



# Multiscale dilated denoising convolution with channel attention mechanism for micro-seismic signal denoising

Jianxian Cai<sup>1</sup> · Zhijun Duan<sup>1</sup> · Li Wang<sup>1</sup> · Juan Meng<sup>1</sup> · Zhenjing Yao<sup>1</sup>

Received: 5 April 2023 / Accepted: 9 January 2024 / Published online: 14 February 2024  
© The Author(s) 2024

## Abstract

Denoising micro-seismic signals is paramount for ensuring reliable data for localizing mining-related seismic events and analyzing the state of rock masses during mining operations. However, micro-seismic signals are commonly contaminated by various types of complex noise, which can hinder micro-seismic accurate P-wave pickup and analysis. In this study, we propose the Multiscale Dilated Convolutional Attention denoising method, referred to as MSDCAN, to eliminate complex noise interference. The MSDCAN denoising model consists of an encoder, an improved attention mechanism, and a decoder. To effectively capture the neighborhood features and multiscale features of the micro-seismic signal, we construct an initial dilated convolution block and a multiscale dilated convolution block in the encoder, and the encoder focuses on extracting the relevant feature information, thus eliminating the noise interference and improving the signal-to-noise ratio (SNR). In addition, the attention mechanism is improved and introduced between the encoder and decoder to emphasize the key features of the micro-seismic signal, thus removing the complex noise and further improving the denoising performance. The MSDCAN denoising model is trained and evaluated using micro-seismic data from Stanford University. Experimental results demonstrate an impressive increase in SNR by 11.237 dB and a reduction in root mean square error (RMSE) by 0.802. Compared to the denoising results of the DeepDenoiser, CNN-denoiser and Neighbor2Neighbor methods, the MSDCAN denoising model outperforms them by enhancing the SNR by 2.589 dB, 1.584 dB and 2dB, respectively, and reducing the RMSE by 0.219, 0.050 and 0.188, respectively. The MSDCAN denoising model presented in this study effectively improves the SNR of micro-seismic signals, offering fresh insights into micro-seismic signal denoising methodologies.

**Keywords** Micro-seismic signals · Signal denoising · Convolutional neural networks · Multiscale dilated convolution · Attention mechanisms

## List of symbols

$A$  The signal or noise  
 $A^*$  The normalized signal or noise

$b_f$  The bias of the last layer  
 $C_{1 \times 1}$  Conventional convolutions with a kernel size of  $1 \times 1$   
 $d$  The dilated rate of the dilated convolution kernel  
 $D_{f=1}, D_{f=6}, D_{f=12}, D_{f=18}$  Dilated convolutions with a kernel size of  $1 \times 3$  and dilated factors of 1, 6, 12, and 18  
 $F$  The input to the last layer  
 $F_{sq}(\cdot, \theta)$  The compression operation  
 $F_{ex}(\cdot, \theta)$  The extraction operation  
 $F_{scale}$  The weighting operation  
 $f$  The normalization function Sigmoid  
 $k$  The convolution kernel size  
 $N$  The lengths of the data  
 $s$  The dilated convolution size  
 $SM$  Soft-additive activation function

These authors contributed equally to this work.

✉ Jianxian Cai  
cjxlaq@163.com  
Zhijun Duan  
dzj\_0108@163.com  
Li Wang  
wangli0817@163.com  
Juan Meng  
mengjuan@cidp.edu.cn  
Zhenjing Yao  
yaozhenjing000@126.com

<sup>1</sup> Hebei Ke Laboratory of Seismic Disaster Instrument and Monitoring Technology, Langfang 065201, China

$u_c$	The characteristic channel of the micro-seismic signal and noise
$w_i$	The weight at the location of the convolution kernel
$W_1, W_2$	The weights of the two fully connected layers
$W_f$	The weight matrix
$X_1, X_2, X_3, X_4$	The features of each branch
$y^{j-1}$	The feature map of the previous layer
$y_i, y$	The original signals
$z_i, z$	The signals after denoising
$\alpha$	The mean of the signal or noise
$\mathcal{E}$	The standard deviation of the signal or noise
$\gamma$	Activation function
$\lambda$	The scale of the batch normalization layer
$\eta$	Shift parameters of the batch normalization layer
*	The convolution operations
{+}	An element-by-element addition operation
{•}	A fusion operation of different channels

### Abbreviations

CNN	Convolutional neural network
DC	Dilated convolution
DMS	Multiscale dilated convolution
EMD	Empirical mode decomposition
Improved SE	Improved attention mechanism
LR	Learning rate
ODC	One-dimensional convolution
PSNR	P-wave signal-to-noise ratio
$r$	Correlation coefficient
RMSE	Root mean square error
SNR	Signal-to-noise ratio
STFT	Short-time Fourier transform
Up	Up-sampling
WT	Wavelet transform

## Introduction

The micro-seismic monitoring system is an equipment system that integrates micro-seismic sensors and data recording equipment to detect and record micro-vibration signals, which can collect full waveform data of micro-vibrations in underground rocks and strata Ma et al. (2021), Khan et al. (2023). By denoising micro-seismic signals, information such as micro-seismic P-wave arrival time and micro-seismic source localization can be obtained to achieve accurate

monitoring and early warning Ahmad and Takeshi (2021), Tarek et al. (2021), Xu et al. (2023). Therefore, how to effectively eliminate the noise, improve the signal-to-noise ratio (SNR), and promote the practical application of micro-seismic monitoring systems is one of the important topics worth of study.

At present, there are two main categories of micro-seismic signal denoising methods: traditional methods and deep learning (DL) models. Traditional methods mainly include wavelet transform (WT) Yuan et al. (2020), short-time Fourier transform (STFT) Mao (2022), and variational modal decomposition (VMD) methods Liu et al. (2022). These ideas were developed by Qi and Wang (2020) proposed a local polynomial Fourier transform (LPFT) method that can efficiently describe the instantaneous frequency variations of local high-order polynomial fits and obtain high spectral and energy-concentration results. Lin et al. (2022) proposed extending the time-domain synchronous compressed WT to the spatial domain to accurately characterize spatially varying signals for seismic noise suppression. Liu et al. (2023) proposed a method based on S-transform and improved VMD to improve the time-frequency domain resolution and seismic reflection performance through inverse spectral deconvolution. These methods rely heavily on manually designed features and rules, which limits their ability to extract features from micro-seismic signals.

In recent years, deep learning has used massive data for learning and training and has made greater progress in feature extraction, achieving "train once, process many" Mumuni and Mumuni (2022); Yang et al. (2023). Compared to traditional signal denoising methods, deep learning-based denoising methods have obvious advantages in terms of SNR, correlation coefficient, and robustness Karniadakis et al. (2021); Muther et al. (2023). This approach has found widespread application in the realm of signal processing as Zhu et al. (2019) used a U-shaped convolutional neural network to denoise seismic data, which created a nonlinear mapping between noise and denoised seismic data by combining depth-weighted information, resulting in excellent denoising results. However, it could not be applied to the task of denoising micro-seismic signals from different regions. To this end, Dong et al. (2020) used a convolutional neural network model to effectively suppress random noise in each region and recover meaningful seismic events. However, the model needs to adjust the network parameters according to the characteristics of random noise in different regions, which leads to poor generalization ability to remove different regions. Therefore, Saad et al. (2022) use unsupervised DL and attention networks to remove unwanted noise from seismic data. The advantage of this algorithm is that it does not require any a priori information about the input data, which improves the generalization ability of the model for denoising. However, the sensory field of ordinary

convolution in the model is small, and the neighborhood feature-capturing ability is poor. In response to this, Dong et al. (2022) introduced dilated convolution based on the generative adversarial network (GAN) to design a denoising method with a weak dependence on real noise data, but the method has a weak ability to extract multiscale features of micro-seismic signals. In addition, Wang et al. (2023) applied a self-supervised approach to denoising seismic data using the Neighbor2Neighbor strategy, which allows model training without clean seismic data. However, it did not enhance the ability to capture micro-seismic signal features.

The development of micro-seismic signals denoising algorithms faces the following challenges: (1) Micro-seismic signal characteristics are complex. The combination of micro-seismic data reflects changes in kinematic and dynamical features of the wavefield, forming micro-seismic signal neighborhood features and multiscale feature relationships with each other. However, the small receptive field of ordinary convolution can only extract a single feature of the micro-seismic signal and cannot capture the neighborhood features and multiscale features of the micro-seismic signal, which brings difficulties in accurately eliminating the noise interference of micro-seismic signals. Zhang et al. (2020). (2) The acquisition of micro-seismic signals is significantly influenced by complex environments, such as periodic industrial disturbance noise, impulse noise from on-site mechanical or anthropogenic vibration, and irregular background noise in the time domain. These noises make it difficult for the model to adequately extract micro-seismic signal features. Addressing these complex environmental challenges is critical to achieving robust and accurate denoising results. However, the current algorithms have problems such as insufficient micro-seismic signal feature extraction and insufficient research on complex noise interference, which lead to poor elimination of complex noise interference and low SNR of the denoised micro-seismic signal.

Motivated by the above analysis, this study tries to address the problem that current micro-seismic signal denoising methods are unable to effectively remove complex noise interference. In this study, we propose a novel denoising method called MSDCAN. The MSDCAN denoising model is framed by convolutional self-coding, which consists of an encoder, an improved attention mechanism (SE), and a decoder. The encoder is responsible for extracting micro-seismic neighborhood features and multiscale features, the improved SE is responsible for extracting micro-seismic signals significantly and ignoring noise interference, and the decoder is responsible for the micro-seismic signal detail information. Considering micro-seismic neighborhood features and multiscale features, the encoder contains an initial dilated convolution (DC) block and a multiscale dilated convolution (DMS) block. DC block and DMS block improve the encoder by employing dilated convolution,

which can prevent micro-seismic neighborhood features and multiscale features from disappearing in forward propagation by expanding the receptive field. In addition, the SE is improved and introduced into the denoising model to optimize the significant feature extraction of micro-seismic signals and eliminate complex noise interference.

In summary, our advantages are summarized as follows:

1. The MSDCAN denoising model uses an encoder-improved SE-decoder network structure. The encoder focuses on capturing micro-seismic signal neighborhood features and multiscale features, the improved SE focuses on micro-seismic signal salient feature extraction, and the decoder focuses on recovering micro-seismic signal detail features.
2. Design the encoders based on dilated convolution to prevent micro-seismic neighborhood features and multiscale features from disappearing in forward propagation by expanding the receptive field.
3. The SE is improved and introduced between the encoder and decoder to optimize the feature extraction of micro-seismic signals. The improved SE retains the advantages of selectively enhancing micro-seismic signal features with fewer parameters, making the network model easier to train and achieving better denoising.

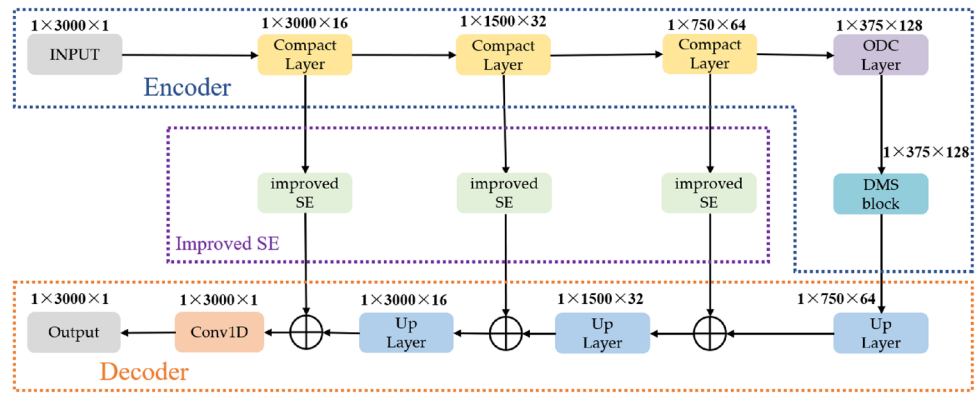
The remainder of this study is organized as follows. Section 2 provides an in-depth understanding of the proposed MSDCAN denoising model, discussing its overall structure and main modules. Section 3 focuses on the dataset and its preprocessing and evaluation metrics. Section 4 presents the analysis of the experimental results. Finally, Section 5 concludes the paper.

## Method

### Structure of the MSDCAN denoising model

A self-coding convolutional network possesses the ability to acquire the characteristics of input data, reconstruct its structure to compress the input and subsequently employ a decoder to reconstruct and output the compressed features. Essentially, it represents a neural network grounded in the backpropagation algorithm, aligning the target output with the input. In this study, a self-coding convolutional neural network serves as the framework, and Fig. 1 illustrates the structure of the MSDCAN denoising model. The MSDCAN denoising model comprises three key components: an encoder, an improved SE, and a decoder. The encoder incorporates both a DC block and a DMS block, responsible for extracting micro-seismic signal features. Meanwhile, the improved SE assigns feature weights to

**Fig. 1** MSDCAN denoising model structure



**Table 1** The details of the MSDCAN denoising model

Name	Details	Size
Input		3000
Compact layer	[Conv1x5 Str.1]x2;Maxpooling1x2	3000
Compact layer	[Conv1x5 Str.1]x2;Maxpooling1x2	1500
Compact layer	[Conv1x5 Str.1]x2;Maxpooling1x2	750
ODC layer	[Conv1x5 (dilated=2) Str.1]x2	375
DMS block	Conv1x5 (dilated=1,6,12,18)	375
UP layer	Conv.Trans1x5 Str.2;[Conv1x5 Str.1]x2	750
UP layer	Conv.Trans1x5 Str.2;[Conv1x5 Str.1]x2	1500
UP layer	Conv.Trans1x5 Str.2;[Conv1x5 Str.1]x2	3000
Output	Conv1x1 Str.1	3000

various noisy micro-seismic signals, enabling the precise removal of noise. The decoder, designed with an Up block based on transposed convolution, is tasked with restoring the denoised micro-seismic signal. Detailed information about the MSDCAN denoising model is presented in Table 1.

**Encoder layer**

Traditional convolutional neural networks exclusively acquire knowledge from micro-seismic signals using standard convolution techniques, limiting their ability to capture

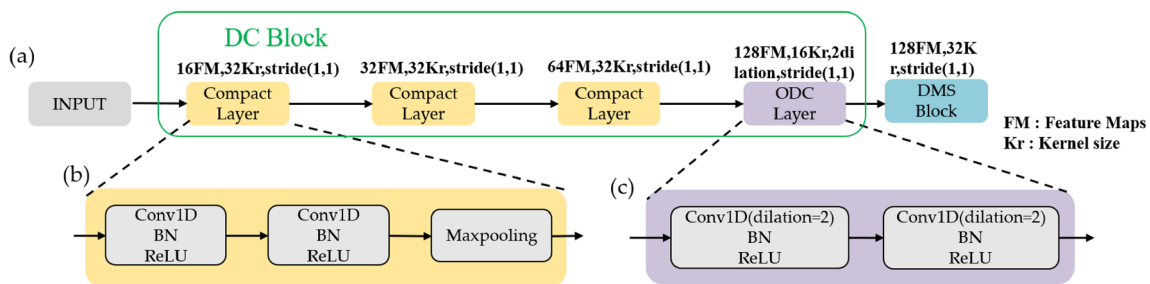
features of specific scales. Nonetheless, micro-seismic signals resulting from rock bursts exhibit intricate attributes and diverse scale variations, thereby impacting the denoising capabilities of conventional convolutional methods. Hence, this study enhances the structure of the encoder layer by introducing dilated convolution, as depicted in Fig. 2. The encoding layer is composed of a DC block and a DMS block. The one-dimensional micro-seismic signal first undergoes extraction of micro-seismic neighborhood information via the DC block, followed by the extraction of multiscale features from the micro-seismic signal through the DMS block.

