applied sciences

MDPI

Article

Seismic Data Denoising Based on Wavelet Transform and the Residual Neural Network

Tianwei Lan, Zhaofa Zeng *, Liguo Han and Jingwen Zeng

College of Geoexploration Science and Technology, Jilin University, Changchun 130026, China
* Correspondence: zengzf@jlu.edu.cn

Abstract: The neural network denoising technique has achieved impressive results by being able to automatically learn the effective signal from the data without any assumptions. However, it has been found experimentally that the performance of the method using neural networks gradually decreases with increasing pollution levels when processing contaminated seismic data, and how to improve the performance will become the direction of further development of the method. As a traditional method widely used for tainted seismic data, the wavelet transform can effectively separate the signal from the noise. Thus, we propose a method combining wavelet transform and a residual neural network that achieves good results in suppressing random noise data.

Keywords: wavelet transform; noise suppression; residual neural network; random noise; wavelet decomposition level

check for updates

Citation: Lan, T.; Zeng, Z.; Han, L.; Zeng, J. Seismic Data Denoising Based on Wavelet Transform and the Residual Neural Network. *Appl. Sci.* **2023**, *13*, 655. https://doi.org/10.3390/app13010655

Academic Editors: Dongyuan Shi, Woon-Seng Gan and Jihui Aimee Zhang

Received: 2 December 2022
Revised: 25 December 2022
Accepted: 28 December 2022
Published: 3 January 2023

1. Introduction

Seismic signals will inevitably be affected by instruments, environments, and other factors, resulting in the need for seismic noise attenuation. Generally, seismic noise may be broken down into two distinct categories: coherent noise and random noise. Coherent noise has apparent kinematic characteristics and certain regularity. In contrast, random noise has strong randomness, and its existence is an important factor affecting the signal-to-noise ratio (SNR) of seismic data. To suppress the noise, scholars have proposed various denoising methods, such as F-x deconvolution [1,2], empirical modal decomposition (EMD) [3,4], Time-frequency peak filtering method [5], median filtering method [6], and so on. Denoising seismic data in virtue of the f-x approach is hindered by the fact that it presumes the signal to be linear and stationary, which is at odds with the nonstationary nature of seismic records; in contrast, EMD captures the nonstationary nature of the signal since it is data-driven, but mode mixing can cause the time-frequency distribution to be incorrect. Owing to the nonlinearity of the signal, the frequency peak filtering based on time-frequency analysis entails a loss of amplitude when denoising; the median filtering method is presented as a simple method of removing noise from a nonstationary signal, raising distortions in the signal shape related to the filter length. Subsequently, the sparse transform emerged as a method that can distinguish between noise and valid signals and has been extensively investigated in seismic random noise suppression. With the sparse transform-based denoising method, it is assumed that the seismic data are sparse in some transform domains, after transforming the seismic data to the sparse transform domain, thresholding is applied to the coefficients, and then the processed coefficients are inverse-transformed back to the original domain. Classic sparse transforms, including Fourier transform [7,8] and wavelet transform [9,10], are commonly used as sensitive and effective methods. Curvelet [11], Contourlet [12], and Shearlet [13] transforms were later advanced to compensate for the wavelet transform's lack of directionality and anisotropy. However, none of the aforementioned methods are data-driven and swayed by the characteristics of their own sparse transformation bases; furthermore, despite their increased speed, all methods still require estimating the noise level and choosing the filtering threshold. To eliminate these drawbacks and better represent the sparsity of seismic data, dictionary learning [14,15] started to play a role in seismic data denoising. Although dictionary learning eliminates the above-mentioned drawbacks to some extent, some of the overall data are missed when learning is performed using a limited number of random patches.

The aforementioned techniques for suppressing noise in seismic data have their pros and cons, so how to improve the SNR of noisy data with a better approach is always a concern. In computing [16], medicine [17], and image processing [18], the machine learning method has emerged as a leading research path since the turn of the 21st century, in tandem with the advancement of high-performance computing capabilities. Under this influence, the neural network has been introduced to seismic data processing and has obtained favorable results.

