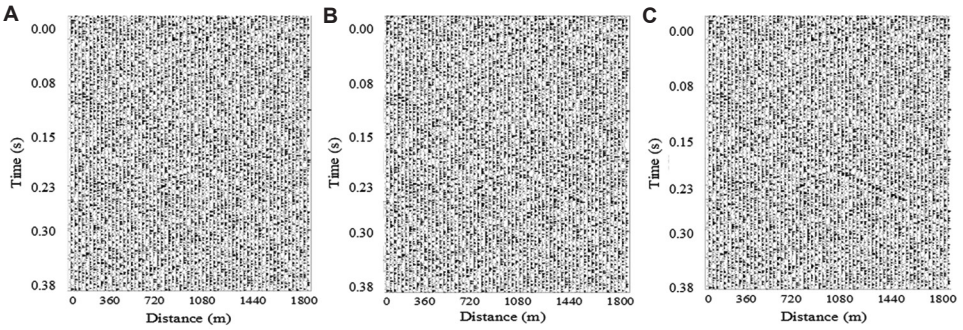

Figure 11. Residual results of the three strategies. (A) Strategy 1. (B) Strategy 2. (C) Strategy 3

exhibiting consistently favorable metric values compared to alternative approaches.

Figure 12A and B present the comparative evaluation metrics across the simulated dataset after completing training. The results reveal Strategy 1's consistent advantage in signal preservation, with both alternative strategies demonstrating relatively reduced performance across the measured parameters.

2.5. GEB-MPU-Net denoising principle

The process of removing noise from seismic recordings that contain a combination of signals and noise is referred

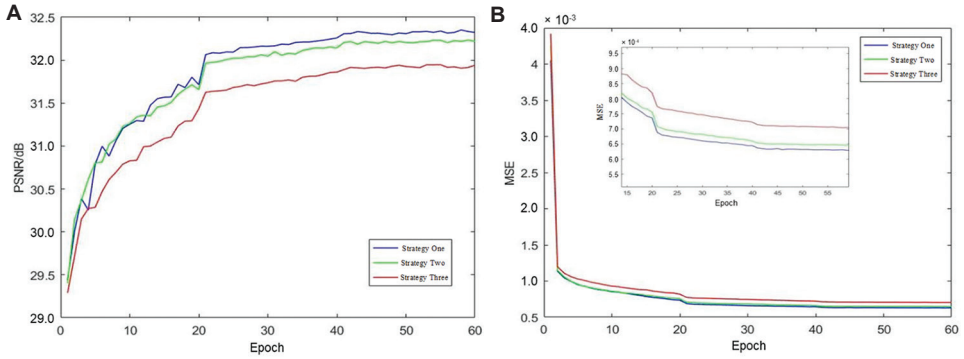

Figure 12. Influence of three strategies on denoising. (A) The PSNR performance curve on the test set. (B) The MSE performance curve on the test set
Abbreviations: MSE: Mean squared error; PSNR: Peak signal to noise ratio

Table 1. Comparison of parameters after denoising of three strategies

Strategies	Signal-to-noise ratio (dB)	Mean squared error
Noisy signal	−5.1753	0.0161
	−8.1856	0.0321
	−9.9465	0.0482
Strategy 1	11.8588	3.1806 e−04
	9.1082	5.9920 e−04
	7.2832	9.1215 e−04
Strategy 2	11.5535	3.4123 e−04
	8.6248	6.6975 e−04
	6.7186	0.0010
Strategy 3	11.5279	3.4324 e−04
	8.1443	7.4810 e−04
	6.0058	0.0012

to as random noise suppression in seismic data. Equation I represents the noisy seismic data.

$$d = x + n \quad \text{(I)}$$

where d denotes the noisy data, x represents the clean seismic signal, and n signifies the random noise. In this study, we employed a noise learning technique by inputting the noisy seismic data into the GEB-MPU-Net neural network. Through the process of residual learning, the network was trained to predict the noise, which was then subtracted from the input noisy seismic records to yield denoised seismic data. The specific procedure is outlined in Equations II and III.

$$Nt = R\,(d;\theta) \quad \text{(II)}$$

$$\hat{x} = d - Nt \quad \text{(III)}$$

where Nt represents the predicted noise output by the network, R denotes the residual mapping process, θ encompasses the parameters of the network, including weights ω and biases b, and $\hat{x}$ signifies the predicted seismic record. We utilized the MSE between the pure noise and the predicted noise as the loss function to optimize the parameters. The formulation for the loss function is shown in Equation IV.

$$L_{loss}(\theta) = \frac{1}{M}\sum_{i=1}^{M}\left\| R\,(d_i;\theta) - n_i \right\|^2_{GEB-MPU-Net} \quad \text{(IV)}$$

where M is the number of samples in the training set, $\|\cdot\|^2_{GEB-MPU-Net}$ denotes the Frobenius norm, d_i represents the noisy seismic data, and n_i is the pure noise. The objective of the network's continuous training is to minimize the loss function, a non-negative real-valued function. A smaller loss indicates a reduced error between the predicted noise and the actual noise, leading to denoised seismic records that closely approximate the ideal seismic records.