**DC block**

The DC block is composed of three compact layers and an one-dimensional dilated convolution(ODC) layer (refer to Fig. 2). Each compact layer comprises two one-dimensional convolution layers and a Maxpooling layer, enabling the extraction of abstract features from the micro-seismic signal through convolution operations. This can be represented by the following equation Tang et al. (2021):

$$x(i) = \gamma \left( \sum_{i=0}^{m-1} w_i y^i \right) \tag{1}$$

where  $w_i$  represents the weight  $i$  at the location of the convolution kernel;  $y^i$  is the feature map of the previous layer,



**Fig. 2** a Encoder structure; b compact layer structure; c ODC layer structure

which serves as the input for the current layer;  $\gamma$  represents the activation function;  $m$  represents the number of compact layers;  $x$  represents the output feature map obtained from the preceding compact layer.

The study chooses Rectified Linear Unit (ReLU) Glorot et al. (2011) as the activation function with an output of  $\max(0, x)$ , providing stability to the network. Furthermore, to expedite the training process, a normalization layer known as Batch Normalization (BN) Ioffe and Szegedy (2015) is incorporated. BN normalizes the intermediate outputs within a mini-batch, reducing the internal covariate shift and improving the overall stability and convergence of the network. This can be represented by the following Eq. Saad et al. (2022):

$$B(x) = \lambda \frac{x - \mu(x)}{\sqrt{\sigma^2(x)}} + \eta \tag{2}$$

where  $\lambda$  and  $\eta$  represent the scale and shift parameters of the BN layer, respectively;  $B$  represents the output of BN;  $x$  represents the output feature map obtained from the preceding convolution layer;  $\mu$  represents mean value of the preceding convolution layer;  $\sigma$  represents variance of the preceding convolution layer. These parameters are learned by each layer in the training process.

The micro-seismic signal contains valuable neighborhood information that encapsulates important characteristics such as the homophase axis and phase changes. By leveraging the inherent properties of this neighborhood information, it becomes easier to distinguish meaningful data from noise interference. To enhance the denoising performance, the study introduces an ODC layer that capitalizes on the ability to extract more features by exploiting the neighborhood information of the micro-seismic signal. The ODC layer, as depicted in Fig. 2c, is composed of two one-dimensional dilated convolution layers. These layers are designed to effectively capture and process the neighborhood information. The feature map of the ODC layer is mathematically represented as follows Shi et al. (2020):

$$X(i) = \sum_s x[i + d \times s - 1]k[s] \tag{3}$$

where  $d$  is the dilated rate of the dilated convolution kernel. For this study, a specific dilated rate of  $s = 2$  is chosen.  $k$  represents the convolution kernel size utilized within the dilated convolution operation;  $s$  denotes the dilated convolution size;  $x$  represents the output feature map obtained from the preceding compact layer.

### DMS block

The dilated convolution operation alone is limited in its ability to sample the dilated part, resulting in discontinuous extraction of micro-seismic information and suboptimal denoising performance. To address this limitation, a DMS block is designed to extract multiscale information from micro-seismic signals by simultaneously convolving multiple dilations with varying dilated rates. The structure of the DMS block is illustrated in Fig. 3. The input to the DMS block is the feature map  $X$  obtained from the preceding DC block. This feature map  $X$  is decomposed into four distinct parts:  $X_1, X_2, X_3, X_4$ . These four parts are obtained through different convolutions: an ordinary  $1 \times 3$  convolution, a  $13 \times 1$  dilated convolution with a dilation rate of 6, a  $25 \times 1$  dilated convolution with a dilation rate of 12, and a  $37 \times 1$  dilated convolution with a dilation rate of 18, respectively. This decomposition strategy is employed to make extracted feature information correlated. To capture the complex features of the micro-seismic signal more comprehensively, the features obtained with a dilation rate of 1 are further overlaid onto the feature maps with dilated rates of 6, 12, and 18. This layer-by-layer superimposition ensures that the branches with dilated rates of 6, 12, and 18 not only contain features with larger dilated rates but also incorporate features with smaller dilated rates. The resulting superimposed feature maps represent the learned features of the micro-seismic signal achieved by the DMS block. The expressions for the features extracted from the four branches are presented in Eqs. 4–7, Song et al. (2022):

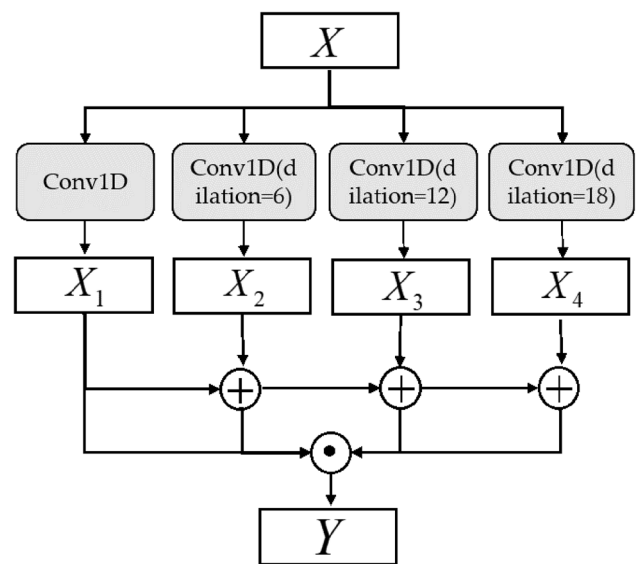


Fig. 3 Structure of the DMS block

$$X_1 = X * C_{1 \times 1} * D_{f=1} \tag{4}$$

$$X_2 = X * C_{1 \times 1} * D_{f=6} \tag{5}$$

$$X_3 = X * C_{1 \times 1} * D_{f=12} \tag{6}$$

$$X_4 = X * C_{1 \times 1} * D_{f=18} \tag{7}$$

where  $X_1, X_2, X_3$  and  $X_4$  are the features of each branch, respectively, and the symbol  $*$  is the convolution operation,  $C_{1 \times 1}$  denotes conventional convolutions with a kernel size of  $1 \times 1$ .  $D_{f=1}, D_{f=6}, D_{f=12}$  and  $D_{f=18}$  denote dilated convolutions with a kernel size of  $1 \times 3$  and dilated factors of 1, 6, 12 and 18.

The four branch outputs are fused by stacking layer-by-layer Chen et al. (2017) for the four branch features:

$$Y = \{X_1, X_1 + X_2, X_1 + X_2 + X_3, X_1 + X_2 + X_3 + X_4\} \tag{8}$$

where the symbol  $\{+\}$  indicates an element-by-element addition operation and  $\{\bullet\}$  indicates a fusion operation of different channels.

### Improved SE

The micro-seismic signal contains various complex noises that seriously affect the denoising effect. To address this issue, the study leverages the channel attention mechanism, which enables the assessment of the relative importance of micro-seismic signal characteristics and noise. By assigning appropriate weights to these characteristics, an improved attention mechanism(improved SE) structure is designed (see Fig. 4b). In this improved SE structure, a one-dimensional convolutional layer instead of a fully connected layer is employed. This design choice offers several advantages. Firstly, it reduces the computational complexity of the model while preserving the spatial structure of the data. Additionally, this ensures that important information such as seismic phase and polarization, which are crucial for accurate analysis, is not lost due to distortion of the micro-seismic signal. Furthermore, considering that micro-seismic signal values are bipolar, i.e., they can be both positive and negative, the study adopts the LeakyReLU activation function. This activation function is well-suited for handling bipolar signals and helps preserve the polarity information during the denoising process.

The improved SE mechanism consists of three main operations: compression  $F_{sq}(\bullet, \theta)$ , extraction  $F_{ex}(\bullet, \theta)$ , and weighting  $F_{scale}$ , each playing a crucial role in enhancing the denoising effect of micro-seismic signals. Firstly, the compression operation  $F_{sq}(\bullet, \theta)$  compresses the features obtained from the coding layer. This compression operation converts the feature maps

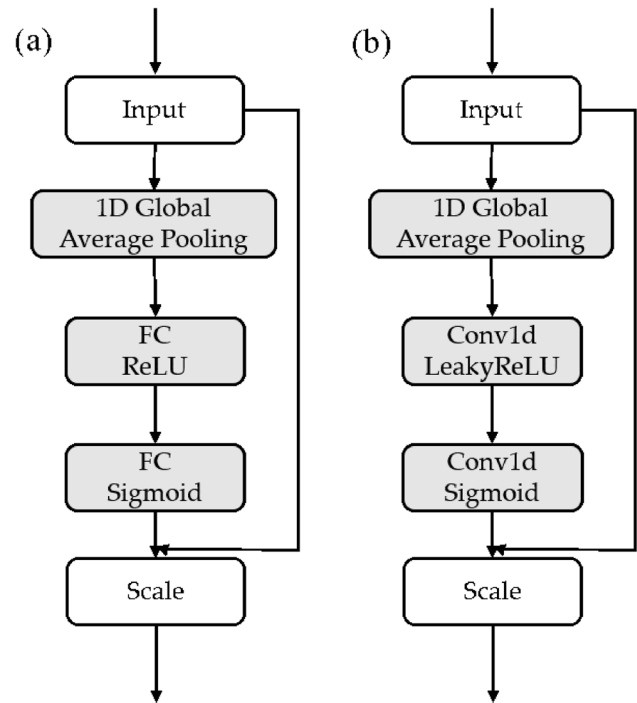


Fig. 4 Improved SE and original SE block a Original SE block b Improved SE block

into real numbers, thereby capturing the global perceptual field of the micro-seismic signal. By obtaining this global perceptual field, the model gives us a holistic understanding of the micro-seismic signal. Next, the extraction operation  $F_{ex}(\bullet, \theta)$  is performed, which involves adding a convolutional layer both above and below the LeakyReLU activation function. This operation assigns weights to each feature channel, allowing the model to prioritize the importance of different channels in the micro-seismic signal. This step facilitates the extraction of salient signal features while suppressing the influence of noise. Finally, the weighting operation  $F_{scale}$  combines the output weights obtained from the extraction operation with the original feature map. This operation applies the obtained weights channel by channel, enabling the distribution of weights across the micro-seismic signal and noise features. This weight distribution emphasizes the micro-seismic signal features while disregarding the noise features, ultimately enhancing the denoising effect of the micro-seismic signal. The mapping relationship among these operations is presented in Eqs. 9 to 10 Duan et al. (2023):

$$F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \tag{9}$$

$$F_{ex}(z, W) = f[W_2 f(W_1, z)] \tag{10}$$

$$F_{scale}(u_c, s_c) = u_c \cdot s_c \tag{11}$$

where  $u_c$  represents the characteristic channel of the micro-seismic signal and noise;  $W_1$  and  $W_2$  are the weights of the two one-dimensional convolutional layers; the LeakyReLU activation function is applied to introduce nonlinearity to the outputs of the two one-dimensional convolutional layers.  $f$  represents the normalization function Sigmoid, which normalizes the weight parameters to [0, 1].

### Decoder layer

To generate a noise-reduced micro-seismic signal of the same dimensions as the input, the MSDCAN denoising model incorporates a decoder layer, which consists of three up-sampling (Up) layers and one one-dimensional convolutional layer. This structure aims to reconstruct the compressed feature map obtained from the intermediate layer back into the original signal. The structure of the decoder layer is illustrated in Fig. 5a, and the Up layer is shown in Fig. 5b within the orange box. The Up layer comprises a transposed convolutional layer and two convolutional layers. The MSDCAN method utilizes three Up operations to reduce the dimensionality of the feature map and restore it to the size of the original signal. In the Up process, the output of the intermediate layer serves as the input and is combined with the output of the improved SE through transposed convolution. The resulting feature map is then convolved with a layer that has a similar number of neurons as the input. The softmax activation function is employed to obtain the output Saad et al. (2022):

$$O = SM(W_f \times F + b_f) \tag{12}$$

where  $W_f$  and  $b_f$  are the weight matrix and bias of the last layer, and  $F$  denotes the input to the last layer.  $SM$  is a soft-additive activation function with the following output Saad et al. (2022):

$$SM(x) = \frac{x}{\sum_{j=1} x} \tag{13}$$

## Data preprocessing and evaluation indicators

### Dataset and preprocessing

The study endeavors to assess the performance of the MSDCAN model using the Stanford Earthquake Dataset (STEAD), which comprises data from 225 stations sourced from 20 diverse seismic networks across the globe Mousavi et al. (2019). The waveform records encompass observations from an array of seismometer variants, encompassing 600 waveforms acquired by broadband seismometers and 400 waveforms captured by short-period seismometers. To exemplify the efficacy of the MSDCAN denoising model in processing micro-seismic signals, a comprehensive dataset of 33,800 signal readings originating from micro-seismic events, featuring magnitudes ranging from M0.5 to 3.0 and sampled at 100 Hz, was meticulously curated for this study. Figure 6 shows comparison of micro-seismic data before and after preprocessing.

When acquiring micro-seismic signals, one cannot escape the influence of both human and instrumental interferences, including issues such as zero drift and ultra-low-frequency disruptions. An excessive amount of zero drift has the potential to oversaturate the neural network’s activation function, thereby suppressing the genuine micro-seismic signals and failing to extract their meaningful characteristics. In this study, we opt for the band-pass filtering technique, configuring the low-pass filter frequency at 1, the high-pass filter frequency at 20, and the filter order at 4. This selection effectively eliminates the aforementioned sources of interference. Taking into account that Gaussian white noise can mitigate overfitting in the denoising network and augment its capacity

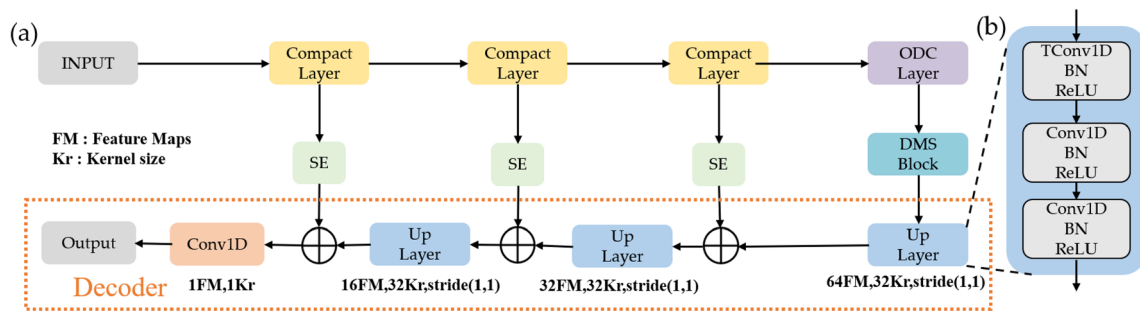
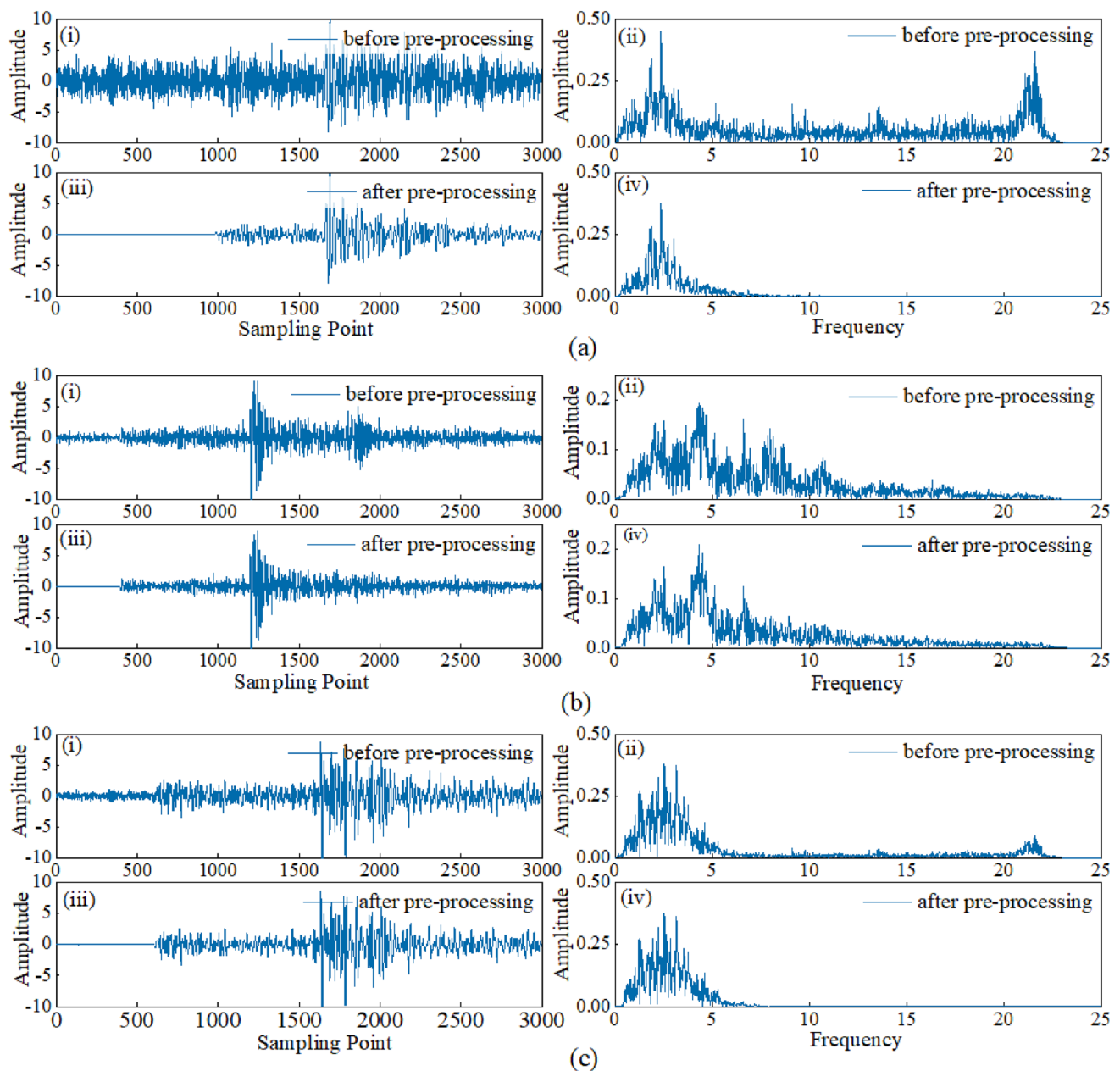