Generative adversarial networks (GANs) [19], Resnet [20], and U-Net [21] are commonly used networks. A GAN employs a method in which generators and discriminators contest each other, with generators generating data to cheat discriminators, and discriminators differing between generated and labeled data allowing generators to produce more compliant data; however, subject to pattern collapse, GAN networks may struggle to deliver the expected data. Resnet, DnCNN, and U-net all belong to convolutional neural networks (CNNs). The CNN-based noise suppression method for seismic data involves recovering the effective signal by extracting features from the dataset. Compared with traditional denoising methods, the above noise-suppression approaches avoid threshold selection and defects of the methods themselves, achieving better denoising results. Still, with the increase in noise level, restricted by their structures, neural networks are left with limited denoising performances. Other authors have offered some potential answers to this difficult problem. As an illustration, the seismic data denoising method combining U-net and STFT has been highly successful [22,23]. However, even a simplified U-net has a much larger number of parameters than a Resnet with fewer residual blocks [24]. This may be a hindrance to the practical application of neural nets in seismic data denoising. Therefore, as a simpler and more stable denoising structure, Resnet is adopted. To further enhance neural network noise-reduction capabilities, inspired by the work of other scholars, this paper proposes a seismic signal denoising method combining wavelet transform and a neural network (DWT-Resnet). Utilizing the fact that a valid signal differs from noise after the wavelet transform, the transformed signal serves as the input of the neural network, removing noise from the signal with a clean signal as the labeled data. Experiments show that the method proposed in this paper achieves superior performance in denoising compared to the method using neural networks alone.

The main contributions of this study consist of:

- Proposing a denoising method for seismic data based on wavelet transform and residual neural network.
- Analyzing the performance of residual neural networks for noisy removal under different decomposition levels of different wavelet bases
- An analysis of the number of layers of the residual neural network on the noise removal performance in the wavelet transform-based neural network denoising method is presented.

The rest of the paper is organized as follows: Section 2 provides a detailed description of the method, evaluation criteria, and model. Experimental results of our approach on simulated seismic data versus real seismic data are presented and analyzed in Sections 3 and 4. Finally, Section 5 presents the conclusion, discusses this paper, and provides a discussion of future directions.

2. Methods and Evaluation Index

2.1. DWT-Resnet Denoising Principle

Figure 1 shows the network structure of DWT-Resnet. The noise-bearing seismic data were decomposed using 1D wavelets, ranking the decomposed high-frequency and low-frequency coefficients in order to form a dataset conforming to the 1D Resnet input, and then the clean seismic data were taken as the training labels of the Resnet, which were trained for denoising. The trained network was applied to the test set using the same wavelet-based transform and decomposition levels so that the denoising results of noise-bearing seismic data were directly obtained.

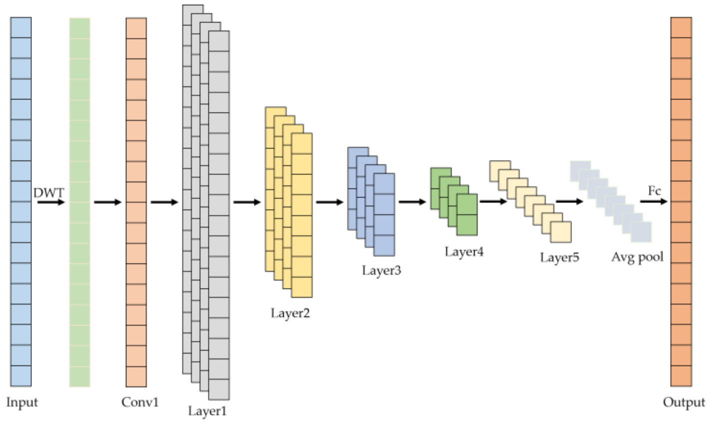


Figure 1. The network structure of DWT-Resnet.

The mathematical principle of the DWT-Resnet method is described next.

As a transform domain analysis method similar to Fourier transform, wavelet transform separates the effective signal from the noise by doing the inner product of the basis function and signal. For a certain given wavelet $\Phi(x)$, the wavelet basis function is:

$$\Phi_{a,b}(x) = \frac{1}{\sqrt{a}}\Phi(\frac{x-b}{a}) \tag{1}$$

where a is the scale parameter, b is the time window parameter, $a, b \in \mathrm{R}^+$. That is, the wavelet basis function $\Phi_{a,b}(x)$ is the shift and scale transformation of the initial wavelet function $\Phi(x)$. Therefore, for the signal $f(x)$, its wavelet transform equation is:

$$WT(a,b) = \int_{-\infty}^{+\infty} f(x)\overline{\Phi(\frac{x-b}{a})dx} \tag{2}$$

where $\overline{\Phi_{a,b}(x)}$ is the complex conjugate of the signal $\Phi_{a,b}(x)$.