The comprehensive dataset from the 1994 Canadian reverse masking experiment, named Model94_shots.segy, served as the training data for the network. This dataset comprises 277 shots, each containing 480 recording channels, with a channel spacing of 15 m and a shot interval of 90 m. Following manual processing, the dataset demonstrated a high SNR and has been widely recognized as a representative clean record. Users can add varying levels of Gaussian white noise or real noise according to their requirements. After normalization, the data were segmented into 64 × 64 patches using a sliding window with a step size of 32, resulting in a total of 41,776 samples, with 31,656 samples designated for training and 10,120 samples for testing. The Adam optimization algorithm was utilized during training, with MSE as the loss function. The learning rate gradually decreased from 2×10^{-4} to 2×10^{-6}, the batch size was set to 8, and the number of training epochs was established at 60. The experiments were conducted within the PyTorch deep learning framework (version 2.2.0), operating on a Windows 11 system. The computations

were performed on a server equipped with an Intel(R) Core(TM) i5-12500H processor, 16 GB of RAM, CUDA 11.6, and an NVIDIA RTX 3050 Ti graphics card. The specific experimental steps were as follows:

(i) Step 1: Prepare the seismic signal dataset and conduct preprocessing
(ii) Step 2: Introduce noise into the clean seismic records and train the network using GEB-MPU-Net
(iii) Step 3: Adjust the network hyperparameters to ensure that the network output closely approximates the added noise
(iv) Step 4: Subtract the predicted noise output from the noisy records to obtain the denoised seismic records
(v) Step 5: Evaluate the trained network using both the noisy synthetic seismic records and actual seismic data
(vi) Step 6: Illustrate the frequency-wavenumber spectra of the denoised synthetic recordings and compare the time-frequency domain waveforms of the single-channel records
(vii) Step 7: Analyze the denoising performance in comparison to time-frequency analysis (TFPF), conventional U-Net, residual dense network (RDNet), residual dense block U-Net (RDBU-Net), and MPU-Net.

3. Experimental results

3.1. Synthetic records processing results

The clear synthetic seismic record comprises four distinct cross-seismic events distributed across 61 channels, with each trace containing 384 temporally sampled points at 1 ms intervals, as shown in Figure 13A. The Ricker wavelets exhibited characteristic dominant frequencies of 40 Hz and 60 Hz, representing typical exploration scenarios. The synthetic noisy record shown in Figure 13B was generated

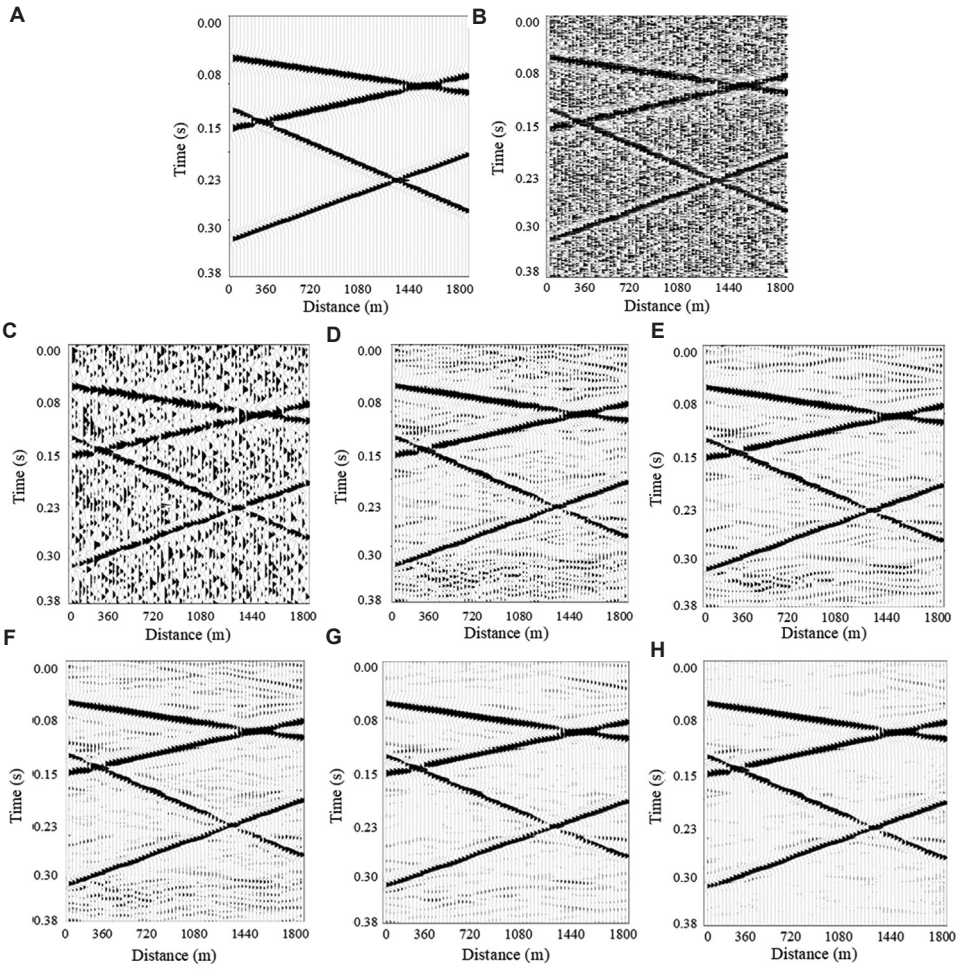

Figure 13. Denoising results of synthetic seismic records. (A) Pure record. (B) Noisy record. (C) TFPF. (D) U-Net. (E) RDNet. (F) RDBU-Net. (G) MPU-Net. (H) GEB-MPU-Net

