ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

Seismic signal denoising using variational mode decomposition and a denoising convolutional neural network

Shengrong Zhang[1,2], Liang Zhang[1]*, and Xuesha Qin[3]

[1]The Key Laboratory of Advanced Manufacturing Technology of the Ministry of Education, Guizhou University, Guiyang, Guizhou, China

[2]School of Computer and Electronics Information, Guangxi University, Nanning, Guangxi, China

[3]China-ASEAN School of Economics, Guangxi University, Nanning, Guangxi, China

Abstract

Effectively recovering signals buried in noise remains a challenging topic in seismic data denoising. Many conventional methods often fail to accurately capture the characteristics of seismic signals. To address this issue, this study proposed an effective method called variational mode decomposition (VMD)–denoising convolutional neural network (DnCNN). The method first applies VMD to decompose the originally noisy signal into multiple intrinsic mode functions (IMFs) with band-pass characteristics, thereby achieving effective decoupling of different frequency components and noise separation. Selected IMFs are then combined into a multi-channel input and fed into the DnCNN for end-to-end modeling and denoising reconstruction. By decomposing the noisy signal into IMFs corresponding to specific frequency bands and learning them through DnCNN, the network can better extract features within each frequency band. Serving as a front-end filter, the VMD module enhances the network's ability to represent effective frequency components, suppresses high-frequency random noise, and improves the resolution of weak signals. Experimental results demonstrated that the proposed method effectively captures signal characteristics and recovers signals from both real and synthetic seismic data. In conclusion, the proposed VMD–DnCNN method provides a robust and efficient solution for seismic signal denoising.

Keywords: Variational mode decomposition; Denoising convolutional neural network; Intrinsic mode functions; Recover weak signals; Seismic denoising

***Corresponding author:**
Liang Zhang
(liangzhang@gzu.edu.cn)

Citation: Zhang S, Zhang L, Qin X. Seismic signal denoising using variational mode decomposition and a denoising convolutional neural network. *J Seismic Explor.* doi: 10.36922/JSE025260030

Received: June 29, 2025

Revised: August 12, 2025

Accepted: August 13, 2025

Published online: September 4, 2025

1. Introduction

Seismic signals are often characterized by non-stationary properties and are susceptible to various external interferences during acquisition, such as complex mixed noise caused by exploration instruments, wind, and transportation activities.[1,2] These types of noise can significantly degrade the quality of subsequent imaging and interpretation processes. To better extract geological information from seismic data, it is crucial to effectively isolate signals from noise. Moreover, the accurate recovery of weak signals can further enhance geological exploration efforts.[3-5] Therefore, many researchers have investigated

effective seismic signal recovery under low signal-to-noise ratio (SNR) conditions.[6] Currently, denoising methods can be broadly categorized into four groups: Time–frequency analysis methods, decomposition-based methods, low-rank-based methods, and deep learning methods.

Time–frequency analysis methods aim to exploit the differences in time–frequency distributions between useful seismic reflections and noise. By applying time–frequency transformations, seismic data can be represented in a joint time–frequency domain, enabling the separation and suppression of noise from signal components. For example, the wavelet transform[7] achieves denoising by decomposing the signal into different frequency bands, retaining the dominant frequency components associated with the signal while removing high-frequency components typically attributed to noise. The short-time Fourier transform[8] utilizes its localized time–frequency resolution to expand non-stationary seismic signals in the time–frequency domain, allowing for clearer distinction between signal and noise. Likewise, the S-transform[9] constructs a two-dimensional time–frequency representation and leverages the seismic signal's concentration and continuity in local frequency content to facilitate signal–noise separation. However, these methods often suffer from difficulties in identifying optimal basis functions and are highly sensitive to thresholding strategies. The performance of these approaches is strongly dependent on the threshold level and the specific selection scheme, where improper thresholds may lead to significant degradation in denoising quality and signal preservation. These limitations are not unique to the aforementioned methods, but are also observed in other time–frequency analysis techniques, such as the seislet transform,[10] curvelet transform,[11] and contourlet transform.[12]

Decomposition-based methods aim to extract intrinsic structures from noisy signals by separating effective components from noise interference and subsequently reconstructing the denoised seismic signal. For example, empirical mode decomposition[13,14] is an adaptive and data-driven technique for processing non-stationary signals. It decomposes the original signal into a set of intrinsic mode functions (IMFs) with localized time-frequency characteristics. By analyzing the frequency and energy features of each component, noise-dominated modes can be identified and discarded, followed by signal reconstruction for denoising. Variational mode decomposition (VMD)[15] formulates a variational optimization problem to decompose the original signal into a set of band-limited sub-signals (mode components). Noise-dominated components are recognized and removed based on their frequency and energy characteristics, achieving efficient denoising. On the other hand, singular value decomposition[16] decomposes the seismic data matrix into ordered components according to their energy. Principal components represent the effective signal, while low-energy components correspond to noise, and denoising is performed through reconstruction. The major challenge of decomposition-based methods lies in the mode mixing phenomenon, where certain modes contain both noise and signal components, complicating their separation.

Low-rank-based methods exploit the strong structural properties of seismic noise signals in time and space domains. When arranged as matrices or tensors, seismic signals typically exhibit low-rank characteristics, whereas noise is random and high-rank. Low-rank decomposition techniques can therefore extract the structured signal components while suppressing high-rank noise. For example, principal component analysis[17] projects seismic data onto a set of principal components, retaining the first few components that contain the main information and discarding the subsequent minor components, thus achieving denoising. Cadzow filtering[18] constructs a Hankel matrix from the signal and iteratively applies singular value decomposition, low-rank approximation, and Hankel structure reconstruction to preserve the primary signal components while suppressing noise. Both methods treat the effective signal as a low-rank structure and extract meaningful cycles by reducing the rank. However, a notable limitation of these approaches is that their performance heavily depends on the prior assumptions regarding the rank.

Deep learning methods[19,20] essentially construct end-to-end mapping functions that automatically learn the relationship between noisy and clean signals, thereby achieving noise suppression and signal recovery. Previous studies have systematically demonstrated how neural networks can be applied to seismic signal denoising, successfully employing deep learning for this purpose.[21] For example, edge-feature-guided wavelet U-Net[22] integrates wavelet transforms to design dual decoders aimed at edge detection, capturing shape and edge information of effective signals. Other studies introduced the representation of seismic data in the time–frequency domain as input to neural networks,[23] enabling the model to learn the characteristics of seismic signals in the time–frequency space more effectively, and proposed an identification-based denoising approach.[24] Noise locations are first identified, and then the network performs targeted denoising, thereby more accurately removing noise while avoiding the inadvertent removal of useful signals. Some transformer-based methods innovatively combined sparse channel-wise attention transformers with diffusion models through a seismic prior extraction network, achieving efficient and high-quality seismic data interpolation.[25]

In addition, a transformer-based seismic data denoising model has been introduced, incorporating a novel self-supervised pretraining strategy to effectively capture long-range dependencies and improve noise attenuation while preserving weak signals.[5] To enhance seismic data denoising performance, most methods involve modifications to neural network architectures; however, these improvements often come at the cost of increased computational time.

In summary, traditional seismic denoising methods often suffer from limitations related to performance and parameter tuning. Whether it is time–frequency methods, decomposition-based methods, or low-rank methods, parameter adjustments are typically required according to the noise intensity. A common challenge lies in how to properly decompose the signal in a way that removes noise while preserving the useful signal. This leads to these methods lacking adaptability to varying noise levels. For deep learning methods, when noise levels are excessively high, it is often difficult to effectively capture the characteristics of seismic signals. The construction of end-to-end mappings can introduce bias, resulting in poor learning of weak signals and consequently unsatisfactory denoising performance. Therefore, effectively addressing the limitations of these approaches constitutes a major challenge in seismic signal processing and is the central focus of this study.

Decomposition-based and deep learning-based methods each have their advantages and drawbacks. Deep learning methods can effectively handle non-Gaussian and nonlinear noise, whereas decomposition-based methods rely primarily on frequency decomposition and are less capable of adapting to complex background noise. Furthermore, these approaches can identify low-amplitude reflections that decomposition methods may mistakenly remove as noise, especially for weak high-frequency reflections. However, deep learning implicitly models frequency information and lacks explicit frequency band control. In contrast, decomposition-based methods explicitly separate different frequency bands, facilitating the removal of band-specific noise and improving the preservation of waveform structures, particularly the low-frequency primary components.

Motivated by the strong learning capabilities of deep learning models and the intrinsic decomposition principles, this study proposed a VMD–denoising convolutional neural network (DnCNN) framework. This approach leverages the advantages of deep learning models to compensate for the shortcomings of traditional methods, while utilizing the strengths of decomposition models to complement deep learning models' limitations. VMD effectively decomposes seismic data to extract intrinsic features, enabling DnCNN[26] to learn the characteristics of different IMFs more effectively. Through training, DnCNN can mitigate the mode mixing problem inherent in VMD by continuously learning which IMFs are useful and which should be discarded. Experiments conducted on both synthetic and field seismic data demonstrated that this method not only effectively suppresses noise but also outperforms several traditional denoising techniques in terms of denoising performance.