The seismic data containing noise y, can be regarded as a superposition of clean data and random noise, i.e.,

$$y = z + v \tag{3}$$

By taking the seismic data containing noise through the wavelet transform as the input of the neural network, the clean data z are predicted by building a mapping of $R(\mathrm{DWT}(y); P)$, i.e.,

$$R(\mathrm{DWT}(y); P) \approx z \tag{4}$$

where $P = \{\mathrm{W}, \mathrm{B}\}$ is the parameter of the neural network and $DWT(y)$ is the result of the noise-containing data after the wavelet transform. The loss function of the neural network is:

$$l(P) = \sum_{k=1}^{Q} \|R(\mathrm{DWT}(y_k); P) - z_k\|_F^2 \tag{5}$$

where $\mathrm{DWT}(y)$ indicates the wavelet domain noise data y_k acquired by using the one-dimensional (1D) wavelet transform, z_k represents the clean data obtained by feeding y_k into the neural network and training, Q is the batch size. Normally, the smaller the $l(P)$, the better the network optimization, and the higher the denoising capability. Since different wavelet bases with different decomposition levels and different layers of Resnet have different effects on the DWT-Resnet method, these factors will be explored in Sections 3 and 4.

2.2. Denoising Performance Evaluation Index

SNR is a popular metric for gauging the effectiveness of seismic data processing. In this paper, the signal-to-noise ratio is applied to evaluate the noise rejection ability of DWT-Resnet. The equation is as follows:

$$\mathrm{SNR} = 10\log_{10}\left\{\frac{\|z\|_F^2}{\|z-\tilde{z}\|_F^2}\right\} \tag{6}$$

where z is the original seismic data used as labels and $\tilde{z}$ is the denoised seismic data.

2.3. Training Model and Parameter Setting

Since this paper deals with 1D seismic data, considering the amount of data and the performance of the machine itself, Resnet18 is the chosen network, based on Resnet18, changing the convolutional layer, pooling layer, and batchnormal to 1D, while keeping the rest of the settings unchanged. As for the network parameter settings, according to the amount of data to be processed, the batch size is set to 100 and the initial learning rate range is $[10^{-6}, 9\times10^{-6}]$. During the training process, the learning rate decreases to 90% of the original rate every 10 epochs.

The experiments in this paper were run on a machine equipped with an Intel(R) Xeon(R) CPU E5-2678 v3 @ 2.50 GHz and an NVIDIA GeForce RTX 2080 Ti.

3. Synthetic Seismic Data Processing Results

We first used synthetic seismic data to test the denoising performance of the proposed method.

For the original simulated data, we added different levels of noise to the data to test the DWT-Resnet18 performance, forming SNRs from low to high of −7.512 dB, −5.467 dB, −2.756 dB, 4.074 dB, and 18.074 dB, respectively. The original data consisted of 8100 seismic traces, with 512 sample points per seismic trace and a 0.001 s sampling period. Therefore, for the obtained noisy data and original simulated data, the corresponding first 80% of the data were extracted, respectively, with the mirror flip and up–down flip for data enhancement, are made into the training data, and the remaining 20% of the data 10% as the validation set 10% as the test set.

Five commonly used wavelet bases haar, db, sym, coif, bior, rbio, were selected to verify the noise-reducing performance of DWT-Resnet with different wavelet bases and different numbers of decomposition levels. The related outcomes of the trials were listed in Table 1. The red-marked data in the table represent the best denoising outcomes, while the blue-marked data represent the second-best denoising results. As can be seen in Table 1, except for the rbio wavelet, the best denoising results were achieved in decomposing levels 2 and 3 when processing seismic data with different SNRs. The rbio wavelet offers the best denoising impact of all wavelets when processing seismic data with an SNR of 18,074 dB, but only when the wavelet's decomposition level is 5. There is a non-significant difference between the best denoising result and the second-best denoising result in the denoised seismic data, with a difference of 0.116 dB for decomposition level 3 and 0.108 dB for decomposition level 4. Overall, the optimal denoising results differ from the suboptimal denoising results when dealing with high SNR seismic data, while the optimal denoising results are similar to the suboptimal denoising results when dealing with low SNR seismic data. Therefore, considering this situation, it can be said that 2- and 3-level decomposition denoising can be preferred when different wavelet basis functions are used to denoise this synthetic seismic data.

Table 1. DWT-Resnet18 denoising results for seismic data with different noise levels, using different wavelet basis functions, and at different wavelet decomposition levels.

