applied sciences

MDPI

Article

Denoising Method for Seismic Co-Band Noise Based on a U-Net Network Combined with a Residual Dense Block

Jianxian Cai, Li Wang *, Jiangshan Zheng, Zhijun Duan, Ling Li and Ning Chen

Hebei Key Laboratory of Seismic Disaster Instrument and Monitoring Technology, Institute of Disaster Prevention, Langfang 065201, China
* Correspondence: wangli_8701@163.com

Abstract: To address the problem of waveform distortion in the existing seismic signal denoising method when removing co-band noise, further improving the signal-to-noise ratio (SNR) of seismic signals and enhancing their quality, this paper designs a seismic co-band denoising model Atrous Residual Dense Block U-Net (ARDU), which uses a U-shaped convolutional neural network (U-Net) as a basic framework and combines atrous convolution and the residual dense block (RDB). In the ARDU model, atrous convolution is connected with residual dense blocks to form the feature extraction unit of the model encoder. Among them, the residual dense blocks can deepen the network's depth and enhance the feature extraction ability of the network on the premise of mitigating the gradient-vanishing and gradient-exploding problem. Atrous convolution can enlarge receptive fields, reduce waveform distortion, and protect effective signals without increasing network parameters. To test the denoising performance of the ARDU model, the Stanford Global Seismic dataset was used to construct a training set and a test set and the model was trained and tested on it. The experimental results of the ARDU model for different types of seismic co-band noise showed that this model can effectively remove seismic co-band noise, protect effective signals, improve the SNR of seismic signals, and enhance the quality of seismic signals. To further verify the denoising effect of the model, this model was compared with the wavelet threshold denoising U-Net model and the denoising residual dense block (DnRDB) model, and the results showed that the ARDU model has the best SNR, *r* (correlation coefficient), and root-mean-square error (RMSE) and the least distortion of the seismic signal waveform.

Keywords: denoising; co-band noise; residual dense block; atrous convolution; waveform distortion

check for updates

Citation: Cai, J.; Wang, L.; Zheng, J.; Duan, Z.; Li, L.; Chen, N. Denoising Method for Seismic Co-Band Noise Based on a U-Net Network Combined with a Residual Dense Block. *Appl. Sci.* **2023**, *13*, 1324. https://doi.org/10.3390/app13031324

Academic Editor: Jong Wan Hu

Received: 16 November 2022
Revised: 12 January 2023
Accepted: 14 January 2023
Published: 19 January 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Earthquakes not only threaten people's lives but also cause huge economic losses. Earthquake early warning, as an effective means to reduce earthquake damage, is mainly achieved by monitoring earthquake signals. High-quality earthquake signals can ensure their accuracy for early warning. However, seismic signals are easily contaminated by various random noises in the process of generation, transmission, and acquisition, especially when the random noises and seismic signals share the same frequency band. Co-frequency band noise seriously affects the quality of seismic signals and reduces the accuracy of earthquake early warning. Therefore, it is necessary to denoise seismic signals, especially to remove the co-frequency band noise in them [1].

Denoising methods for seismic co-frequency band noise mainly include traditional methods and methods based on deep learning [2]. Traditional methods mainly include Fourier transform [3–5], wavelet threshold denoising [6–13], and empirical mode decomposition (EMD) [14]. Fourier transform [3–5] removes noise by analyzing the difference of frequency components between noise and signal, but this method can only obtain the frequency components of the signal as a whole and cannot accurately predict the time when each frequency component appears; as a result, signals with a great difference in the time domain may have the same spectrogram, which limits dealing with non-stationary signals.

Wavelet threshold denoising [6–12] transforms signals from the time domain to the time-frequency domain through wavelet transform, which is suitable for processing non-stationary signals and has a reasonable resolution in both time and frequency domains. However, this method needs to select the number of wavelet decomposition layers and also select appropriate thresholds to quantify the wavelet coefficients of each layer after decomposition, resulting in an algorithm that relies on threshold selection and limited noise reduction.

EMD is a self-adaptive signal analysis method proposed by Huang et al. [14] in 1998. This method can decompose any type of signal into a finite number of intrinsic mode functions (IMFs) in theory, which is suitable for the analysis of non-linear and non-stationary signal sequences. In the process of denoising, a seismic signal with noise is decomposed into a limited number of IMFs and a clean seismic signal is obtained by removing the noise's IMFs. EMD is mainly based on the timescale characteristics of the signal itself and does not need to set the basic function or threshold in advance, but it also has some problems, such as mode aliasing and difficulty in selecting an IMF. To sum up, traditional seismic-signal-denoising methods require more manual participation and also need to select different threshold functions or basis functions for different types of noise, which leads to problems, such as difficulty in guaranteeing the accuracy of denoising and high computational cost.

In recent years, deep learning technology has continued to evolve, and it has achieved excellent results in image classification, image noise reduction, speech noise reduction, and other fields [15–22] with its self-adaptive feature learning and strong classification ability. Therefore, geophysicists began to use deep learning technology to self-adaptively reduce the noise in the co-frequency band in seismic signals. In 2018, Yu et al. [23] used a convolutional neural network (CNN) to denoise seismic signals. However, the co-band noise in a complex environment has the characteristics of no fixed waveform, non-stationary feature, high energy, and the same frequency band range as effective signals, which leads to poor denoising ability of the common CNN for co-band noise, serious signal waveform distortion after denoising, and a low SNR. In 2018, Jin et al. [24] adopted the residual learning strategy; this method used noise as the label during model training to make the model self-adaptively learn the noise in the signals and finally achieved the purpose of denoising by removing the noise recognized by the model. Although this method mitigates the problems of gradient vanishing and gradient exploding caused by deepening the network, the features of noise are more complex than seismic signals and the denoising ability of the model is limited, resulting in serious waveform distortion and a low SNR after denoising. In 2019, Zhu et al. [25] used the symmetric encoder–decoder structure of the U-Net model to denoise seismic signals. The model has a strong ability of feature extraction, which can extract seismic signal feature maps of different scales and fuse the signal feature maps extracted by the encoder with the decoder through feature concatenating [26]. However, the down-sampling process of the model encoder is realized by reducing the feature map of seismic signals, which inevitably has the problem of losing the details of the seismic signals. In 2021, Gao et al. [27] designed the denoising residual dense block (DnRDB) model. This model deepens the depth of the network using dense skip connections to enhance the multiplexing of signal features and further mitigate the problem of signal detail loss during the denoising process of seismic signals. However, the size of the input signal feature map of the model is single, especially when removing non-stationary, high-energy co-band noise without a fixed waveform, and the signal feature that the model can learn is limited, resulting in waveform distortion of the denoised seismic signals.

In summary, to mitigate the waveform distortion of seismic signals after denoising, further improve the SNR of seismic signals, and improve the quality of seismic signals, this paper takes the U-Net as the basic framework, combines atrous convolution and residual dense blocks, and designs a seismic co-band denoising model ARDU. The ARDU model connects atrous convolution and residual dense blocks to form a feature extraction unit of the model encoder. This feature extraction unit can deepen the network's depth, expand

the convolution receptive field, and obtain more abundant signal features, thereby reducing waveform distortion, protecting effective signals, and improving the quality of seismic signals. The model was trained and tested on the Stanford Global Seismic dataset, and the results showed that the ARDU model can effectively remove seismic noise in the co-frequency band, protect effective signals, improve the SNR of seismic signals, and enhance the quality of seismic signals. Compared with wavelet threshold denoising, the U-Net model, and the DnRDB model, the ARDU model has the best overall denoising effect and the least waveform distortion of seismic signals.