2. Methods

In decomposition-based methods, mode mixing often occurs, resulting in decomposed modes that may contain noise components, which is unavoidable. Whether the decomposition extracts components from high frequency to low frequency, optimizes for band-limited signals so that each mode concentrates on a specific frequency band, or employs other decomposition techniques, it is essentially impossible to prevent noise from being introduced into the decomposed components. However, deep learning networks possess strong learning capabilities and can progressively distinguish between noise and useful signals through continuous training. Therefore, deep learning networks were combined with traditional methods to form a VMD–DnCNN denoising architecture.

Seismic signals are often represented as seismic profiles. Let the function $c(x, y)$ denotes the ideal noise-free signal and $n(x, y)$ represents the noise component, where x is the sampling time and y is the trace number. The observed noisy data can then be expressed as:

$$f(x,y) = c(x,y) + n(x,y) \tag{I}$$

The 2D VMD method was applied to decompose the noisy seismic data. The formulation of 2D VMD is as follows:

$$\min_{u_k,\omega_k} \sum_{k=1}^{K} \| \nabla \left(u_k(x,y) \cdot e^{-j\omega_k x} \right) \|_2^2, \; s.t. \sum_{k=1}^{K} u_k(x,y) = f(x,y) \tag{II}$$

$$f(x,y) = \sum_{k=1}^{3} u_k(x,y) \tag{III}$$

where, $u_k(x, y)$ denotes the K decomposed mode, ω_k is its center frequency along the x direction, and ∇ represents the two-dimensional gradient operator, which measures the smoothness or bandwidth of the mode in the frequency domain. The term $e^{-j\omega_k x}$ performs a frequency shift of the mode to concentrate it around the low-frequency baseband, facilitating a unified calculation of the mode's bandwidth.

After decomposing the signal into three modes based on bandwidth, these modes were combined with the

original signal as a four-channel data input to the network, allowing the deep learning model to learn the process of reconstructing the original signal from the three modes during training. By decomposing the data through VMD, different band-limited signal components were formed, significantly reducing the learning burden on the deep learning network. Compared to the original single-channel structure, where only noisy data are input to the network, this approach alleviates the heavy learning load and makes it easier for the network to capture features of both the seismic signal and noise.

When inputting the decomposed modes into the network, the model first learns the characteristics of different frequency bands. Furthermore, since the primary noise frequency bands have already been separated through decomposition, the network can more readily identify the main noise components during training. Even when non-noise dominant frequency bands are mixed with some noise, the network can leverage the features learned from the noise-dominant bands to recognize and denoise these components effectively. Moreover, effective signal information may also be present within the primary noise frequency bands. In this case, the useful signal features learned from the non-noise-dominant bands help the network capture and preserve valid signal components within the noise-dominant bands.

As a denoising network, DnCNN has been widely used for processing seismic signals. The architecture of DnCNN is illustrated in Figure 1. The network first passes the input through a convolutional layer, followed by a rectified linear unit (ReLU) activation function. Sixty-four convolutional kernels are used to extract preliminary low-level seismic features, including local waveform shapes, edges, and frequency components. The ReLU activation enhances the nonlinear representation capability, helping to distinguish seismic signal structures from high-frequency noise. The middle part of the network consists of 18 repeated blocks, each comprising a convolutional layer, batch normalization, and a ReLU activation. Batch normalization balances feature distributions across different batches, suppressing outliers. Through these 18 repeated operations, the network progressively extracts abstract seismic features, enabling it to ignore unstructured noise while preserving waveforms with reflective patterns. Finally, the output passes through a convolutional layer that compresses the 64 deep feature channels back to the original channel number, yielding the final denoised result.

The input structure of DnCNN was modified to better enable it to learn seismic signal information. By applying VMD to decompose noisy signals, the noise is separated into three modal components: High-frequency IMF, mid-frequency IMF, and low-frequency IMF. These three modes approximately correspond to high-, mid-, and low-frequency seismic phases and noise structures, facilitating the network's ability to distinguish and process noise in different frequency bands. Each of the three IMFs from VMD concentrates around a certain center frequency, which is dynamically adjusted during iterations to ensure each mode focuses on a specific frequency band.

In practice, IMF1 captures high-frequency details, often containing seismic noise and sharp reflections; IMF2 captures mid-frequency seismic phases, which include seismic signals but may also contain noise; IMF3 corresponds to low-frequency main structures, encompassing the primary seismic phases and reflection interfaces. Thus, decomposing the seismic signal into these three frequency bands effectively separates high-frequency disturbances, mid-frequency seismic phases, and low-frequency structural components.

The overall denoising process is illustrated in Figure 2. The VMD–DnCNN denoising procedure consists of two steps: The first step is model training, and the second step is seismic data denoising. In the first step, the selected synthetic data are segmented by time windows, followed by data selection using a Monte Carlo strategy. This produces the training labels for the network. Gaussian noise is then added to the labels to generate the noisy samples. These samples are decomposed using 2D VMD into three IMFs, which serve as the channels of the input samples. Subsequently, the time-domain channels of the labels are concatenated with the IMF channels of the

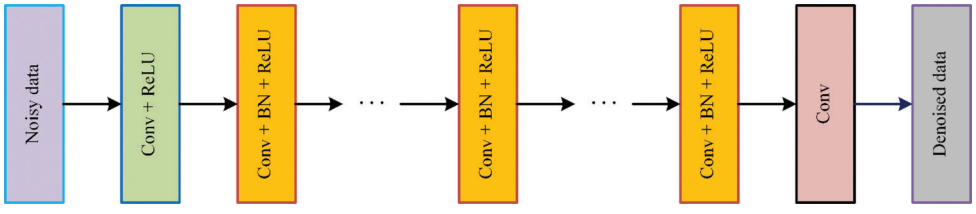

Figure 1. Structure of a denoising convolutional neural network.
Abbreviations: BN: Batch normalization; Conv: Convolutional layer; ReLU: Rectified linear unit.

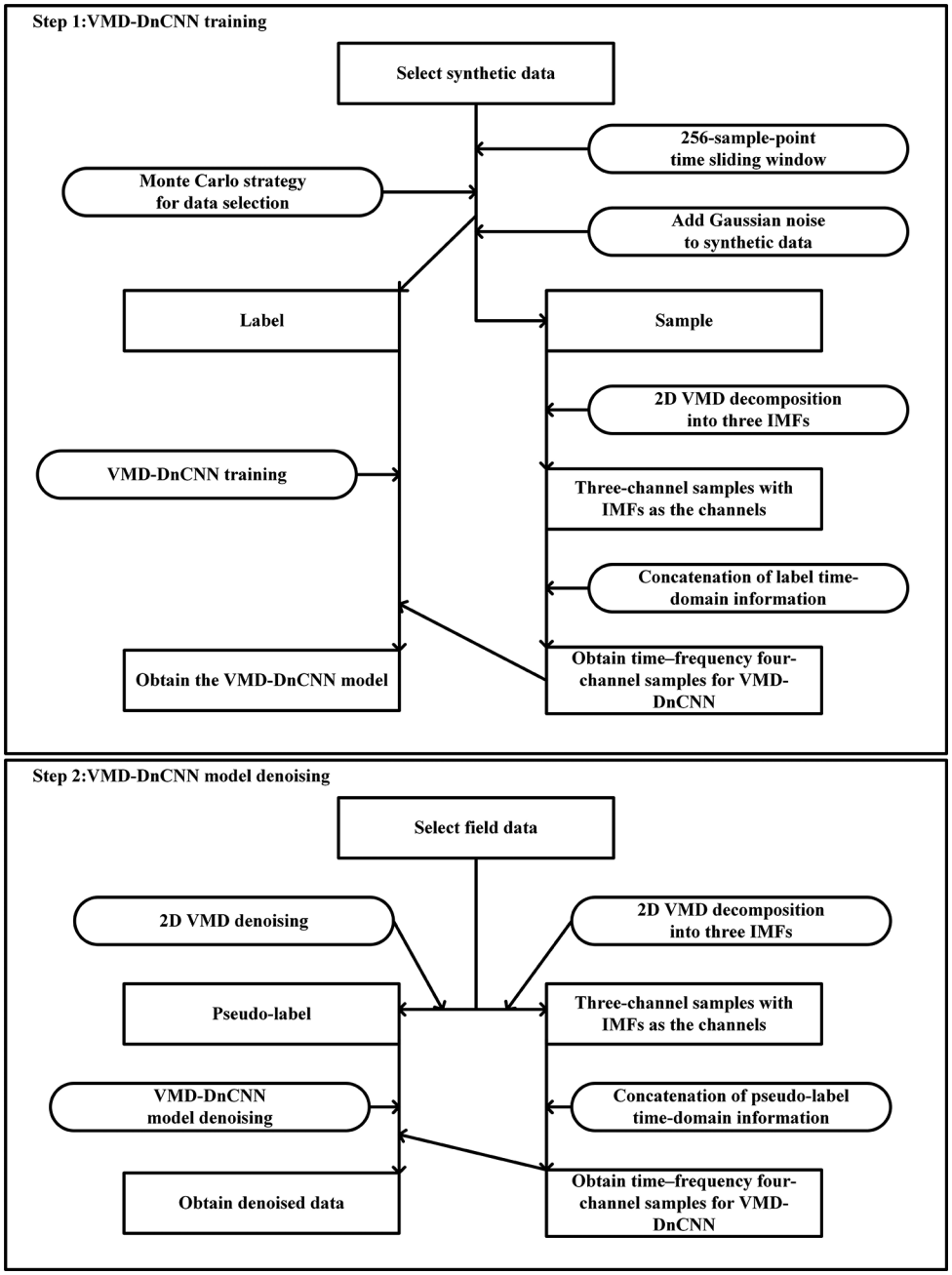

Figure 2. Main workflow of the proposed method.
Abbreviations: DnCNN: Denoising convolutional neural network; IMF: Intrinsic mode function; VMD: Variational mode decomposition.

