

FEW-SHOT SEISMIC DATA DENOISING VIA AUGMENTATION WITH AN IMPROVED DIFFUSION MODEL

LIQUAN ZHAO^{1,2}, PENG LIU² AND TIE ZHONG^{1,2,*}

¹Key Laboratory of Modern Power System Simulation and Control
and Renewable Energy Technology, Ministry of Education

²School of Electrical Engineering
Northeast Electric Power University

No. 169, Changchun Road, Jilin 132012, P. R. China

{ zhaoliquan; 2202300390 }@neepu.edu.cn; *Corresponding author: zht@neepu.edu.cn

Received September 2025; revised December 2025

ABSTRACT. *In seismic exploration, enhancing the signal-to-noise ratio (SNR) of seismic data is crucial for effective interpretation. Although numerous studies have demonstrated that supervised deep learning methods outperform traditional approaches, the denoising performance of supervised deep learning is highly dependent on the size and diversity of the training dataset. However, acquiring a sufficient amount of real seismic noise data remains challenging, which restricts its practical applications. To address this limitation, firstly, this paper proposed a data augmentation strategy based on the denoising diffusion probabilistic model, in which the reverse prediction network is improved to reduce the reliance of supervised deep learning on large-scale training datasets, using distributed acoustic sensing data as an example. In this study, the progressive noise injection and denoising technique of the enhanced diffusion model is employed to construct a latent mapping between seismic data and Gaussian noise. This method is able to generate enough synthetic noise data with a probability distribution resembling that of actual seismic noise. Secondly, a dual-encoder complementary noise suppression network is proposed by integrating dilated convolutions and wavelet transforms, which allows comprehensive extraction of both deep and shallow features, followed by intelligent feature fusion through an attention mechanism. Experimental evaluations on both synthetic and field seismic data demonstrate that the proposed strategy effectively suppresses complex noise while recovering a weak signal.*

Keywords: Data augmentation, Denoising diffusion probabilistic model (DDPM), Deep learning, Seismic data denoising

1. Introduction. This technique uses underground observation and analysis of the propagation patterns of artificially generated seismic waves to infer subsurface geological structures and resource distribution [1]. However, the gathered seismic data is frequently tainted by significant noise, which significantly hampers subsequent applications. This is because complex exploration environments and equipment have an impact on the data acquired, and random noise is complex and unpredictable [2,3]. For many geophysical applications [4,5], increasing the SNR of seismic data is essential. However, there are still two significant issues with this process: a lack of training data with noisy-clean pairs and inadequate denoising performance [6,7].

Among the existing methods for enhancing seismic data SNR, traditional approaches primarily rely on filtering techniques and signal decomposition [8,9]. Band-pass filtering separates effective signals from linear noise via a two-dimensional Fourier transform [10], which performs well for dipping noise (e.g., surface waves) but suffers from the loss

of low-frequency components in the effective signal. Wavelet threshold filtering can effectively extract frequency-domain signals, but risks signal distortion and edge artifacts [11]. In recent years, methods for reducing seismic noise attenuation have been proposed [12,13]. However, due to the limitations of linear low-rank assumptions and the non-linear complexity of seismic data, these methods struggle to recover weak signals and ensure the authenticity of the denoised results when handling complex seismic data. Zhao et al. utilized the DnCNN network to train a model for seismic data denoising [14]. Consequently, researchers have continuously innovated network architectures to improve seismic data SNR further. A novel residual learning network (ResNet) was proposed to effectively address nonstationary noise in field seismic data [15]. Zhong et al. introduced a multi-scale structure based on CNN and designed a U-Net [16]. Zhang et al. incorporated generative adversarial networks (GANs) into seismic data denoising and developed MSRD-GAN [17], which produces more realistic denoising results, particularly excelling under low signal-to-noise ratio conditions. The emergence of Transformer networks and Mamba [18,19] enabled the handling of more complex seismic data, albeit with overly complex network structures. Wang et al. integrated an edge mechanism into the U-Net and proposed EFGW-UNet [20].

These techniques provide satisfactory denoising results, but training these network structures requires a lot of data. However, obtaining high-quality seismic noise-clean data pairs is always difficult due to the time and effort required for data collection and annotation [21], which inevitably restricts the use of deep learning-based seismic denoising techniques. Although algorithms and techniques for denoising small-sample seismic data [22,23] are possible, they have drawbacks like poor generalization ability, a propensity for overfitting, and model instability. As a result, having enough training data is essential [24]. Geometric transformations (flipping, rotation, scaling), the addition of Gaussian noise, and elastic deformation are examples of traditional data augmentation techniques [25-27], but they are ineffective at increasing sample diversity. Dong et al. proposed the use of a WGAN model trained on real seismic noise data, with the aim of generating sufficient quantities of noise patches [28]. The synthetic data was subsequently used to reconstruct the seismic data. Although WGAN mitigates the mode collapse problem inherent in conventional GANs through Wasserstein distance metrics, the generator may still produce similar samples when dealing with complex data distributions or insufficient training samples, thereby limiting the diversity of augmented data. Furthermore, WGAN-generated samples tend to concentrate predominantly in high-density regions of the data distribution while demonstrating limited capacity for generating edge or rare samples – precisely those critical for enhancing model robustness through data augmentation.

Diffusion models have recently gained widespread attention due to their high generation quality and stability [29,30]. Unlike conventional approaches, diffusion model-based augmentation tends to produce samples that cover the entire data distribution, rather than focusing solely on high-density regions, thereby facilitating the augmentation of edge samples. Consequently, this paper proposes a diffusion model-based seismic noise data augmentation framework, utilizing 600 real seismic noise patches (64×64) as the training set for the diffusion model. Additionally, we optimize the prediction network in the reverse diffusion process to enhance the noise prediction capability, enabling the model better to learn the probability distribution of seismic noise data and generate sufficient training samples.

Fine-scale details are particularly susceptible to noise interference, while semantic-level structures may be misinterpreted under low signal-to-noise ratio conditions. We propose a dual-encoder framework within the MW-UNet architecture, where two distinct yet

complementary pathways extract and encode information from different perspectives. The main contributions can be summarized as follows.

1) This study proposes an improved diffusion model to address the problem of insufficient training data in seismic data denoising. By leveraging a limited amount of seismic data collected in exploration, the model generates sufficient training samples. Specifically, the reverse prediction network of the diffusion model is enhanced to improve noise estimation during the reverse denoising process. After training, the model can accept random Gaussian noise as input and gradually eliminate it through the reverse process to generate synthetic samples. The proposed method improves upon existing data augmentation techniques by enhancing the prediction network in the reverse process, resulting in superior sample quality, enhanced generalization capability, and greater diversity in generated outputs.

2) We present MW-UNet, a novel architecture that combines complementary detail feature extraction and primary feature extraction for seismic noise attenuation using diffusion-model-augmented data. We employ two encoders operating at different resolutions to extract both deep and shallow information features. The method effectively suppresses noise while preserving weak signals and maintaining seismic signal continuity, even when processing complex seismic data.

3) The proposed multi-scale module, combined with a convolutional block attention module (CBAM), selectively enhances the most informative features while diminishing the influence of secondary characteristics, significantly improving noise suppression performance for complex seismic data. Furthermore, our wavelet-enhanced (WE) feature extraction module integrates the wavelet transform with convolutional layers, enabling effective decomposition of frequency-domain information. This hybrid approach aims to extract critical frequency-domain features and enhance the recovery of weak seismic signals.