2. The ARDU Denoising Model

Based on the U-Net, the feature extraction unit of the model encoder is composed of atrous convolution and residual dense blocks. The designed ARDU model consists of an input layer, a middle layer, and an output layer. The structure is shown in Figure 1. As shown in this figure, the input to the ARDU model is a seismic signal with noise and the output is the seismic signal after denoising. The red dashed-line box contains the encoder part of the middle layer, and the green dashed-line box contains the decoder part of the middle layer. The ARDU model first uses the input layer to extract the preliminary features of the seismic signal with noise; then extracts the signal feature maps of different scales through the encoder; restores the seismic signal feature maps through the decoder, meanwhile concatenating the feature maps obtained by the encoder and the decoder; and finally performs denoising.

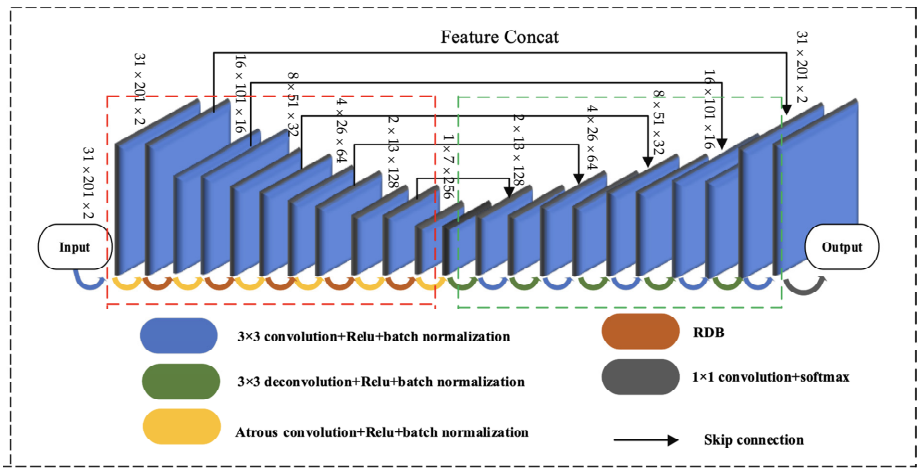


Figure 1. ARDU model structure diagram.

2.1. Input Layer of the ARDU Model

In the time-frequency domain, the seismic signal with noise can be expressed as:

$$Y(t,f) = X(t,f) + N(t,f) \tag{1}$$

In Formula (1), $Y(t,f)$ represents the seismic signal with noise, $X(t,f)$ represents the clean seismic signal, and $N(t,f)$ represents noise.

The input layer of the ARDU model consists of a convolution layer with a 3×3 convolution kernel, a rectified linear unit (ReLU) activation function, and batch normalization (BN), whose main function is to perform preliminary signal feature extraction of the seismic signal with noise, $Y(t,f)$. The ReLU activation function is shown in Formula (2), and its function image is shown in Figure 2. Here, x represents the seismic signal features extracted from the convolution layer. The ReLU activation function is simple in form, it only needs to judge whether the input is greater than zero, and its calculation speed is fast. It can mitigate the gradient-vanishing problem in the deep network and enable the deeper network to be better trained, thus fully maximizing the performance of the network. The main purpose

of batch normalization is to transform the input data into a normal distribution, mitigate gradient vanishing, and speed up network training.

$$f(x) = \max(0, x) = \begin{cases} x & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (2)$$

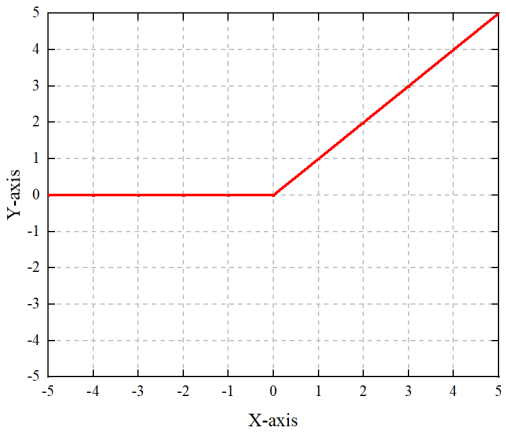


Figure 2. ReLU function.

2.2. Middle Layer of the ARDU Model: The Role of Atrous Convolution

The middle layer of the ARDU model is a symmetrical U-shaped structure, which is composed of an encoder, a decoder, and feature concatenating.

The main function of the encoder is to extract seismic signal features through down-sampling. It is composed of five groups of feature extraction units and the atrous convolution of the bottom layer. Each group of feature extraction units is composed of atrous convolution, an activation function, batch normalization, and a residual dense block, in which the activation function and batch normalization are the same as those of the input layer.

The down-sampling of the encoder is realized by reducing the characteristic image of the seismic signal, which inevitably loses the detailed information about signal characteristics. The atrous convolution can expand the receptive field of the convolution layer without increasing the network parameters to obtain more feature information about the seismic signals [28]. In the ARDU model, after many experimental results, the denoising effect is the best when the dilated coefficient of the atrous convolution is set to 2, so the dilated coefficient of the atrous convolution used by the model encoder was set to 2. The atrous convolution of the bottom layer is also used to expand the receptive field of the convolution kernel, enhance the neighborhood information learning of the seismic signal feature map, and contact the signal context information so as to better restore the structural feature of the seismic signal. After atrous convolution, batch normalization and the activation function are also added, and their effect is consistent with the input layer of the model.

The implementation principle of atrous convolution is shown in Figure 3. Figure 3a shows a common convolution with a receptive field size of 3×3. In Figure 3b, atrous convolution with a dilated coefficient r of 2 is added on the basis of the common convolution with a 3×3 receptive field, and the field of the convolution kernel is increased to 5×5. Let the convolution kernel size of the convolution layer be $k \times k$; when the dilated coefficient of the atrous convolution is r, the calculation form of its receptive field size v is shown in Formula (3).

$$v = [(r-1) \times (k-1) + k] \times [(r-1) \times (k-1) + k] \quad (3)$$

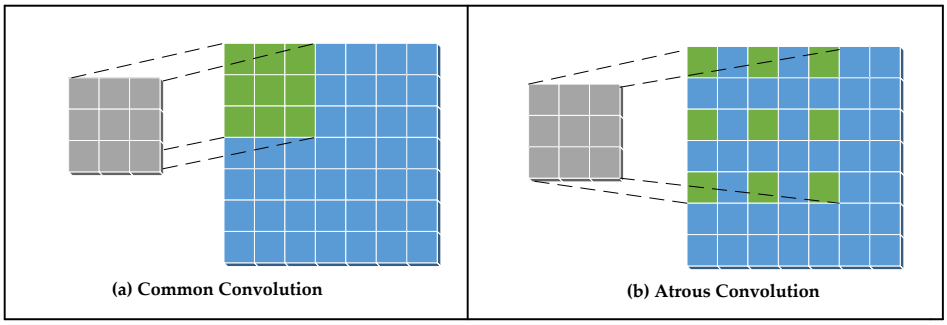


Figure 3. Different convolution operations.

The residual dense block is mainly composed of dense connections and global residual learning, and the structure is shown in Figure 4. The dense skip connection connects each current convolutional layer to all convolutional layers, realizing the interconnection between all layers, as shown by the blue arrows in Figure 4. This method deepens the depth and width of the network to a certain extent, integrates the information about all layers as far as possible, forms a continuous memory mechanism, and enhances the feature extraction ability of the model. The global residual learning gives the ability of backoff to the residual dense block by adding a directly connected skip connection between input and output, as shown by the red arrow in Figure 4. In this way, the residual learning can automatically select whether to perform feature extraction through the three convolution layers according to the effect of gradient descent, directly skip the convolution layer and select the skip connection to perform feature flow, or combine the features of the three convolution layers and the skip connection to perform output, thereby mitigating the problems of gradient vanishing and gradient exploding caused by deepening of the number of network layers and making the network easier to train. For each unit of the residual dense block, as shown in the green dashed-line box in Figure 4, batch normalization and the activation function are added after the convolution layer so that the input of each layer has a stable distribution, which can effectively avoid gradient diffusion and improve the generalization ability and convergence speed of the network.