Abbreviations: GEB-MPU-Net: Group enhanced convolutional blocks-multi-stage progressive U-Net; MPU-Net: Multi-stage progressive U-Net; RDBU-Net: Residual dense block U-Net; RDNet: Residual dense network; TFPF: Time-frequency analysis; U-Net: U-shaped convolutional network

through additive Gaussian white noise at a level of 85% contamination of the pristine dataset. The noisy record was processed through multiple denoising approaches, encompassing conventional TFPF, established neural networks (U-Net, RDNet), and advanced architectures (RDBU-Net, MPU-Net, GEB-MPU-Net). Figure 13C-H displays the processed outputs from each denoising method, with corresponding residual patterns shown in Figure 14. Figure 13C presents the results obtained by TFPF, demonstrating partial noise attenuation while retaining visible signal components in the corresponding residuals, as shown in Figure 14A. This outcome highlights the method's fundamental limitation in achieving complete signal-noise separation, with discernible seismic events persisting in the residual domain. Figures 13D and 14B present the U-Net processed results and the corresponding residual record, demonstrating significantly improved noise suppression capabilities compared to conventional approaches. The neural network output exhibits enhanced signal clarity while effectively attenuating both random and coherent noise components. Figure 13E and F present the processed outputs from RDNet and RDBU-Net, and Figure 14C and D show the corresponding residuals, respectively, demonstrating superior noise suppression compared to the baseline U-Net architecture. Both advanced networks showed progressively improved seismic event visibility, with enhanced signal-background differentiation by the processed records.

Figures 13G and 14E illustrate the denoising outcome and residuals of MPU-Net, where the seismic events are more distinct, and the noise is substantially suppressed, although some distortion occurs at the intersection points of the events. Figure 13H presents the GEB-MPU-Net processed results, demonstrating exceptional signal clarity and waveform coherence in the reconstructed seismic record. The corresponding residuals in Figure 14F show negligible seismic event remnants, indicating near-complete signal preservation and noise separation.

Figure 15A-H presents comparative frequency-wavenumber transformations of the pure record, noisy record, and all processed outputs. This comprehensive spectral analysis enables detailed evaluation of wavenumber-frequency characteristics across different denoising approaches. Figure 15C shows the frequency-wavenumber spectrum following TFPF, revealing characteristic spectral overlap between residual noise components and preserved signal energy. While the method demonstrated partial noise suppression in certain frequency-wavenumber bands, significant signal-noise ambiguity persisted across critical regions of the spectrum. Figure 15D-F demonstrates that U-Net, RDNet, and RDBU-Net achieve substantial random noise suppression in the wavenumber-frequency domain. However, these architectures showed limited effectiveness against persistent low-frequency noise components, revealing a common challenge in neural network-based seismic

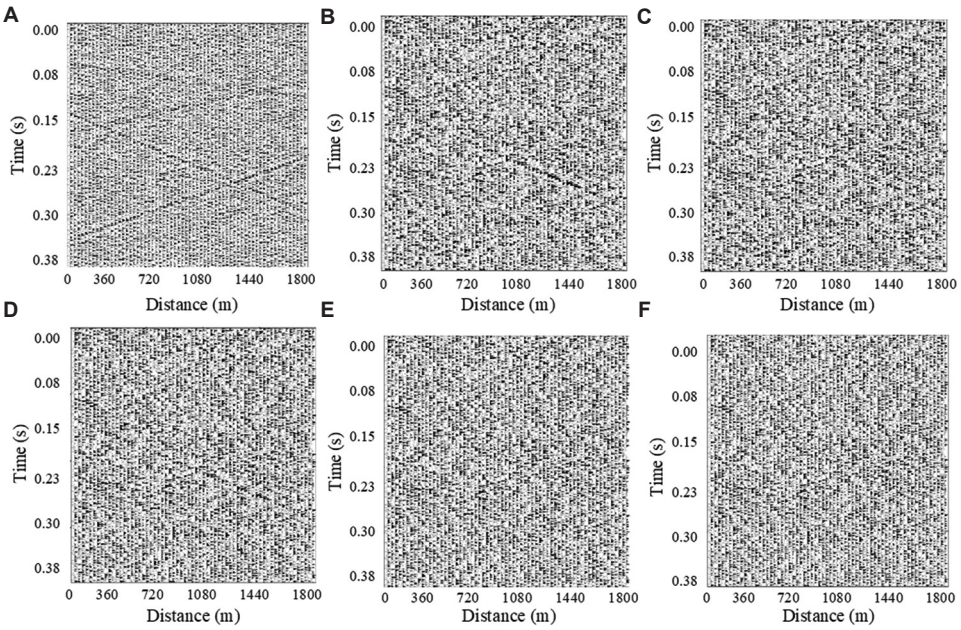

Figure 14. Residual results of synthetic seismic records. (A) TFPF. (B) U-Net. (C) RDNet. (D) RDBU-Net. (E) MPU-Net. (F) GEB-MPU-Net
Abbreviations: GEB-MPU-Net: Group enhanced convolutional blocks-multi-stage progressive U-Net; MPU-Net: Multi-stage progressive U-Net; RDBU-Net: Residual dense block U-Net; RDNet: Residual dense network; TFPF: Time-frequency analysis; U-Net: U-shaped convolutional network