samples to form the final time–frequency training samples required by VMD–DnCNN. By training the network with these labels and samples, the final model is obtained. In the second step, the seismic data to be denoised are first processed by 2D VMD to obtain pseudo-labels and seismic data decomposed into three IMF modes. The pseudo-label's time-domain data and the three IMF modes are combined to form a four-channel seismic input. Finally, these data are denoised by the trained VMD–DnCNN model.

3. Network training

The performance of neural network models is highly dependent on the quality of the training dataset; therefore, constructing a high-quality dataset and applying proper preprocessing are particularly critical. Given that the recordings before the onset of direct waves primarily consist of background noise without useful seismic information, this study removed certain non-informative data during

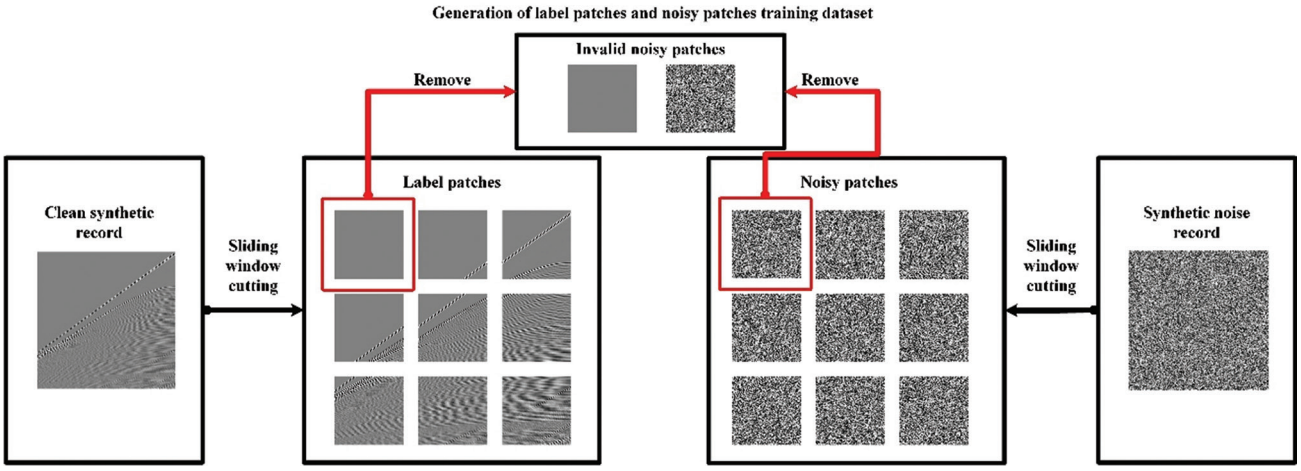

Figure 3. Training dataset generation process.

the preprocessing stage to enhance training efficiency. The overall data generation process is illustrated in Figure 3. A Monte Carlo strategy[27] was adopted to eliminate invalid synthetic data, incorporating non-zero label filtering and effective fluctuation filtering mechanisms. Specifically, silent segments before the arrival of direct waves, smooth sections, and other training-irrelevant samples were excluded. Samples with zero-valued amplitudes were directly discarded, and samples with a standard deviation $<10^{-3}$ were also removed. After filtering, the data underwent Max–Abs normalization to standardize the seismic signal amplitude range to the interval [−1,1]. This normalization accelerates model convergence during training and ensures amplitude consistency across different traces and samples, thereby reducing model bias. The training parameters for all models are listed in Table 1. All models used the ADAM optimizer.[28] The sample length was set to 256 points, with one sample selected every 128 points. When the remaining trace length was less than 256 points, it was padded forward to meet the required length. The initial learning rate was set to 10^{-4} and decayed by a factor of 10 every 40 epochs. The total number of training epochs was set to 200.

The synthetic seismic data used for training was derived from the 2007 British Petroleum (BP) Anisotropic Velocity Benchmark, a two-dimensional synthetic dataset released by BP. The dataset consists of 1,641 shot gathers, each containing 800 seismic traces. Each trace has 1,151 sampling points, with a sampling rate of 125 Hz and a sampling interval of 8 ms, resulting in a trace duration of 9.208 s. Due to the similarity between adjacent shot gathers, a subset of 20 gathers was selected for training by sampling five consecutive shots every 500 gathers. Specifically, the selected training gathers were 1–5, 501–505, 1,001–1,005, and 1,501–1,505. Shot gathers 10, 510, 1,010, and 1,510 were used for denoising evaluation. This sampling strategy,

Table 1. Training parameters of convolutional neural network-based methods

Hyperparameter	DnCNN	U-Net	VMD–DnCNN
Optimizer	ADAM	ADAM	ADAM
Patch size	256	256	256
Batch size	100	100	100
Epoch	200	200	200
Learning rate range	$[10^{-4}, 10^{-7}]$	$[10^{-4}, 10^{-7}]$	$[10^{-4}, 10^{-7}]$
Input channels	1	1	4

Abbreviations: DnCNN: Denoising convolutional neural network; VMD: Variational mode decomposition.

which spans different and dispersed seismic environments, facilitates the model to learn more representative and generalizable feature representations. Gaussian noise with a mean of 0 and a standard deviation of 0.3921 was added to the synthetic data to simulate noisy conditions. According to the characteristics of the Gaussian distribution, 99.7% of the SNR values of the added noise lie within the range [−3,3]. A total of 98,900 samples were generated from the 20 shot gathers. For all models, 80% of the samples were used for training and the remaining 20% for validation.

4. Synthetic experiment

In this section, the denoising performance of the proposed VMD–DnCNN model was evaluated using synthetic seismic data. As shown in Figure 4, Gaussian noise was added to shot gathers 10, 510, 1,010, and 1,510 to achieve SNRs of 6 dB, 0 dB, −5 dB, and −10 dB, respectively. To better assess the denoising effectiveness of VMD–DnCNN, comparative analyses were conducted with several typical seismic denoising methods, including both conventional and deep learning-based approaches. It was observed that when the noise level corresponded to an SNR of 6 dB, some