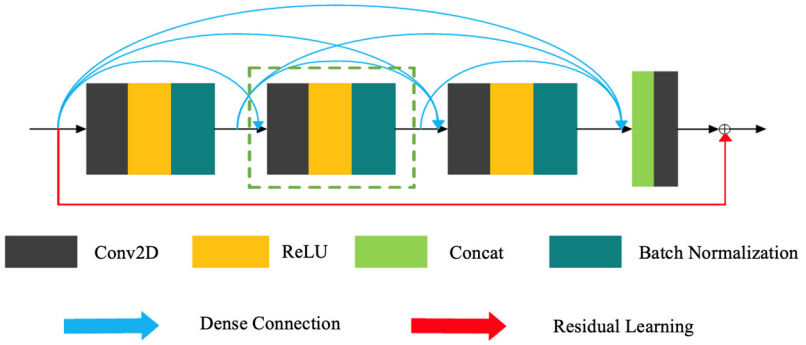


Figure 4. Structure of a residual dense block.

The main function of the decoder is to restore the feature information and distribution law of the seismic signal through up-sampling. It consists of five groups of feature restoration units. Each group of feature restoration units includes convolution and deconvolution. After convolution and deconvolution, activation functions and batch normalization are added, whose functions are the same as those of the input layer.

In a common CNN, the input of the convolutional layer only comes from the output of the previous layer and the feature map has the same size. The feature map of a single scale cannot fully describe the feature information and distribution law of the signal. The network maps out deeper features through convolution and activation, and the features expressed by layers of different depths are different. Therefore, the ARDU model performs multi-scale

signal feature concatenating between the encoder and the decoder and fuses the signal feature map extracted by the encoder with the decoder to generate a new signal feature map. This improves the learning ability of the ARDU model for seismic signal features.

2.3. Output Layer of the ARDU Model

The output layer of the ARDU model consists of convolution with a 1×1 convolution kernel and the softmax function. The output layer uses convolution to adjust the number of signal feature maps input by the middle layer and outputs the time-frequency mask $\widehat{M}(t, f)$ of the denoised signal with the softmax function.

As a supervised learning algorithm, the ARDU model is inspired by the mask [28–30] in the supervised speech denoising algorithm. In this paper, the ideal amplitude mask (IAM) is used as the training target of the network model to obtain the time-frequency relationship between the seismic signal with noise and the clean seismic signal. The predicted mask can be used to estimate the frequency band distribution of the noise and seismic signals more accurately, and it can effectively reduce the waveform distortion of seismic signals and significantly suppress noise. The mask used in this paper is shown in Formula (4), where the mask $M(t, f)$ represents the proportion of the clean seismic signal in the seismic signal with noise, ranging from 0 to 1.

$$M(t, f) = \frac{|X(t, f)|}{|X(t, f)| + |N(t, f)|} \tag{4}$$

The time-frequency mask $\widehat{M}(t, f)$ predicted by the model is calculated and passed through the noise-containing seismic signal $Y(t, f)$ to obtain a clean seismic signal, as shown in Formula (5).

$$\widehat{X}(t, f) = \widehat{M}(t, f)Y(t, f) \tag{5}$$

Finally, the time-domain seismic signal $\widehat{X}(t)$ is obtained with inverse short-time Fourier transform (STFT-1), as shown in Formula (6):

$$\widehat{X}(t) = STFT^{-1}\left\{\widehat{X}(t, f)\right\} \tag{6}$$

3. Data Preprocessing and Evaluation Metric

3.1. Data Preprocessing

Training and testing the ARDU model require a large amount of data, so the Stanford Global Seismic dataset *STEAD* (Stanford Earthquake Dataset) [30], which contains 1 million three-component seismic signals, was used in this paper. These 1 million three-component seismic signals come from 450,000 earthquakes from January 1984 to August 2018, providing reliable seismic signals and noise labels for supervised learning algorithms.

In this paper, 130,000 seismic signals and noise signals were selected from STEAD and the randomly selected data were cropped. The first 3–5 s of the arrival time of the P wave given by the dataset labels were randomly taken as the initial point of cropping, and the 30 s backward from that point were cropped because the 30 s include the main waveform of the event, and the signal sampling rate was set to 100 Hz. To prevent the model from overfitting, the cropped seismic and noise signals were normalized with Formula (7), and finally, the cropped seismic signals and noise were randomly synthesized into 130,000 seismic signals with a low SNR and noise, which were randomly divided into a training set, a validation set, and a test set according to the ratio of 6:2:2 for model training, validation, and testing. In Formula (7), v_i represents the amplitude value of the seismic event signal.

$$v_i = \frac{v_i}{\max|v_i|} \tag{7}$$

A CNN has a stronger feature extraction ability for 2D data, so to fully use the denoising performance of the model, short-time Fourier transform (STFT) was applied to

the synthetic seismic signals with noise to obtain 2D time-frequency domain signals. The STFT is shown in Formula (8), where $z(k)$ is the source signal at time k and $g^*(kT - mT)$ is the window function. In Formula (8), m represents the translation of the time window, n represents the number of sampling points in the time window, T represents the sampling period in the time domain, and F represents the sampling period in the frequency domain.

$$STFT_Z(m,n) = \sum_{k=-\infty}^{\infty} z(k)g^*(kT - mT)e^{-j2\pi(nF)k} \tag{8}$$

3.2. Evaluation Metrics

The signal-to-noise ratio (SNR), r (correlation coefficient), and root-mean-square error (RMSE) were used to objectively evaluate the denoising performance of the seismic signal denoising and ensure the reliability and accuracy of the denoising experiments. The calculation formulas of the three evaluation metrics are shown in Formulas (9)–(11).

$$\text{SNR} = 10\log_{10}(\frac{\sigma_s}{\sigma_n}) \tag{9}$$

$$r(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var[X]Var[Y]}} \tag{10}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(X'(i) - X(i))}{n}} \tag{11}$$

The SNR represents the ratio of the clean signal to the noise, in dB. The larger the SNR, the smaller the proportion of noise, where σ_s represents the energy of the signal and σ_n represents the energy of the noise. In addition, r represents the correlation degree between the denoised signal and the label signal, and the closer the value is to 1, the greater the overall correlation is between the denoised seismic signal and the seismic signal waveform with labels, where X represents the reference label signal, Y represents the denoised seismic signal processed by the model, $Cov(X,Y)$ represents the covariance of the reference label signal and the denoised signal, $Var[X]$ represents the variance of the reference label signal, and $Var[Y]$ represents the variance of the denoised signal. The RMSE describes the error of each sampling point between the model output signal and the label signal. The smaller the RMSE is, the more similar the denoised seismic signal is to the label signal. The waveform distortion degree after signal denoising can be described from the deviation of each sampling point. Here, $X'(i)$ represents the seismic signal that is output by the model fitting, $X(i)$ represents the reference label seismic signal, and n is the number of signal sampling points.

4. Experiments and Analysis

Based on the ARDU denoising model in Figure 1, the basic experimental flow framework is shown in Figure 5. The experimental process mainly included three steps:

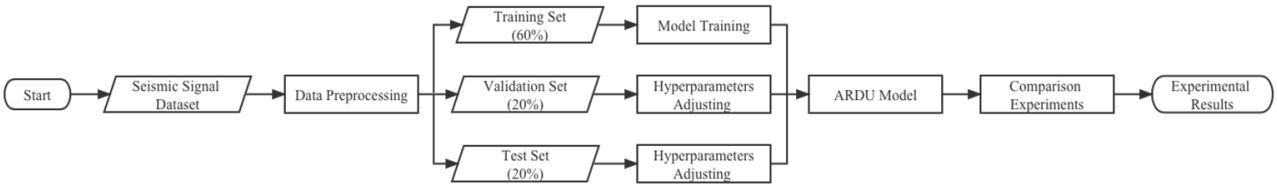


Figure 5. Basic experimental flow framework.

(1) Data preprocessing. The data were preprocessed with cropping, correction, translation, inversion, filtering, and signal and noise synthesis, and a training set, validation set, and test set were created.