Input SNR(dB)			**18.074**			
Wavelet decomposition level Wavelet name	haar	db5	sym4	coif3	bior1.5	rbio3.1
Level 1 output SNR(dB)	26.390	26.761	26.836	27.184	26.734	27.151
Level 2 output SNR(dB)	27.679	28.551	28.608	28.229	27.791	28.418
Level 3 output SNR(dB)	26.828	26.971	27.004	26.196	26.794	28.875
Level 4 output SNR(dB)	26.218	25.899	26.236	26.558	26.622	28.991
Level 5 output SNR(dB)	26.426	26.314	26.371	26.496	26.698	28.983
Input SNR(dB)			4.074			
Wavelet decomposition level Wavelet name	haar	db8	sym7	coif4	bior6.8	rbio3.1
Level 1 output SNR(dB)	19.365	19.105	18.923	19.582	19.161	19.230
Level 2 output SNR(dB)	20.494	20.897	20.700	20.684	20.575	21.337
Level 3 output SNR(dB)	20.279	21.491	21.213	21.363	21.621	21.466
Level 4 output SNR(dB)	19.997	20.965	21.060	20.940	20.686	21.345
Level 5 output SNR(dB)	19.875	21.121	20.547	20.806	20.619	21.298
Input SNR(dB)			−2.756			
Wavelet decomposition level Wavelet name	haar	db4	sym7	coif2	bior2.6	rbio6.8
Level 1 output SNR(dB)	15.734	15.572	15.712	15.448	15.618	15.795
Level 2 output SNR(dB)	17.339	17.652	17.416	17.542	17.473	17.588
Level 3 output SNR(dB)	17.086	17.899	17.882	17.943	17.751	18.035
Level 4 output SNR(dB)	15.517	16.024	16.392	16.187	16.047	16.280
Level 5 output SNR(dB)	15.221	15.920	15.998	15.826	15.704	15.799
Input SNR(dB)			−5.467			
Wavelet decomposition level Wavelet name	haar	db7	sym7	coif3	bior1.5	rbio4.4
Level 1 output SNR(dB)	13.946	13.861	13.923	13.971	14.078	14.051
Level 2 output SNR(dB)	15.316	15.569	15.400	15.579	15.401	15.530
Level 3 output SNR(dB)	15.114	16.059	16.131	15.804	15.781	15.941
Level 4 output SNR(dB)	13.095	13.624	13.589	13.820	13.573	13.764
Level 5 output SNR(dB)	13.033	13.704	13.018	13.366	12.847	13.052
Input SNR(dB)			−7.512			
Wavelet decomposition level Wavelet name	haar	db5	sym8	coif5	bior6.8	rbio4.4
Level 1 output SNR(dB)	13.113	12.821	12.972	13.043	13.104	12.990
Level 2 output SNR(dB)	13.743	13.903	13.943	14.014	14.015	14.037
Level 3 output SNR(dB)	13.444	14.321	14.296	14.233	14.338	14.264
Level 4 output SNR(dB)	11.240	11.896	11.792	11.756	11.424	11.995
Level 5 output SNR(dB)	11.137	11.089	10.936	11.183	10.709	11.200

Figure 2 displays the ideal denoising results of each wavelet basis with the input of different SNRs seismic data obtained using DWT-Resnet18 denoising with multiple wavelet bases for different SNRs seismic data, providing deeper insight into the outcomes obtained by each wavelet. It can be seen that the denoising effects of different wavelet bases vary widely when processing high SNR seismic data, but the denoising effect of haar wavelet and bior wavelet is especially different from the other three. As the SNR decreases, the denoising results of haar wavelet and other wavelets gradually decrease from 1.312 dB to 0.595 dB, while the denoising results of other wavelets tend to overlap, i.e., the denoising effects of different wavelets will tend to be the same as the noise is further enhanced.

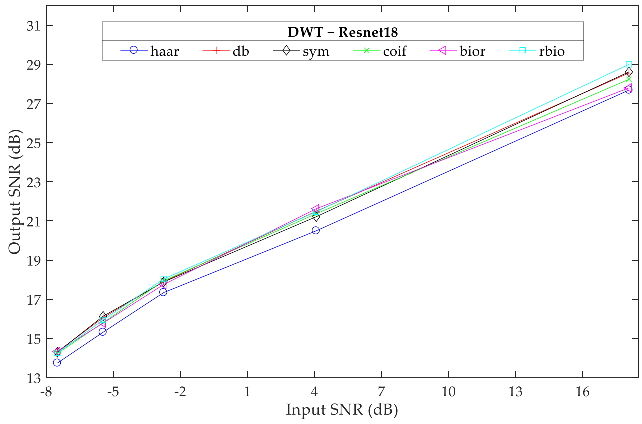


Figure 2. Trend of the best denoising results of DWT−Resnet18 with varying SNRs of noisy seismic data under different wavelet basis functions.

Previous experiments were conducted under Resnet18, but as to whether Resnet18 can achieve optimal denoising results is yet to experimented with. Therefore, taking the seismic data with the SNR of 4.074 dB as an example, we verified the denoising effect of the Resnet network under different layers by increasing or decreasing the number of layers of Resnet18 with a 3-level wavelet decomposition using sym7 (Figure 3).