Fig. 5 (a) The structure of the MSDCAN (b) The structure of the Up layer



**Fig. 6** Comparison of micro-seismic data before and after preprocessing. **a–c** are the three samples before and after signal preprocessing. For each subfigure, (i) denotes the noisy signal, (ii) is the spectrogram of the noisy signal, (iii) denotes the spectrogram corresponding to the

preprocessed signal, and (iv) denotes the spectrogram corresponding to the spectrogram of the preprocessed signal. The SNRs before and after preprocessing are **a** (i):  $-6.632\text{dB}$ , **a** (iii):  $6.785\text{dB}$  **b** (i):  $0.432\text{dB}$ , **b** (iii):  $3.156\text{dB}$ ; **c** (i):  $1.522\text{dB}$ , **c** (iii):  $4.086\text{dB}$

for generalization, the sample data undergoes a preliminary noise preprocessing step. In order to evaluate the effectiveness of the preprocessing step, Fig. 6 shows some of the results obtained from the preprocessed training set. As can be seen from Fig. 6(ii) and (iv), there is no spectral aliasing of the spectrum after data preprocessing.

Furthermore, considering the substantial variation in amplitude values observed in the original micro-seismic

signals, the absolute amplitude values alone do not provide a reliable feature for waveform recognition. What's more, limiting the signal data to a smaller range helps prevent data overflow or accuracy issues with floating point representations. Therefore, the sample data were normalized in order to speed up convergence during network training and to improve numerical stability. The normalization process, as depicted in Eq. 14 Wang et al. (2022),  $A^*$  is the normalized



signal,  $A$  is the original signal,  $A_{\max}$  is the maximum value of the original signal, and  $A_{\min}$  is the minimum value of the original signal. The dataset is then randomly partitioned into training data, validation data, and test data in a ratio of 7:2:1, respectively.

$$A^* = 2 * \frac{A - A_{\min}}{A_{\max} - A_{\min}} - 1 \quad (14)$$

## Evaluation indicators

To provide a quantitative assessment of the denoising performance achieved by the MSDCAN denoising model, three evaluation metrics, namely signal-to-noise ratio (SNR), root mean squared error (RMSE), and correlation coefficient ( $r$ ), are employed. These metrics enable an objective evaluation of the denoising effectiveness. The calculations for these metrics are outlined in Eqs. 15–17 Cai et al. (2023).

$$\text{SNR} = 10 \log_{10} \left[ \frac{\sum_{i=1}^N y_i^2}{\sum_{i=1}^N (z_i - y_i)^2} \right] \quad (15)$$

$$\text{RMSE} = \frac{1}{N} \sqrt{\sum_{N=1}^N |y_i - z_i|^2} \quad (16)$$

$$r(z, y) = \frac{\text{Cov}(z, y)}{\sqrt{\text{Var}[z]\text{Var}[y]}} \quad (17)$$

where  $y_i$  and  $y$  represent the original signals,  $z_i$  and  $z$  denote the signals after denoising, and  $N$  represents the length of the data, indicating the total number of samples present in the micro-seismic signal.

The larger the SNR in the indicator, the more information the true micro-seismic signal contains in the signal, and the better the denoising effect. The RMSE of the indicator is the error between the output signal of the calculation model and each sampling point of the label signal. The closer its value approaches 0, the closer the denoised signal is to the original signal, and the better the denoising effect is. The closer the  $r$  of the indicator is to 1, the greater the correlation between the denoised micro-seismic signal and the real micro-seismic signal, and the better the denoising effect. This means that the location of the denoised micro-seismic signal almost does not change. Three indicators can characterize the degree of preservation of the original signal after denoising of micro-seismic signals from different angles. But from the above three formulas, it can be seen that the calculation of SNR, RMSE, and  $r$  needs to be combined with the simulation signal, because only the simulation signal can accurately know the pure signal. Therefore, in the

subsequent denoising of actual data, the P-wave signal-to-noise ratio (PSNR) is used to evaluate the denoising effect of the MSDCAN denoising model, as detailed in Sect. 4.6.1 of the paper.

## Experiments

The MSDCAN denoising model is implemented using the PyTorch deep learning framework and trained and tested on a single NVIDIA GeForce RTX 3060 GPU with 6 GB of GPU memory. The utilization of parallel GPU operations significantly enhances gradient computation efficiency and accelerates the training of deep neural networks. By leveraging parallel GPUs, the computational burden of model operations is alleviated, resulting in increased processing speed. This approach finds practical application in the field of mine micro-seismic analysis, enabling efficient model training and testing.

In this section, we validate the MSDCAN denoising model using three distinct datasets: the Stanford University micro-seismic dataset, Beijing micro-seismic real-world data, and Shanxi mine micro-seismic real-world data. To assess the influence of model depth on noise reduction, as discussed in Sect. 4.1, we devise three model configurations and compare their performance on the Stanford University micro-seismic dataset, thereby selecting the optimal layer count. To demonstrate the effectiveness of introducing blocks into the MSDCAN denoising model, we compare and evaluate them on the Stanford University micro-seismic dataset in Sect. 4.2. Following this, in Sects. 4.3 and 4.4, we conduct a thorough analysis and comparison of MSDCAN's performance against existing methods, considering varying signal-to-noise ratios (SNRs) and types of noise. Furthermore, in Sect. 4.5, we put the MSDCAN denoising model to practical use by applying it to actual micro-seismic signals from Beijing and Shanxi mines, thus highlighting its real-world applicability and generalization capabilities. Finally, in Sect. 4.6, we discuss the limitations of the MSDCAN denoising model and propose future work directions.

### Optimization of the number of model layers

The number of layers in the MSDCAN denoising model significantly influences the denoising efficacy. Therefore, the study devised three block structures to optimize the layer count, configuring the codec layers to be 2, 3, and 4. In other words, the encoder hosts a corresponding number of tight blocks, and the decoder incorporates a matching number of up-sampling blocks-2, 3, and 4, respectively. Likewise, the improved attention blocks also align with this layer count-2, 3, and 4. Consequently, the corresponding neural networks are denoted as MSDCAN-CA2, MSDCAN-CA3,

and MSDCAN-CA4, respectively. The training and validation losses of the three network models are shown in Fig. 7, with the blue, green and red curves corresponding to the experimental results of the three network models. As seen in Fig. 7, the training loss of the MSDCAN-CA3 network decreases swiftly, along with the validation loss, all without any signs of overfitting. The loss value of the MSDCAN-CA3 network drops to 0.03 when the epoch is greater than 440 and remains at this level thereafter, largely stable, with the lowest loss values compared to the other two networks, averaging 6.1–11% lower than the MSDCAN-CA2 training loss, 8.1–14% lower than the validation loss, and with oscillations in MSDCAN-CA2. The training loss is on average 2–7% lower than that of MSDCAN-CA4, the validation loss is on average 2–3% lower, and MSDCAN-CA4 shows significant oscillation. The model parameter sizes of the three networks are shown in Table 2, with MSDCAN-CA3 having 1780k fewer model parameters compared to MSDCAN-CA4 and 433k more model parameters compared to MSDCAN-CA2.

MSDCAN-CA3 has the best denoising accuracy with a model parameter of 586k, which is equivalent to approximately 0.57 MB. It not only improves the denoising accuracy but also greatly saves training resources, which meets the low-power hardware equipment requirements of computer equipment at the mine-quake site where the model parameters need to be compressed within 2 MB. This is because the number of model layers is too small, the number of model parameters is small, the model is unable to mine the potential higher-order relationships among the nodes in the micro-seismic signals, the mapping ability is weakened, and the SNR is reduced; when the number of model layers is too high, the number of model parameters is high and the model can perfectly map between the training samples and the target perfectly, but this mapping lacks the ability to

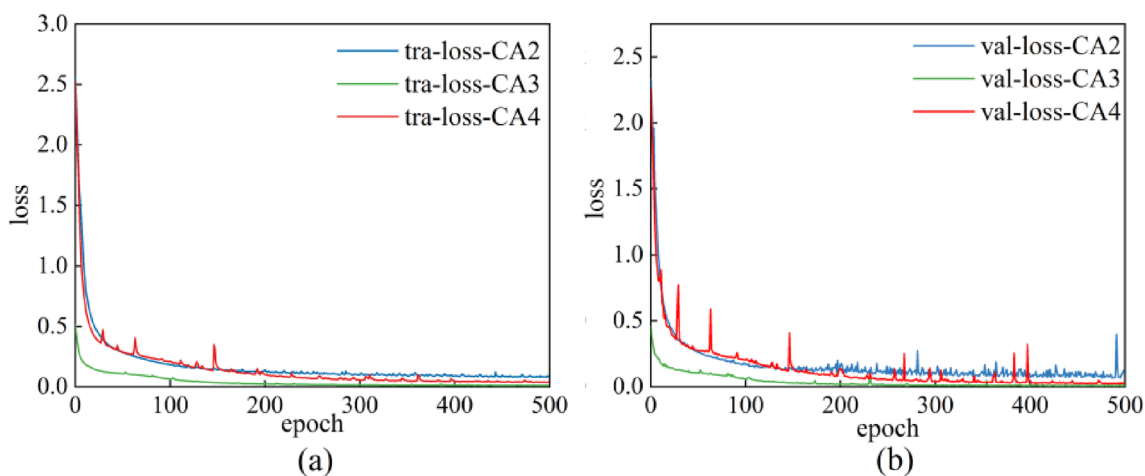
**Table 2** Optimization analysis of MSDCAN denoising model with different layers

	SNR/dB	Model parameters
MSDCAN-CA2	11.575	153k
MSDCAN-CA3	13.258	586k
MSDCAN-CA4	13.153	2366k

generalize and the model computation slows down. Therefore, in this study, MSDCAN-CA3 is chosen as the neural network model structure for the denoising task.

### Cross-validation experiments

To accurately evaluate the generalization performance of the MSDCAN denoising model, a 5-fold cross-validation was employed. The entire dataset was evenly split into 5 subsets, with each subset taking a turn as the test set while the remaining 4 subsets served as training sets for model training. Within the training sets, data were randomly divided into 80% for training and 20% for validation. The training data were used to build the optimal classification model, while the validation data were used to refine the network structure. Over the course of 500 epochs, the model was trained on the training data and evaluated on the validation data at each epoch. The model with the highest classification accuracy on the validation data was preserved and subsequently tested on the test set. At the conclusion of each fold, the denoising signal SNR, RMSE, and  $r$  of the models were computed, and the final result was determined by averaging the outcomes of the 5-fold cross-validation. Table 3 showcases the results of the 5-fold cross-validation, revealing that the SNR ranges from 12.782 to 13.763 dB, RMSE varies



**Fig. 7** Loss curves for different depths of the MSDCAN denoising model **a** training loss **b** validation loss

**Table 3** Cross-validation experiment results

Fold	SNR	RMSE	<i>r</i>
1k-fold	12.782	0.414	0.933
2k-fold	13.734	0.232	0.966
3k-fold	13.009	0.329	0.942
4k-fold	13.763	0.227	0.969
5k-fold	13.002	0.413	0.955
Mean	13.258	0.323	0.953

between 0.232 and 0.414, and the correlation coefficient falls within the range of 0.933–0.969.

### Ablation experiments

To verify the impact of Multiscale Dilated Convolution (DMS) block, the Initial Dilated Convolution (DC) block, and the Improved Attention Mechanism (Improved SE) block in the MSDCAN denoising model on model performance, ablation experiments and analysis were conducted, including: original self-coding network (ablation model CAE), retaining the DC and improved SE blocks (ablation model DCAN), retaining two blocks: DC and DMS (ablation model MSDCN), and retaining two blocks: improved SE and DMS (ablation model MSCAN) as well as preserving the DC, DMS, and original SE blocks (ablation model MSDCAN-original SE). The experimental results are shown in Table 4.

From Table 4, it can be seen that the average SNR of the ablation model CAE is the lowest, indicating that the three blocks play an important role in the denoising effect of the MSDCAN model. Compared with the MSDCAN model, the average SNR of the ablation model DCAN (lacking DMS) decreased by 0.965 dB, the average SNR of the ablation model MSDCN (lacking improved SE) decreased by 0.441 dB, and the average SNR of the ablation model MSCAN (lacking DC) decreased by 0.238 dB. This indicates that deleting any of the DMS, improved SE, and DC blocks will significantly reduce the denoising effect of the network. Compared with the original SE, the improved SE increased the average SNR by 1.724 dB, indicating that the improved SE has a significant denoising effect on the MSDCAN denoising model.

**Table 4** Results of the ablation experiments on the MSDCAN network

	DMS	SE	DC	SNR/dB	RMSE	<i>r</i>
CAE	x	x	x	9.56	0.488	0.91
DCAN	x	✓	✓	12.268	0.353	0.947
MSDCN	✓	x	✓	12.792	0.358	0.942
MSCAN	✓	✓	x	12.995	0.319	0.949
MSDCAN	✓	✓	✓	13.258	0.323	0.953
MSDCAN-original SE	✓	Original	✓	11.534	0.542	0.87

Three sets of parameters were chosen for training after conducting multiple tests on the training dataset, and the outcomes were compared. The learning rate (LR), Batch\_Size, and epochs for the MSDCAN denoising model were fine-tuned and selected. In each experiment, all the parameters, except for the one being tested, were held constant.

The learning rate, a critical hyperparameter in the realm of deep learning, plays a pivotal role in deciding whether and when the model can effectively converge toward a minimum. This study explored several common learning rates, specifically LR = 0.01, LR = 0.001, LR = 0.0001, and LR = 0.00001, and the results presented in Table 5 reveal that optimal values for both SNR and *r* are achieved, while the RMSE value is minimized when LR = 0.0001. This is because a learning rate that is too high causes the network to fail to converge and the model accuracy to decrease; a learning rate that is too small prolongs the network convergence time and reduces the speed of model training. Therefore, the study chooses LR = 0.0001 as the learning rate for the training of the MSDCAN denoising model.