(2) Model training, validation, and testing. The training set was used to train the model, learn the features of the seismic signal and noise, and fit the model parameters. The validation set was used to validate the model, adjust the hyperparameters of the model, and preliminarily evaluate the denoising performance of the model. The test set was used to test the denoising performance of the ARDU model and verify the feasibility of the denoising model.

(3) Comparison of denoising effects of different algorithms. The test set data were used to compare the denoising effect of the ARDU model and other denoising methods, and the SNR, *r*, and RMSE were calculated to evaluate the denoising performance of the model.

4.1. ARDU Model Training and Test Analysis

Model training used the Adam random optimization algorithm to optimize parameters, the default learning rate was 0.001, *batch_size* was set to 20, and NVIDIA GEFORCE RTX3090 was used to train, validate, and test the model. During training, the model used the feedback of the value of the cross-entropy loss function to continuously adjust the hyperparameters, such as the weight and bias of each layer, to further reduce the difference between the predicted value of the current network and the label value in order to minimize the difference. The cross-entropy loss function is shown in Formula (12), and the Adam algorithm was used as the optimizer for the iterative solution. The function of the Adam algorithm is to optimize the calculation process of the loss function to accelerate the convergence of the model. The decline curve of the training loss function is shown in Figure 6. About 65 batches (Epoch) were trained over 25 h, then the value of the loss function tended to be stable, and the model training was completed by default.

$$Loss = -\sum_{i=1}^{2}\sum_{j=1}^{n} M(t,f)\cdot \log(\widehat{M}(t,f)) \tag{12}$$

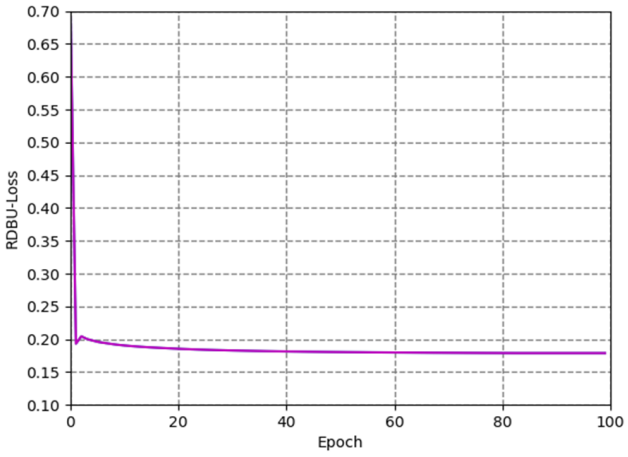
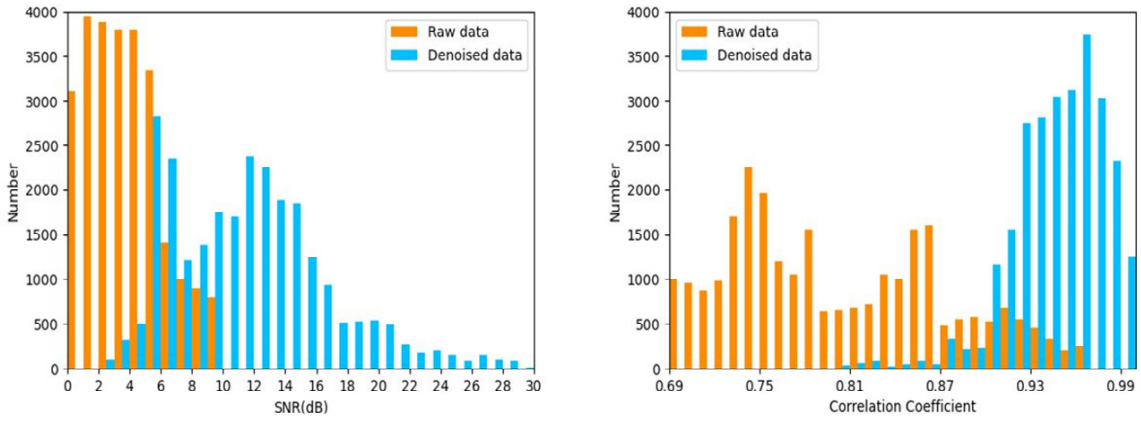


Figure 6. Decline curve of the loss function.

To preliminarily validate the feasibility of the denoising of the ARDU model, the model was tested on the prepared test set, and the test results are shown in Figure 7. The orange bars in the figure represent the seismic signal with noise before denoising, and the blue bars represent the seismic signal after denoising. According to the calculation in Figure 7a, the average SNR of the signal in the test set before denoising was 3.77 dB and the average SNR of the signal after denoising was 11.59 dB. The improvement in the SNR shows that the ARDU model can effectively remove the noise in seismic signals and enhance signal quality. According to Figure 7b, the correlation coefficient of the signal in the test set before

denoising was mainly in the interval [0.69, 0.87] and the average correlation coefficient value was 0.8024. After denoising with the ARDU model, the correlation coefficient interval is increased to [0.87, 0.99], the average correlation coefficient value is 0.9484, and the correlation coefficient is close to 1. This shows that the ARDU model can protect effective signals and reduce waveform distortion. To sum up, the ARDU model has the feasibility of denoising seismic noise in the co-frequency band and it can significantly improve the quality of seismic signals.

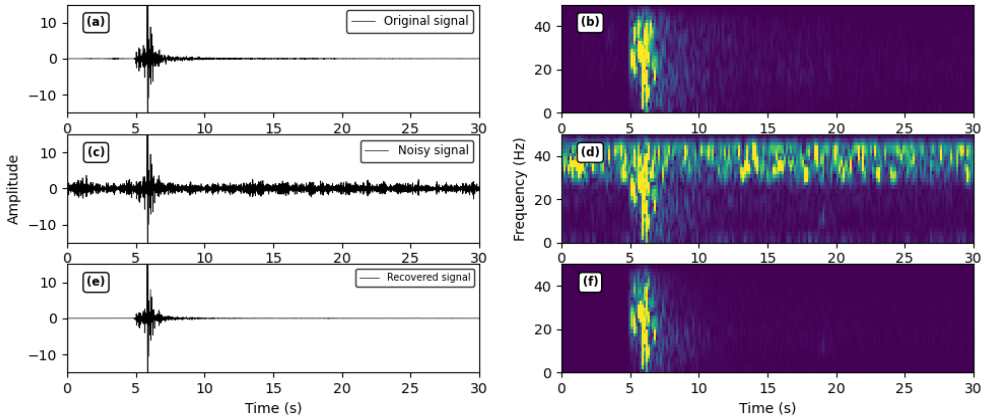


(a) Comparison of SNR Before and After Denoising **(b) Comparison of Correlation Coefficient Before and After Denoising**

Figure 7. Improvement of the SNR and correlation coefficient in test set.