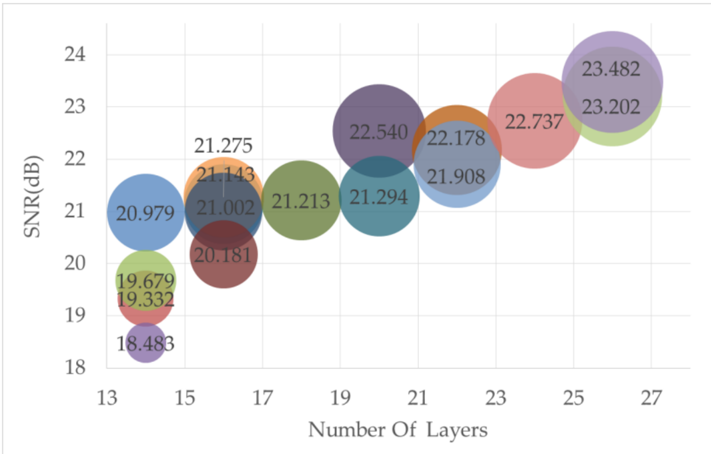


Figure 3. Denoising results of seismic data with SNR of 4.074 dB for DWT−Resnet18 networks with different numbers of layers for a fixed wavelet type and wavelet decomposition levels.

One convolutional layer and four layers of two convolutional residual blocks make up the Resnet18 architecture shown in Figure 1. In this thesis, the increase or decrease of the number of layers in the Resnet18 network is concentrated on the layers with residual blocks. Resnet14 means removing any layer of Resnet18. When the number of Resnet layers is 14, no matter which layer is removed, the denoising effect will be worse, especially the removal of layer4, which leads to the worst denoising result, only 18.483 dB; while layer16 is the removal of one of the residual blocks in any layer when the number of Resnet layers is 16, the removal of one residual block in layer4 has the greatest effect on the network denoising. Inspired by the above experiments, the subsequent increase in the number of network layers is either by adding a new layer or by increasing the number of residual blocks in the layer. Following layer4, two further layers, layers 5 and 6, are brought in, where layer5 includes 2 residual blocks, each with 1024 channels, and so on layer6 is a layer composed of 2 residual blocks containing 2048 channels. Then the 20-layer Resnet includes two cases, one is to increase the number of residual blocks of layer4, and the other is to add a layer5 containing 1 residual block, the SNR of the latter is 22.540 dB, which is 1.246 dB higher than the former, making the neural network denoising effect has been greatly improved. Inspired by the above experimental results, better results were obtained when the number of layers of the Resnet network was increased to 26, i.e., layer5 containing 4 residual blocks linked after layer4. Although another combination of the 26-layer network is to link layer5 and layer6 in sequence after layer4, when its denoising effect is the best, the training time also increases due to the increase of network layers, and it takes 8251.736 s to train the 26-layer network. However, in terms of SNR, the 26-layer network with layer6 is only 0.280 dB better than the 26-layer network with only layer5, this is obviously a less suitable denoising solution, so the layer5 with 4 residual blocks is used as the optimal denoising network.

Next, with the results in Table 1, the denoising performance of DWT-Resnet26 was tested at different SNRs, choosing different wavelet bases, using the number of decomposition levels for the best denoising results in DWT-Resnet18. Following denoising results were obtained through experiments (Table 2, Figure 4).

Table 2. DWT−Resnet26 denoising results obtained by using different wavelet bases for seismic data with different SNRs.

Input SNR(dB)			**18.074**			
Wavelet name	haar	db5	sym4	coif3	bior1.5	rbio3.1
output SNR(dB)	27.657	29.694	29.574	29.443	28.1281	28.256
Input SNR(dB)			4.074			
Wavelet name	haar	db8	sym7	coif4	bior6.8	rbio3.1
output SNR(dB)	22.251	23.473	23.202	23.465	23.304	21.331
Input SNR(dB)			−2.756			
Wavelet name	haar	db4	sym7	coif2	bior2.6	rbio6.8
output SNR(dB)	17.423	18.035	18.329	18.606	18.129	18.617
Input SNR(dB)			−5.467			
Wavelet name	haar	db7	sym7	coif3	bior1.5	rbio4.4
output SNR(dB)	15.565	16.040	15.938	15.944	15.778	16.247
Input SNR(dB)			−7.512			
Wavelet name	haar	db5	sym8	coif5	bior6.8	rbio4.4
output SNR(dB)	13.336	14.640	14.224	14.336	14.551	14.433