The rest of this article is organized as follows. Section 2 briefly introduces the augmented model and the denoising network architecture. Section 3 describes the process of noise data augmentation and provides a comparative analysis of the augmented data. Section 4 compares the denoising performance of the proposed model by processing synthetic and field records. Section 5 discusses in detail the strengths and limitations of the performance method. Finally, Section 6 concludes the paper.

2. Method. To address the challenge of insufficient seismic data, we proposed an improved DDPM to learn the distribution characteristics of seismic noise from limited seismic datasets, thereby generating sufficient synthetic seismic data as illustrated in Figure 1. It is a genus of plants. For effective noise attenuation and weak signal recovery in complex seismic data, we propose MW-UNet, which hierarchically extracts signal features through two encoders – a high-level encoder and a low-level encoder. A feature fusion module is used to combine these features to generate a semantically rich feature representation.

2.1. Seismic noise data augmented model. Based on the traditional DDPM framework, we improve the prediction network to strengthen its noise estimation capability. In the improved DDPM, the distribution difference between the predicted noise in the reverse process and the actual added noise in the forward process is minimized. The loss function of DDPM is as follows:

$$L_{simple} = \|\epsilon_t - \epsilon_\theta(x_t, t)\|^2 \quad (1)$$

The loss function of the DDPM can be seen that by learning the gradient field of the data distribution, the error of the predicted noise is reduced, and stable generation is

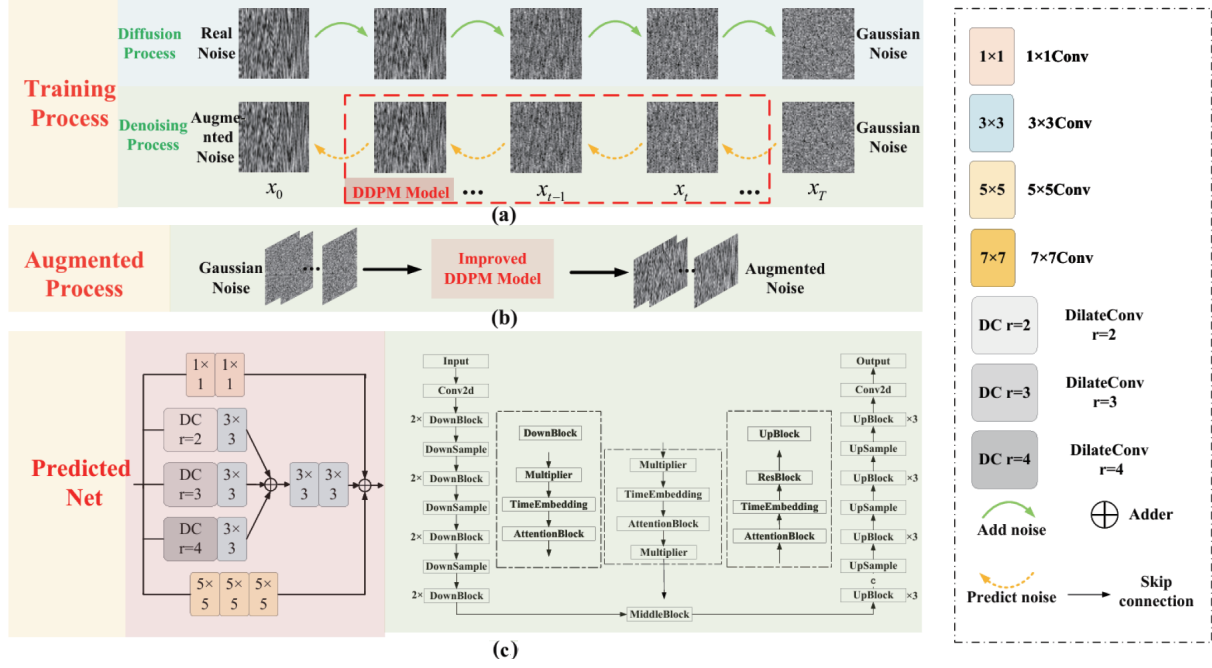


FIGURE 1. Work flow of augmentation process: (a) Training process of DDPM; (b) augmentation process using reverse denoising process; (c) improved predicted U-Net architecture

achieved through an iterative process. This design gives it significant advantages in terms of generation quality, stability, and theoretical interpretability.

The seismic data generation process consists of two stages: model training and data generation. Model training procedure is shown in Figure 1(a). First, input data are sampled from a dataset of 600 seismic noise records. Gaussian noise conforming to a standard distribution is then incrementally added to the input data over multiple timesteps, ultimately yielding fully perturbed seismic noise data x_t . A network based on U-Net was reconstructed to predict the Gaussian noise in the reverse process. Figure 1(c) shows the U-Net network. The gradient information is computed based on the loss function between the noise added in the forward process and the noise predicted in the reverse process, followed by updates to the network parameters via backpropagation.

The improved reverse prediction network replaces the ResBlocks in the original down sampling pathway with the multi-scale module illustrated in Figure 1(c). While the receptive field of the original ResBlocks remains limited (even when stacked), the multi-scale module incorporates parallel convolutions with varying dilation rates ($r = 2, 3, 4$), significantly expanding the network's contextual perception range. This enriched feature representation provides stronger structural priors for the diffusion process, enabling more accurate restoration of realistic semantic and structural consistency in the denoised predictions at each step.

The process begins by sampling random noise from a standard Gaussian distribution shown in Figure 1(b). The trained improved diffusion model then progressively denoises the data, utilizing the improved U-Net to predict and remove noise iteratively. After T iterations, the generated samples correspond to the distribution of the original seismic data, as formulated in Equation (2).

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{\beta_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_{\theta}(x_t, t) \right) + \sqrt{\frac{1 - \bar{\alpha}_{t-1}}{1 - \bar{\alpha}_t}} \beta_t z \quad (2)$$

In the formula, $\epsilon_\theta(x_t, t)$ represents the Gaussian noise predicted by the improved U-Net, and z represents the standard Gaussian noise. Samples similar to the probability distribution of the input data are obtained by removing the noise through multiple steps.

2.2. Noise attenuation network architecture. We propose a dual-encoder design within the MW-UNet architecture shown in Figure 2 to effectively address the dual challenge of preserving fine-grained image details while capturing high-level semantic structures.