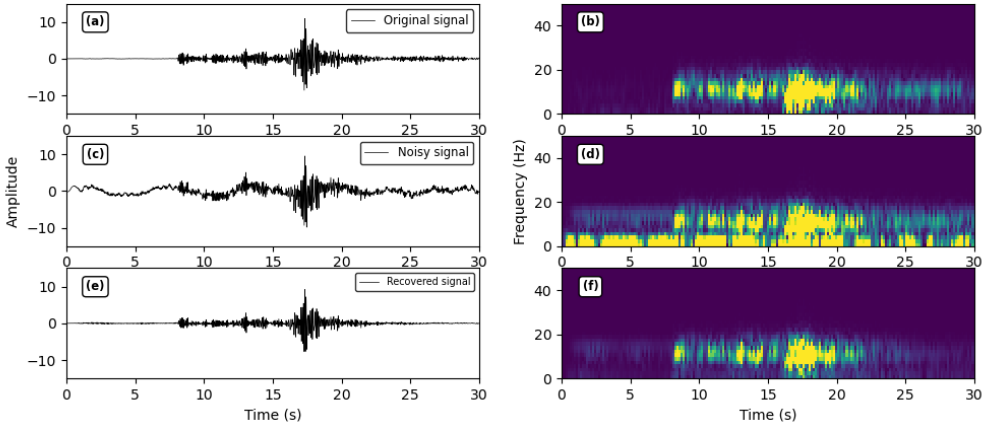
4.2. Denoising Experiments and Analysis of Different Types of Co-Band Noise

According to the frequency band range of seismic signals, co-frequency band noise was divided into medium-high-frequency noise, low-frequency noise, and mixed-frequency band noise, and the denoising performance of these three kinds of co-band noises was tested using the ARDU model to verify the effectiveness of the model for different noises. The test results are shown in Figures 8–11, where (a), (c), and (e), respectively, represent the time-domain seismic signals with labels, seismic signals with noise, and the denoised seismic signals, while (b), (d), and (f), respectively, represent the time-frequency domain signals of (a), (c), and (e). Figure 8 shows the medium-high-frequency noise. The frequency of the medium-high-frequency noise was relatively high, and the overlapping range with the frequency band of the seismic signals was small. After denoising, the SNR of the signals increased by 12.45 dB, r increased by 25.18%, and the RMSE reduced by 76.16%. Figure 9 shows the low-frequency noise. The low-frequency noise in the co-band seriously overlapped with the frequency band of the seismic signals, and the noise and seismic signals were poorly separated. Compared to medium-high-frequency noise, denoising is more difficult. After denoising, the SNR of the signal increased by 10.41 dB, r increased by 30.03%, and the RMSE reduced by 70.68%. Figures 10 and 11 shows the mixed-frequency band noise. The mixed-frequency band noise contained multiple frequency components, and it seriously overlapped with the frequency band of the seismic signals, so it was difficult to reduce the noise. After denoising, the SNR of the signal increased by 9.91 dB, r increased by 27.34%, and the RMSE reduced by 62.27%. To sum up, the ARDU model can efficiently remove different types of noise, improve the SNR of the signals, protect effective signals, and reduce waveform distortion.



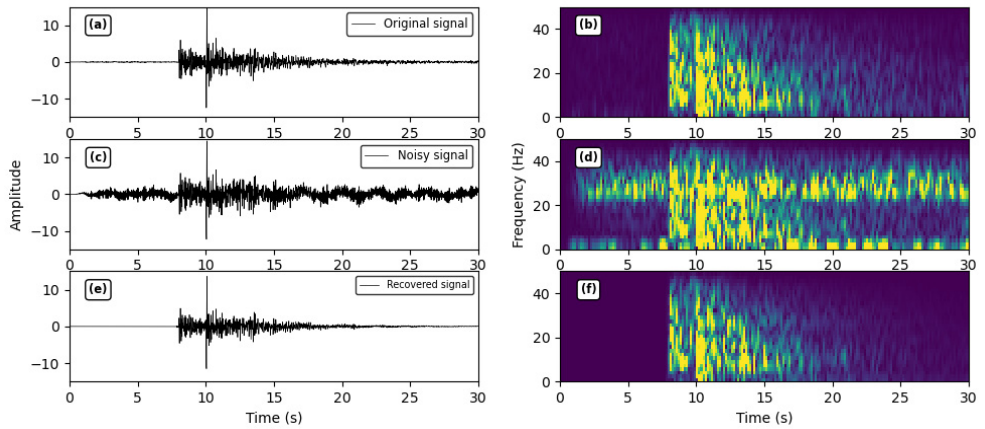
SNR = 2.15 dB, r = 0.7851, RMSE = 0.7807 Before Denoising,
SNR = 14.60 dB, r = 0.9828, RMSE = 0.1861 After Denoising

Figure 8. Co-band denoising of medium-high-frequency noise and seismic signals.



SNR = 1.00 dB, r = 0.7427, RMSE = 0.8912 Before Denoising,
SNR = 11.41 dB, r = 0.9657, RMSE = 0.2613 After Denoising

Figure 9. Co-band denoising of low-frequency noise and seismic signals.



SNR = 0.37 dB, r = 0.7237, RMSE = 0.8911 Before Denoising,
SNR = 11.51 dB, r = 0.9738, RMSE = 0.3516 After Denoising

Figure 10. Co-band denoising of band-limited noise and seismic signals.

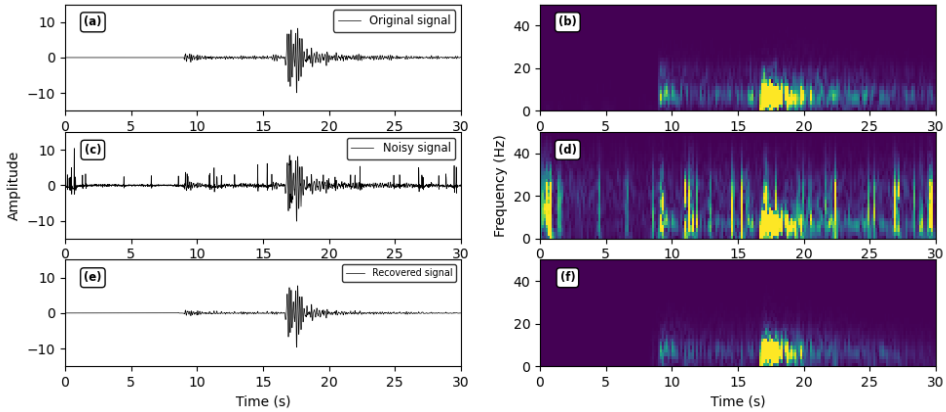


Figure 11. Co-band denoising of spike pulse noise and seismic signals.

4.3. Comparative Experiments and Analysis

To further verify the effectiveness of the ARDU model, it was compared with wavelet threshold denoising, the DnRDB model, and the U-Net model and the experimental results were quantified using the SNR, *r*, and RMSE. All methods adopted the same experimental data; the selected experimental data were the seismic signals with a high SNR, medium-high-frequency noise, low-frequency noise, and mixed-frequency band noise; the length of the selected signals contained 3000 sampling points; and the seismic signals and the three kinds of noise were synthesized into the noise-containing seismic signals with SNRs of −6 dB, 0 dB, and 6 dB, respectively. The denoising results using different algorithms are shown in Tables 1–3 and in Figures 12–14. Among them, the figures show the denoising results of different denoising methods for seismic signals containing different levels of mixed-band noise; (a) and (b) are the time domain and time-frequency domain seismic signals with labels; (c) and (d) are the time domain and time-frequency domain seismic signals with noise; and (e) and (f), (g) and (h), (i) and (j), and (k) and (l) are denoising signals in the time domain and time-frequency domain of the wavelet threshold, DnRDB model, U-Net model, and ARDU model, respectively.

Table 1. Denoising results of seismic signals with an SNR of −6 dB using different methods.

Denoising Method	Seismic Signals with an SNR of −6 dB		
	Signal-to-Noise Ratio (SNR)	Correlation Coefficient (*r*)	Root-Mean-Square Error (RMSE)
Wavelet	0.35 dB	0.5143	0.9325
DnRDB	2.04 dB	0.6247	0.8249
U-Net	5.27 dB	0.7756	0.6451
ARDU	**9.08 dB**	**0.9265**	**0.4321**

Note: Bold text indicates the best performance.

Table 2. Denoising results of seismic signals with an SNR of 0 dB using different methods.

Denoising Method	Seismic Signals with an SNR of 0 dB		
	Signal-to-Noise Ratio (SNR)	Correlation Coefficient (*r*)	Root-Mean-Square Error (RMSE)
Wavelet	4.11 dB	0.7899	0.6892
DnRDB	5.66 dB	0.8619	0.6203
U-Net	8.16 dB	0.8872	0.4687
ARDU	**10.99 dB**	**0.9514**	**0.3335**

Note: Bold text indicates the best performance.