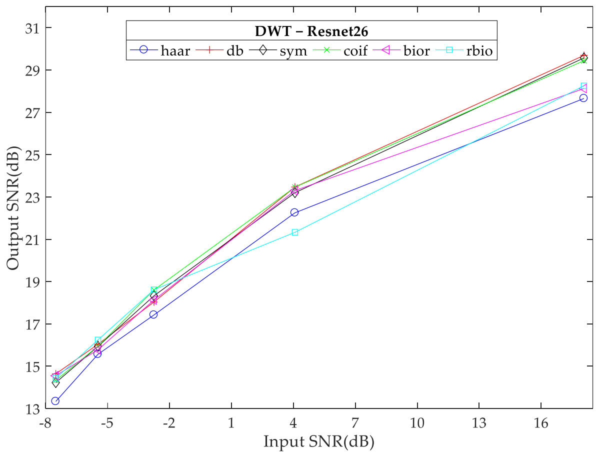


Figure 4. Trends of the best denoising results of DWT−Resnet26 with different SNRs seismic data under different wavelets.

From Table 2 and Figure 4, it is observed that when the number of layers of the residual neural network is increased to 26, the denoising differences are more obvious when different wavelet bases are processed with seismic data at high input SNR; haar performance results have been poor, and rbio does not improve its denoising performance significantly at low SNR, with the increase of noise level, its denoising effect is similar to the performance of several wavelets except haar. Db and coif remain relatively stable.

Comparing the best denoising results of Resnet18 and Resnet26, DWT-Resnet18 and DWT-Resnet26 (Figure 5), it is clear that the residual neural network denoising method combined with wavelets performs better. DWT-Resnet26 achieves the best performance at higher SNR, while DWT-Resnet18 and DWT-Resnet26 obtain similar denoising results when the SNR is low. In contrast, the denoising effect of Resnet26 is conversely inferior to that of Resnet18 when processing data without wavelet transform. A possible cause of the conjecture is the deepening of the network model [25], which causes overfitting, but wavelet-transformed data are not affected by overfitting.

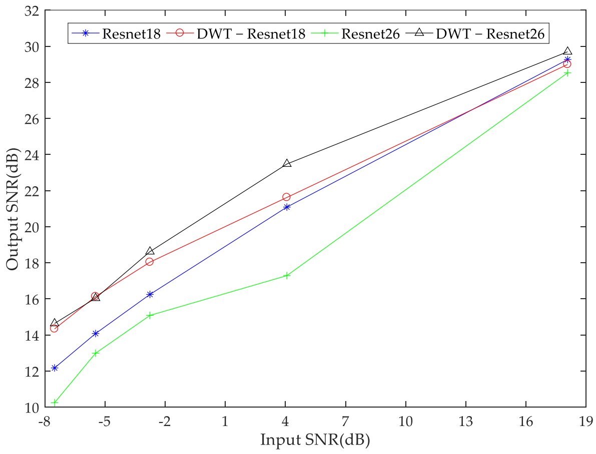
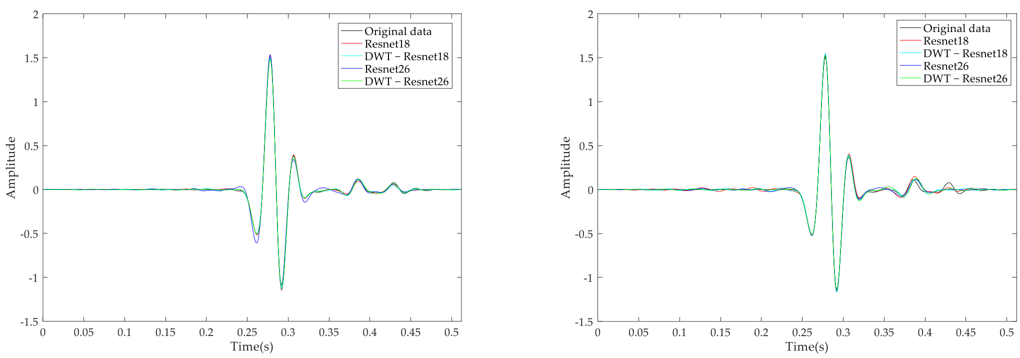


Figure 5. Comparison of denoising results using Resnet18, Resnet26, DWT−Resnet18, DWT−Resnet26 with different input SNRs.

To visualize the denoising effect of each method, Figure 6 shows the comparison of single-channel clean seismic data and the corresponding single-channel data after denoising with Resnet18, DWT-Resnet18, Resnet26, and DWT-Resnet26 at different SNRs.

Results point out that although DWT-Resnet improves the denoising ability of Resnet, even DWT-Resnet is unable to adequately match the original seismic data when the SNRs of the input seismic data decline.