The Batch\_Size plays a crucial role in optimizing the model and determining the training speed. To expedite the training process in a gradient descent algorithm, it is customary to use a Batch\_Size that is a power of 2. In this study, Batch\_Sizes of 16, 32, 64, 128, and 256 were evaluated for the MSDCAN denoising model, and the outcomes are presented in Table 6. When its value is 128, the model SNR reaches its maximum value, and after reaching the maximum SNR, the accuracy begins to decrease as the number increases. This is because too small a batch can make training longer and less efficient; too large a batch can lead to a decrease in the model's ability to generalize. Therefore, the study selects 128 for the Batch\_size.

**Table 5** Influences of learning rates

LR	SNR	RMSE	<i>r</i>
LR = 0.01	10.693	0.414	0.938
LR = 0.001	11.734	0.383	0.946
LR = 0.0001	13.258	0.323	0.953
LR = 0.00001	10.763	0.427	0.939

**Table 6** Influences of batch\_size

Batch_size	SNR	RMSE	<i>r</i>
Batch_size=16	10.966	0.427	0.912
Batch_size=32	12.873	0.381	0.947
Batch_size=64	12.932	0.394	0.938
Batch_size=128	13.258	0.323	0.953
Batch_size=256	11.825	0.415	0.932

**Table 7** Influences of Epoch

Epoch	SNR	RMSE	<i>R</i>
epoch = 400	11.923	0.358	0.925
epoch = 500	13.258	0.323	0.953
epoch = 600	12.372	0.349	0.945

To develop an effective denoising model, the training process requires an appropriate choice of epochs. An epoch represents a single training iteration, encompassing both forward and backward propagation for all batches. In this study, epochs of 400, 500, and 600 were assessed for the MSDCAN denoising model, and the findings are presented in Table 7.

When epoch = 500, the values of both SNR and *r* are maximized and the value of RMSE is minimized. This is because, at epoch=400, the model does not learn enough to capture the micro-seismic signal features, resulting in a low SNR, RMSE, and *r* that are not close enough to 1, making the model denoising ineffective; at epoch=600, the model training falls into the overfitting phenomenon, resulting in the model denoising effect is not as effective as at epoch=500. Therefore, the paper chooses epoch=500 as the epoch for MSDCAN denoising model training.

In the training process, an initial learning rate was set, typically denoted as  $\eta$ . To strike a balance between training accuracy and speed, a given value decay method was employed to decay the learning rate throughout training, with a minimum learning rate specified as  $\eta_{min}$ . To expedite model convergence, the Adam optimizer Kingma and Ba (2014) was utilized. Adam offers high computational efficiency and adapts the learning rate in the iterative solving process. In the training phase, a total of 500 epochs were performed, and a batch size of 128 was employed. These settings enable the model to learn from the training data effectively, refining its performance over multiple iterations.

## Comparative experiments

The study used 3600 test signals to compare the performance of MSDCAN with three other denoising methods: DeepDenoiser Zhu et al. (2019), CNN-denoiser Zhang et al. (2020) and Neighbor2Neighbor Wang et al. (2023).

Three of these methods, DeepDenoiser, CNN-denoiser, and Neighbor2Neighbor, were recreated using PyTorch and subsequently trained using the training dataset as outlined in Sect. 3.1. Evaluation of all methods took place on the test dataset detailed in Sect. 3.1. In light of the increasing complexity of noise in micro-seismic signals, which corresponds to a greater proportion of noisy waveforms and a reduction in the conspicuousness of seismic phases, the presence of micro-seismic signals within the signal waveform remains uncertain. Consequently, for a comprehensive assessment of the efficacy of the denoising models in handling intricate noise scenarios, the study conducted a thorough comparison and analysis of three denoising methods by synthesizing micro-seismic signals characterized by varying SNRs. Figures 8a–d present the denoising outcomes at SNRs of 2 dB, 0 dB, – 2 dB, and – 6 dB, respectively.

In Fig. 8a, when the SNR is 2 dB, all four denoising methods can eliminate the noise in the micro-seismic signals, but the MSDCAN denoising model has a better denoising effect. The denoising efficacy of the DeepDenoiser, CNN-denoiser, and Neighbor2Neighbor methods noticeably declines as the SNR decreases. Moving to 8b, at an SNR of 0 dB, MSDCAN successfully preserves valuable micro-seismic information. Conversely, the CNN-denoiser method may inadvertently retain some noise as useful signals, Deepdenoiser may lead to the loss of micro-seismic information, and Neighbor2Neighbor may result in a denoised signal waveform that substantially deviates from the original, causing signal distortion. In 8c, at an SNR of – 2 dB, the MSDCAN model effectively removes various types of noise while maximally retaining valuable micro-seismic information. On the other hand, CNN-denoiser, DeepDenoiser, and Neighbor2Neighbor methods preserve some impulse noise information as useful signals, leading to the misidentification of micro-seismic events and P-wave arrivals. Lastly, in Fig. 8d, at an SNR of – 6 dB, the DeepDenoiser method inadvertently removes micro-seismic information as noise, resulting in signal distortion. In contrast, MSDCAN, CNN-denoiser, and Neighbor2Neighbor methods do not exhibit significant distortion. However, the CNN-denoiser and Neighbor2Neighbor methods still exhibit residual noise, resulting in minimal distortion.

Figure 9 presents the fluctuations in performance metrics across the four methods at varying SNRs. Specifically, Fig. 9a, b, and c depict the results for SNR, RMSE, and *r*, respectively. As can be clearly seen from Fig. 9a, the MSDCAN, CNN-denoiser, DeepDenoiser, and Neighbor2Neighbor methods yield average SNRs of 13.258 dB, 8.585 dB, 10.714 dB, and 11.768 dB, respectively. As depicted in Fig. 9b, the average RMSEs for these four methods are 0.323, 0.500, 0.572, and 0.505, respectively. Furthermore, Fig. 9c reveals that the average *r* stands at 0.953, 0.903, 0.877, and 0.952 for the MSDCAN, CNN-denoiser, DeepDenoiser, and

Neighbor2Neighbor methods, respectively. These findings underscore the supremacy of the MSDCAN model, as it boasts the highest SNR, the lowest RMSE that approximates 0, and  $r$  approaching 1. This underscores the robustness of the MSDCAN denoising model against random noise in micro-seismic signals. It showcases an adaptive capacity to remove noise across varying SNRs, thereby significantly elevating the SNR and overall quality of micro-seismic signals.

The above experimental results show that the signal-to-noise ratio and correlation coefficient of the micro-seismic signals after denoising using Deepdenoiser, CNN-denoiser, and Neighbor2Neighbor methods are improved, and the RMSE is reduced, but all of them are inferior to the MSDCAN denoising model. This is because the Deepdenoiser method exhibits a fixed time-frequency resolution, which is attributed to the utilization of short-time Fourier transform (STFT) time windows. As a result, it is difficult to adequately recognize and preserve micro-seismic information, leading to suboptimal denoising results. On the other hand, the CNN-denoiser model introduces slight distortion as it faces the challenge of extracting subtle information from micro-seismic signals, such as frequency and amplitude along the same phase axis. In the case of the Neighbor2Neighbor method, it produces smooth signals but can produce severe distortions due to its limited sensitivity in distinguishing micro-seismic signals from noise. In contrast, the MSDCAN denoising model uses dilated convolution to extract neighborhood features from micro-seismic signals. This approach utilizes a wider range of contextual information and can effectively distinguish between micro-seismic signal components and noise components. As a result, the denoised micro-seismic signal achieves the highest SNR and exhibits excellent denoising results.

### Different types of noise testing

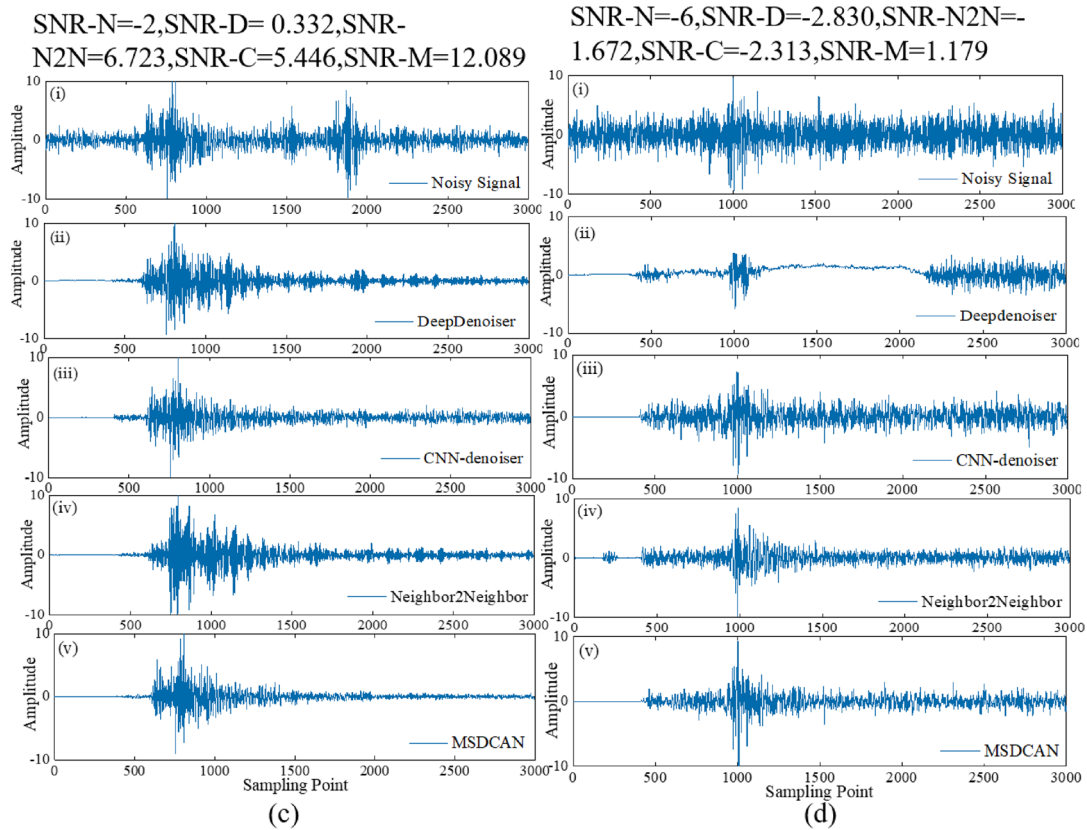
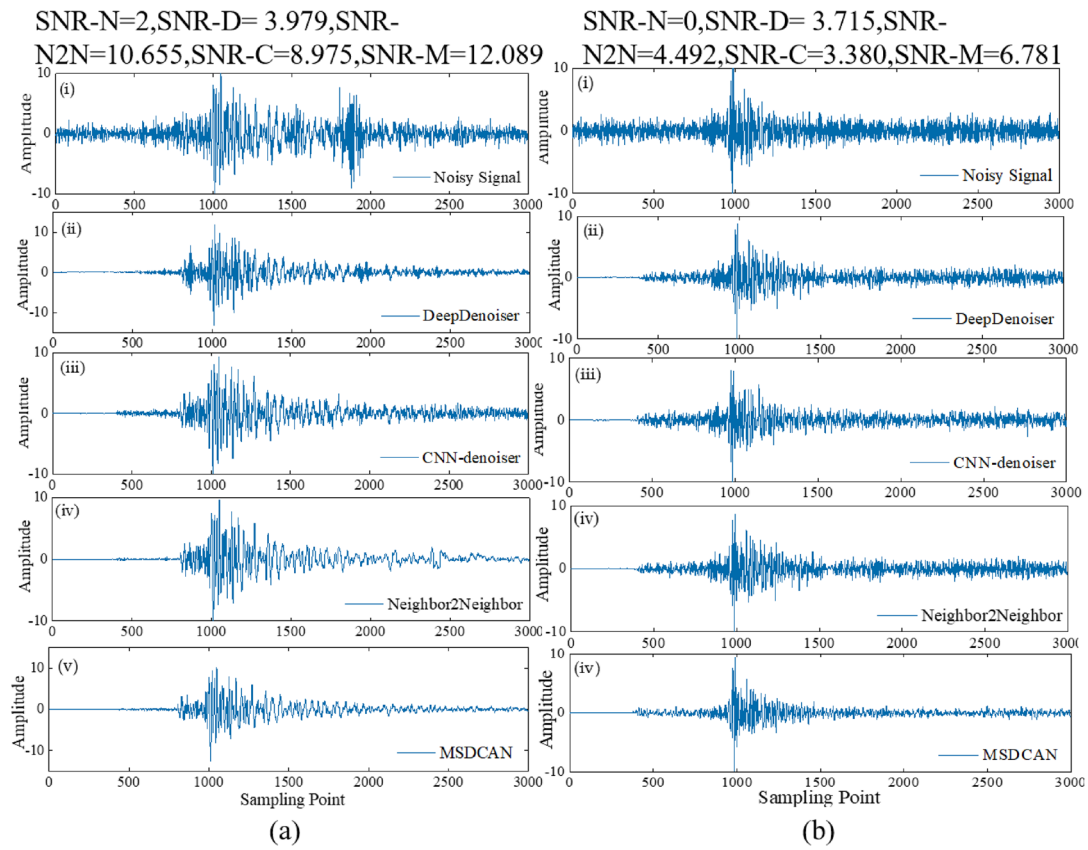
Considering that micro-seismic signals are often affected by background noise, impulse noise, periodic noise, and mixed noise, in order to evaluate the denoising performance of the MSDCAN denoising model for different types of noise, SNR, RMSE, and  $r$  were used as evaluation indicators. Four denoising methods, MSDCAN, DeepDenoiser, CNN-denoiser, and Neighbor2Neighbor, were used to denoise 3200 micro-seismic signals, including background noise, impulse noise, periodic noise and mixed noise. The quantitative results of denoising are shown in Table 8, and the qualitative results of denoising are shown in Figs. 10, 11, 12 and 13.

Regarding background noise, as demonstrated in Table 8, we computed the SNR, RMSE, and  $r$  for the denoising outcomes obtained from the four methods. The SNRs corresponding to the MSDCAN model, DeepDenoiser,

CNN-denoiser, and Neighbor2Neighbor are 7.601 dB, 4.542 dB, 4.937 dB, and 5.601 dB, respectively. The RMSE values are 0.573, 0.874, 0.8, and 0.573, while the  $r$  values are 0.886, 0.719, 0.815, and 0.847, respectively. Notably, the denoising results from the MSDCAN denoising model exhibit the highest SNR, the lowest RMSE, and a  $r$  closest to 1. The marked enhancement in SNR, significant reduction in RMSE, and substantial improvement in correlation coefficient indicate the effectiveness of the MSDCAN denoising model in suppressing background noise while faithfully recovering micro-seismic signals. To provide a more intuitive assessment of the denoising effect, we generated waveform curves for the noise signal and the denoised signals using various denoising methods, as illustrated in Fig. 10. Post-denoising, the MSDCAN denoising model yields a relatively smooth waveform, whereas the DeepDenoiser method leads to significant waveform distortion. The CNN-denoiser method results in a rough waveform after denoising, with more residual noise both preceding the arrival of the P-wave and at the conclusion of the micro-seismic event. Although the Neighbor2Neighbor method effectively eliminates noise and produces a smooth signal post-denoising, signal distortion arises from the seamless transition to the original signal. This distortion can impede the accurate identification of subsequent micro-seismic events.

From the experimental results, it can be seen that the MSDCAN denoising model shows the best performance in dealing with the background noise in micro-seismic signals, which proves that the MSDCAN denoising model can eliminate complex noise interference. Deepdenoiser methods, CNN-denoiser methods, and Neighbor2Neighbor methods are less capable of selectively suppressing noise because they do not selectively enhance micro-seismic signal features, leading to difficulties in background noise removal. In contrast, the MSDCAN denoising model excels in removing background noise and better distinguishing between signal and noise due to the combination of an improved SE that focuses more attention on the signal or important information.