Table 3. Denoising results of seismic signals with an SNR of 6 dB using different methods.

Denoising Method	Seismic Signals with an SNR of 6 dB		
	Signal-to-Noise Ratio (SNR)	Correlation Coefficient (*r*)	Root-Mean-Square Error (RMSE)
Wavelet	8.17 dB	0.9173	0.6514
DnRDB	8.27 dB	0.8487	0.4504
U-Net	12.99 dB	0.9638	0.3238
ARDU	**14.53 dB**	**0.9772**	**0.2472**

Note: Bold text indicates the best performance.

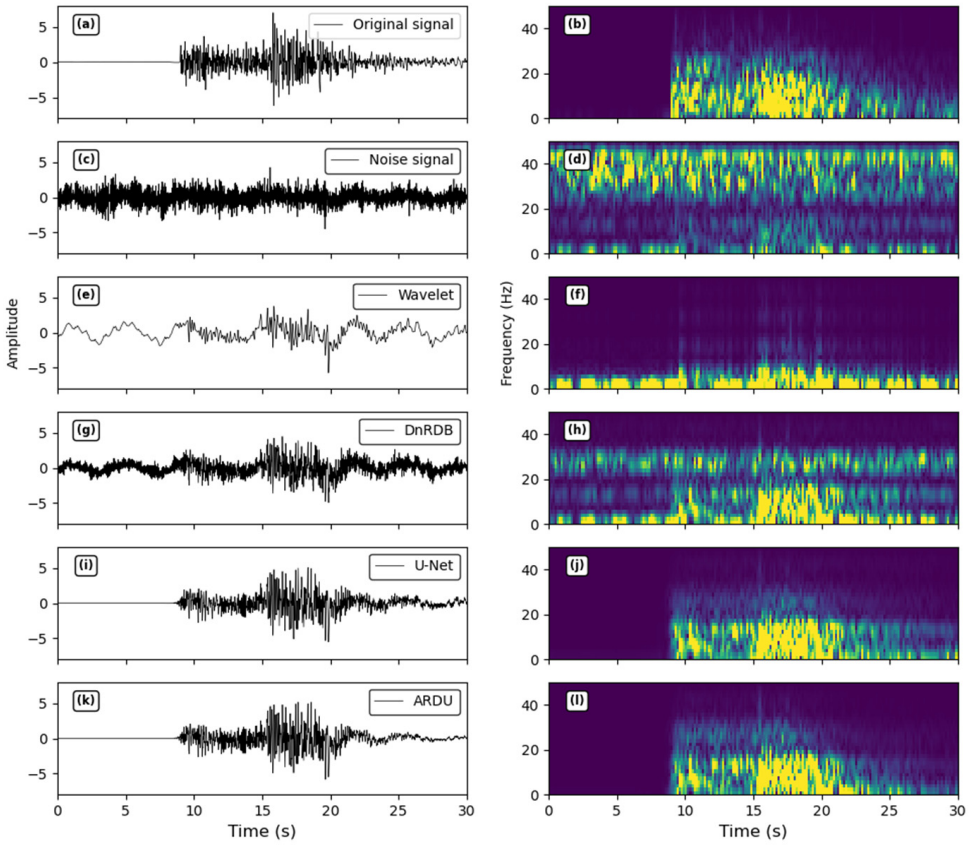


Figure 12. Denoising results of seismic signals with an SNR of −6 dB using different methods.

It can be seen from Table 1 and Figure 12 that when dealing with noise-containing seismic signals with an SNR of −6 dB, the SNR, *r* , and RMSE of the ARDU model improved the most and the effect was the best among several denoising methods, followed by the U-Net model. The SNR and *r* of the ARDU model were 3.81 dB and 16.29% higher than those of the U-Net model, respectively, which indicates that the ARDU model can effectively remove co-band noise and improve the quality of seismic signals. The RMSE reduced by 33.02%, and it was the lowest, which shows that the degree of signal waveform distortion after denoising with the ARDU model is the smallest.

It can be seen from Table 2 and Figure 13 that when dealing with seismic signals with an SNR of 0 dB, the ARDU model had the greatest improvement in the SNR, *r*, and RMSE of the seismic signals with noise and the denoising effect was the best. Compared to the U-Net model, the SNR and *r* of the ARDU model improved by 2.83 dB and 7.24%, respectively, and the RMSE reduced by 28.85%, which indicates that the quality of the signals significantly improves after denoising with the ARDU model and the signal distortion is the smallest.

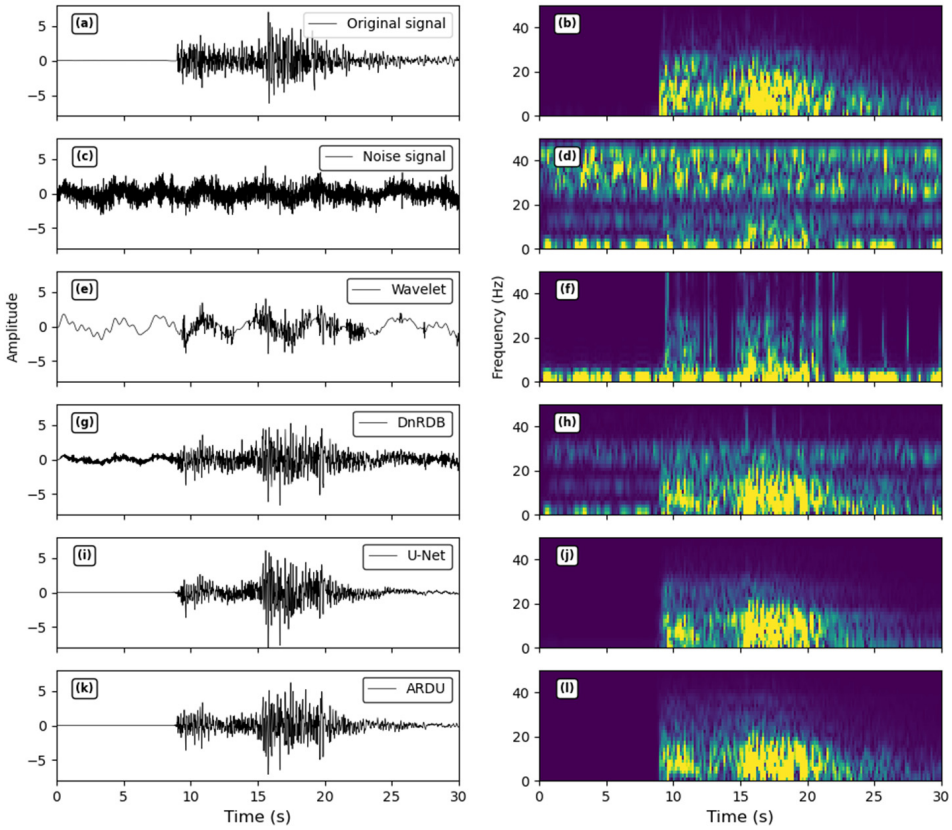


Figure 13. Denoising results of seismic signals with an SNR of 0 dB using different methods.