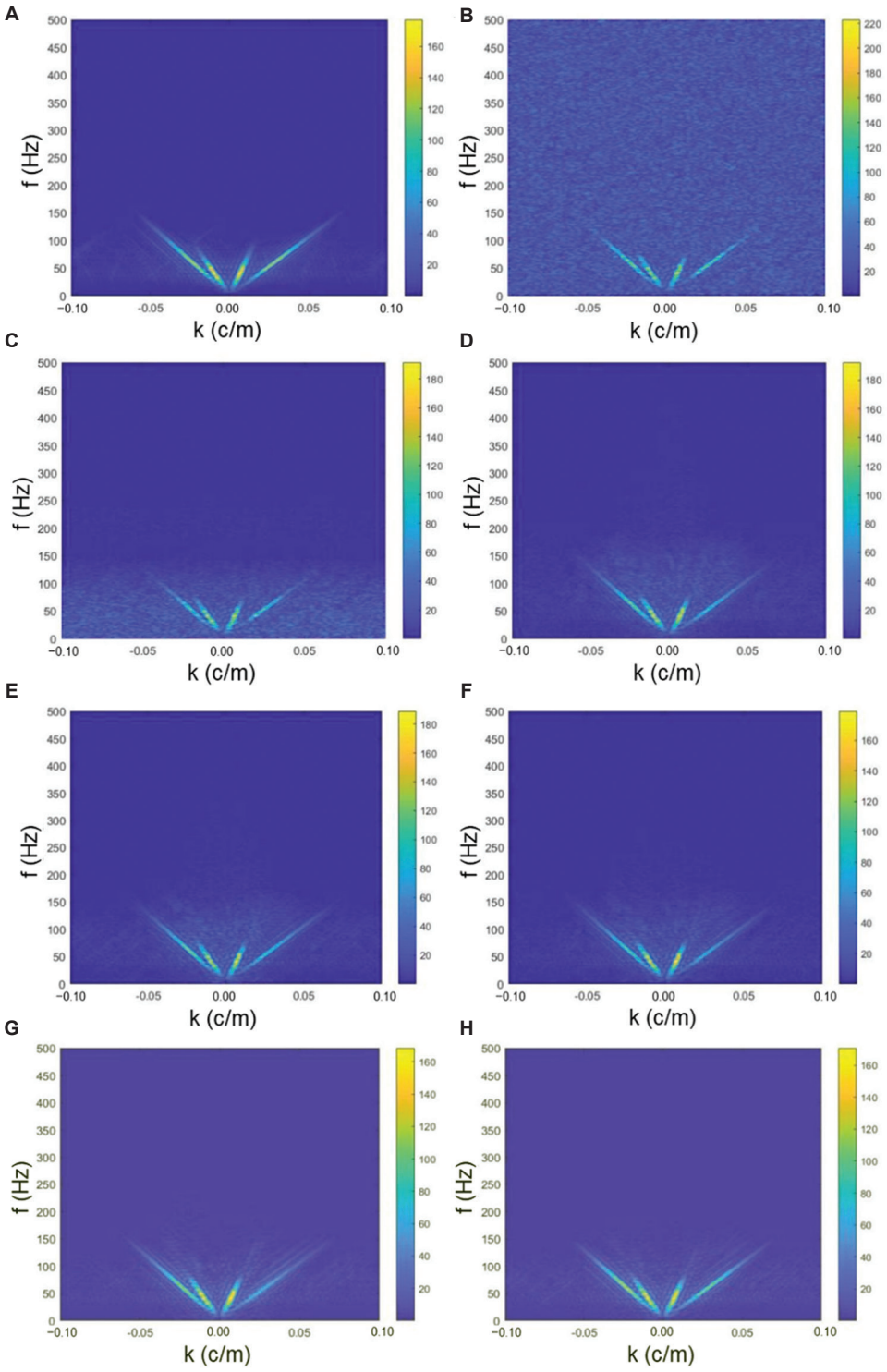

Figure 15. Frequency-wavenumber spectra of synthetic seismic records. (A) Pure record. (B) Noisy record. (C) TFPF. (D) U-Net. (E) RDNet. (F) RDBU-Net. (G) MPU-Net; (H) GEB-MPU-Net

Abbreviations: GEB-MPU-Net: Group enhanced convolutional blocks-multi-stage progressive U-Net; MPU-Net: Multi-stage progressive U-Net; RDBU-Net: Residual dense block U-Net; RDNet: Residual dense network; TFPF: Time-frequency analysis; U-Net: U-shaped convolutional network

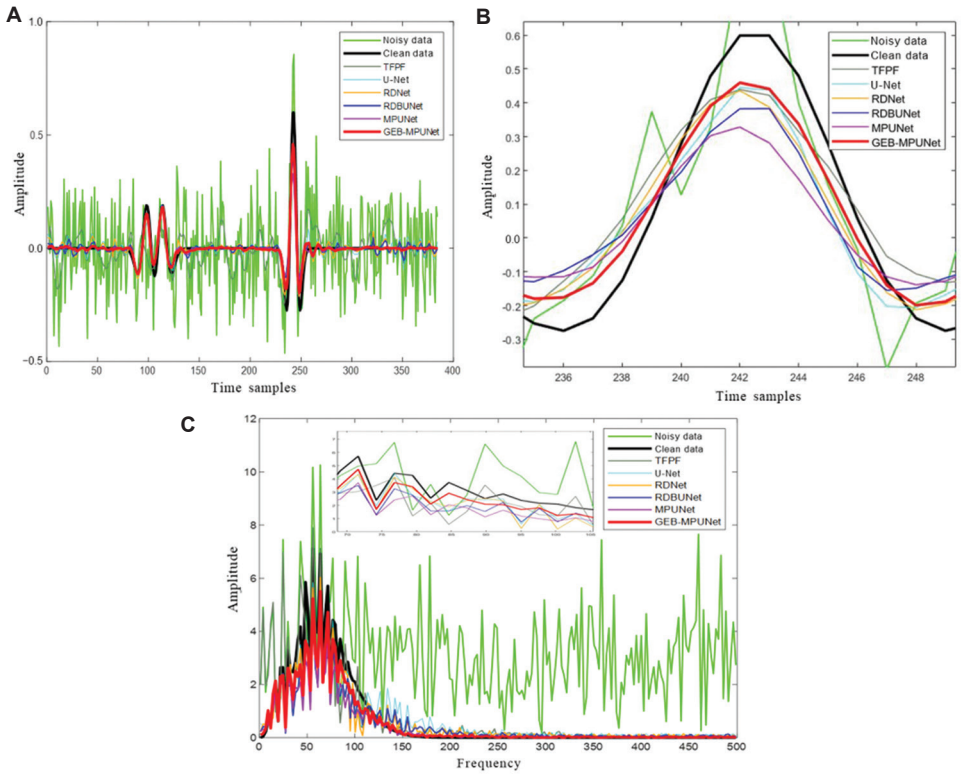

Figure 16. Single trace record comparison. (A) Time-domain waveform. (B) Enlarged view of the last peak in (A). (C) Frequency-domain waveform Abbreviations: GEB-MPU-Net: Group enhanced convolutional blocks-multi-stage progressive U-Net; MPU-Net: Multi-stage progressive U-Net; RDBU-Net: Residual dense block U-Net; RDNet: Residual dense network; TFPF: Time-frequency analysis; U-Net: U-shaped convolutional network