For impulse noise, as detailed in Table 8, we computed the SNR, RMSE, and  $r$  for the denoising outcomes obtained from each of the four methods. The SNR values corresponding to the MSDCAN model, DeepDenoiser, CNN-denoiser, and Neighbor2Neighbor are 18.393 dB, 16.919 dB, 10.594 dB, and 16.393 dB, respectively. The RMSE values are 0.151, 0.184, 0.380, and 0.350, while the  $r$  values are 0.994, 0.986, 0.944, and 0.961, respectively. Notably, the denoising results obtained from the MSDCAN denoising model exhibit the highest SNR, the lowest RMSE, and a  $r$  closest to 1. The remarkable increase in SNR, significant reduction in RMSE, and substantial improvement in  $r$  indicate the MSDCAN denoising model's capacity to effectively suppress impulse noise while preserving P-wave information. To provide a



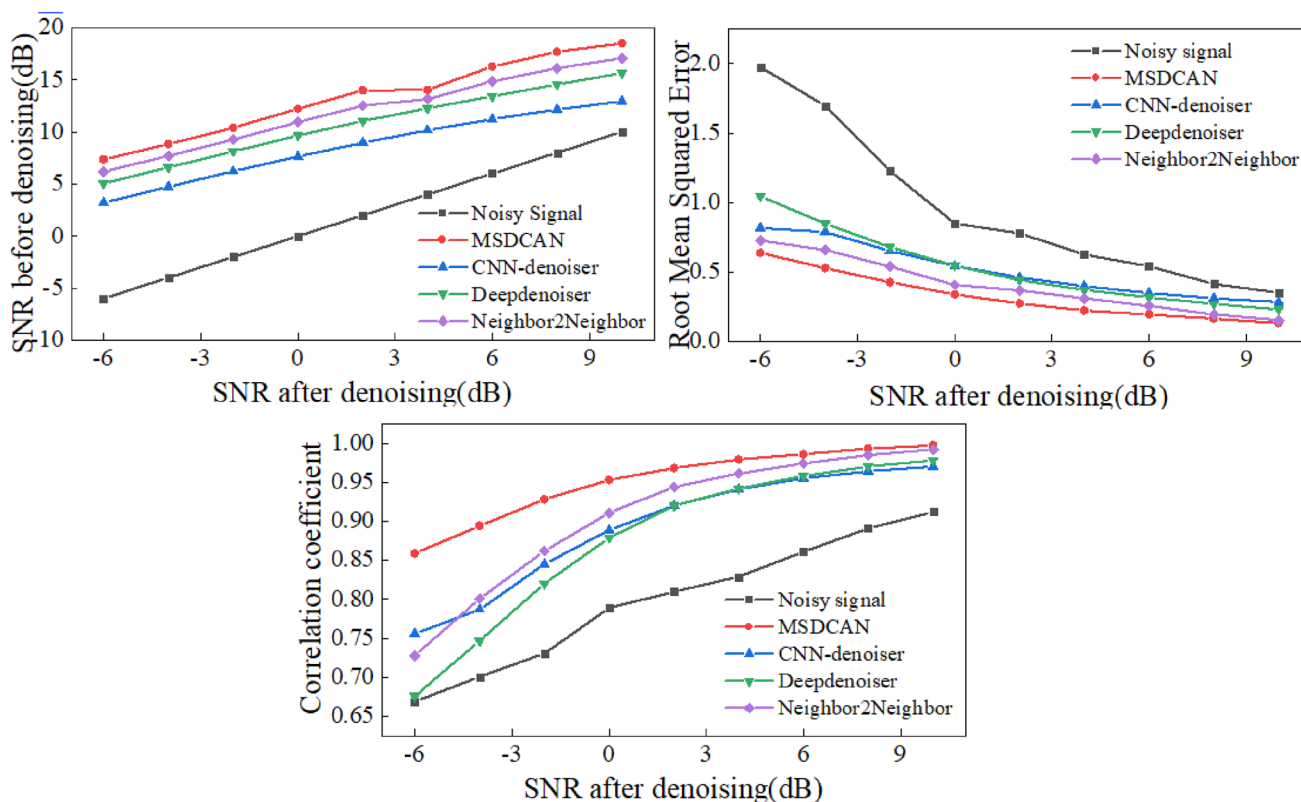
**Fig. 8** Denoising results of micro-seismic signals with different SNRs; **a-d** are four examples for the denoised signals corresponding to the DeepDenoiser method, CNN-denoiser method, Neighbor2Neighbor method and MSDCAN method. SNR-N, SNR-D, SNR-C, SNR-N2N and SNR-M are the SNRs for the noisy, DeepDenoiser, CNN-denoiser, Neighbor2Neighbor and the MSDCAN signals. The SNR values in dB. For each subfigure, (i) denotes the noise signal, (ii) denotes the denoised signal corresponding to the DeepDenoiser method, (iii) represents the denoised signal corresponding to the CNN-denoiser method, (iv) is the denoised signal corresponding to the Neighbor2Neighbor method, and (v) is the denoised signal corresponding to the MSDCAN method

more intuitive assessment of the denoising effect, we generated waveform curves for the noise signal and the denoised signals using various denoising methods, as depicted in Fig. 11. The denoising results of all four methods exhibit no significant distortion. However, the MSDCAN denoising model stands out with the most effective noise suppression. While the DeepDenoiser, CNN-denoiser, and Neighbor2Neighbor methods can effectively remove high-amplitude pulse components, low-amplitude noise persists before the arrival time of the P-wave, impacting the accuracy of subsequent P-wave detection.

From the experimental results, it can be seen that the MSDCAN denoising model shows the best performance in dealing with impulse noise in micro-seismic signals, which

proves that the MSDCAN denoising model can eliminate the interference of impulse noise. Impulse noise is usually a waveform with a false amplitude caused by mechanical or anthropogenic vibrations in the field and exists mainly before the arrival time of the P-wave. Its amplitude and frequency are very similar to the arrival time of the seismic phase, the only difference being the polarization direction information of both. DeepDenoiser and Neighbor2Neighbor methods have a small sensory field and limited feature extraction capability because they use ordinary convolution to capture the detailed information of micro-seismic signals, which makes it challenging to identify the polarization information of micro-seismic signals and the denoising effect is poor. The CNN-denoiser method does not extract the complex features of micro-seismic signals because it uses a single-scale dilated convolution to capture the neighborhood features of micro-seismic signals without considering the multiscale features of micro-seismic signals, leading to difficulties in removing the impulse noise. In contrast, the MSDCAN denoising model uses different dilated rates to design multiscale dilated convolution, which can capture different ranges of multiscale information, which makes the MSDCAN denoising model the most effective for denoising.

For periodic noise, as delineated in Table 8, we computed the SNR, RMSE, and r for the denoising results achieved



**Fig. 9** Performance comparison between DeepDenoiser, CNN-denoiser, Neighbor2Neighbor and MSDCAN **a** improvement of SNR; **b** improvement of RMSE; **c** correlation coefficient

**Table 8** Table with merged rows and columns using tabularx

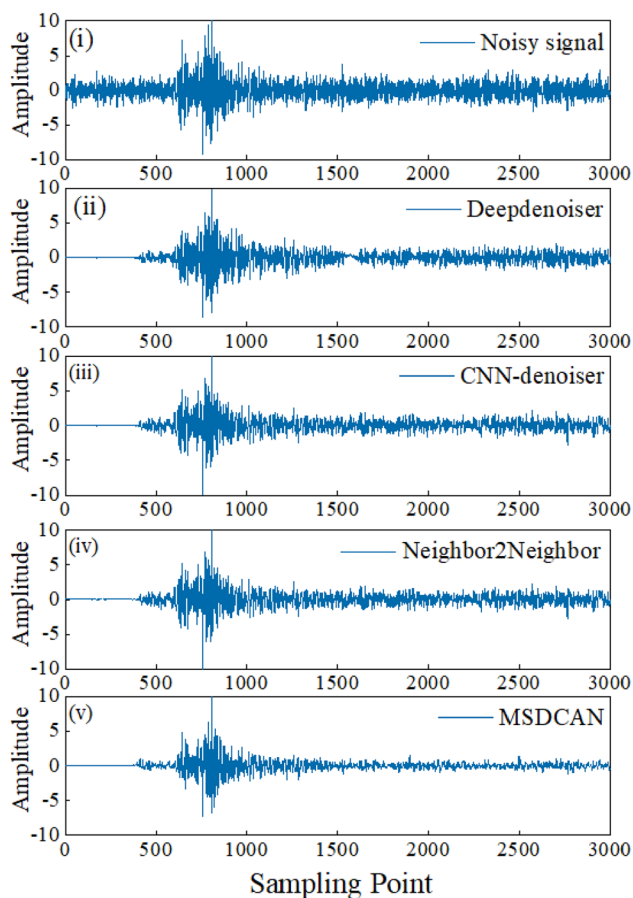
Type of noise	SNR/RMSE/r				
	Before denoising	DeepDenoiser	CNN-denoiser	Neighbor2 Neighbor	MSDCAN
Background noise	2/1.181/0.739	4.542/0.874/0.719 (↑ 2.542/ ↓ 0.307/ ↓ 0.02)	4.937/0.800/0.815 (↑ 3.03/ ↓ 0.381/ ↑ 0.076)	5.601/0.608/0.847 (↑ 3.601/ ↓ 0.573/ ↑ 0.108)	7.601/0.573/0.886 (↑ 5.601/ ↓ 0.608/ ↑ 0.147)
Impulse noise	2/1.181/0.733	16.919/0.184/0.986 (↑ 14.919/ ↓ 0.997/ ↑ 0.253)	10.594/0.380/0.944 (↑ 8.594/ ↓ 0.801/ ↑ 0.211)	16.393/0.350/0.961 (↑ 14.393/ ↓ 0.831/ ↑ 0.228)	(18.393/0.151/0.994) (↑ 16.393/ ↓ 1.035/ ↑ 0.261)
Periodic noise	2/1.506/0.622	17.456/0.175/0.988 (↑ 15.456/ ↓ 1.331/ ↑ 0.366)	14.686/0.236/0.978 (↑ 11.686/ ↓ 1.27/ ↑ 0.356)	16.787/0.268/0.986 (↑ 14.787/ ↓ 1.238/ ↑ 0.364)	(18.787/0.138/0.994) (↑ 16.787/ ↓ 1.368/ ↑ 0.372)
Mixed noise	2/1.181/0.739	3.760/0.936/0.789 (↑ 1.760/ ↓ 0.245/ ↑ 0.05)	8.024/0.534/0.903 (↑ 6.024/ ↓ 0.647/ ↑ 0.164)	8.289/0.0.793/0.915 (↑ 6.289/ ↓ 0.388/ ↑ 0.176)	10.289/0.408/0.934 (↑ 8.289/ ↓ 0.773/ ↑ 0.195)

through each of the four methods. The SNR values corresponding to the MSDCAN model, DeepDenoiser, CNN-denoiser, and Neighbor2Neighbor are 18.787 dB, 17.456 dB, 14.686 dB, and 16.787 dB, respectively. The RMSE values are 0.138, 0.175, 0.236, and 0.268, respectively, while the  $r$  values are 0.994, 0.988, 0.978, and 0.986, respectively. It is noteworthy that the denoising outcomes yielded by the MSDCAN denoising model exhibit the highest SNR, the lowest RMSE, and a correlation coefficient closest to 1. The substantial increase in SNR, significant reduction in RMSE, and considerable enhancement of  $r$  underscore the MSDCAN denoising model's ability to eliminate periodic noise while fully preserving micro-seismic events. To provide a more visually intuitive assessment of the denoising effect, we generated waveform curves for the noise signal and the denoised signals using various denoising methods, as depicted in Fig. 12. All four methods prove effective in removing periodic noise from the signal. From the experimental results, it becomes evident that the MSDCAN denoising model effectively denoises the signal before the arrival time of the P-wave without introducing any artifacts, resulting in a smooth waveform. In contrast, the DeepDenoiser method, CNN-denoiser, and Neighbor2Neighbor methods still exhibit numerous artifacts in the signal after denoising, particularly before the arrival time of the P-wave. This results in an uneven signal that cannot accurately detect the P-wave.

From the experimental results, it can be seen that the MSDCAN denoising model improves the SNR and  $r$  of the denoised signal and reduces the RMSE when dealing with periodic noise in micro-seismic signals. This proves that the MSDCAN denoising model can eliminate the interference of periodic noise. Since periodic noise usually repeats signal variations at fixed time intervals within a certain frequency range, DeepDenoiser methods, CNN-denoiser methods, and Neighbor2Neighbor methods are unable to focus attention on the localized region of periodic noise, and it is difficult to accurately isolate the micro-seismic signal from the periodic noise, which results in the inability to reduce the interference of the periodic noise on the micro-seismic signal. However, the MSDCAN denoising model, due to the incorporation of an improved SE, is not only more localized in dealing with periodic noise, but also better distinguishes the frequency bands where the periodic noise overlaps with micro-seismic signals, allowing for better handling of periodic noise.

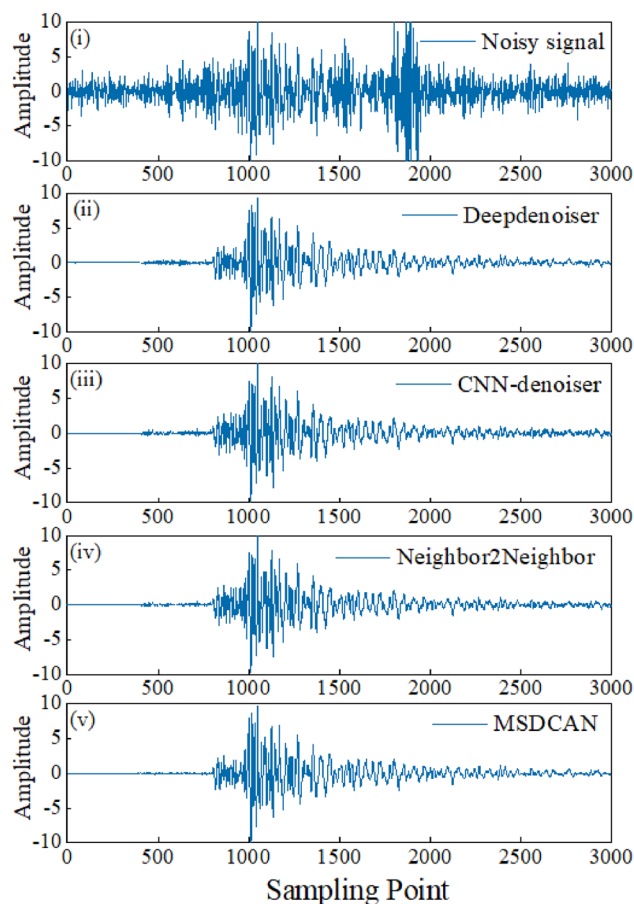
For mixed noise, as indicated in Table 8, we computed the SNR, RMSE, and  $r$  for the denoising results obtained through each of the four methods. The SNR values corresponding to the MSDCAN model, DeepDenoiser, CNN-denoiser, and Neighbor2Neighbor are 10.289 dB, 3.76 dB, 8.024 dB, and 8.289 dB, respectively. The RMSE values are 0.408, 0.936, 0.534, and 0.793, respectively, while the  $r$  values are 0.934, 0.789, 0.903, and 0.915, respectively.