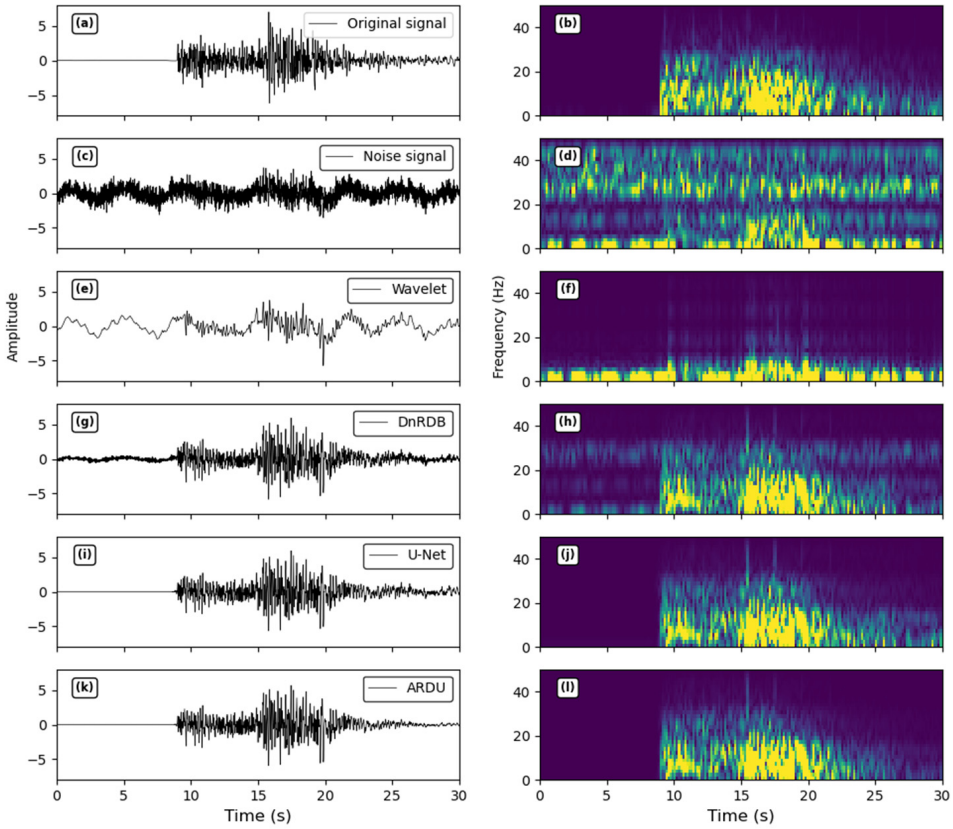


Figure 14. Denoising results of seismic signals with an SNR of 6 dB using different methods.

It can be seen from Table 3 and Figure 14 that when dealing with seismic signals with an SNR of 6 dB, the ARDU model had the best overall denoising effect of the co-band noise, followed by the U-Net model. Compared to the U-Net model, the SNR and *r* of the ARDU model improved by 1.54 dB and 1.39%, respectively, while the SNR and *r* did not improve much, but the RMSE reduced by 23.66%, which indicates that the ARDU model can significantly mitigate the problem of signal waveform distortion after denoising and further improve the quality of seismic signals.

To further verify the effectiveness of the ARDU model on the noise reduction in actual seismic signals, comparison experiments were conducted between the ARDU model and the DnRDB model, the U-Net model, and the RDBU model. The average signal-to-noise ratio was used to evaluate the noise reduction performance of the model, which was obtained using a blind estimation algorithm of the signal-to-noise ratio. Experimental results are shown in Table 4 and Figure 14. Among them, the picture shows the denoising results of four noise reduction methods on the mixed-frequency noise in the same frequency band. Here, (a) and (b) are the actual seismic signals in the time domain and the time-frequency domain, and (c) and (d), (e) and (f), (g) and (h), and (i) and (j) are denoising signals in the time domain and time-frequency domain of the DnRDB model, the U-Net model, the RDBU model, and the ARDU model, respectively.

Table 4. Denoising results of actual seismic signals using the ARDU model.

Denoising Method	Actual Seismic Signals of Jiuzhaigou in Wenchuan	
	Pre-Noise Reduction (SNR)	**After Noise Reduction (SNR)**
DnRDB	2.33 dB	5.92 dB
U-Net	2.33 dB	7.61 dB
RDBU	2.33 dB	8.04 dB
ARDU	**2.33 dB**	**8.82 dB**

Note: Bold text indicates the best performance.

It can be seen from Table 4 that when denoising actual seismic signals, the ARDU model had the greatest improvement in the signal-to-noise ratio. The average signal-to-noise ratio of the dataset after denoising was 8.82 dB, which was 2.9 dB, 1.21 dB, and 0.78 dB higher than that of the DnRDB model, the U-Net model, and the RDBU model, respectively. As can be seen from Figure 15, the ARDU model can effectively remove the energy of noise and protect the detailed features of seismic signals. In conclusion, the ARDU model has a good noise reduction effect on actual seismic signals, significantly improves the signal-to-noise ratio of signals, and is conducive to improving the accuracy of subsequent seismic signal processing, such as seismic phase identification.

In summary, the denoising effect of the ARDU model is the best among several denoising methods, which shows that the ARDU model connects atrous convolution with residual dense blocks to form a new feature extraction unit of the encoder, which can deepen the network's depth, expand the receptive field of the network, reduce waveform distortion, protect effective signals, and improve the quality of seismic signals.

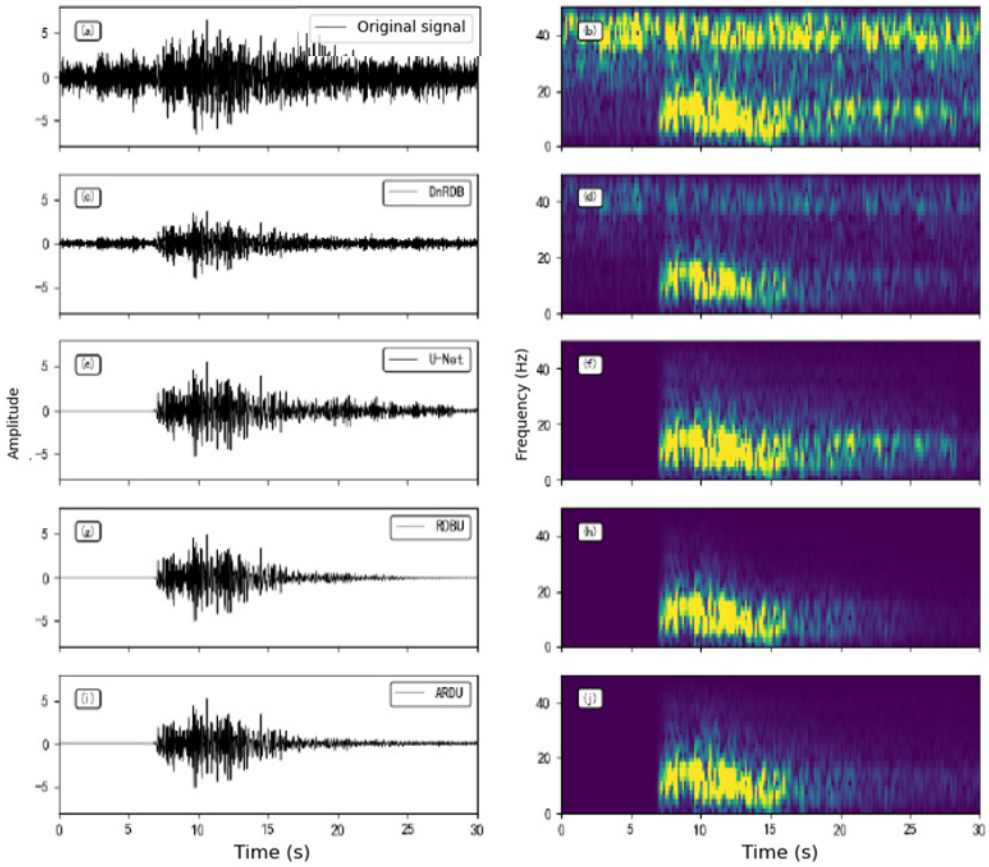


Figure 15. Denoising results of actual seismic signals by different denoising methods.