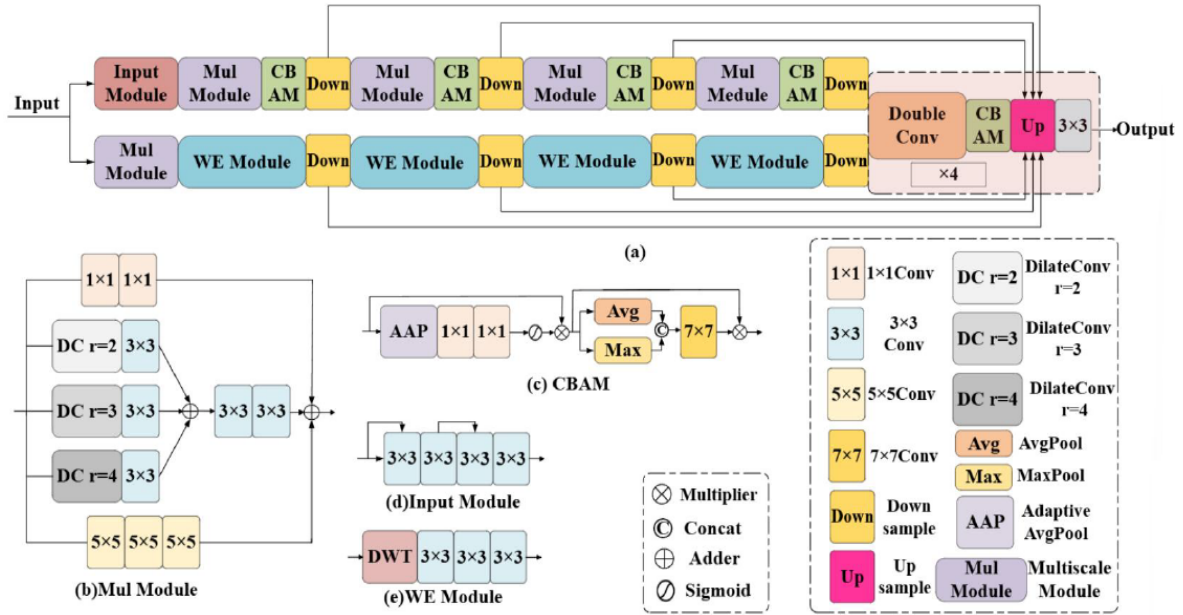


FIGURE 2. Network structure of the MW-UNet: (a) Overall structure of MW-UNet; (b)-(e) structure of multi-scale module, CBAM, input module, and WE module

1) The high-level encoder uses a multi-scale module shown in Figure 2(b). It also uses CBAM shown in Figure 2(c). These components capture semantic context and global architectures. They provide the network with a broader understanding of spatial dependencies. The multi-scale module excels at noise suppression. It extracts features at different receptive fields. This enhances the network's ability to represent multi-scale noise and structural details.

The multi-scale module has three parallel branches. The first branch consists of two 1×1 convolutions, forming a bottleneck structure that preserves information while improving training stability in deep networks. The second branch uses three 3×3 dilated convolutions with dilation rates ($r = 2, 3, 4$) to efficiently capture multi-scale contextual information while maintaining local detail perception. The module achieves exponential expansion of the receptive field without increasing parameters, ensuring computational efficiency and feature continuity, a key factor for its effectiveness in denoising tasks. Two subsequent 3×3 standard convolutions refine the fused features, eliminating potential discontinuities and artifacts caused by dilated convolutions while enhancing local consistency. The third branch contains three standard 5×5 convolutions to supplement mid-scale feature extraction in the spatial dimension. These three branches cover feature extraction from pixel-level to image-patch levels. The final output combines features from all branches, enabling comprehensive extraction of both global and local features while suppressing most noise in complex seismic data.

However, features extracted by different branches of the multi-scale module may exhibit information overlap, particularly in smooth background regions and noisy areas that generate similar responses. Additionally, some high-frequency noise and true edges may show comparable activation intensities in multi-scale features, requiring dynamic discrimination.

The CBAM addresses these challenges through its dual attention mechanism, which consists of channel attention and spatial attention components that optimize features from different dimensions. The channel attention mechanism adaptively recalibrates channel weights, suppressing noise-dominated channels while enhancing useful signal channels. The spatial attention component focuses on noise-dense regions while protecting detail-rich areas. This combined approach ensures effective noise suppression while maintaining structural integrity and semantic continuity in the processed seismic data.

2) The low-level encoder focuses on extracting local features such as textures, edges, and high-frequency components through the integration of wavelet transform and WE modules. This branch serves as a detail preserver, maintaining structural integrity even under severe noise conditions. The input first undergoes processing through the multi-scale module to suppress the majority of noise, after which the features are fed into the WE module.

The wavelet-enhanced convolution (WE) represents a key innovation in the low-level encoder, combining wavelet transform with convolutional operations to achieve more efficient separation of noise from useful signals and recovery of weak signals. While traditional convolution operates solely in the spatial domain, WE module integrates the Haar wavelet transform (Equation (3)) to enable simultaneous feature analysis in both frequency and spatial domains.

$$\begin{aligned} f_{LL} &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, & f_{LH} &= \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}, \\ f_{HL} &= \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}, & f_{HH} &= \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \end{aligned} \quad (3)$$

where f_{LL} denotes the low-pass filter, f_{LH} denotes the vertical high-pass filter, f_{HL} denotes the horizontal high-pass filter, and f_{HH} denotes the global high-pass filter. The Haar wavelet transform decomposes data into low-frequency approximations and high-frequency details through simple matrix operations, providing the mathematical foundation for frequency-domain processing in the WE module. The multi-receptive field properties of convolution are complementary to the multi-receptive field properties of wavelet transform, thus preventing the loss of frequency-domain information that occurs with conventional convolution approaches.

MW-UNet utilizes skip connections at multiple levels and a shared decoder that combines features extracted by both branches, enabling complementary learning. The high-level encoder focuses on extracting more abstract semantic and contextual information, where multi-scale receptive fields and the attention mechanism (CBAM) help identify overall shapes and semantic relationships. The low-level encoder simultaneously preserves local high-frequency features, maintaining the original details even under strong noise interference. This complementary interaction enables the model to perform accurate denoising while preserving structural integrity and semantic correctness, a capability particularly valuable in high-noise environments.

3. Statistical Analysis of Augmented Data. While SDL-based denoising techniques require extremely realistic datasets to optimize network parameters, DDPM typically requires seismic noise data as the training set.

3.1. Noise datasets augmentation using DDPM. Using DAS data processing as an example, the enhanced DDPM was trained on approximately 600 64×64 noise patches extracted from the original shot gathers. Approximately 18,000 augmented noise patches were generated by synthesizing additional noise patches using the trained diffusion model. For the ensuing CNN training, these produced noise samples created an improved noise dataset. The augmented noise's probability distribution closely matched that of the real noise in the original shot gathers, and the generated samples' high quality and diversity satisfied the deep learning training requirements thanks to the synthetic dataset. Real and augmented noise patch examples are shown in Figure 3. Their energy distributions and vibration patterns are strikingly similar.

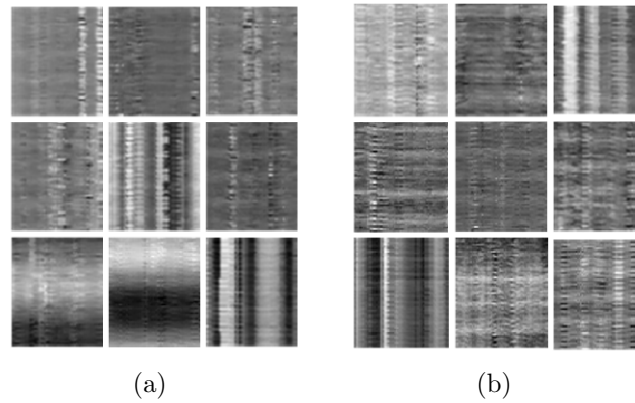


FIGURE 3. Samples of real and augmented data: (a) Real seismic noise patches; (b) augmented noise patches

3.2. Statistical analysis of augmented data. We created 18,000 synthetic noise samples and trained a WGAN network using 1,000 real seismic noise patches of size 64×64 to assess the enhanced diffusion model's augmentation capabilities. Likewise, using the specific parameters provided in Table 1, the traditional DDPM was trained to generate an additional 18,000 noise samples. The similarity between augmented and real noise data for the three approaches was evaluated using statistical features and probability density distributions. Additionally, 1,000 patches of augmented and real noise were chosen at random for statistical comparison. The WGAN-augmented data in Figure 4(a) exhibited a steeper tail decay and a sharper main peak, indicating potential issues with limited diversity and mode collapse. The improved DDPM showed a main peak position closer to the real data than the conventional DDPM, indicating a higher degree of similarity to actual seismic noise. Regarding statistical features, which are illustrated in Figure 4(b), the enhanced DDPM generated data with mean values that were more in line with the actual noise