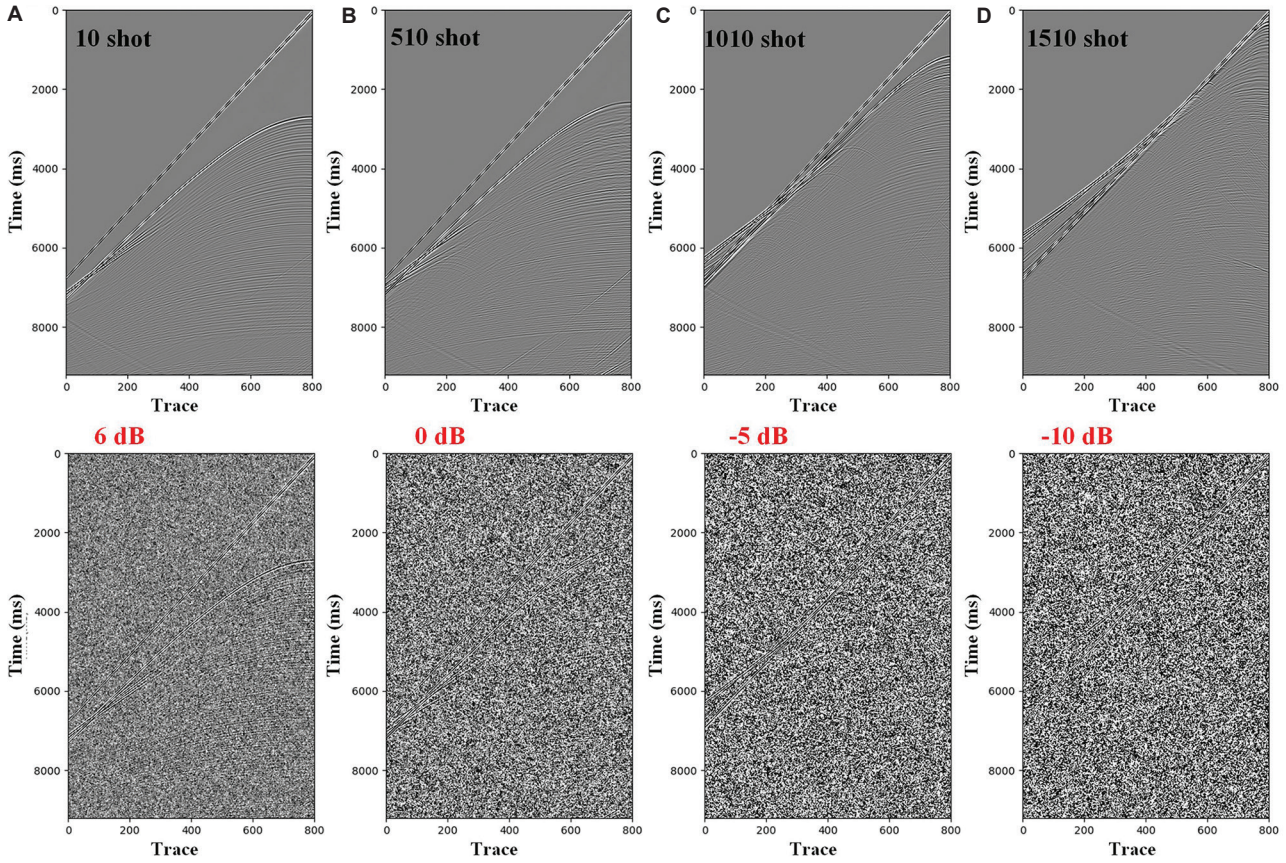

Figure 4. Noise-free data and the corresponding data with added Gaussian noise under varying conditions: (A) 6 dB, (B) 0 dB, (C) –5 dB, and (D) –10 dB.

seismic reflections were still barely visible. However, as the noise increased to 0 dB, most of the seismic signals were overwhelmed by noise, and only a small portion remained observable. At −5 dB, only the strong direct wave energy was distinguishable. When the noise level reached −10 dB, almost no valid seismic information was visually identified.

To evaluate the denoising performance of the proposed VMD–DnCNN model, it was compared with several widely used seismic denoising methods, including traditional techniques, such as the wavelet transform and VMD, as well as deep learning-based approaches, such as the DnCNN and U-Net model.[29] For parameter settings, the number of IMFs in the VMD was set to three, and a six-level Daubechies-4 (db4) wavelet was used for the wavelet transform. Thresholds in both the VMD and wavelet methods were adaptively adjusted based on the complexity of the noise in the seismic data. To ensure experimental fairness, all denoising methods were implemented and tested on an Nvidia GeForce RTX 4060 Ti GPU with 16 GB of video memory. In addition, to quantitatively assess the denoising performance of each method, four commonly used evaluation metrics for seismic signal quality were adopted: SNR, root mean square error (RMSE), peak SNR (PSNR), and structural similarity index measure (SSIM). The specific formulations of these two metrics are defined as follows:

$$SNR = 10\log_{10}\left(\frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\left(X(i,j)-\bar{X}\right)^2}{\sum_{i=1}^{N}\sum_{j=1}^{M}\left(\hat{Y}(i,j)-X(i,j)\right)^2}\right) \quad \text{(IV)}$$

$$RMSE = \sqrt{\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(\hat{Y}(i,j)-X(i,j)\right)^2} \quad \text{(V)}$$

$$SSIM = \frac{\left(2\mu_X\mu_Y + C_1\right)\left(2\sigma_{XY} + C_2\right)}{\left(\mu_X^2 + \mu_Y^2 + C_1\right)\left(\sigma_X^2 + \sigma_Y^2 + C_2\right)} \quad \text{(VI)}$$

$$PSNR = 10\log_{10}\left(\frac{MAX_X^2}{MSE}\right) \quad \text{(VII)}$$

where X represents the original seismic signal; $\hat{Y}$ represents either the noisy seismic data or the denoised seismic data, which are used to calculate the SNR and

RMSE for the original signal or the denoised result, respectively; M and N represent the number of receivers (traces) and the number of sampling points per receiver, respectively; μ_X and μ_Y are the mean amplitudes of X and Y, respectively; σ_X^2 and σ_Y^2 are the variances; σ_{XY} is the covariance between X and Y; the constants C_1 and C_2 are small positive numbers introduced to avoid division by zero; MAX_X^2 is the squared maximum amplitude of the reference signal; and MSE is the mean squared error between X and Y. The SSIM ranges from −1 to 1, with higher values indicating greater structural similarity between the two signals. The PSNR is expressed in decibels (dB), and higher values indicate better reconstruction quality with smaller differences from the reference signal.

A comparative analysis of denoising performance was conducted using the synthetic seismic data with added Gaussian noise (Figure 4). The denoising results of various methods are presented under different SNRs. To better illustrate the robustness of each method under varying noise levels, a performance comparison curve was additionally generated by adding Gaussian noise ranging from −10 dB to 6 dB in 1 dB increments to shot gather 1,510. This allows for the evaluation of how effectively each method performs against different noise intensities.

First, the visual denoising results of each method across different synthetic shot gathers were examined (Figure 5). When the noise level was 0 dB, traditional methods such as the wavelet and VMD models left noticeable background noise, and much of the noise remained entangled with the seismic signals. In contrast, deep learning methods like DnCNN and U-Net effectively suppressed background noise, although some seismic signals were inadvertently removed. Nevertheless, continuous signals across traces were better preserved. At −5 dB and −10 dB noise levels, the VMD method showed even more background noise residue, while the wavelet method exhibited increasingly severe artifacts. Under these conditions, DnCNN and U-Net also struggled to remove background noise effectively, and the retained signals were heavily contaminated by residual noise. In particular, at −10 dB, it became nearly impossible to identify any valid seismic information using these methods. In contrast, the VMD–DnCNN model demonstrated superior denoising performance. At 0 dB, although some background noise remained, it preserved more continuous and weak signals than other methods, maintaining better signal continuity. At −5 dB, it successfully removed most background noise while retaining the underlying seismic signals, achieving a near-complete recovery. Even at −10 dB, the VMD–DnCNN model was still capable of recovering meaningful signals; although some noise remained, the overall clarity of the signal was significantly better than with other methods. From this comparison, it is evident that VMD–DnCNN consistently outperformed other methods under both strong and weak noise conditions. Notably, in scenarios with severe noise contamination where other methods failed to recover seismic signals effectively, VMD–DnCNN retained finer details of the signal. Within the red rectangles in Figure 5, it can be seen that both the strong direct arrivals and the weaker reflections following the direct waves were better restored using VMD–DnCNN compared to other methods. Even under strong noise conditions, the proposed method was able to improve SNR to over 20 dB, whereas the performance of other methods—especially deep learning models—declined under such noisy scenarios. As shown in Table 2, the quantitative results also confirm that VMD–DnCNN achieved the highest improvements in both SNR and RMSE metrics compared to the other benchmark methods.

The importance of computational efficiency in seismic signal processing should not be overstated. Therefore, the computational performance of different denoising methods was analyzed by evaluating their performance under various noise levels. The SNR improvements achieved by each method across different noise intensities are illustrated in Figure 6, and the detailed performance metrics are summarized in Table 3. All methods were evaluated under identical denoising environments with consistent hyperparameter settings, including the number of training epochs, sample length, and learning rate. The training times required for VMD–DnCNN, DnCNN, and U-Net were 0.504 h, 0.498 h, and 2.144 h, respectively, while the traditional methods did not require any pretraining. In terms of average inference time per denoising task, the VMD–DnCNN, DnCNN, U-Net, VMD, and wavelet models required 2.178 s, 2.1012 s, 3.343 s, 66.26 s, and 1.291 s, respectively. Their corresponding average SNR improvements were 29.05 dB, 11.53 dB, 12.83 dB, 9.34 dB, and 8.95 dB. Both VMD–DnCNN and DnCNN consisted of 20 convolutional layers, while U-Net contained 19 convolutional layers along with four downsampling and four upsampling operations. Although U-Net slightly outperformed DnCNN in terms of denoising results, it required nearly four times the training time of DnCNN. While VMD achieved slightly better denoising results than the wavelet method, its time and memory consumption were significantly higher compared to the other methods, making it less practical for large-scale applications. Furthermore, DnCNN and VMD–DnCNN exhibited similar inference speeds and were slightly faster than U-Net. Although their inference times were longer than those of the wavelet method, they remained within