5. Conclusions

To solve the problem of signal waveform distortion of existing seismic-signal-denoising methods in removing the co-frequency band noise, further improve the SNR of seismic signals, and enhance the quality of seismic signals, this paper took the U-Net as the basic framework, combined atrous convolution and residual dense blocks, and designed a denoising ARDU model for seismic co-band noise. The ARDU model connects atrous convolution and residual dense blocks to form a feature extraction unit of the model's encoder. Among them, the residual dense blocks can deepen the network's depth and enhance the feature extraction ability of the network on the premise of mitigating gradient vanishing and gradient exploding. Atrous convolution can expand the receptive field without increasing the network parameters so as to obtain more abundant signal features, reduce waveform distortion, protect effective signals, and improve the quality of seismic signals. To test the feasibility and effectiveness of the ARDU model, the Stanford Global Seismic dataset was used to construct a training set and a test set and the model was trained and tested on it. The denoising results were compared with the signals containing labels, and the SNR, r, and RMSE were calculated to quantify the denoising performance of the model. Finally, the following conclusions are drawn: (1) The ARDU model can effectively remove the co-band noise in seismic signals, improve the SNR of seismic signals, and enhance the quality of seismic signals, and (2) compared with wavelet threshold denoising, the DnRDB model, and the U-Net model, the ARDU model has the least distortion and the best denoising effect.

Author Contributions: Formal analysis, J.C. and J.Z.; investigation, N.C.; methodology, J.C.; software, L.L.; writing—original draft, J.C., L.W. and J.Z.; writing—review and editing, Z.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Langfang Science and Technology Research and Development Plan Self-funded Project (grant No. 2022011037), the Fundamental Research Funds for the Central Universities (grant no. ZY20180112), the research and practice project of higher education reform in Hebei Province from 2019 to 2020 (grant no. 2019GJJG471), Scientific Research Plan Projects for the National Key Research and Development Programme of China (grant no. 2018YFC1503801), the Scientific Research Project Item of Hebei Province Education Department (grant no. QN2018317), the second batch of new engineering research and practice projects of the Ministry of Education (grant no. E-SXWLHXLX20202607), and the 2020 education research and teaching reform project of the College of Disaster Prevention Technology (grant no. JY2020A12).

Data Availability Statement: Data source: ref. [30].

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gao, H. Study of Seismic Data Residual Statics and AI Denoising. Ph.D. Thesis, University of Science and Technology of China, Hefei, China, 2018.
2. Wang, F. Research on Seismic Data Denoising and Reconstruction Using Deep Learning. Ph.D. Thesis, Zhejiang University, Hangzhou, China, 2020.
3. Sacchi, M.D.; Ulrych, T.J.; Walker, C.J. Interpolation and extrapolation using a high-resolution discrete Fourier transform. *IEEE Trans. Signal Process.* **1998**, *46*, 31–38. [CrossRef]
4. Zhang, Z.; Liu, C. Application of Fractional Fourier Transform in seismic data denoising. *Glob. Geol.* **2014**, *17*, 110–114.
5. Zhai, M.Y. Seismic data denoising based on the fractional Fourier transformation. *J. Appl. Geophys.* **2014**, *109*, 62–70. [CrossRef]
6. Cao, S.; Chen, X. The second-generation wavelet transforms and its application in denoising of seismic data. *Appl. Geophys.* **2005**, *2*, 70–74. [CrossRef]
7. Gaci, S. The Use of Wavelet-Based Denoising Techniques to Enhance the First-Arrival Picking on Seismic Traces. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 4558–4563. [CrossRef]
8. Mousavi, S.M.; Langston, C.A.; Horton, S.P. Automatic microseismic denoising and onset detection using the synchrosqueezed continuous wavelet transform. *Geophysics* **2016**, *81*, V341–V355. [CrossRef]
9. Chen, W.; Song, H. Automatic noise attenuation based on clustering and empirical wavelet transform. *J. Appl. Geophys.* **2018**, *159*, 649–665. [CrossRef]
10. Shan, H.; Ma, J.; Yang, H. Comparisons of wavelets, contourlets and curvelets.in seismic denoising. *J. Appl. Geophys.* **2009**, *69*, 103–115. [CrossRef]
11. Golestani, A.; Kolbadi, S.S.; Heshmati, A.A. Localization and de-noising seismic signals on SASW measurement by wavelet transform. *J. Appl. Geophys.* **2013**, *98*, 124–133. [CrossRef]
12. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.-C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **1998**, *454*, 903–995. [CrossRef]
13. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
14. Roth, H.R.; Yao, J.; Lu, L.; Stieger, J.; Burns, J.E.; Summers, R.M. *Detection of Sclerotic Spine Metastases via Random Aggregation of Deep Convolutional Neural Network Classifications*; Springer International Publishing: Berlin/Heidelberg, Germany, 2015; Chapter 1, pp. 3–12. [CrossRef]
15. Perol, T.; Gharbi, M.; Denolle, M. Convolutional Neural Network for Earthquake Detection and Location. *Sci. Adv.* **2018**, *4*, e1700578. [CrossRef] [PubMed]
16. Zhang, H.; Ma, C.; Pazzi, V.; Zou, Y.; Casagli, N. Microseismic Signal Denoising and Separation Based on Fully Convolutional Encoder ecoder Network. *Appl. Sci.* **2020**, *10*, 6621. [CrossRef]
17. Simard, P.Y.; Steinkraus, D.; Platt, J.C. Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis. In Proceedings of the 7th International Conference on Document Analysis and Recognition (ICDAR 2003), Edinburgh, UK, 6 August 2003.
18. Taigman, Y.; Yang, M.; Ranzato, M.A.; Wolf, L. DeepFace: Closing the Gap to Human-Level Performance in Face Verification. In Proceedings of the IEEE Conference on Computer Vision & Pattern Recognition, Columbus, OH, USA, 23–28 June 2014.
19. Maas, A.; Le, Q.V.; O'neil, T.M.; Vinyals, O.; Nguyen, P.; Ng, A.Y. Recurrent Neural Networks for Noise Reduction in Robust ASR. In Proceedings of the Conference of the International Speech Communication Association, Portland, OR, USA, 9–13 September 2012.
20. Zhang, K.; Zuo, W.; Chen, Y.; Meng, D.; Zhang, L. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE Trans. Image Process.* **2016**, *26*, 3142–3155. [CrossRef] [PubMed]
21. Yu, S.; Ma, J.; Wang, W. Deep learning for denoising. *Geophysics* **2019**, *84*, V333. [CrossRef]
22. Jin, Y.; Wu, X.; Chen, J.; Han, Z.; Hu, W. Seismic data denoising by deep-residual networks. In *SEG Technical Program Expanded Abstracts*; Society of Exploration Geophysicists: Houston, TX, USA, 2018.

23. Zhu, W.; Mousavi, S.M.; Beroza, G.C. Seismic Signal Denoising and Decomposition Using Deep Neural Networks. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 9476–9488. [CrossRef]
24. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015.
25. Gao, Z.; Zhang, S.; Cai, J.; Hong, L.; Zheng, J. Research on Deep Convolutional Neural Network Time-Frequency Domain Seismic Signal Denoising Combined with Residual Dense Blocks. *Front. Earth Sci.* **2021**, *9*, 571. [CrossRef]
26. Yu, F.; Koltun, V. Multi-Scale Context Aggregation by Dilated Convolutions. *arXiv* **2015**, arXiv:1511.07122.
27. Xu, Y.; Du, J.; Dai, L.R.; Lee, C.H. An Experimental Study on Speech Enhancement Based on Deep Neural Networks. *IEEE Signal Process. Lett.* **2013**, *21*, 65–68. [CrossRef]
28. Huang, P.S.; Kim, M.; Hasegawa-Johnson, M.; Smaragdis, P. Deep learning for monaural speech separation. In Proceedings of the IEEE International Conference on Acoustics, Florence, Italy, 4–9 May 2014.
29. Weninger, F.; Hershey, J.R.; Le Roux, J.; Schuller, B. Discriminatively trained recurrent neural networks for single-channel speech separation. In Proceedings of the Signal & Information Processing, Atlanta, GA, USA, 3–5 December 2014.
30. Mousavi, S.M.; Sheng, Y.; Zhu, W.; Beroza, G.C. STanford EArthquake Dataset (STEAD): A Global Data Set of Seismic Signals for AI. *IEEE Access* **2019**, *7*, 179464–179476. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.