TABLE 1. Hyper-parameter setting of DDPM and improved DDPM

Hyper-parameter	DDPM	Improved DDPM
Optimizer	ADAM	ADAM
Patch size	64×64	64×64
Batch size	32	32
Number of epochs	1000	1000
Num_timesteps	2000	2000
Learning rate range	$[10^{-4}, 0.02]$	$[10^{-4}, 0.02]$

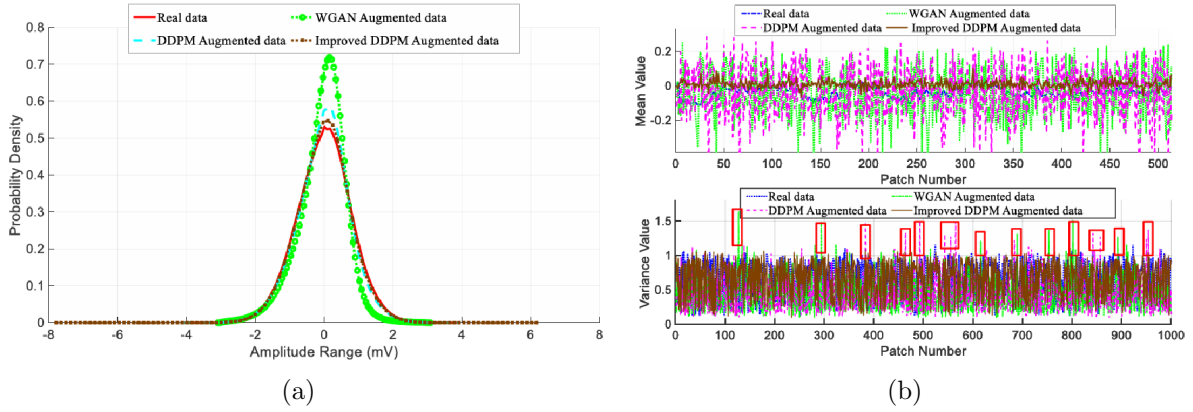


FIGURE 4. Analysis in statistical properties: (a) Comparison in probability density; (b) comparison in mean and variance

samples. The traditional DDPM and WGAN showed wider distribution ranges and larger fluctuations (shown in the red box) in the variance comparison. On the other hand, the enhanced DDPM successfully balanced stability and diversity by achieving a very consistent amplitude fluctuation range. Overall, the suggested approach showed better modeling capability and higher augmentation quality than the other methods, outperforming them in both realism and distribution consistency.

3.3. Denoising results on synthetic record. Vertical seismic profiling (VSP) data acquisition has utilized DAS, a new acquisition technology. We created one noise dataset with actual seismic noise and three additional noise datasets produced by various augmentation techniques to assess the effectiveness of these techniques. Each dataset included 18,000 64×64 noise patches for fairness. Both the U-Net and WM-UNet networks were used for denoising synthetic records to ensure the persuasiveness of the results; Table 2 provides a summary of the network parameters.

TABLE 2. Training parameters of the CNN-based networks

Hyper-parameter	DnCNN	U-Net	MW-UNet
Optimizer	ADAM	ADAM	ADAM
Patch size	64×64	64×64	64×64
Batch size	128	128	128
Number of epochs	60	60	60
Input channels	1	1	1
Learning rate range	$[10^{-3}, 10^{-5}]$	$[10^{-3}, 10^{-5}]$	$[10^{-3}, 10^{-5}]$

The synthetic record was processed using denoising models that had been trained using both real and augmented noise. To better examine the precision and applicability of the suggested augmentation techniques, comparative studies were carried out.

To recover effective signals and suppress noise in the simulated records, it is important to note that the noise datasets enhanced by all three techniques can be utilized as training sets. However, there are minor variations between the three augmentation techniques when compared to the outcomes produced using actual noise data shown in Figure 5(b). Visual examination shows that the residual noise (second row in Figure 5) shows different levels of signal leakage under the same denoising network. While the improved DDPM-augmented data yields residual noise that more closely resembles the results obtained with real noise

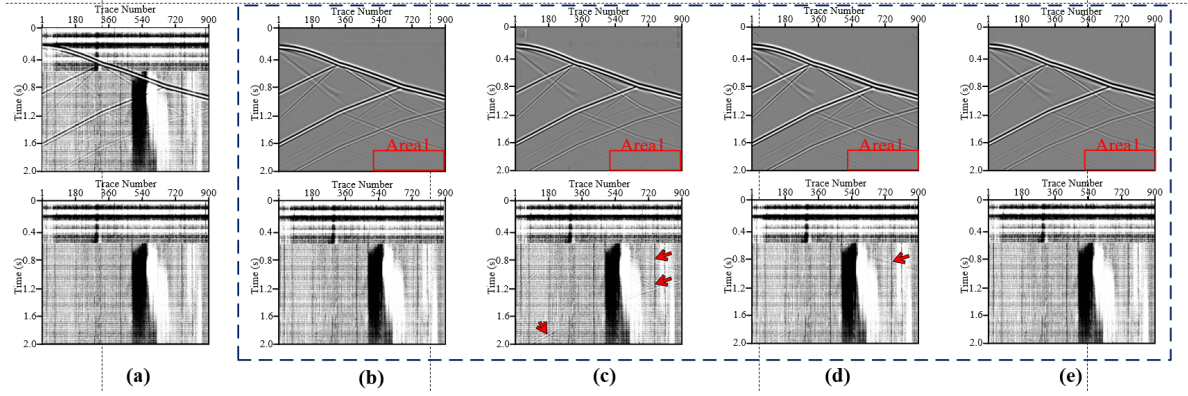


FIGURE 5. Comparisons for denoising results and filtered noise: (a) Noisy synthetic record and noise data: The blue dashed box represents the results of U-Net; (b)-(e) denoising results of real data, WGAN augmented data, DDPM augmented data, and improved DDPM augmented data

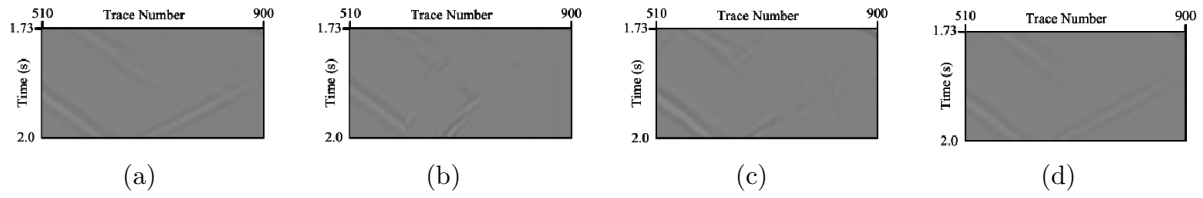


FIGURE 6. The crimson rectangular areas in Figure 5: (a)-(d) Enlarged results of real data, WGAN augmented data, DDPM augmented data, and improved DDPM augmented data, respectively

data, the red arrows for the U-Net residual noise results (second row in Figure 5) show that signal leakage is more noticeable when using WGAN-augmented or conventional DDPM-augmented data.

The red-boxed areas were enlarged in order to look more closely at these variations, as seen in Figure 6. The simulated record processed by WM-UNet is shown in Area 1, where it is clear that the real data and the improved DDPM-augmented data both recover weak signals more successfully than the other two approaches.