**Fig. 10** Denoising performance in the presence of background noise. (i) indicates the noisy signal, (ii) represents the denoised signal corresponding to the DeepDenoiser method, (iii) denotes the denoised signal corresponding to the CNN-denoiser method, (iv) is the denoised signal corresponding to the Neighbor2Neighbor method, and (v) is the denoised signal corresponding to the MSDCAN method

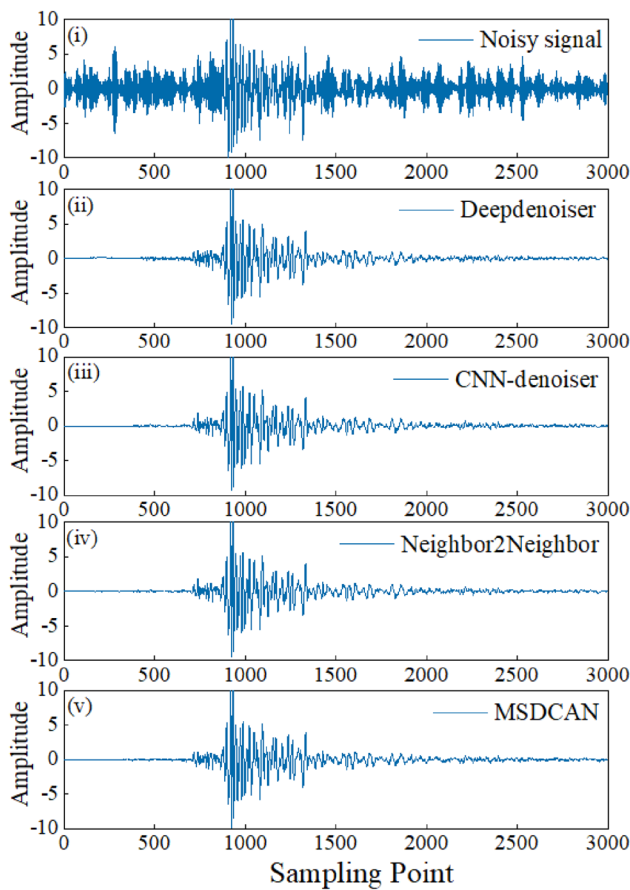
Notably, the denoising outcomes achieved by the MSDCAN denoising model exhibit the highest SNR, the lowest RMSE, and a  $r$  closest to 1. The substantial enhancement in SNR, marked reduction in RMSE, and significant improvement in  $r$  underscore the ability of the MSDCAN denoising model to eliminate mixed noise. To provide a more visually intuitive assessment of the denoising effect, we have generated waveform curves for the noise signal and the denoised signals using various denoising methods, as shown in Fig. 13. After denoising, the MSDCAN model smoothens the signal and notably enhances the SNR. In contrast, DeepDenoiser and CNN-denoiser exhibit a significant amount of noise residue, particularly when the signal's end oscillations have a high amplitude. The Neighbor2Neighbor method still retains a considerable amount of minor noise after denoising. The rationale behind these results lies in the nature of mixed noise,



**Fig. 11** Denoising performance in the presence of impulsive noise. (i) indicates the noisy signal, (ii) denotes the denoised signal corresponding to the DeepDenoiser method, (iii) represents the denoised signal corresponding to the CNN-denoiser method, (iv) denotes the denoised signal corresponding to the Neighbor2Neighbor method, and (v) denotes the denoised signal corresponding to the MSDCAN method

which encompasses a combination of pulse noise and irregular noise within micro-seismic signals.

From the experimental results, it can be seen that the MSDCAN denoising model improves the SNR and  $r$  of the denoised signal and reduces the RMSE when dealing with mixed noise in micro-seismic signals. This proves that the MSDCAN denoising model can eliminate the interference of mixed noise. Mixed noise generally includes impulse noise and irregular noise in micro-seismic signals, and the DeepDenoiser method is weak in removing noise because the receptive field is too small to preserve detailed features. Mixed noise contains different types of noise at different times and frequencies, and the CNN-denoiser method and the Neighbor2Neighbor method are weak in denoising because they cannot adjust the attention to the different components of the micro-seismic signal to accommodate the complex noise. In contrast, the MSDCAN model excels



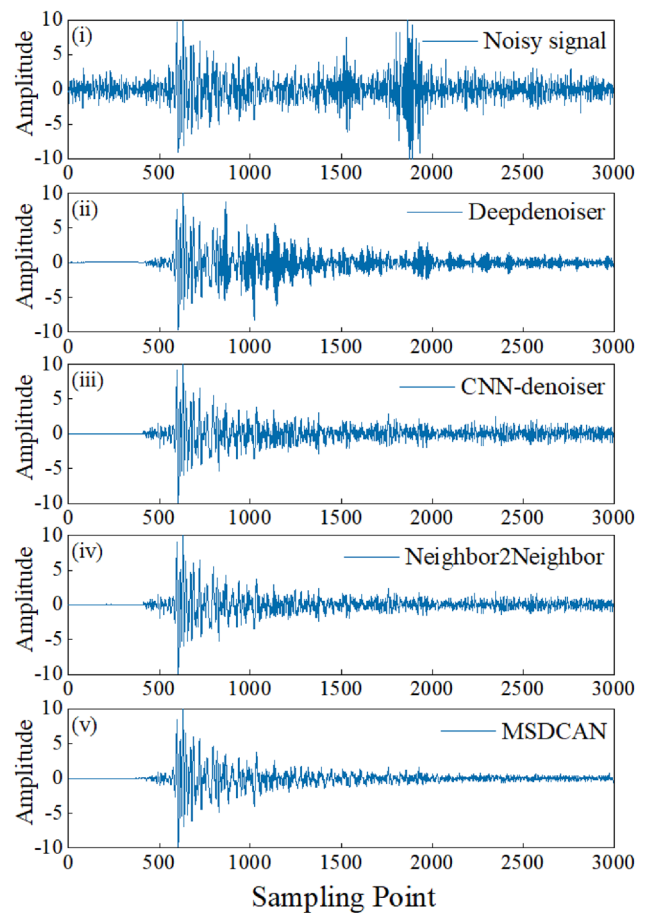
**Fig. 12** Denoising performance in the presence of cyclic noise. (i) indicates the noisy signal, (ii) is the denoised signal corresponding to the DeepDenoiser method, (iii) denotes the denoised signal corresponding to the CNN-denoiser method, (iv) denotes the denoised signal corresponding to the Neighbor2Neighbor method, and (v) denotes the denoised signal corresponding to the MSDCAN method

at retaining detailed information about the signal while eliminating different types of noise with significant denoising effects due to the use of dilated convolution and improved SE.

**Practical applications**

**Application of micro-seismic denoising at Beijing stations**

In order to further investigate the practical applicability and versatility of the MSDCAN denoising model, we applied it to micro-seismic events recorded at Beijing stations from 2011 to 2017. For model validation, a random selection of 348 waveform data samples from micro-seismic events was employed. We followed the data preprocessing methodology outlined in Sect. 3.1 to prepare the micro-seismic data. In situations where the restoration of the actual signal is unattainable, we employed the P-wave signal-to-noise ratio(PSNR)



**Fig. 13** Denoising performance in the presence of mixed noise. (i) indicates the noisy signal, (ii) is the denoised signal corresponding to the DeepDenoiser method, (iii) represents the denoiser signal corresponding to the CNN-denoiser method, (iv) denotes the denoised signal corresponding to the Neighbor2Neighbor method, and (v) denotes the denoised signal corresponding to the MSDCAN method

as an evaluation metric to scientifically assess the quality of signal denoising before and after applying the model in practical scenarios. The PSNR serves as a robust indicator, offering a precise measure of the distinction between the signal proximate to the P-wave and the background noise. A higher PSNR signifies enhanced signal quality following denoising. The calculation formula for the PSNR is presented in Eq. 18:

$$PSNR = 10 \lg \left( \frac{\sum_{i=p}^{p+l} x_i^2}{\sum_{i=p-l}^p x_i^2} \right) \tag{18}$$

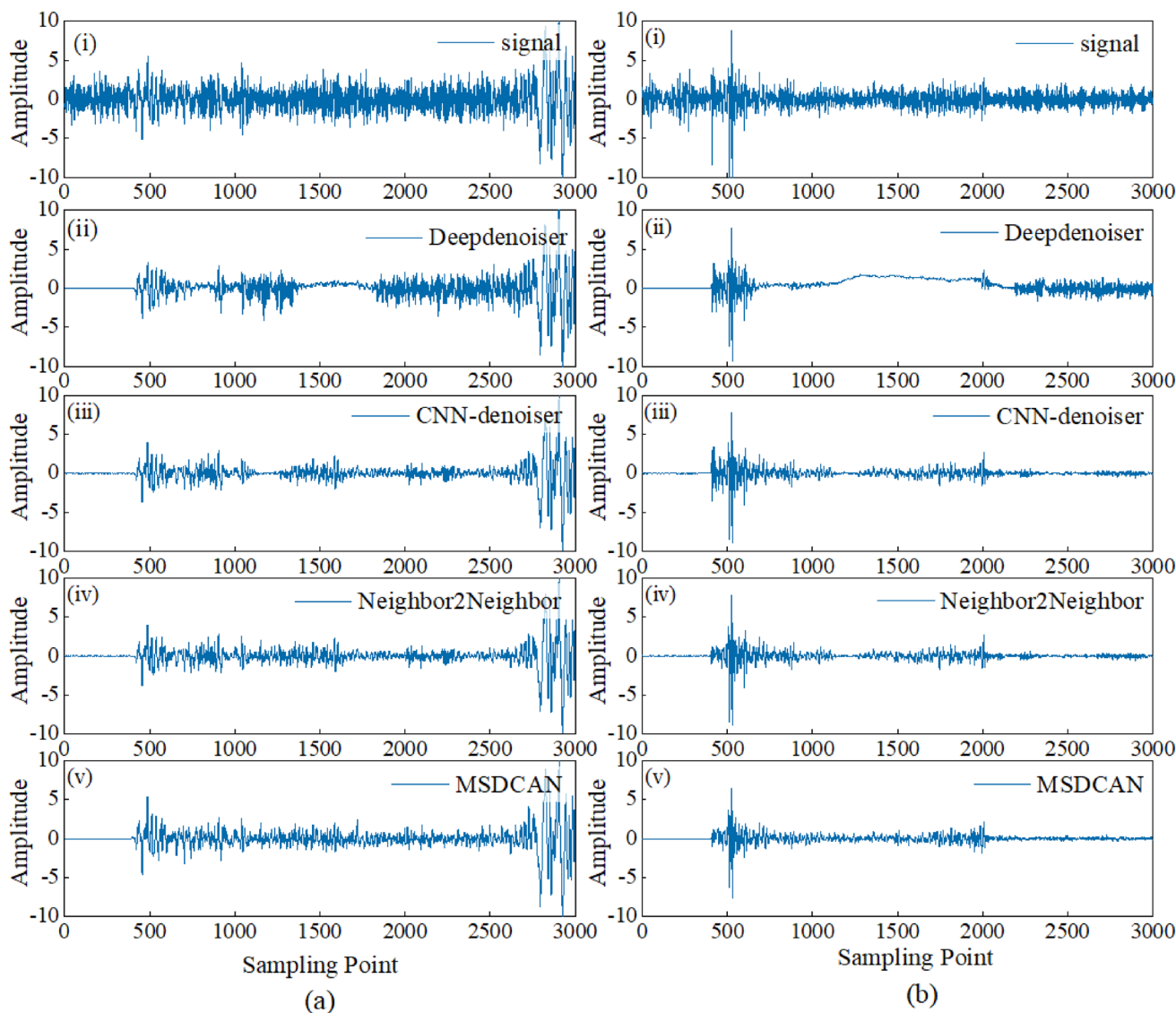
where p represents the exact P-wave arrival time, l is the calculated length, and  $x_i$  is the micro-seismic signal. The PSNR can better reflect the difference between the signal near the P-wave and the background noise.

**Table 9** Average PSNR

Type	Average PSNR	Timing/s
Original signal	1.268	–
DeepDenoiser	12.864	0.0819
CNN-denoiser	26.376	0.0719
Neighbor2Neighbor	30.112	0.0921
MSDCAN	35.478	0.0871

Table 9 presents the average PSNR results for the MSD-CAN model, DeepDenoiser, CNN-denoiser, and Neighbor2Neighbor. It is evident that when compared to the DeepDenoiser, CNN-denoiser, and Neighbor2Neighbor methods, the MSDCAN model exhibits an impressive increase in PSNR by 22.614 dB, 9.102 dB, and 5.366 dB, respectively. Table 9 also highlights that the Neighbor2Neighbor method requires the most extended processing time, whereas the MSDCAN denoising model operates efficiently within 0.0808 s, satisfying the demands of real-time detection.

Furthermore, as depicted in Fig. 14, it is evident that the DeepDenoiser, CNN-denoiser, and Neighbor2Neighbor



**Fig. 14** Results comparison of different denoising methods on micro-seismic signals from Beijing stations. **a–b** are two examples of the denoised signals corresponding to the DeepDenoiser method, the CNN-denoiser method, the Neighbor2Neighbor method and the MSDCAN method. For each subfigure, (i) denotes the noise signal,

(ii) denotes the denoised signal corresponding to the DeepDenoiser method, (iii) is the denoised signal corresponding to the CNN-denoiser method, (iv) represents the denoiser signal corresponding to the Neighbor2Neighbor method, and (v) represents the denoiser signal corresponding to the MSDCAN method

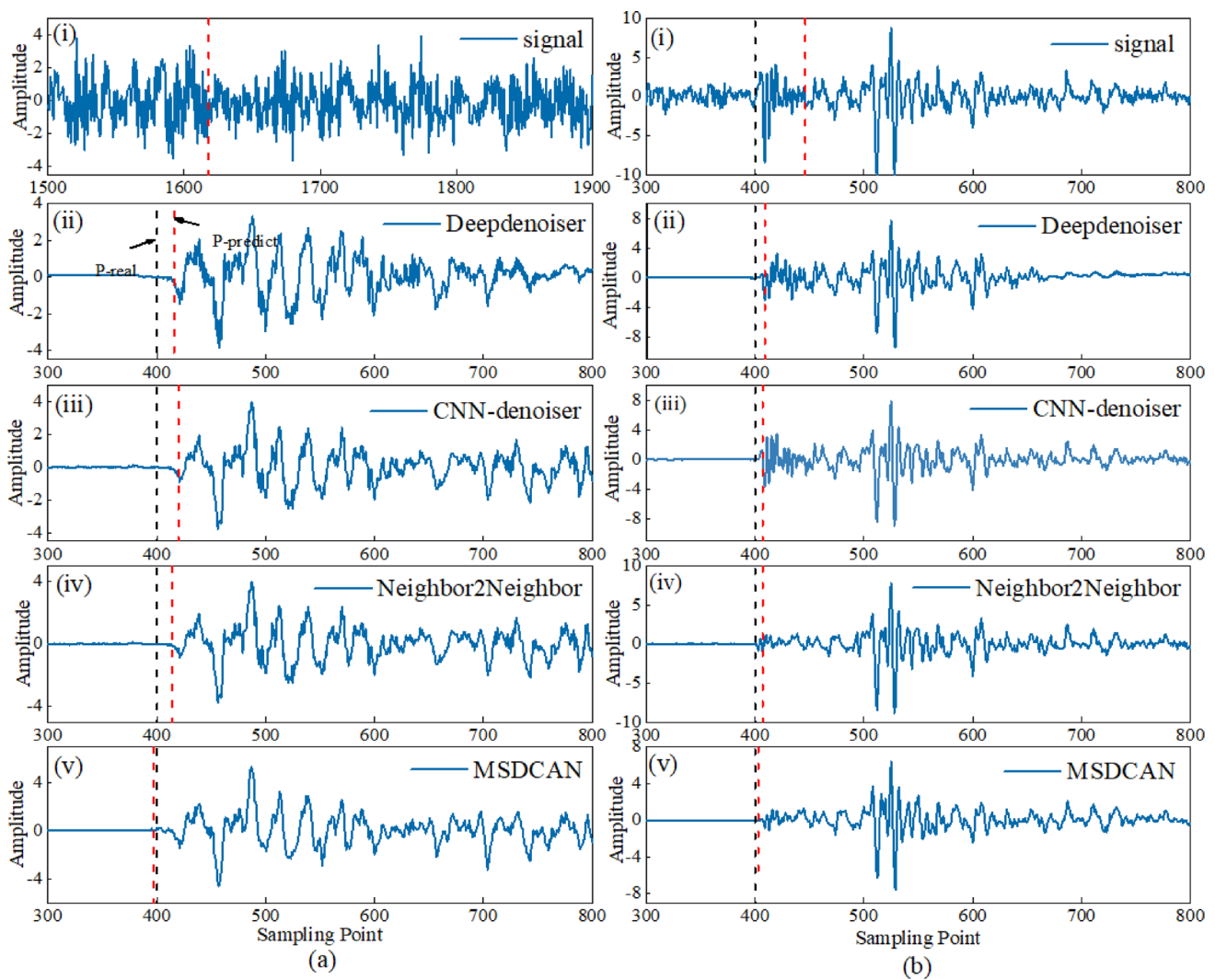
methods introduce severe signal distortion. In contrast, the MSDCAN denoising model not only effectively suppresses noise but also preserves the complete characteristics of P-wave arrivals post-denoising. The preservation results in significant amplitude changes, minimal signal distortion, and a remarkable enhancement in SNR. The Deepdenoiser method of denoising appears to have a naked-eye detectable waveform distortion caused by improperly set window lengths, as shown in Fig. 14b (ii); the CNN-denoiser method denoising then leads to a certain degree of loss of small-size texture details; the Neighbor2Neighbor method still leaves non-negligible noise after denoising. Overall, the MSDCAN denoising model shows a smoother and more stable performance in the processing of the P-wave of the micro-seismic signal and the starting and stopping of the micro-seismic event while removing the complex noise, making the overall frequency variation of the signal uniform, avoiding side effects such as spectral aliasing, and eliminating the interference of the complex noise on the micro-seismic P-wave pickups.