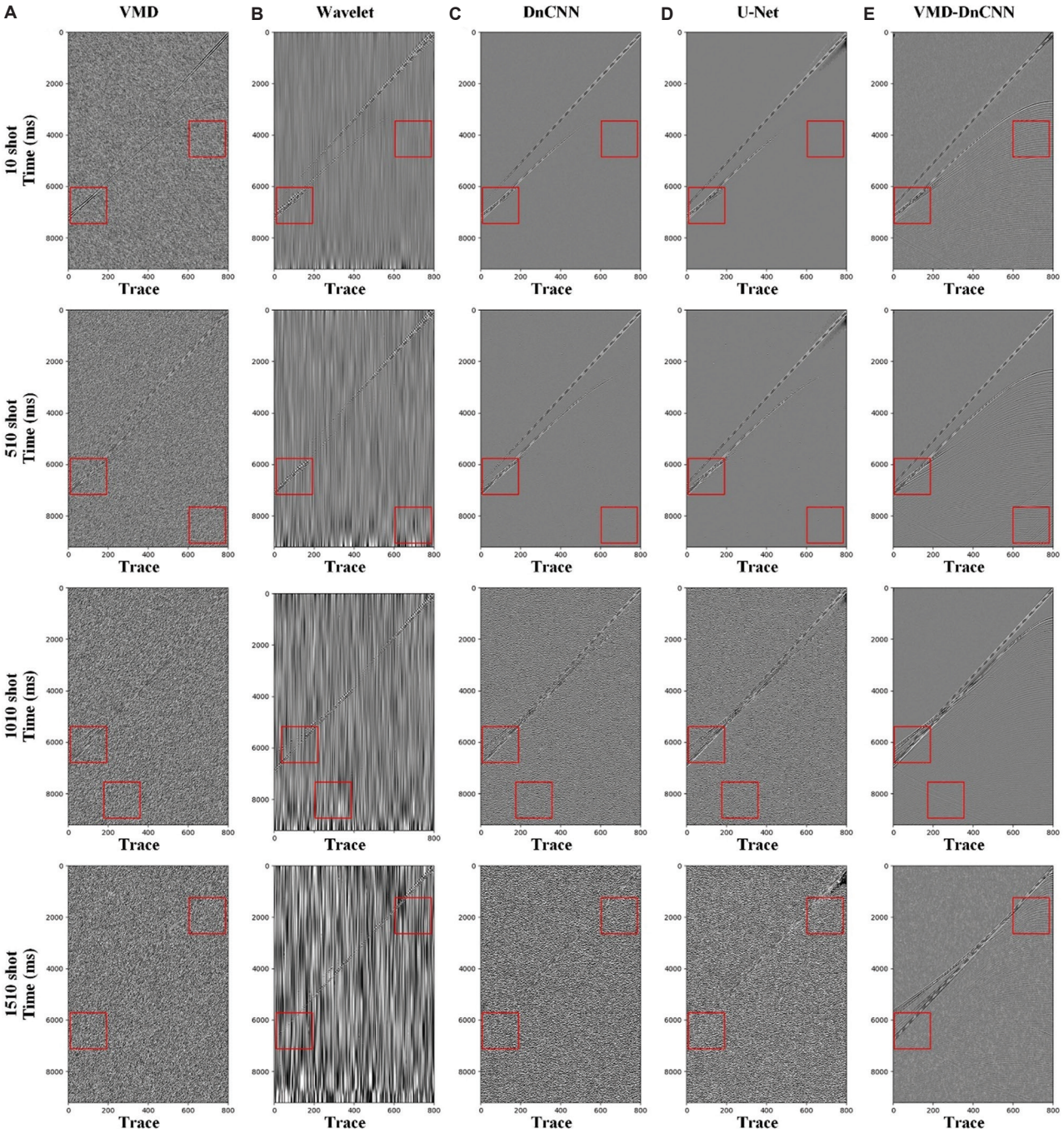

Figure 5. Denoising performance of different methods under different SNR conditions. Panels A–F display, from top to bottom, the denoised results corresponding to varying SNR levels for each respective method: (A) VMD, (B) wavelet, (C) DnCNN, (D) U-Net, and (E) VMD–DnCNN.
Abbreviations: DnCNN: Denoising convolutional neural network; SNR: Signal-to-noise ratio; VMD: Variational mode decomposition.

an acceptable range for practical use. Although deep learning methods required a one-time model pretraining phase, the training time was relatively acceptable. Once trained, the model could be reused without retraining, making the cost of pretraining negligible in the long term.

The robustness of each method across 17 different SNR levels was also analyzed using the performance curves, as shown in Figure 6. While all methods demonstrated a generally linear increase in SNR with decreasing noise, VMD–DnCNN achieved optimal denoising performance within the noise levels that were close to those used during

training. Even outside that range, although it did not always yield the highest SNR, it still consistently outperformed the other methods in denoising effectiveness. This phenomenon occurs because, outside the training interval, the frequencies of the high-, mid-, and low-frequency modes decomposed by 2D-VMD vary in response to differing noise intensities. Consequently, the network applies denoising based on signal feature frequencies learned during training, which may not accurately correspond to the characteristics of the input data.

Table 2. Result of processing at different SNRs by deep learning methods

Noisy record (dB)	Parameter	VMD	Wavelet	DnCNN	U-Net	VMD–DnCNN
6	SNR	12.40	12.87	19.45	19.77	23.07
	RMSE	0.6166	0.5842	0.2736	0.2639	0.1806
	PSNR	54.54	55.22	61.73	62.32	65.41
	SSIM	0.9992	0.9990	0.9998	0.9998	0.9999
0	SNR	7.85	9.14	17.90	18.15	27.03
	RMSE	1.0524	0.9072	0.3307	0.3216	0.1156
	PSNR	51.28	51.40	60.08	48.38	69.29
	SSIM	0.9981	0.9977	0.9997	0.9943	0.9999
−5	SNR	5.67	5.23	4.19	5.81	24.95
	RMSE	1.4332	1.5068	1.6990	1.4099	0.1556
	PSNR	47.87	46.99	45.76	48.38	66.71
	SSIM	0.9944	0.9939	0.9920	0.9943	0.9999
−10	SNR	0.80	2.25	−5.73	−3.47	23.23
	RMSE	2.8440	2.4054	6.0296	4.6496	0.2148
	PSNR	42.05	42.93	35.39	37.23	63.91
	SSIM	0.9817	0.9851	0.9173	0.9238	0.9998

Abbreviations: DnCNN: Denoising convolutional neural network; PSNR: Peak signal-to-noise ratio; RMSE: Root mean square error; SNR: Signal-to-noise ratio; SSIM: Structural similarity index measure; VMD: Variational mode decomposition.

As a frequency-domain analysis can better highlight seismic signal characteristics, this section further analyzes the denoising results using frequency–wavenumber (F–K) spectra, as shown in Figure 7. The denoising results of shot gather 510 were used as an example for detailed analysis. As shown in Figure 7A, the left panel shows the noise-free data, while the right panel demonstrates the data after adding Gaussian noise. Figure 7B–F displays the denoised results from different methods on the left, and the corresponding removed noise components on the right. As shown in Figure 7A, the dominant frequency of the original seismic signal lay in the range from 20–50 Hz. After adding noise, the frequency content shifted significantly into the 70–80 Hz range, completely overwhelming the original signal. By examining the denoising results, both the VMD and wavelet methods were found to remove portions of the strong direct wave signals. Specifically, VMD tended to misclassify parts of the original signal as noise, and its output was centered around 50 Hz. The wavelet method partially restored the original 20–50 Hz range in some regions, but also retained components in the higher-frequency band, indicating inconsistency in noise suppression. Among deep learning-based approaches, both U-Net and DnCNN exhibited better overall denoising performance. However, they tended to suppress certain high- and low-frequency components

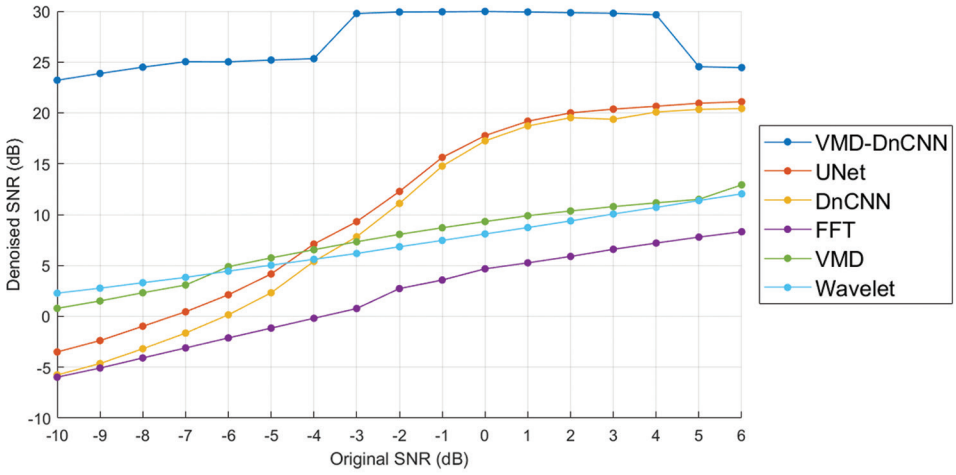

Figure 6. Denoising performance curves of the VMD, wavelet, DnCNN, U-Net, FFT, and VMD–DnCNN methods under different SNR conditions. Abbreviations: DnCNN: Denoising convolutional neural network; FFT: Fast Fourier transform; SNR: Signal-to-noise ratio; VMD: Variational mode decomposition.