Single-trace waveforms were examined in the time and frequency domains to provide a more logical comparison of the simulated record results derived from the three augmented datasets, as illustrated in Figure 7. The findings show that denoising results from the enhanced DDPM-augmented data are more in line with those from real data.

4. Denoising Performance Comparison. To assess the denoising performance of MW-UNet, we compared it with CNN-based networks, specifically DnCNN and U-Net, as well as traditional techniques such as WT and BPF. Following the extraction of wavelet coefficients, the processed signal was decomposed into 15 levels in WT using the db5 wavelet function. The passband for the BPF was set between 30 and 70 Hz. Table 2 provides a summary of the DnCNN and U-Net network parameters. All CNN-based techniques were trained using forward-modeled data as the signal set and enhanced DDPM-augmented noise data as the noise set in order to guarantee an equitable comparison.

4.1. Denoising results on synthetic records. The denoising results for the synthetic record are shown in Figure 8. Figures 8(b) and 8(c) demonstrate how the traditional techniques struggle to suppress complex noise, leaving a significant amount of residual

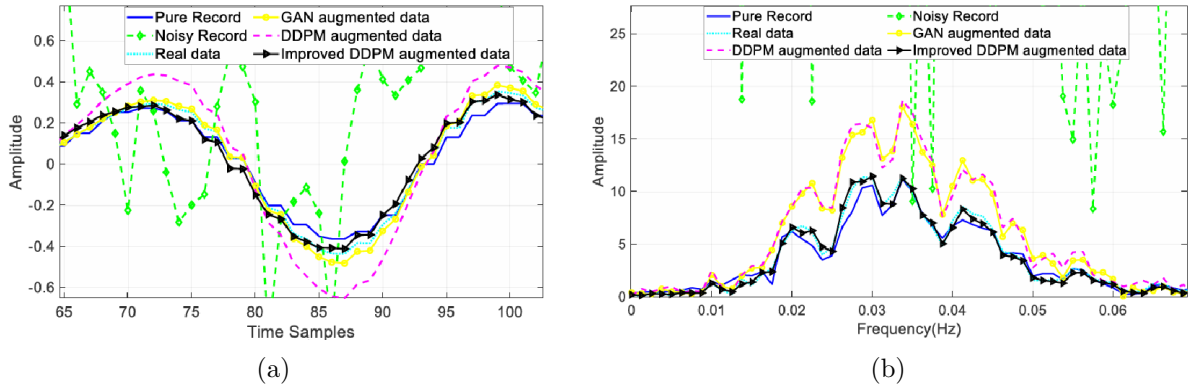


FIGURE 7. Comparisons of trace records: (a) Temporal results; (b) spectral results

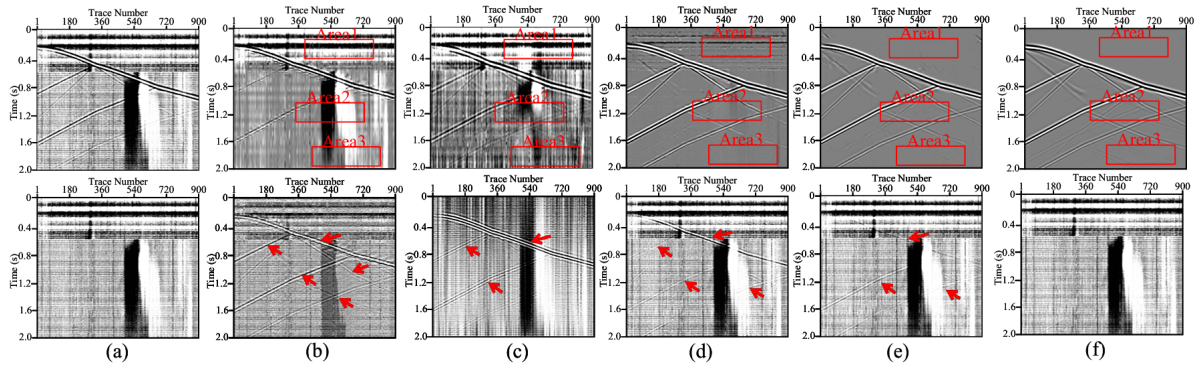


FIGURE 8. Comparisons for denoising results and filtered noise: (a) Noisy synthetic record and noise data; (b)-(f) denoising results of WT, BPF, DnCNN, U-Net, and MW-UNet

noise in the output and making it difficult to recover the tainted signals. CNN-based networks, on the other hand, can typically recover legitimate signals while suppressing the majority of complex noise. Nonetheless, some background noise is still present in the DnCNN results (top of Figure 8(d)). Furthermore, as shown by the red arrows in Figure 8, there are varying degrees of signal leakage in the residual noise for both conventional methods and CNN-based networks (DnCNN and U-Net), with the conventional methods exhibiting the most severe leakage.

As seen in Figure 9, we enlarged the red-boxed areas to analyze the denoising outcomes further. The enhanced view of Area 1 clearly demonstrates that MW-UNet suppresses complex noise with noticeably improved performance. While MW-UNet can successfully restore weak signals, DnCNN and U-Net show insufficient capacity to do so in Areas 2 and 3. The F-K domain analysis shown in Figure 10 confirms that MW-UNet performs noticeably better at denoising than DnCNN, U-Net, and conventional methods, providing a clearer picture of weak-signal recovery and signal leakage within the red-boxed areas.

Moreover, we also select a trace record for a detailed comparison. As shown in Figure 11, MW-UNet's denoised results match the clean signals the best, followed by U-Net and DnCNN, while the conventional methods perform the worst when single-trace records are compared. The accuracy and efficacy of signal recovery were quantitatively assessed using the signal-to-noise ratio (SNR) and root mean square error (RMSE), which are defined

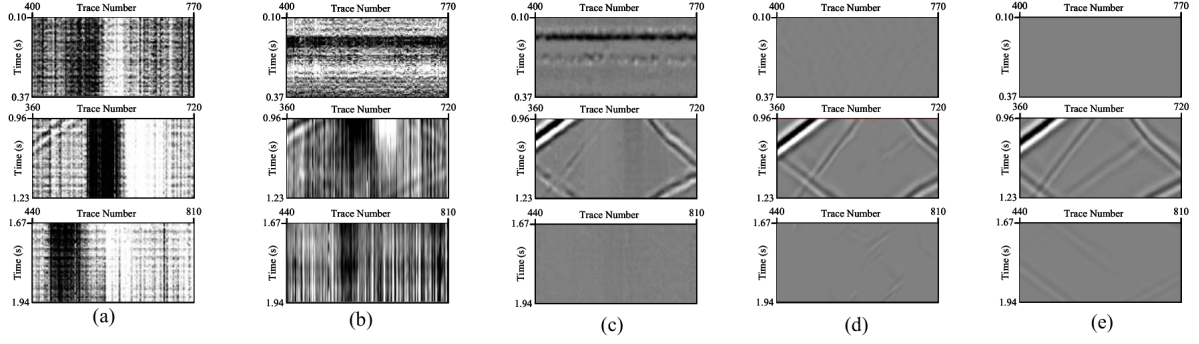


FIGURE 9. The crimson rectangular areas in Figure 8. The figures in the first, second, and third rows are from Area 1, Area 2, and Area 3, respectively. (a)-(e) Enlarged results of WT, BPF, DnCNN, U-Net, and MW-UNet, respectively.