Enhancing signal quality profoundly impacts subsequent micro-seismic data processing. The accurate determination of P-wave arrival time forms the bedrock for micro-seismic data positioning and source parameter calculations. In case of featuring robust SNRs, automatic pickup algorithms adeptly discriminate between the arrival time of P-wave and the pre-noise segment, facilitating precise results. Conversely, when dealing with low SNR signals, these automatic pickup algorithms may exhibit notable discrepancies. However, by applying the MSDCAN denoising model to micro-seismic signals, the SNR can be significantly augmented. This augmentation, in turn, enhances the quality of P-wave arrival time pickup, ultimately refining the accuracy of micro-seismic positioning. Figure 15a and b, respectively, illustrates the picking results for two micro-seismic signals showcased in Fig. 15. The black and red dashed lines represent the outcomes of expert manual picking and STA/LTA automatic picking, respectively. As shown in Fig. 15a, it becomes apparent that the original, undenoised micro-seismic signals differ from those denoised using four distinct methods by 12.18 s, 0.16 s, 0.2 s, 0.15 s, and 0.03 s concerning the first break pickup results of P-wave, as determined through the STA/LTA method and expert manual picking, respectively. Meanwhile, the results in Fig. 15b indicate discrepancies of 0.46 s, 0.09 s, 0.07 s, 0.05 s, and 0.03 s in the initial P-wave arrival time. These findings underscore that the application of the MSDCAN denoising model significantly elevates the quality of micro-seismic signals and diminishes errors in P-wave arrival time pickup.

### Application of micro-seismic denoising in Shanxi mines

The MSDCAN model was employed to denoise micro-seismic events recorded in a Shanxi mine in 2020. The mine encompassed a coal seam with an inclination ranging from 0° to 8° and a thickness varying between 2.58 m to 2.62 m, with an average thickness of 2.6 m. The coal seam was located beneath a top slab surrounded by mudstone, which is categorized as a soft rock formation. High-precision micro-seismic sensors were deployed at a burial depth of 25–50 m, and the data were sampled at a frequency of 1 kHz. A total of 346 micro-seismic data instances were obtained for analysis. Prior to denoising, the micro-seismic data underwent resampling and preprocessing procedures based on the methods outlined in section 3.1. The denoising outcomes achieved using the MSDCAN model, as well as the other two methods, are presented in Table 10. Notably, the application of the MSDCAN model yielded substantial enhancements in signal quality. The MSDCAN model outperformed the other methods in terms of the SNR, as evaluated by the PSNR index. This outcome signifies that the MSDCAN model effectively preserves micro-seismic information to a greater extent, resulting in superior micro-seismic signal quality.

Figure 16 illustrates the comparative results of time-domain and spectral characteristics for various denoising algorithms applied to micro-seismic signals post-denoising. Fig. 16 a and c depicts time-domain plots, presenting time-amplitude coordinates, while Fig. 16 b and d display spectral plots using frequency amplitude coordinates. Upon examining the experimental outcomes, it becomes evident that DeepDenoiser denoising introduces significant distortion in the time-domain features, along with spectral aliasing in the spectral characteristics. In terms of time-domain characteristics, CNN-denoiser demonstrates some denoising capabilities; however, it still exhibits some residual noise, with varying levels of effectiveness against different noise levels. Spectrally, it is observed that certain high-frequency components of the signal vary in the range of 10–25 Hz, lacking distinct primary frequency features, and thus, its practicality is compromised. Neighbor2Neighbor, on the other hand, proves effective in denoising micro-seismic signals containing noise, as indicated by the time-domain features. However, the peak features in the spectrum are not as pronounced. In contrast, MSDCAN excels in the removal of different types of noise, maximizing the restoration of micro-seismic information. Spectrally, there is an absence of spectral aliasing, and its denoising capability remains unaffected by varying noise profiles. It effectively suppresses high-frequency components in the ranges of 0–10 Hz and 10–25 Hz, showcasing well-defined peak features in the



**Fig. 15** Results comparison of the MSDCAN algorithm with the DeepDenoiser method, the CNN-denoiser method, and the Neighbor-2Neighbor method. **a–b** Zoomed-in plots of P-wave arrival times for the 2 examples in Fig. 14. For each subfigure, (i) denotes the noisy signal, (ii) denotes the denoised signal corresponding to the Deep-

denoiser method, (iii) is the denoised signal corresponding to the CNN-denoiser method, (iv) is the denoised signal corresponding to the Neighbor2Neighbor method, and (v) is the denoised signal corresponding to the MSDCAN method

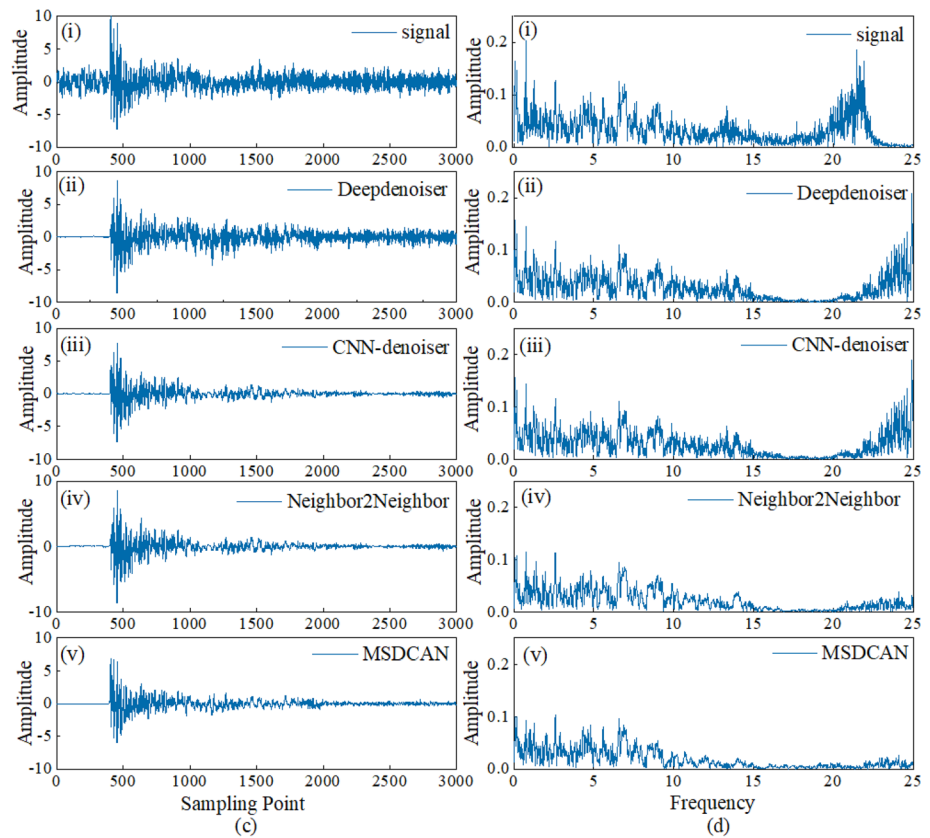
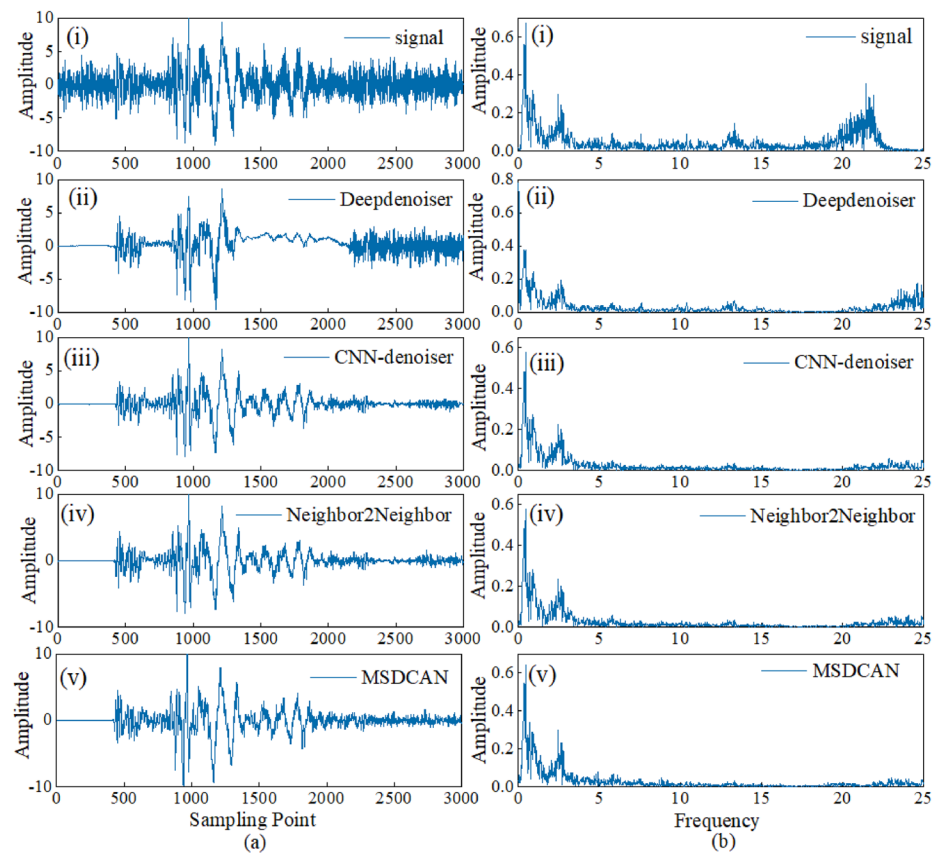
**Table 10** Average PSNR for different types

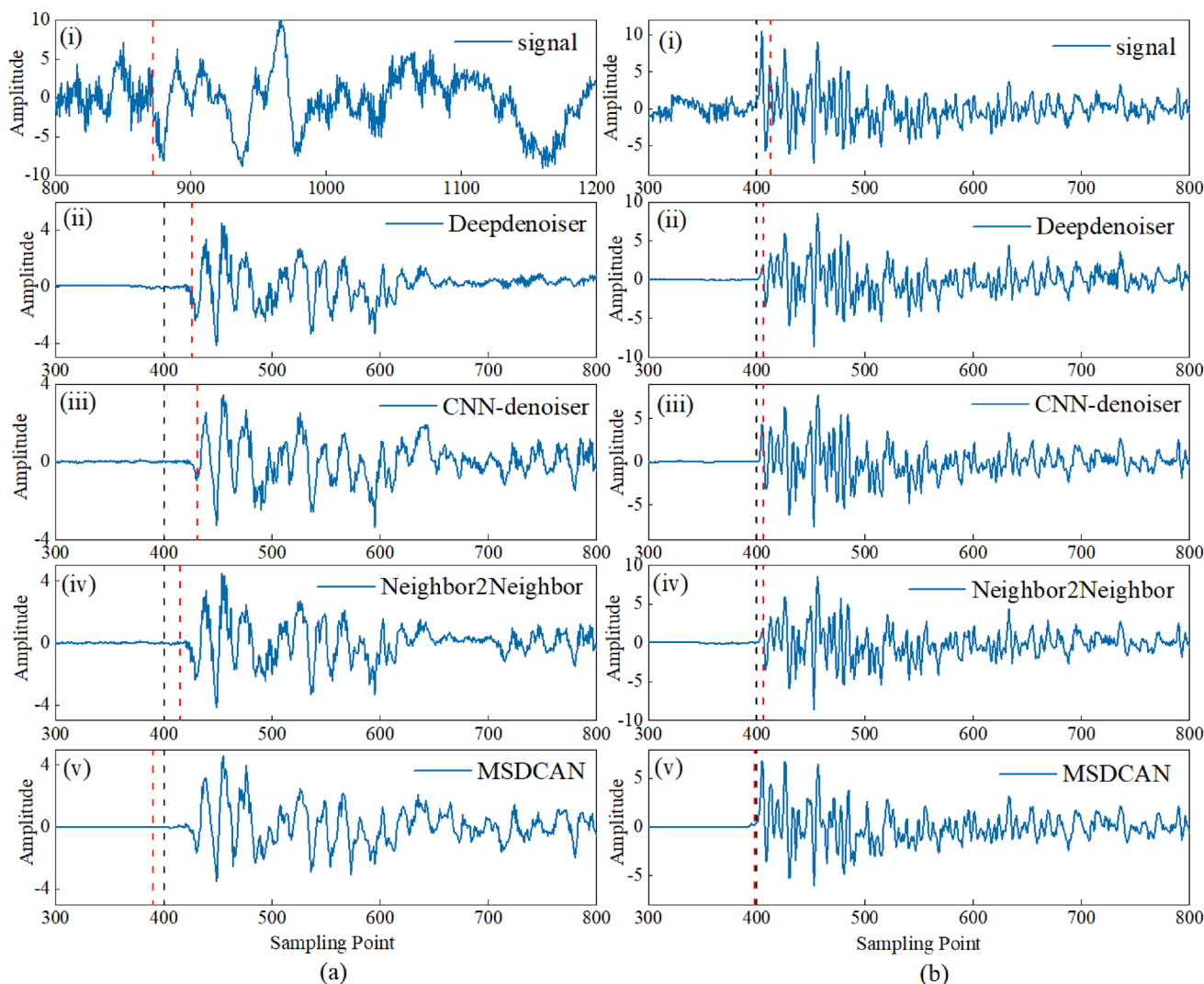
Type	Average PSNR
Original signal	2.105
DeepDenoiser	7.899
CNN-denoiser	13.442
Neighbor2Neighbor	18.617
MSDCAN	29.269

spectrum. Overall, it proficiently extracts time-frequency domain information from authentic micro-seismic signals.