processing. Comparative analysis of Figure 15G and H reveals GEB-MPU-Net's enhanced spectral fidelity, with its frequency-wavenumber transform exhibiting closer alignment to the noise-free reference than MPU-Net's output. This improved spectral reconstruction demonstrates the architecture's advanced noise suppression while maintaining critical signal components across wavenumber-frequency domains.

To evaluate GEB-MPU-Net's performance, a random single trace (Trace 45) was selected from the noise-free reference dataset (Figure 13A) for detailed time-frequency analysis. The time-domain comparing waveforms are presented in Figure 16A. An enlarged view of the last peak of the time domain waveform is provided in Figure 16B. Lastly, the frequency domain comparing waveforms can be observed in Figure 16C. The results indicate that the waveform generated by GEB-MPU-Net closely resembles that of the pure signal, demonstrating that the seismic signal recovered through GEB-MPU-Net denoising is the most complete and cleanest, with a significant advantage in preserving effective signal amplitude.

The study employed standard quantitative metrics to assess denoising effectiveness and signal preservation across different methods. Comparative analysis reveals GEB-MPU-Net's superior performance in both noise suppression and amplitude retention relative to alternative approaches, as documented in Table 2.

Table 2. Comparison of parameters after denoising with different methods

Methods	Parameters	
	Signal-to-noise ratio (dB)	Mean squared error
Noisy signal	−8.1856	0.0321
TFPF	0.1597	0.0047
U-Net	5.4996	0.0014
RDNet	6.8447	0.001
RDBU-Net	7.4174	8.84 e−04
MPU-Net	8.3013	7.22 e−04
GEB-MPU-Net	9.1082	5.99 e−04

Abbreviations: GEB-MPU-Net: Group enhanced convolutional blocks-multi-stage progressive U-Net; MPU-Net: Multi-stage progressive U-Net; RDBU-Net: Residual dense block U-Net; RDNet: Residual dense network; TFPF: Time-frequency analysis; U-Net: U-shaped convolutional network.

3.2. Field data processing results

Figure 17A shows part of an acquired field seismic dataset collected under forested terrain conditions, comprising