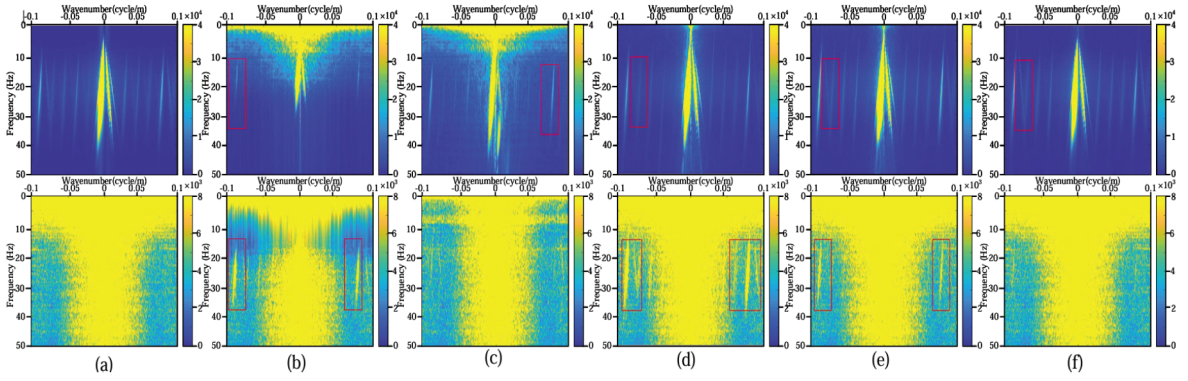


FIGURE 10. F-K domain analysis of results in Figure 8: (a) F-K spectrum of pure synthetic record and noise data; (b)-(f) F-K spectra of denoising results (top subplots) and filtered noise (bottom subplots) for WT, BPF, DnCNN, U-Net, and MW-UNet

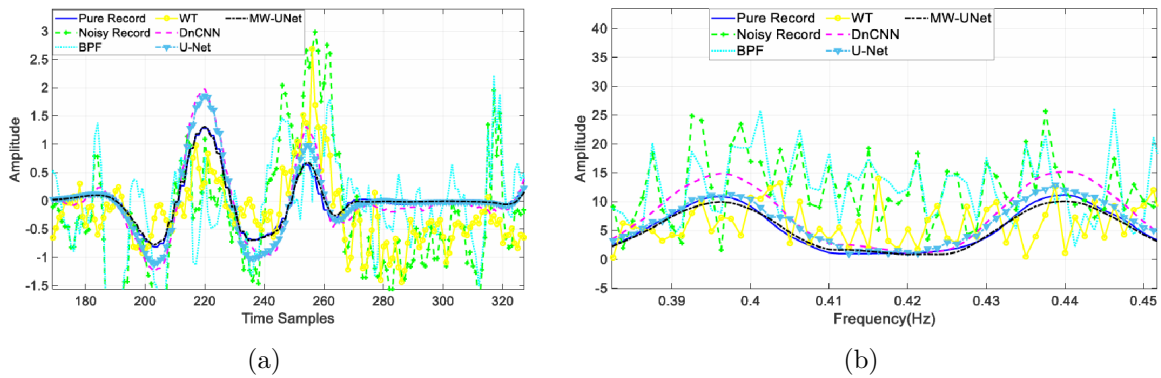


FIGURE 11. Comparisons of trace records: (a) Temporal results; (b) spectral results

as follows:

$$SNR(dB) = \frac{10 \log_{10} \sum_{i=1}^N s_i^2}{\sum_{i=1}^N (d_i - s_i)^2} \quad (4)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (d_i - s_i)^2}{N}} \quad (5)$$

In the formulas, s_i denotes the clean signal, d_i represents the denoised (predicted) signal, and N is the number of sampling points. A higher SNR value indicates better denoising performance, whereas a lower RMSE value reflects higher prediction accuracy. Table 3 shows the comparative results. MW-UNet has the best SNR and RMSE values, demonstrating its superior noise suppression ability and ability to recover legitimate signals reliably.

TABLE 3. SNR and RMSE

Method	WT	BPF	DnCNN	U-Net	MW-UNet
SNR	-2.2739	5.9602	14.6744	16.5213	20.2336
RMSE	0.7236	0.2741	0.1954	0.0835	0.0564

4.2. Denoising results on real records. We used the techniques on actual DAS records in Figure 12(a) for analysis in order to further guarantee the dependability of the denoising performance. Figures 12(b) and 12(c) demonstrate that WT and BPF's capacity to suppress noise is still restricted. Within the remaining noise in Area 4, a significant amount of signal leakage is evident. Signal leakage is also a feature of CNN-based techniques like DnCNN and U-Net.

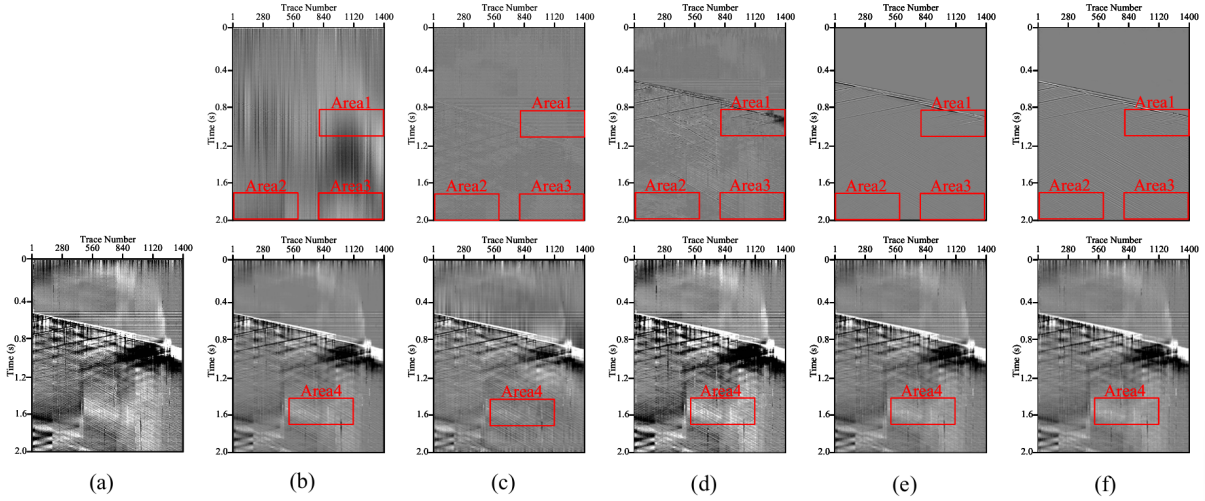


FIGURE 12. Comparison of field DAS data denoising results: (a) Field DAS record; (b)-(f) denoising results (top subplots) and filtered noise images (bottom subplots) of WT, BPF, DnCNN, U-Net, and MW-UNet, respectively

As seen in Figure 13, Area 1, Area 2, and Area 3 in the red rectangular boxes were enlarged to allow for a thorough comparison of denoising details. The close-up of Area 1 demonstrates that the noise-contaminated signals cannot be fully recovered by either the DnCNN method or the conventional methods. It is evident in Areas 2 and 3 that while the U-Net approach successfully reduces noise, MW-UNet performs better when it comes to recovering faint signal details.