(**a**) (**b**)

Figure 6. *Cont.*

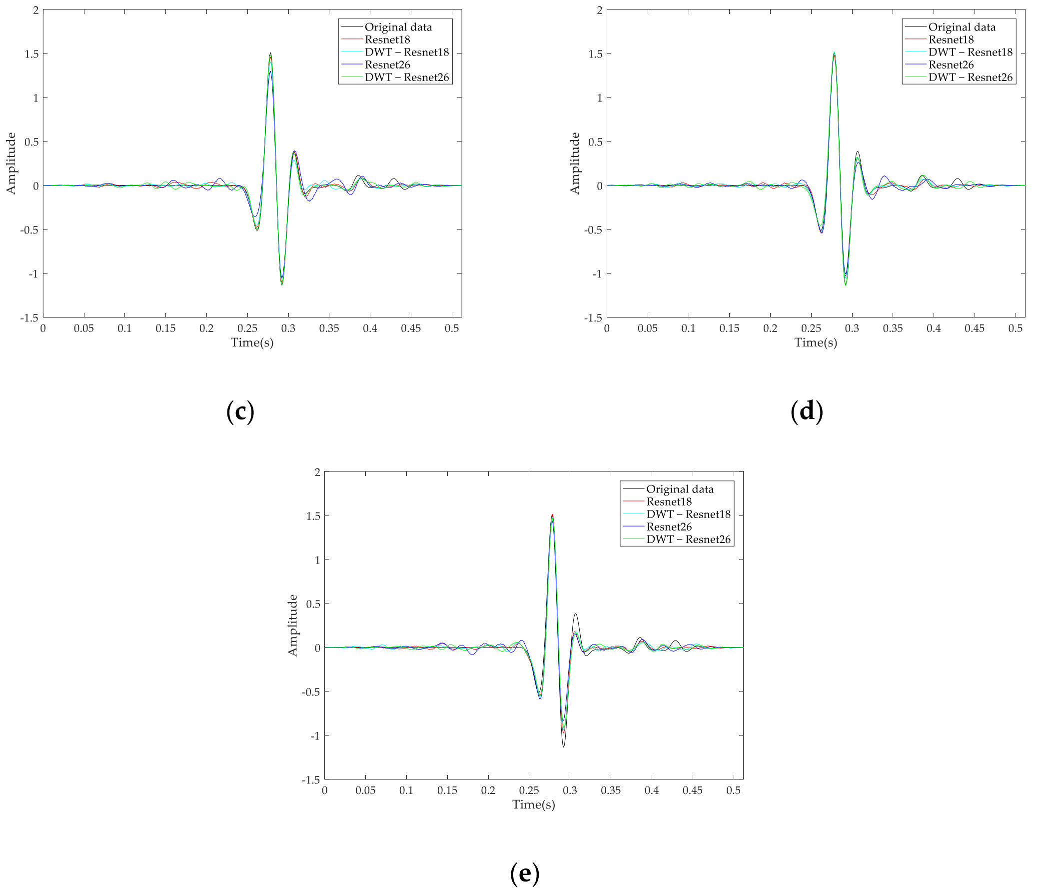


Figure 6. (**a**–**e**) illustrates the single−channel data compared with the clean single−channel data after denoising with four denoising methods at SNR ranked from high to low.

4. Real Seismic Data Processing Results

Real data consists of 14,735 seismic traces, with 256 sample points per seismic trace and a 0.002 s sampling period. The process of making the training set, test set, and validation set is the same as making the synthetic seismic dataset. For verifying the feasibility of the method, the results of denoising with different wavelet bases at different numbers of decomposition layers with different input SNRs were measured. The experiments demonstrate that in most cases, after the wavelet decomposition at level 2 or 3, optimum denoising results can be obtained by residual neural network processing. When the input SNR is −5.203 dB and −2.273 dB, the optimal denoising result mostly appears in the wavelet decomposition of level 1 and 2. However, in case the optimal denoising result appears in the wavelet decomposition of level 1, it is not much different from the result of the level 2 or 3 wavelet decomposition, which is within 0.5 dB.

Notice that unlike the simulated data, where the optimal wavelet for each wavelet basis is essentially fixed even at different noise levels, this property of the real data is related to the data itself, where the simulated data have zero data except where seismic waves are not present, while the real data still have non-zero irrelevant data where there are no seismic waveforms, even though they are not visible to the naked eye.

Figure 7 plots the trend of the denoising capability of Table 3. Except for the substantial discrepancies at high SNRs for haar, db, sym, coif, and bior wavelets, it is evident that the difference between the Resnet18 denoising outcomes under various wavelets reduces as the noise level increases. The overall denoising effect under rbio is stable, and as the input SNR decreases, its denoising effect gradually outperforms the denoising effect of Resnet18 under other wavelet bases.