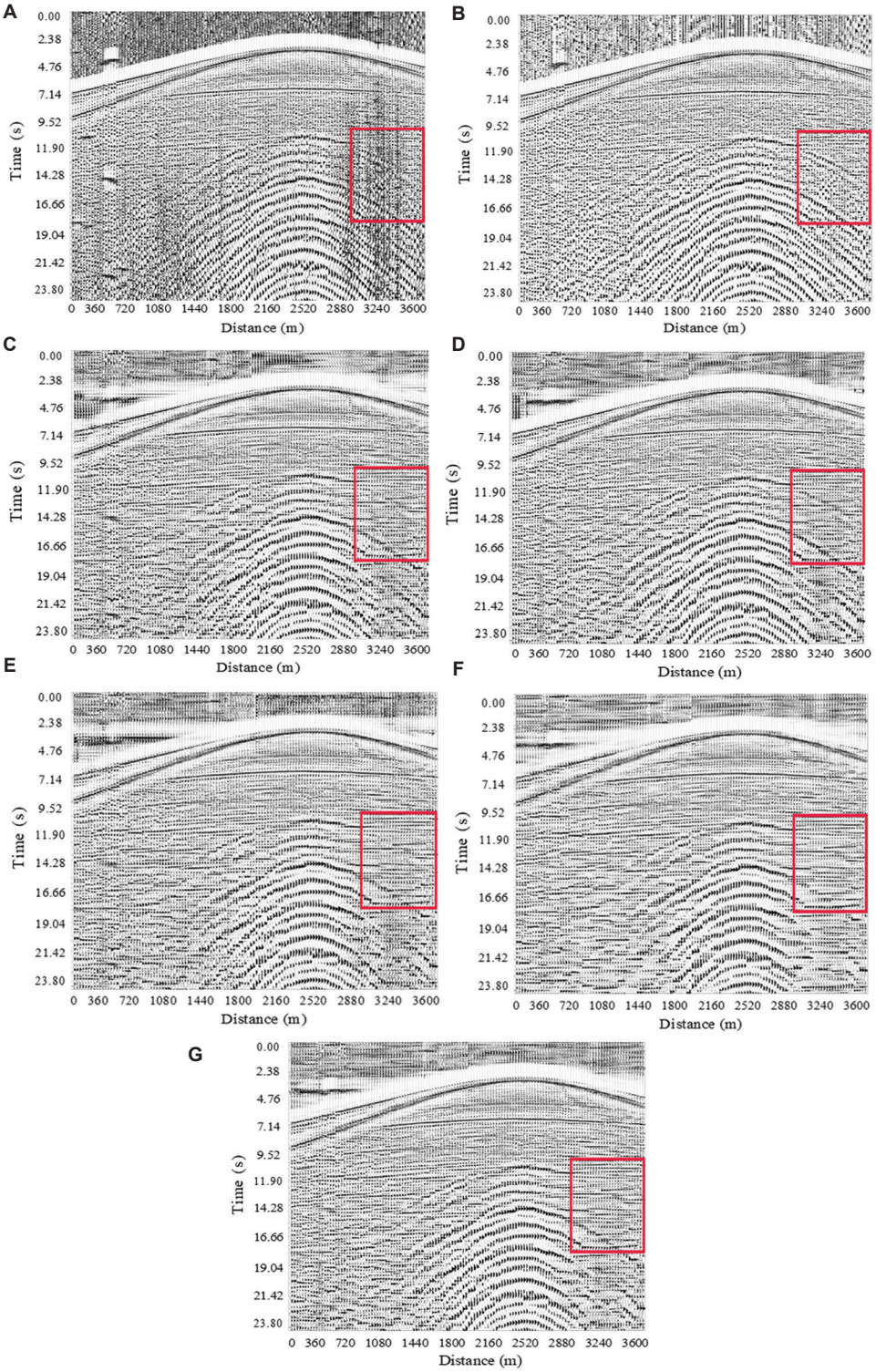

Figure 17. Denoising results of the field seismic record. (A) Field seismic data. (B) TFPF. (C) U-Net. (D) RDNet. (E) RDBU-Net. (F) MPU-Net. (G) GEB-MPU-Net

Abbreviations: GEB-MPU-Net: Group enhanced convolutional blocks-multi-stage progressive U-Net; MPU-Net: Multi-stage progressive U-Net; RDBU-Net: Residual dense block U-Net; RDNet: Residual dense network; TFPF: Time-frequency analysis; U-Net: U-shaped convolutional network

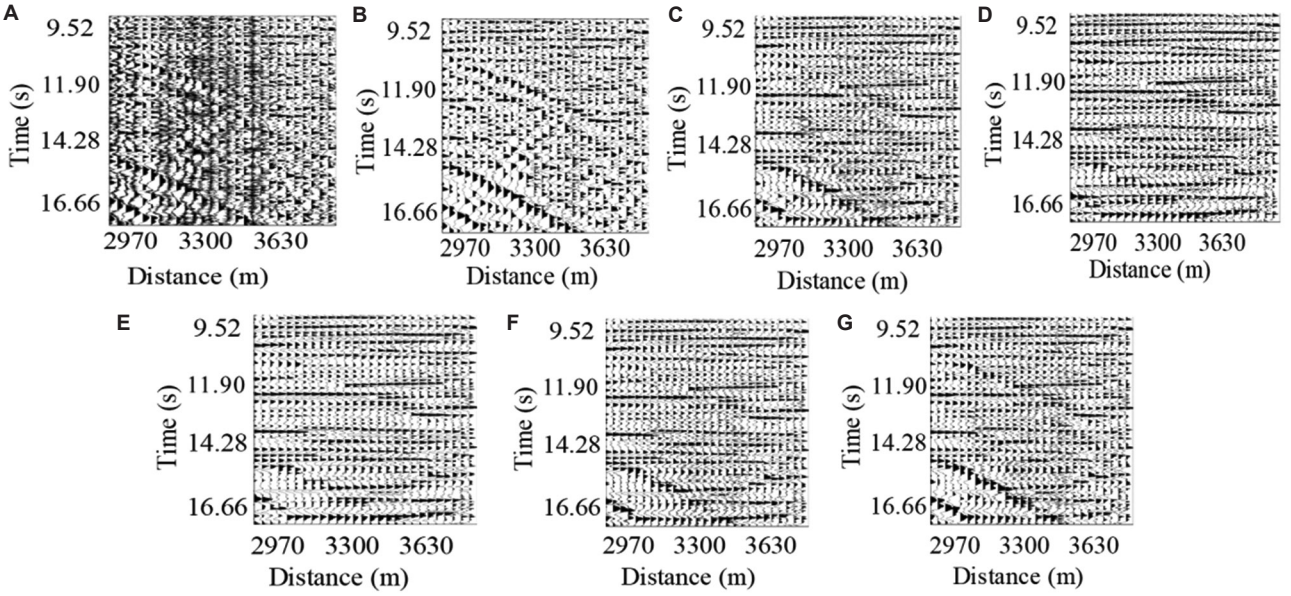

Figure 18. Enlarged comparison of red boxes in Figure 17. (A) Field seismic data. (B) TFPF. (C) U-Net. (D) RDNet (E) RDBU-Net. (F) MPU-Net. (G) GEB-MPU-Net
Abbreviations: GEB-MPU-Net: Group enhanced convolutional blocks-multi-stage progressive U-Net; MPU-Net: Multi-stage progressive U-Net; RDBU-Net: Residual dense block U-Net; RDNet: Residual dense network; TFPF: Time-frequency analysis; U-Net: U-shaped convolutional network