Table 3. Computer performance analysis

Hyperparameter	VMD	Wavelet	DnCNN	U-Net	VMD–DnCNN
Average processing time (s)	66.260	1.291	2.012	3.343	2.178
Training time (h)	0	0	0.498	2.144	0.501
Average improved SNR (dB)	9.34	8.95	11.53	12.83	29.05
Memory cost (MB)	125.73	7.63	0.89	22.52	0.89

Abbreviations: DnCNN: Denoising convolutional neural network; SNR: Signal-to-noise ratio; VMD: Variational mode decomposition.

of the original signal, resulting in the recovered signals being mainly concentrated around the 30–40 Hz range. In contrast, the VMD–DnCNN method demonstrated superior performance by preserving both high- and low-frequency information. Within the red rectangular regions in Figure 7, VMD–DnCNN was observed to more effectively restore the seismic signal in areas with wider frequency separation, retaining more detailed seismic features. Overall, the VMD–DnCNN method demonstrated superior capability in restoring the signal's frequency content to its original distribution, thus offering

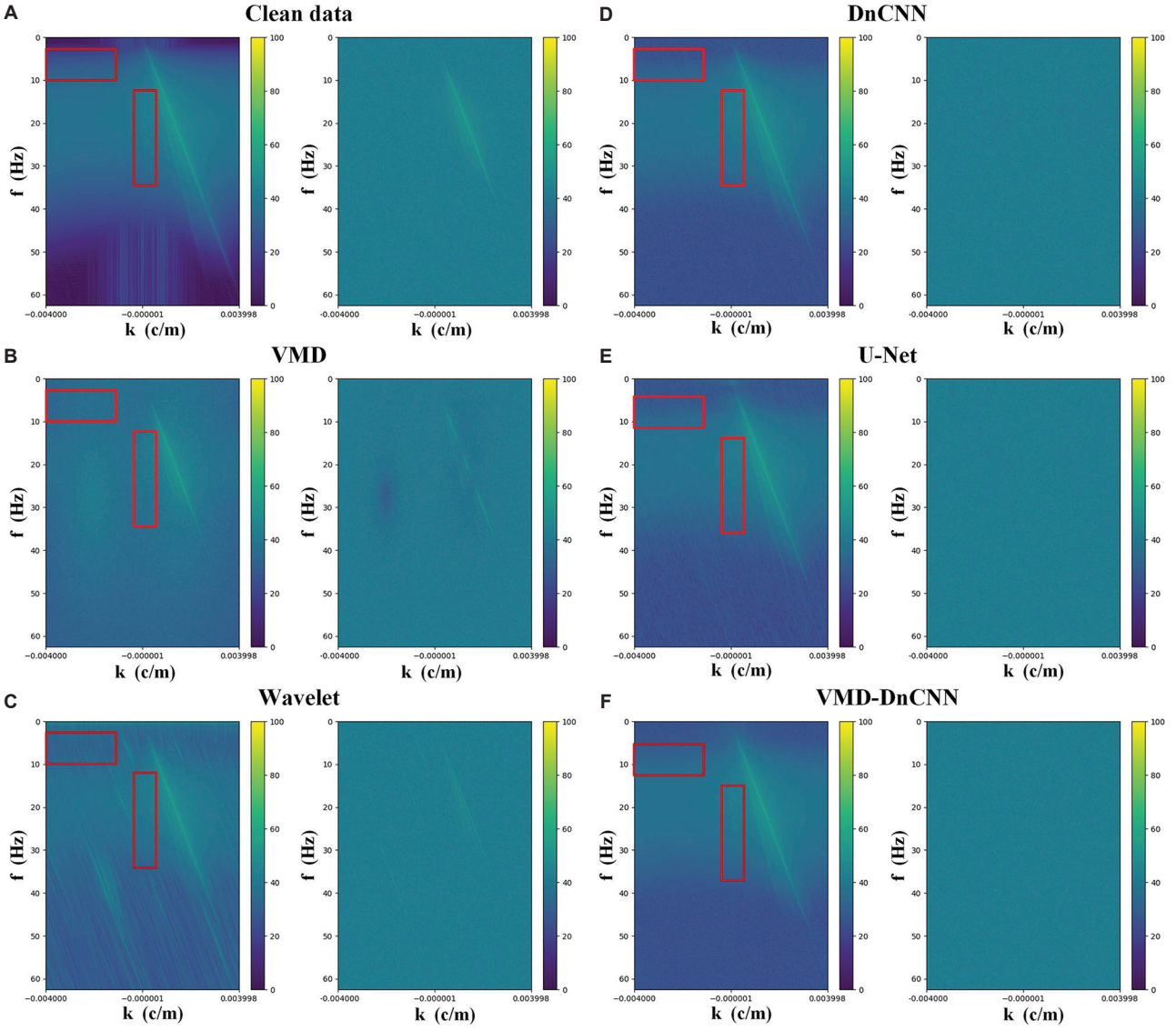

Figure 7. Frequency–wavenumber (F–K) spectrum analysis of noisy shot gathers 510 under different denoising methods: (A) clean F–K spectra, (B) VMD, (C) wavelet, (D) DnCNN, (E) U-Net, and (F) VMD–DnCNN.
Abbreviations: DnCNN: Denoising convolutional neural network; VMD: Variational mode decomposition.

a more accurate and comprehensive reconstruction in the frequency domain.

Subsequently, how the VMD–DnCNN network learns seismic signal features was analyzed by examining feature activation heatmaps at different network layers. The denoising results of shot 1,510 were used as the case study. As illustrated in Figure 8, the heatmaps correspond to the 1st, 4th, 10th, and 16th layers of the network, displayed from left to right. It can be observed that the network predominantly focused on low-frequency components of the seismic signals, with these features being highlighted from the initial layer, albeit in a relatively simplistic form. By the fourth layer, the network began to learn features at even lower frequencies while simultaneously capturing high-frequency components associated with direct arrivals. At the 10th layer, the features remained primarily low-frequency; however, by the 16th layer, the learned features shifted toward higher frequencies, incorporating both low- and high-frequency information. Given that seismic signals primarily consisted of low-frequency components, decomposing the signals into low-, mid-, and high-frequency bands through 2D VMD facilitated more effective feature learning by the network, surpassing the limitations of single time-domain feature extraction.

5. Field data experiment

In this section, the denoising performance of the proposed VMD–DnCNN model was evaluated using real seismic data, specifically marine seismic records. The real seismic data used in this experiment are shown in Figure 9A. By processing the field data, the effectiveness of VMD–DnCNN was demonstrated in real-world scenarios. For comparative analysis, the same baseline methods used in the synthetic data experiments were adopted. The real marine dataset used for evaluation was the 2D Mobil AVO Viking Graben Line 12 dataset. This dataset consists of 1,011 shot gathers, each containing approximately 119 seismic traces with 1,500 sampling points per trace. The sampling interval is 4,000 μs (i.e., 250 Hz sampling rate), resulting in a trace duration of 6 s. Shot gather 1 was selected for denoising analysis in this study. The denoising results of different methods applied to the real marine seismic data are shown in Figure 9.