Figure 17a and b depict the picking outcomes for two micro-seismic signals featured in Fig. 17. The dashed lines, one black and the other red correspond to the results of expert manual picking and STA/LTA automatic picking, respectively. In Fig. 17a, it is evident that there exists a difference of 4.72 s, 0.26 s, 0.31 s, 0.14 s, and 0.08 s between the initial micro-seismic signals without denoising and those subjected to denoising through three different methods, when compared to the P-wave first break picking results obtained using both the STA/LTA method and the expert manual method, respectively. On the other hand, the findings presented in Fig. 17b indicate variations in the initial P-wave arrival time of 0.13 s, 0.06 s, 0.06 s, and 0.02 s, respectively. These observations point to the notable enhancement in the

**Fig. 16** Results comparison of different denoising methods on micro-seismic signals from Shanxi mines. **a** and **c** are the time-domain figures of denoised signals corresponding to different methods; **b** and **d** are the spectral figures of denoised signals corresponding to different methods. For each subfigure, (i) denotes the noise signal, (ii) is the denoised signal corresponding to the DeepDenoiser method, (iii) denotes the denoised signal corresponding to the CNN-denoiser method, (iv) represents the denoiser signal corresponding to the Neighbor2Neighbor method, and (v) represents the denoised signal corresponding to the MSDCAN method





**Fig. 17** Results comparison of the MSDCAN method with the Deep-Denoiser method, the CNN-denoiser method, and the Neighbor-2Neighbor method. **a–b** Zoomed-in plots of P-wave arrival times for the 2 examples in Fig. 16. For each subfigure, (i) represents the noisy signal, (ii) represents the denoised signal corresponding to the Deep-

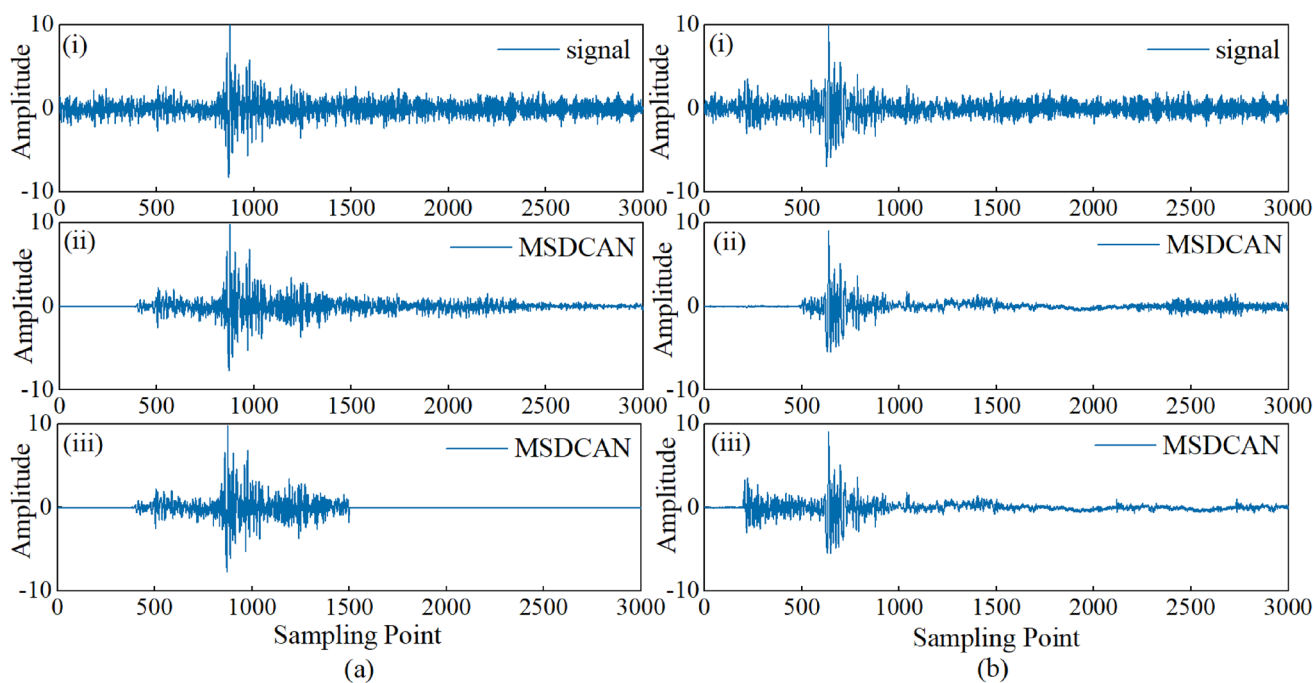
denoiser method, (iii) is the denoised signal corresponding to the CNN-denoiser method, (iv) is the denoised signal corresponding to the Neighbor2Neighbor method, and (v) is the denoised signal corresponding to the MSDCAN method

quality of micro-seismic signals after undergoing MSDCAN denoising, which concurrently leads to a reduction in errors associated with P-wave arrival time picking. The MSDCAN denoising model exerts a substantial influence in the domain of denoising.

**Drawback of MSDCAN denoising model**

The MSDCAN denoising model exhibits exemplary performance when tested on micro-seismic signal datasets with varying SNRs and noise types. However, it is noteworthy that issues may arise due to the selection of expansion factors for the DMS and DC blocks within the MSDCAN denoising model. Figure 18 presents two denoising

scenarios, where improper expansion factor choices result in subpar denoising outcomes. From Fig. 18a, we observe some residual noise occurring before the arrival of the P-wave. To address this issue, we increased the expansion factor using the DC block by a factor of  $s = 5$  to eliminate the remaining noise. This adjustment is depicted in the third row of Fig. 18a. Conversely, Fig. 18b illustrates a situation where the DMS block missed the arrival time of the P-wave during its operation. In such cases, we must reduce the hole factor in the DMS block, which means that the expansion factors for the four branches become 1, 4, 8, and 10. The results of this modification are showcased in the third row of Fig. 18b, where the P-wave signal is accurately restored. While the current MSDCAN denoising model succeeds in eliminating



**Fig. 18** Two examples of inaccurate expansion factors. **a** Residual noise samples. **b** Missing the P-wave samples. For each subgraph, (i) represents noisy data, (ii) represents denoised signals corresponding

to the MSDCAN algorithm, and (iii) represents denoised signals corresponding to appropriate expansion factors

the majority of test signals, further enhancements are necessary to attain a flawless model structure.

Conversely, in the realm of supervised learning, the quantity of data plays a pivotal role, and the availability of extensive training samples is indispensable for optimal model performance. Nevertheless, the acquisition of pristine data for training can prove challenging when dealing with micro-seismic data due to the presence of noise. This impediment often hinders the gathering of sufficiently clean data to support large-scale supervised deep learning endeavors. To address this predicament, unsupervised learning methods come into play, leveraging unlabeled samples for model training. This approach reduces the reliance on pristine micro-seismic data and mitigates the issue of a pronounced decline in denoising efficacy stemming from the scarcity of clean micro-seismic data. In our forthcoming research, we envision expanding our current dataset to encompass genuine micro-seismic data to enhance the efficiency of the micro-seismic signal denoising model. However, this undertaking will necessitate the adaptation of the existing MSDCAN architecture to accommodate the training requirements of the new dataset.

## Conclusions

In order to reduce noise interference in micro-seismic signals and improve data quality and accuracy, the study Multi-scale Dilated Convolutional Attention Network (MSDCAN) denoising model. The denoising performance of the MSDCAN denoising model was studied through model design, simulation, and experiments. The results indicate that:

- (1) The MSDCAN denoising model exhibits strong denoising ability in processing random noise in micro-seismic signals. It skillfully and dynamically removes various types of noise commonly found in these signals, thereby significantly improving the signal-to-noise ratio (SNR) and overall quality of micro-seismic signals.
- (2) The layer optimization experiment shows that in considering the neural network model architecture for denoising, MSDCAN-CA3 is an optimal choice because its performance is comparable to MSDCAN-CA4 in reducing model complexity.
- (3) The results of ablation experiments indicate that multiscale dilated convolution (DMS), improved attention mechanism (improved SE), and initial dilated convolution (DC) all play key roles in the denoising effect of the MSDCAN model. Removing any DMS, improved SE or DC blocks will significantly reduce the net-



work's denoising ability. In addition, the improved SE significantly improves the denoising performance of the MSDCAN model.

- (4) The MSDCAN denoising model is the most effective denoising method, surpassing DeepDenoyer, CNN-denoiser, and Neighbor2Neighbor denoising methods at various SNR levels. It is worth noting that even if SNR is improved, its denoising performance remains consistent.
- (5) The MSDCAN denoising model excels in extracting micro-seismic signal features from background noise, impulse noise, periodic noise, and mixed noise, comprehensively preserving micro-seismic events. Compared with DeepDenoyer, CNN-denoiser, and Neighbor2Neighbor methods, it exhibits superior denoising performance in terms of SNR, root mean square error, and correlation coefficient.
- (6) The successful deployment of the MSDCAN denoising model in real scenarios of Beijing micro-seismic data and Shanxi mining project micro-seismic signal denoising has confirmed its better application effect. In these practical applications, it always exhibits robust denoising performance.

**Author Contributions** Authors' contributions writing—original draft preparation, was contributed by JC and ZD; writing—review and editing, was contributed by LW; supervision was contributed by JM; funding acquisition was contributed by ZY. All authors have read and agreed to the published version of the manuscript.

**Funding** This research was funded by the Open Foundation of the Key Laboratory of Seismic Hazard Instrumentation and Detection Technology of Hebei Province (Grant No. FZ224105); Langfang Science and Technology Bureau, No.2022011037; the Fundamental Research Funds for the Central Universities (Grant No. ZY20215101); Hebei Graduate Innovation Funding Project.

## Declarations

**Conflict of interest** The authors declare no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Ahmad BA, Takeshi T (2021) Machine learning for automatic slump identification from 3D seismic data at convergent plate margins. *Marine Pet Geol.* <https://doi.org/10.1016/j.marpetgeo.2021.105290>
- Cai J, Wang L, Zheng J et al (2023) Denoising method for seismic co-band noise based on a u-net network combined with a residual dense block. *Appl Sci* 13(3):1324. <https://doi.org/10.3390/app13031324>
- Chen LC, Papandreou G, Kokkinos I et al (2017) Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Transact Pattern Anal Mach Intell* 40(4):834–848. <https://doi.org/10.1109/TPAMI.2017.2699184>
- Dong X, Zhong T, Li Y (2020) A deep-learning-based denoising method for multiarea surface seismic data. *IEEE Geosci Remote Sens Lett* 18(5):925–929. <https://doi.org/10.1109/LGRS.2020.2989450>
- Dong X, Lin J, Lu S et al (2022) 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. *Surv Geophys* 43(5):1363–1394. <https://doi.org/10.1007/s10712-022-09702-7>
- Duan R, Chen Z, Zhang H et al (2023) Dual residual denoising autoencoder with channel attention mechanism for modulation of signals. *Sensors* 23(2):1023. <https://doi.org/10.3390/s23021023>
- Glorot X, Bordes A, Bengio Y (2011) Deep sparse rectifier neural networks. In: Proceedings of the fourteenth international conference on artificial intelligence and statistics, JMLR Workshop and conference proceedings, pp 315–323, proceedings.mlr.press/v15/glorot11a.html
- Ioffe S, Szegedy C (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. In: International conference on machine learning, PMLR, pp 448–456, 10.48550/arXiv.1502.03167
- Karniadakis GE, Kevrekidis IG, Lu L et al (2021) Physics-informed machine learning. *Nat Rev Phys* 3(6):422–440. <https://doi.org/10.1038/s42254-021-00314-5>
- Khan M, Xueqiu H, Dazhao S et al (2023) Extracting and predicting rock mechanical behavior based on microseismic spatio-temporal response in an ultra-thick coal seam mine. *Rock Mech Rock Eng* 56(5):3725–3754. <https://doi.org/10.1007/s00603-023-03247-w>
- Kingma DP, Ba J (2014) Adam: a method for stochastic optimization. *arXiv preprint arXiv:1412.6980* 10.48550/arXiv.1412.6980
- Lin H, Xing L, Li Q et al (2022) Spatial-domain synchrosqueezing wavelet transform and its application to seismic ground roll suppression. *IEEE Transact Geosci Remote Sens* 60:1–16. <https://doi.org/10.1109/TGRS.2022.3210606>
- Liu S, Zhou Z, Peng S et al (2022) Improving the resolution of seismic data based on s-transform and modified variational mode decomposition, an application to songliao basin, northeast china. *Acta Geophysica* 70(3):1103–1113. <https://doi.org/10.1007/s11600-022-00781-z>
- Liu W, Liu Y, Li S et al (2023) A review of variational mode decomposition in seismic data analysis. *Surv Geophys* 44(2):323–355. <https://doi.org/10.1109/TGRS.2022.3152984>
- Ma Q, Tian G, Zeng Y et al (2021) Pipeline in-line inspection method, instrumentation and data management. *Sensors.* <https://doi.org/10.3390/s21113862>
- Mao X (2022) A concentrated time-frequency method for reservoir detection using adaptive synchrosqueezing transform. *IEEE Geosci Remote Sens Lett* 19:1–5. <https://doi.org/10.1109/LGRS.2022.3160930>

- Mousavi SM, Sheng Y, Zhu W et al (2019) Stanford earthquake dataset (stead): a global data set of seismic signals for AI. *IEEE Access* 7:179464–179476. <https://doi.org/10.1109/ACCESS.2019.2947848>
- Mumuni A, Mumuni F (2022) Robust appearance modeling for object detection and tracking: a survey of deep learning approaches. *Progr Artif Intell* 11(4):279–313. <https://doi.org/10.1007/s13748-022-00290-6>
- Muther T, Dahaghi AK, Syed FI et al (2023) Physical laws meet machine intelligence: current developments and future directions. *Artif Intell Rev* 56(7):6947–7013. <https://doi.org/10.1007/s10462-022-10329-8>
- Qi P, Wang Y (2020) Seismic time-frequency spectrum analysis based on local polynomial fourier transform. *Acta Geophys* 68:1–17. <https://doi.org/10.1007/s11600-019-00377-0>
- Saad OM, Chen Y, Savvaadis A et al (2022) Unsupervised deep learning for single-channel earthquake data denoising and its applications in event detection and fully automatic location. *IEEE Transact Geosci Remote Sens* 60:1–10. <https://doi.org/10.1109/TGRS.2022.3209932>
- Shi H, Chen J, Si J et al (2020) Fault diagnosis of rolling bearings based on a residual dilated pyramid network and full convolutional denoising autoencoder. *Sensors* 20(20):5734. <https://doi.org/10.3390/s20205734>
- Song Q, Wang M, Lai W et al (2022) Multiscale kernel-based residual CNN for estimation of inter-turn short circuit fault in PMSM. *Sensors* 22(18):6870. <https://doi.org/10.3390/s22186870>
- Tang S, Wang J, Tang C (2021) Identification of microseismic events in rock engineering by a convolutional neural network combined with an attention mechanism. *Rock Mech Rock Eng* 54:47–69. <https://doi.org/10.1007/s00603-020-02259-0>
- Tarek S. I, Tatsunori I, Takeshi T, et al (2021) Extracting high-resolution p-wave reflectivity of the shallow subsurface by seismic interferometry based on autocorrelation of blast mining signals. *Geophys Prospect* pp 1267–1280. 10.1111/1365-2478.13308
- Wang D, Chen G, Chen J et al (2023) Seismic data denoising using a self-supervised deep learning network. *Math Geosci*. <https://doi.org/10.1007/s11004-023-10089-3>
- Wang Z, Li H, Liu S et al (2022) Prediction method of CO<sub>2</sub> injectivity in saline aquifer based on bp neural network. *J Central South Univ Sci Technol* 53(12):4678–4686
- Xu H, Lai X, Shan P et al (2023) Energy dissimulation characteristics and shock mechanism of coal-rock mass induced in steeply-inclined mining: Comparison based on physical simulation and numerical calculation. *Acta Geotechnica* 18(2):843–864. <https://doi.org/10.1007/s11440-022-01617-2>
- Yang Y, Lv H, Chen N (2023) A survey on ensemble learning under the era of deep learning. *Artif Intell Rev* 56(6):5545–5589. <https://doi.org/10.1007/s10462-022-10283-5>
- Yuan Y, Li Y, Zhou S (2020) Multichannel statistical broadband wavelet deconvolution for improving resolution of seismic signals. *IEEE Transact Geosci Remote Sens* 59(2):1772–1783. <https://doi.org/10.1109/TGRS.2020.2997977>
- Zhang H, Ma C, Pazzi V et al (2020) Microseismic signal denoising and separation based on fully convolutional encoder-decoder network. *Appl Sci* 10(18):6621. <https://doi.org/10.3390/app10186621>
- Zhu W, Mousavi SM, Beroza GC (2019) Seismic signal denoising and decomposition using deep neural networks. *IEEE Transact Geosci Remote Sens* 57(11):9476–9488. <https://doi.org/10.1109/TGRS.2019.2926772>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.