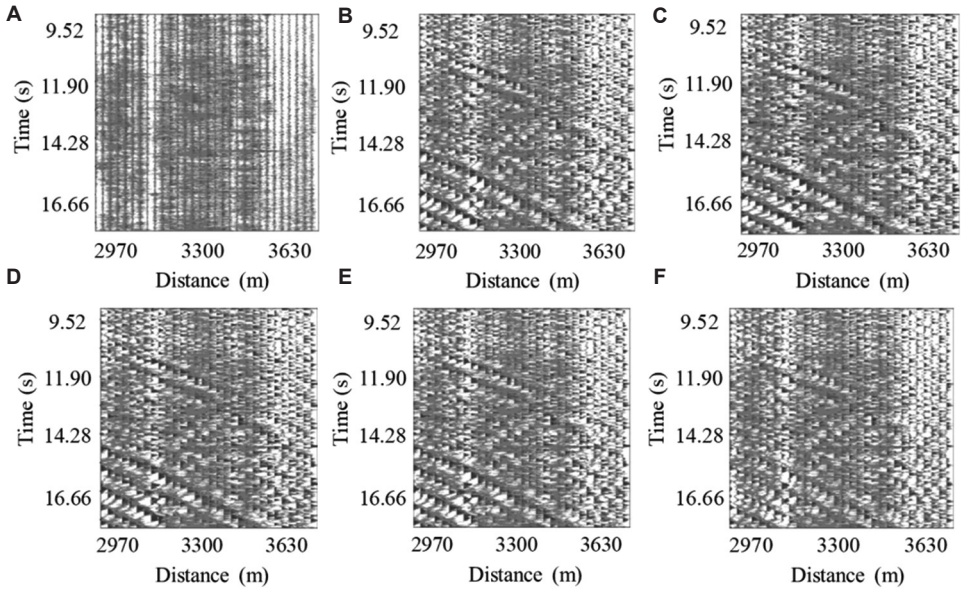

Figure 19. Corresponding residual record of Figure 18. (A) TFPF. (B) U-Net. (C) RDNet. (D) RDBU-Net. (E) MPU-Net. (F) GEB-MPU-Net
Abbreviations: GEB-MPU-Net: Group enhanced convolutional blocks-multi-stage progressive U-Net; MPU-Net: Multi-stage progressive U-Net; RDBU-Net: Residual dense block U-Net; RDNet: Residual dense network; TFPF: Time-frequency analysis; U-Net: U-shaped convolutional network

128 channels with temporal sampling characteristics suitable for detailed subsurface analysis. This record underwent comprehensive processing through multiple denoising approaches, including conventional and deep learning-based methods, with the outputs shown in Figure 17B-G.

The TFPF output in Figure 17B demonstrates partial background noise attenuation while exhibiting characteristic limitations in suppressing surface wave contamination and persistent low-frequency noise components. This performance pattern reflects fundamental constraints of traditional signal processing approaches in complex field environments.

Figure 17C-F presents the processed outputs from U-Net, RDNet, RDBU-Net, and MPU-Net, demonstrating progressive improvements in both noise suppression and signal recovery compared to conventional methods. While these architectures showed enhanced capability in revealing subsurface features, opportunities remain for further resolution enhancement in complex geological settings. Figure 17G demonstrates GEB-MPU-Net's superior processing results, exhibiting comprehensive noise suppression while significantly enhancing seismic reflection continuity and resolution. The output displayed markedly improved signal clarity compared to alternative methods, with well-preserved geological features throughout the profile.

The magnified views of the red boxes in Figure 17 and the corresponding residual records are shown in Figures 18 and 19, respectively. Figures 18B and 19A reveal severe low-frequency noise and surface wave interference that obscure underlying signals. While U-Net processing in Figures 18C and 19B enables initial event detection near surface waves, residual noise contamination remains substantial. As shown in Figures 18D-F and 19C-E, subsequent architectures demonstrate progressive improvements, with RDNet, RDBU-Net, and MPU-Net achieving measurable noise reduction and signal recovery. GEB-MPU-Net emerges as the most effective solution, delivering superior noise suppression and event clarity, as shown in Figure 18G and Figure 19F.

4. Conclusion

Building upon the MPU-Net framework, this study developed GEB-MPU-Net through the systematic integration of GEB following each CAB within the three-stage processing hierarchy. This enhancement establishes a more robust feature learning pipeline while preserving the original network's multi-scale analysis capabilities. The GEB significantly enhanced low-frequency feature representation through three coordinated mechanisms: (i) Strategic channel segmentation enabling specialized frequency processing, (ii) adaptive channel width expansion for comprehensive feature capture, and (iii) intelligent integration of deep-wide channel correlations. This multi-faceted approach optimized information flow across network depths while preserving critical seismic signatures. The GEB module incorporated a novel signal augmentation mechanism to mitigate progressive attenuation of shallow features in deep networks. This design addresses a fundamental limitation in MPU-Net's architecture, where excessive network depth could compromise both denoising accuracy and output stability. The augmentation strategy actively maintains critical near-surface information throughout the processing hierarchy. The GEB module implements residual learning to create direct feature pathways between input and output layers. This architecture strategically combines shallow and deep representations through additive merging, ensuring preservation of critical near-surface features, stable gradient propagation across network depths, and enhanced overall denoising robustness. Experimental results across synthetic and field datasets demonstrated GEB-MPU-Net's consistent advantages over the baseline MPU-Net architecture. The enhanced network exhibited superior signal fidelity through improved amplitude preservation and more effective noise suppression, particularly for random noise components. In addition, the advanced framework yielded clearer seismic reflections with enhanced continuity, enabling more reliable geological interpretation.

Acknowledgments

This research was supported by the Shanxi Key Laboratory of Wireless Communication and Detection, which provided the necessary equipment for this work.

Funding

This work was supported in part by the Natural Science Foundation of Shanxi Province (202103021224012).

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: Guanghui Li
Data curation: Li Wang
Methodology: Guanghui Li, Huiwei Li, Shoufeng He
Software: Huiwei Li, Shoufeng He
Visualization: Guanghui Li, Huiwei Li, Shoufeng He
Writing–original draft: Guanghui Li, Huiwei Li
Writing–review & editing: Guanghui Li, Huiwei Li

Availability of data

Data are available upon request via liguanghui0352@163.com.