Among the traditional methods, VMD effectively removed a significant amount of background noise and successfully separated weak signals from noise; however, some residual artifacts remained, and small portions of continuous strong signals were mistakenly removed as noise. Overall, the performance was relatively good. The wavelet method, on the other hand, tended to remove weak signals during denoising, produce residual artifacts around continuous strong signals, and leave some noise behind. For deep learning methods, both DnCNN and U-Net successfully suppressed background noise. However, the denoising effect of U-Net was inferior to that of DnCNN, as U-Net did not preserve weak signals well and exhibited residual artifacts in regions of continuous strong signals. In contrast, DnCNN showed relatively better overall denoising performance. Although some continuous strong signals were erroneously treated as noise, the continuity of the signals was largely preserved. In terms of weak signal processing, while residual noise still existed, a portion of the weak signals was retained, resulting in comparatively good denoising performance. In the case of VMD–DnCNN, although residual artifacts and incomplete removal of background noise remained, the method achieved a clear separation between weak signals

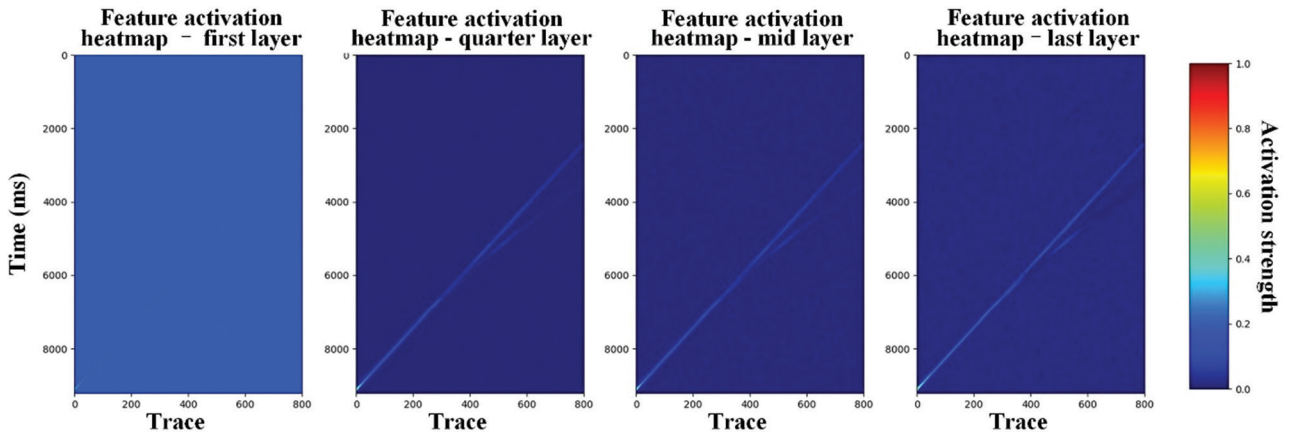

Figure 8. The VMD–DnCNN feature activation heatmap.
Abbreviation: Variational mode decomposition–denoising convolutional neural network.

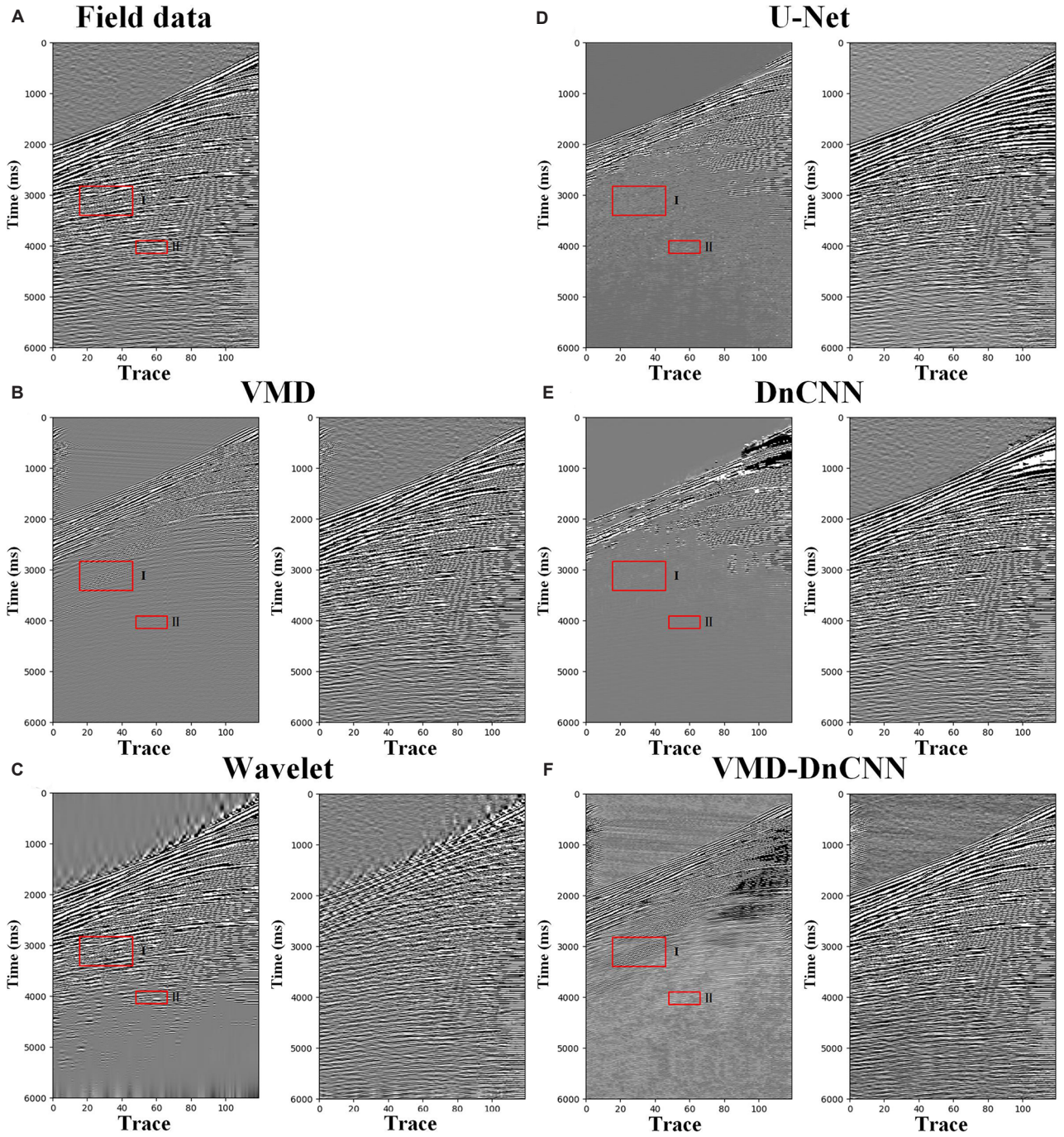

Figure 9. Denoising results on real marine seismic data by different methods: (A) noisy data, (B) VMD, (C) wavelet, (D) DnCNN, (E) U-Net, and (F) VMD–DnCNN.
Abbreviations: DnCNN: Denoising convolutional neural network; VMD: Variational mode decomposition.

and noise while preserving more weak signals overall. By examining the red-highlighted regions of the denoised results in Figure 10, as shown in the enlarged views of areas I and II, the VMD and wavelet methods left most of the noise residuals; U-Net recovered only a small fraction of the signals; DnCNN recovered some components, but with residual noise; and VMD recovered signals well. However, when these results were compared to VMD–DnCNN, the recovered signals by VMD–DnCNN were noticeably clearer.

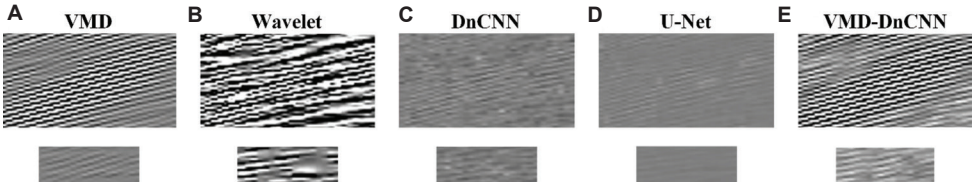

Figure 10. Crimson rectangular areas in Figure 9. From top to bottom are the zoomed-in views of regions I and II from different methods: (A) VMD, (B) wavelet, (C) DnCNN, (D) U-Net, and (E) VMD–DnCNN.
Abbreviations: DnCNN: Denoising convolutional neural network; VMD: Variational mode decomposition.

6. Conclusion

Traditional decomposition methods for seismic signal processing often suffer from the problem of mode mixing, making them difficult to completely prevent noise components from being incorporated into each mode, thereby reducing denoising accuracy. To address this issue, this study proposed a novel seismic signal denoising method that combines VMD with DnCNN—referred to as VMD–DnCNN. Seismic signals are primarily low-frequency, while noise is concentrated in high-frequency bands. Therefore, decomposition through VMD enables the network model to learn the characteristics of both seismic signals and noise more effectively. This method first applies VMD to decompose the noisy signal into three frequency-specific modes, corresponding to high-frequency noise, mid-frequency seismic phases, and low-frequency structural components. These modes, together with the original noisy signal, are used to construct a four-channel input, providing the deep learning network with clearer and more distinguishable frequency information. With this decomposition, the network not only learns to differentiate between effective signal and noise features across frequency bands but also leverages the complementary characteristics of different modes to better detect and preserve weak signals, especially under high-noise conditions. By adapting the DnCNN architecture and expanding the input channels, the VMD–DnCNN model is capable of extracting deep semantic features from seismic data while integrating both time-domain and frequency-domain information. This significantly enhances the model's denoising performance and generalization ability. Experimental results demonstrate that the proposed method outperforms both traditional and standalone deep learning approaches under various noise levels, and it remains effective in preserving fine signal structures even under low SNR conditions.