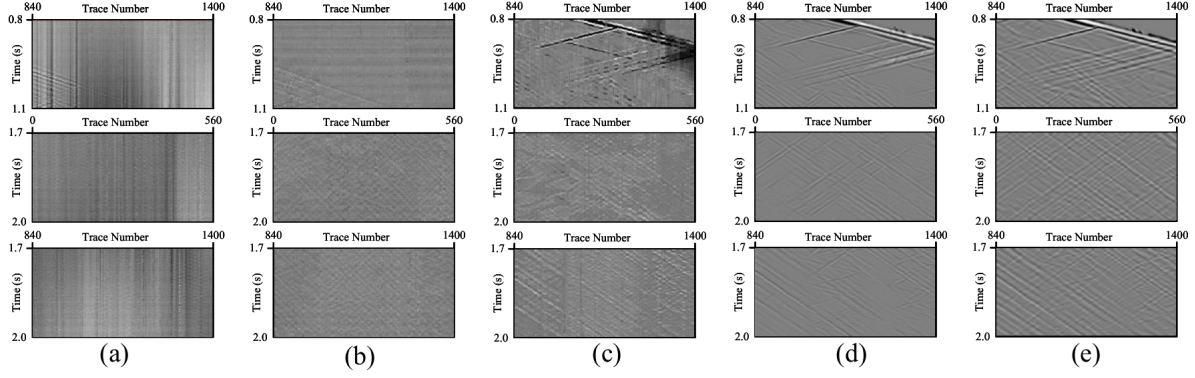


FIGURE 13. The crimson rectangular areas in Figure 12. The figures in the first, second, and third rows are from Area 1, Area 2, and Area 3, respectively. (a)-(e) Enlarged results of WT, BPF, DnCNN, U-Net, and MW-UNet, respectively.

5. Discussion.

5.1. Generalization ability of the improved DDPM-based augmentation method. We experimented with desert seismic data to confirm the suggested augmentation method's capacity for generalization.

Seismic noise samples were enhanced using the same augmentation technique as DAS data. Two training sets were created, one with real data and the other with the enhanced diffusion model. Figure 14 compares the denoising outcomes of models trained on these two datasets. The findings show that the models perform similarly in denoising across training sets, suggesting that the suggested data augmentation technique is still useful for other datasets.

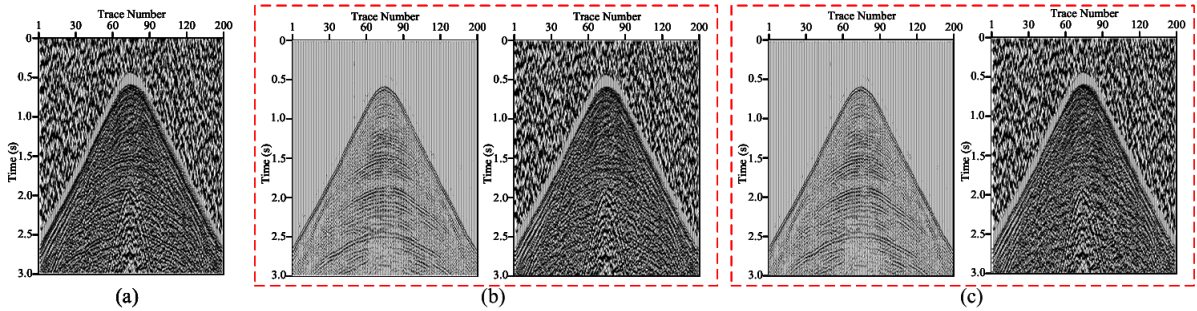


FIGURE 14. Comparison of field desert data denoising results: (a) Field desert record; (b)-(c) denoising results and filtered noise images of real data and improved DDPM augmented data, respectively

5.2. Denoising capability of MW-UNet on other data. Using DnCNN and U-Net as baseline techniques, we also tested the MW-UNet architecture on desert seismic data to assess its denoising performance on other kinds of data. Table 2 contains a list of the network parameters. Figure 15 displays the denoising outcomes of the three CNN-based techniques. As illustrated in Figure 15(a), the DnCNN network shows residual noise when processing background noise because of its comparatively simple structure. As illustrated in Figure 15(b), the U-Net network partially loses shallow-layer signals and surface wave components while suppressing the majority of background noise, leaving a tiny amount of

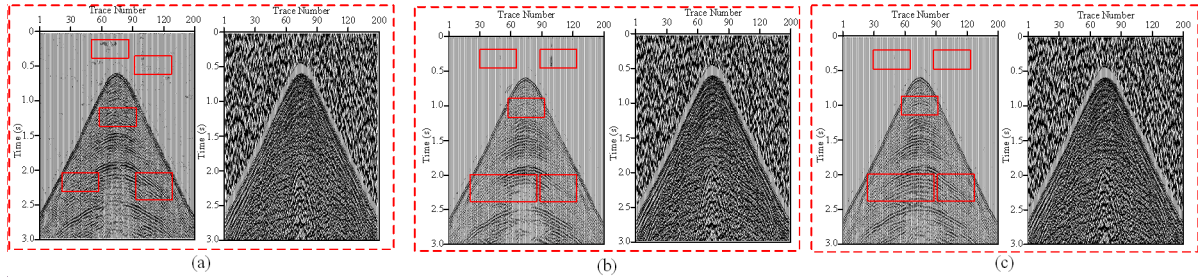


FIGURE 15. Comparison of field desert data denoising results: (a)-(c) Denoising results (top subplots) and filtered noise images (bottom subplots) of DnCNN, U-Net, and MW-UNet, respectively

noise. Conversely, as illustrated in Figure 15(c), MW-UNet performs exceptionally well in recovering shallow-layer signals and maintaining signal continuity, while also successfully suppressing complex noise. These findings show that the suggested denoising technique remains successful when used with different kinds of data.

Although using improved DDPM and DDPM can generate high-quality noisy data, it requires multiple iterations to produce high-quality samples. The current sample must be entered into the prediction network for a complete forward-pass computation at each step, which consumes substantial computational resources and reduces the efficiency of generating sample noisy data.

6. Conclusions. In order to lessen supervised deep learning's reliance on training datasets, this paper proposes a seismic noise data augmentation approach based on improved DDPM. Additionally, we utilize WGAN and traditional DDPM, respectively, to create two more noise datasets. We demonstrate that the synthetic data produced by our enhanced DDPM most closely resembles the actual seismic noise distribution. We test the models' performance on synthetic and field records after training them on the three augmented datasets using both U-Net and MW-UNet to confirm the efficacy of our methodology. The findings validate that better denoising results are obtained with DDPM-augmented data. In order to capture both global and local characteristics, we also suggest an MW-UNet architecture that combines multi-scale analysis and wavelet transform for frequency-domain feature extraction. Even in complex noise environments, this design allows for strong noise suppression and weak signal recovery. Numerous tests using both simulated and actual seismic data attest to the efficacy of our approach. To further meet the demand for high-quality data, DDPM will continue to be optimized in the future, including further improvements to the prediction network and a reduction in the number of reverse steps, to enhance the accuracy and efficiency of data generation. Similarly, the denoising network model will be continuously optimized to improve the SNR of seismic data.

REFERENCES

- [1] F. Cheng, Photonic seismology: A new decade of distributed acoustic sensing in geophysics from 2012 to 2023, *Surveys in Geophysics*, vol.45, no.4, DOI: 10.1007/s10712-024-09840-0, 2024.
- [2] T. Zhong, W. Wang, S. Lu, X. Dong and B. Yang, RMCHN: A residual modular cascaded heterogeneous network for noise suppression in DAS-VSP records, *IEEE Geoscience and Remote Sensing Letters*, vol.20, 2023.
- [3] X. Dong, Y. Li, T. Zhong, N. Wu and H. Wang, Random and coherent noise suppression in DAS-VSP data by using a supervised deep learning method, *IEEE Geoscience and Remote Sensing Letters*, vol.19, 2022.
- [4] O. M. Saad, I. Helmy and Y. Chen, Unsupervised deep-learning framework for 5D seismic denoising and interpolation, *Geophysics*, vol.89, no.4, pp.V319-V330, 2024.