References

1. Lee D, Shin RS, Yeo ME. Denoising sparker seismic data with deep BiLSTM in fractional Fourier transform. *Comput Geosci*. 2024;184:105519.

   doi: 10.1016/j.cageo.2024.105519

2. Geetha K, Kumar MH, Ajitha DK. A novel approach for seismic signal denoising using optimized discrete wavelet transform via honey badger optimization algorithm. *J Appl Geophys*. 2023;219:105236.

   doi: 10.1016/j.jappgeo.2023.105236

3. Zhang S, Zhang D, Yang X, Xu S, Sun Z, Liu Y. Noise reduction method based on curvelet theory of seismic data.

Pet Sci Technol. 2023;41(24):2344-2361.

doi: 10.1080/10916466.2022.2118771

4. Li M, Li Y, Wu N, Wu J, Ma Y, Liu Q. Desert seismic data denoising based on energy spectrum analysis in empirical curvelet domain. *Stud Geophys Geod*. 2020;64(2):373-390.

doi: 10.1007/s11200-019-0476-4

5. Dalai B, Kumar P, Yuan X. De-noising receiver function data using the Seislet transform. *Geophys J Int*. 2019;217(3):2047-2055.

doi: 10.1093/gji/ggz135

6. Hasan MDA, Ahmad ZAB, Leong MS, Hee LM. Automated denoising technique for random input signals using empirical mode decomposition (EMD)-stabilization diagram. *MATEC Web Conf*. 2019;255:01004.

doi: 10.1051/matecconf/201925501004

7. Liu W, Liu Y, Li S, Chen Y. A review of variational mode decomposition in seismic data analysis. *Surv Geophys*. 2023;44(2):323-355.

doi: 10.1007/s10712-022-09742-z

8. Zhao YX, Li Y, Yang BJ. Denoising of seismic data in desert environment based on a variational mode decomposition and a convolutional neural network. *Geophys J Int*. 2020;221(2):1211-1225.

doi: 10.1093/gji/ggaa071

9. Li J, Fan W, Li Y, Li W, Zhang Y, Zeng M. Desert seismic noise suppression based on an improved low-rank matrix approximation method. *J Appl Geophys*. 2020;173:103926.

doi: 10.1016/j.jappgeo.2019.103926

10. Sun F, Zhang Q, Wang Z, Li Y. Compressed sensing with logsum heuristic recover for seismic denoising. *Front Earth Sci*. 2023;11:1285622.

doi: 10.3389/feart.2023.1285622

11. Liu L, Ma J. Structured graph dictionary learning and application on the seismic denoising. *IEEE Trans Geosci Remote Sens*. 2019;57(4):1883-1893.

doi: 10.1109/TGRS.2018.2870087

12. Liu Q, Fu L, Zhang M. Deep-seismic-prior-based reconstruction of seismic data using convolutional neural networks. *Geophysics*. 2021;86(2):V131-V142.

doi: 10.1190/geo2019-0570.1

13. Li X, Hawari KG, Huang F, Liu J, Zhang Y. Review of CNN in aerial image processing. *Imaging Sci J*. 2023;71(1):1-13.

doi: 10.1080/13682199.2023.2174651

14. Wei D, Chen G, Chen J, Li C, Wu S, Zhang Y. Seismic data denoising using a self-supervised deep learning network. *Math Geosci*. 2023;56(3):487-510.

doi: 10.1007/s11004-023-10089-3

15. Li J, Qu R, Lu C. Multiple attention mechanisms-based convolutional neural network for desert seismic denoising. *Pure Appl Geophys*. 2023;180(6):2135-2155.

doi: 10.1007/s00024-023-03255-5

16. Ji G, Wang C. A denoising method for seismic data based on SVD and deep learning. *Appl Sci*. 2022;12(24):12840.

doi: 10.3390/app122412840

17. Guo Z, Zhu S, Chen J, Zhu W. Research on deep convolutional neural network time-frequency domain seismic signal denoising combined with residual dense blocks. *Front Earth Sci*. 2021;9:770748.

doi: 10.3389/feart.2021.681869

18. Cai J, Wang L, Zheng J, Duan Z, Li L, Chen N. Denoising method for seismic co-band noise based on a U-Net network combined with a residual dense block. *Appl Sci*. 2023;13(3):1324.

doi: 10.3390/app13031324

19. Ding M, Zhou Y, Chi Y. Seismic signal denoising using Swin-Conv-UNet. *J Appl Geophys*. 2024;223:105355.

doi: 10.1016/j.jappgeo.2024.105355

20. Zamir SW, Arora A, Khan S, Hayat M, Khan F, Shah M. Multi-stage progressive image restoration. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville, TN, USA, 2021. p. 14821-14831.

doi: 10.1109/CVPR46437.2021.01458

21. Saad MO, Ravasi M, Alkhalifah T. Self-supervised multistage deep learning network for seismic data denoising. *Artif Intell Geophysics*. 2025;6(1):100123.

doi: 10.1016/j.aiig.2025.100123

22. Obou.1016/j.aiig.2025.100123123ing. learning network for seismic data denoising. enoising. ng. sing. ing. ng. l denoising*Geophysics*. 2025;90(3):V205-V219.

doi: 10.1190/geo2024-0456.1

23. Tian C, Yuan Y, Zhang S, Wang X, Li H. Image super-resolution with an enhanced group convolutional neural network. *Neural Netw*. 2022;153:373-385.

doi: 10.1016/j.neunet.2022.06.009

24. Saad OM, Ravasi M, Alkhalifah T. Noise attenuation in distributed acoustic sensing data using a guided unsupervised deep learning network. *Geophysics*. 2024;89(6):V573-V587.

doi: 10.1190/geo2024-0109.1

25. Saad OM, Obou23-0642Bai M, Alkhalifah T, Wang Y, Liu J. Self-attention deep image prior network for unsupervised 3-D seismic data enhancement. *IEEE Trans Geosci Remote Sens*. 2021;60(5):1-14.

doi: 10.1109/TGRS.2021.3108515