Despite its clear advantages in denoising accuracy and robustness, the VMD–DnCNN method has certain limitations in practical applications. First, it relies on a pre-decomposition step, where VMD must be applied to the input signal before feeding data into the network. This preprocessing not only increases the overall computational cost and complexity of the workflow but also makes performance highly sensitive to the VMD parameter settings. Inaccurate parameter choices may lead to suboptimal mode separation, which can adversely affect network learning. Second, as a variational optimization technique, VMD is computationally intensive—especially when dealing with large-scale 2D seismic data—and may significantly increase resource consumption. Moreover, mode mixing remains present to some extent, with noise potentially remaining in effective modes, interfering with the network's learning process. Therefore, future research can explore learnable decomposition mechanisms or end-to-end jointly optimized frameworks that integrate decomposition and network training in a unified architecture, further enhancing the automation and adaptability of seismic denoising.

Acknowledgments

All seismic datasets used in this study are sourced from the open-access resources released by the Society of Exploration Geophysicists (SEG). The authors would like to express their sincere gratitude to BP Exploration Operation Company for providing the high-quality seismic dataset, 2007 BP Anisotropic Velocity Benchmark dataset, to the geophysical research community.

Funding

This work was supported in part by the Deep Earth Probe and Mineral Resources Exploration-National Science and Technology Major Project (grant no.: 2024ZD1002202), in part by the Guizhou Provincial Basic Research Program (Natural Science) (grant no.: QianKeHeJiChu-K[2024] YiBan013), in part by the Guizhou University Basic Research Project (grant no.: GuiDaJiChu[2023]44), and in part by the Guizhou University Talent Introduction Research Project (grant no.: GuiDaRenJiHeZi[2023]10).

Conflict of interest

The authors have approved the submission and declare no conflicts of interest.

Author contributions

Conceptualization: Liang Zhang, Shengrong Zhang
Formal analysis: Xuesha Qin
Investigation: Liang Zhang, Shengrong Zhang
Methodology: Liang Zhang, Shengrong Zhang
Visualization: Xuesha Qin
Writing–original draft: Shengrong Zhang
Writing–review & editing: Liang Zhang, Xuesha Qin

Availability of data

All seismic datasets used in this study are sourced from the open-access resources released by the Society of Exploration Geophysicists (SEG; https://wiki.seg.org/wiki/2007_BP_Anisotropic_Velocity_Benchmark and https://wiki.seg.org/wiki/Mobil_AVO_viking_graben_line_12). The innovative code presented in this paper has been made open source. Interested researchers may download it directly from: https://github.com/zsr981103/VMD-DnCNN.git

References

1. Zhong T, Li F, Zhang R, Dong X, Lu S. Multiscale residual pyramid network for seismic background noise attenuation. *IEEE Trans Geosci Remote Sens*. 2022;60:1-14.

   doi: 10.1109/tgrs.2022.3217887

2. Dong X, Lin J, Lu S, Wang H, Li Y. Multiscale spatial attention network for seismic data denoising. *IEEE Trans Geosci Remote Sens*. 2022;60:1-17.

   doi: 10.1109/tgrs.2022.3178212

3. Cooper HW, Cook RE. Seismic data gathering. *Proc IEEE*. 1984;72:1266-1275.

   doi: 10.1109/PROC.1984.13016

4. Spikes KT, Tisato N, Hess TE, Holt JW. Comparison of geophone and surface-deployed distributed acoustic sensing seismic data. *Geophysics*. 2019;84:A25-A29.

   doi: 10.1190/geo2018-0528.1

5. Wang X, Ma J, Dong X, Cheng M. HCVT-Net: A hybrid CNN-transformer network for self-supervised 3D seismic data interpolation. *J Appl Geophys*. 2025;242:105873.

   doi: 10.1016/j.jappgeo.2025.105873

6. Dong X, Cheng M, Wang H, Li G, Lin J, Lu S. A potential solution to insufficient target-domain noise data: Transfer learning and noise modeling. *IEEE Trans Geosci Remote Sens*. 2023;61:1-15.

   doi: 10.1109/tgrs.2023.3300697

7. Anvari R, Siahsar MAN, Gholtashi S, Kahoo AR, Mohammadi M. Seismic random noise attenuation using synchrosqueezed wavelet transform and low-rank signal matrix approximation. *IEEE Trans Geosci Remote Sens*. 2017;55:6574-6581.

   doi: 10.1109/tgrs.2017.2730228

8. Mousavi SM, Langston CA. Automatic noise-removal/signal-removal based on general cross-validation thresholding in synchrosqueezed domain and its application on earthquake data. *Geophysics*. 2017;82:V211-V227.

   doi: 10.1190/geo2016-0433.1

9. Liu N, Gao J, Zhang B, Wang Q, Jiang X. Self-adaptive generalized S-transform and its application in seismic time-frequency analysis. *IEEE Trans Geosci Remote Sens*. 2019;57:7849-7859.

   doi: 10.1109/tgrs.2019.2916792

10. Neelamani R, Baumstein AI, Gillard DG, Hadidi MT, Soroka WL. Coherent and random noise attenuation using the curvelet transform. *Leading Edge*. 2008;27:240-248.

    doi: 10.1190/1.2840373

11. Starck JL, Candès EJ, Donoho DL. The curvelet transform for image denoising. *IEEE Trans Image Process*. 2002;11:670-684.

    doi: 10.1109/tip.2002.1014998

12. Do MN, Vetterli M. The contourlet transform: An efficient directional multiresolution image representation. *IEEE Trans Image Process*. 2005;14:2091-2106.

    doi: 10.1109/tip.2005.859376

13. Yu S, Ma J. Complex variational mode decomposition for slope-preserving denoising. *IEEE Trans Geosci Remote Sens*. 2017;56:586-597.

    doi: 10.1109/tgrs.2017.2751642

14. Bekara M, Van der Baan M. Random and coherent noise attenuation by empirical mode decomposition. *Geophysics*. 2009;74:V89-V98.

    doi: 10.1190/1.3157244

15. Liu N, Li F, Wang D, Gao J, Xu Z. Ground-roll separation and attenuation using curvelet-based multichannel variational mode decomposition. *IEEE Trans Geosci Remote Sens*. 2021;60:1-14.

    doi: 10.1109/tgrs.2021.3054749

16. Bekara M, Van der Baan M. Local singular value decomposition for signal enhancement of seismic data. *Geophysics*. 2007;72:V59-V65.

    doi: 10.1190/1.2435967

17. Dong X, Zhong T, Li Y. New suppression technology for low-frequency noise in desert region: The improved robust principal component analysis based on prediction of neural network. *IEEE Trans Geosci Remote Sens*. 2020;58:4680-4690.

    doi: 10.1109/tgrs.2020.2966054

18. Trickett S. F-xy cadzow noise suppression. In: *SEG Technical Program Expanded Abstracts*. United States: Society of

Exploration Geophysicists; 2008. p. 2586-2590.

doi: 10.1190/1.3063880

19. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. Vol 1. Cambridge: MIT Press; 2016.

20. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521:436-444.

doi: 10.1038/nature14539

21. Yu S, Ma J, Wang W. Deep learning for denoising. *Geophysics*. 2019;84:V333-V350.

doi: 10.1190/geo2018-0668.1

22. Wang X, Ma J, Dong X, Zhong T, Dong S. EFGW-UNet: A deep-learning-based approach for weak signal recovery in seismic data. *IEEE Trans Geosci Remote Sens*. 2024;62:1-3.

doi: 10.1109/tgrs.2024.3398590

23. Zhu W, Mousavi SM, Beroza GC. Seismic signal denoising and decomposition using deep neural networks. *IEEE Trans Geosci Remote Sens*. 2019;57:9476-9488.

doi: 10.1109/TGRS.2019.2926772

24. Zhang L, Li G, Chen H, *et al*. Identification and suppression of multicomponent noise in audio magnetotelluric data based on convolutional block attention module. *IEEE Trans Geosci Remote Sens*. 2024;62:5905515.

doi: 10.1109/tgrs.2024.3361942

25. Wei X, Zhang C, Jiang B, *et al*. *Efficient Seismic Data Interpolation Via Sparse Attention Transformer and Diffusion Model*. [arXiv Preprint]; 2025.

doi: 10.48550/arxiv.2506.07923

26. Zhang K, Zuo W, Chen Y, Meng D, Zhang L. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Trans Image Process*. 2017;26:3142-3155.

doi: 10.1109/TIP.2017.2662206

27. Yu S, Ma J, Osher S. Monte Carlo data-driven tight frame for seismic data recovery. *Geophysics*. 2016;81:V327-V340.

doi: 10.1190/geo2015-0343.1

28. Kingma DP, Ba J. *Adam: A Method for Stochastic Optimization*. [arXiv preprint]; 2014.

doi: 10.48550/arXiv.1412.6980

29. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015: 18th International Conference, Munich, Germany, Proceedings, Part III*. Vol 18. Springer; 2015. p. 234-241.

doi: 10.1007/978-3-319-24574-4_28