- [5] T. Zhong, M. Cheng, S. Lu, X. Dong and Y. Li, RCEN: A deep-learning based background noise suppression method for DAS-VSP records, *IEEE Geoscience and Remote Sensing Letters*, vol.19, 2022.
- [6] Y. Cui, M. Bai, Z. Zhou and Y. Chen, One-dimensional dictionary learning with variational sparse representation for single-channel seismic denoising, *IEEE Transactions on Geoscience and Remote Sensing*, vol.62, 2024.
- [7] W. Bönke, B. Alaei, A. Torabi and D. Oikonomou, Data augmentation for 3D seismic fault interpretation using deep learning, *Marine and Petroleum Geology*, vol.162, 106706, 2024.
- [8] H. Wang, Y. Chen, B. Liu et al., Nonstationary predictive filtering for seismic random noise suppression – A tutorial, *Geophysics*, vol.86, no.3, pp.W21-W30, 2021.
- [9] N. Liu, F. Li, D. Wang, J. Gao and Z. Xu, Ground-roll separation and attenuation using curvelet-based multichannel variational mode decomposition, *IEEE Transactions on Geoscience and Remote Sensing*, vol.60, pp.1-14, 2022.
- [10] C. Liang, M. P. Thornton, P. Morton, B. J. Hulse, A. Hill and P. Rawlins, Improving signal-to-noise ratio of passive seismic data with an adaptive FK filter, *SEG International Exposition and Annual Meeting*, pp.1703-1707, 2009.
- [11] R. Anvari, M. A. N. Siahpar, S. Gholtashi, A. R. Kahoo and M. Mohammadi, Seismic random noise attenuation using synchrosqueezed wavelet transform and low-rank signal matrix approximation, *IEEE Transactions on Geoscience and Remote Sensing*, vol.55, no.11, pp.6574-6581, 2017.
- [12] M. Bekara and M. Van der Baan, Local singular value decomposition for signal enhancement of seismic data, *Geophysics*, vol.72, no.2, pp.V59-V65, 2007.
- [13] W. Qin, H. Wang, F. Zhang et al., Robust low-rank tensor reconstruction using high-order t-SVD, *Journal of Electronic Imaging*, vol.30, no.6, 063016, 2021.
- [14] Y. Zhao, Y. Li, X. Dong and B. Yang, Low-frequency noise suppression method based on improved DnCNN in desert seismic data, *IEEE Geoscience and Remote Sensing Letters*, vol.16, no.5, pp.811-815, 2019.
- [15] L. Yang, W. Chen, H. Wang and Y. Chen, Deep learning seismic random noise attenuation via improved residual convolutional neural network, *IEEE Transactions on Geoscience and Remote Sensing*, vol.59, no.9, pp.7968-7981, 2021.
- [16] T. Zhong, F. Li, R. Zhang, X. Dong and S. Lu, Multiscale residual pyramid network for seismic background noise attenuation, *IEEE Transactions on Geoscience and Remote Sensing*, vol.60, pp.1-14, 2022.
- [17] J. Zhang, W. Wang, K. Li, H. Wang and Z. Wei, Enhancing seismic data denoising through multiscale analysis across transformer and GAN, *IEEE Geoscience and Remote Sensing Letters*, vol.21, 2024.
- [18] F. Li, H. Liu, W. Wang and J. Ma, Swin transformer for seismic denoising, *IEEE Geoscience and Remote Sensing Letters*, vol.21, 2024.
- [19] X. Wei, X. Liu and H. Zhou, A seismic random noise suppression method based on CNN-Mamba, *Coal Geology & Exploration*, vol.53, no.5, 18, 2025.
- [20] X. Wang, J. Ma, X. Dong, T. Zhong and S. Dong, EFGW-UNet: A deep-learning-based approach for weak signal recovery in seismic data, *IEEE Transactions on Geoscience and Remote Sensing*, vol.62, 2024.
- [21] T. Wang, D. Trugman and Y. Lin, SeismoGen: Seismic waveform synthesis using GAN with application to seismic data augmentation, *Journal of Geophysical Research: Solid Earth*, vol.126, no.4, e2020JB021, 2021.
- [22] Z. Peng, Y. Shao and S. Xia, CCFE: A few-shot learning model for earthquake detection and phase identification, *IEEE Access*, vol.10, pp.124629-124636, 2022.
- [23] B. A. A. Monteiro, H. Oliveira and J. A. dos Santos, Self-supervised learning for seismic image segmentation from few-labeled samples, *IEEE Geoscience and Remote Sensing Letters*, vol.19, 2022.
- [24] H. W. Posamentier, V. Paumard and S. C. Lang, Principles of seismic stratigraphy and seismic geomorphology I: Extracting geologic insights from seismic data, *Earth-Science Reviews*, vol.228, 103963, 2022.
- [25] T. Kumar, R. Brennan, A. Mileo and M. Bendeche, Image data augmentation approaches: A comprehensive survey and future directions, *IEEE Access*, vol.12, pp.187536-187571, 2024.
- [26] D. Wang, G. Chen and Q. Cheng, Seismic data denoising using a self-supervised deep learning network, *Mathematical Geosciences*, vol.56, no.3, pp.487-510, 2024.
- [27] D. Bar-David et al., Elastic deformation of optical coherence tomography images of diabetic macular edema for deep-learning models training: How far to go?, *IEEE Journal of Translational Engineering in Health and Medicine*, vol.11, pp.487-494, 2023.

- [28] X. Dong, J. Lin and S. Lu, Seismic shot gather denoising by using a supervised deep-learning method with weak dependence on real noise data: A solution to the lack of real noise data, *Surveys in Geophysics*, vol.43, no.5, pp.1363-1394, 2022.
- [29] X. Dong, T. Zhong and Y. Li, New suppression technology for low-frequency noise in desert region: The improved robust principal component analysis based on prediction of neural network, *IEEE Transactions on Geoscience and Remote Sensing*, vol.58, no.7, pp.4680-4690, 2020.
- [30] W. W. Godfrey and A. Ratna, Enhancing the video editing capabilities of text-to-video generators using DDPM inversion, *2023 IEEE International Conference on Computer Vision and Machine Intelligence (CVMI)*, Gwalior, India, 2023.

Author Biography



Liquan Zhao is a professor at the School of Electrical Engineering of Northeast Electric Power University, China. He also keeps the position at Key Laboratory of Modern Power System Simulation and Control and Renewable Energy Technology, Ministry of Education. He received his bachelor's degree in Electronic Information Engineering from Harbin University of Science and Technology, China, in 2005, and his doctorate in Information and Communication Engineering from Harbin Engineering University, China, in 2009. After graduating with his doctorate, he taught at Northeast Electric Power University. His main research contents include artificial intelligence, intelligent information processing, and their applications in power systems.



Peng Liu received the B.S. degree in Communication Engineering from Henan University of Engineering, China, in 2022. He is pursuing the M.S. degree in Information and Communication Engineering, Northeast Electric Power University, China. His research interests include digital signal processing and seismic data denoising.



Tie Zhong received the B.S., M.S., and Ph.D. degrees in Communication Engineering from Jilin University, China, in 2006, 2008 and 2016, respectively. He is currently an associate professor in the School of Electrical Engineering, Northeast Electric Power University, China. He also keeps the position at Key Laboratory of Modern Power System Simulation and Control and Renewable Energy Technology, Ministry of Education. His research interests include modern signal processing techniques, weak signal process theory, and intelligent information processing.