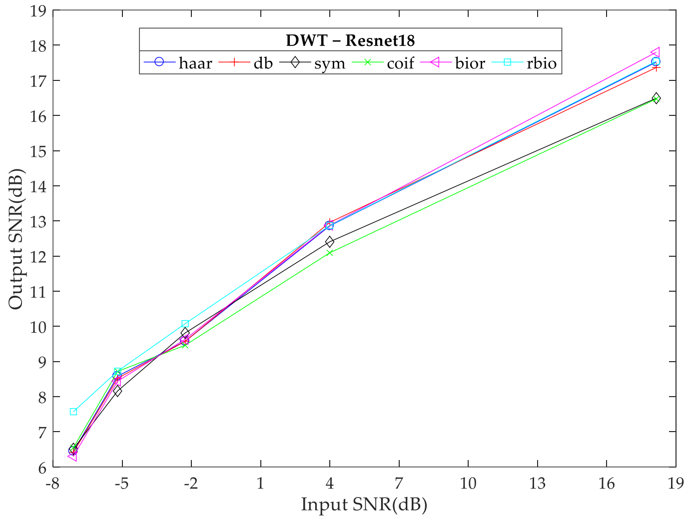


Figure 7. Trend of the best denoising results of DWT−Resnet18 with varying SNRs of real noisy seismic data under different wavelet basis.

To verify the impact of network depth on the denoising effect in the DWT-Resnet denoising method with real seismic data. To verify the effect of the network depth of DWT-Resnet on the real seismic data denoising results. Taking the real seismic data with a SNR of −7.1157 dB as an example, the effect of the number of Resnet layers on the denoising results in the case of a 2-level decomposition using haar is as follows (Figure 8). The image shows that the DWT-Resnet26 still has a remarkable denoising capability.

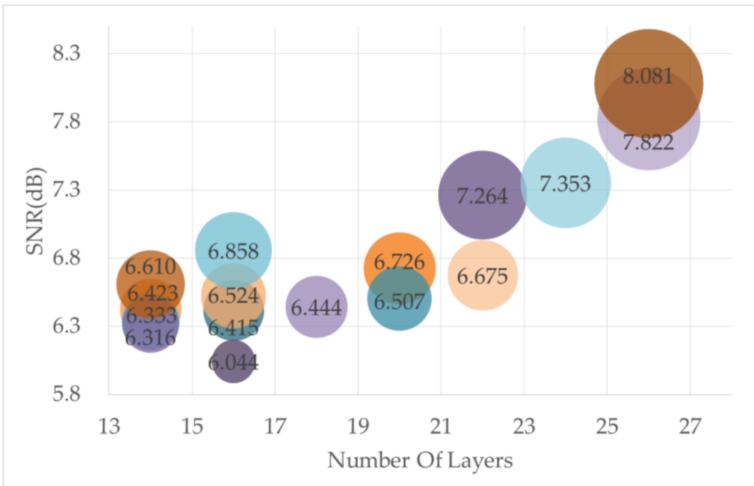


Figure 8. Denoising results of real seismic data with SNR of −7.116 dB for DWT-Resnet networks with a different number of layers for a fixed number of wavelet bases and wavelet decomposition levels.

When increasing the number of layers of the residual neural network to 26, the denoising difference is obvious when different types of wavelets process the real seismic data with high input SNR (Figure 9). Except for the db and haar wavelets, the denoising ability of other wavelets fluctuates somewhat when processing real seismic data, db, and haar behave more consistently when processing seismic data with different SNR compared to other wavelets.

Table 3. DWT-Resnet18 denoising results for real seismic data with different noise levels, using different wavelet basis functions, and at different wavelet decomposition levels.

Input SNR(dB)			**18.148**			
Wavelet decomposition level Wavelet name	haar	db1	sym7	coif2	bior1.1	rbio1.1
Level 1 output SNR(dB)	13.364	17.096	13.581	12.616	11.950	13.781
Level 2 output SNR(dB)	17.519	17.114	15.831	16.470	17.800	17.471
Level 3 output SNR(dB)	17.461	17.365	15.099	15.414	17.496	17.543
Level 4 output SNR(dB)	16.941	16.988	16.492	16.117	17.105	17.103
Level 5 output SNR(dB)	17.096	16.952	15.616	15.806	17.342	16.849
Input SNR(dB)			3.992			
Wavelet decomposition level Wavelet name	haar	db1	sym5	coif1	bior1.1	rbio1.1
Level 1 output SNR(dB)	9.373	9.241	9.75	10.479	9.699	9.51
Level 2 output SNR(dB)	12.857	12.957	12.406	12.092	12.662	12.867
Level 3 output SNR(dB)	12.854	12.927	12.037	11.26	12.876	12.858
Level 4 output SNR(dB)	12.717	12.746	12.118	11.312	12.792	12.72
Level 5 output SNR(dB)	12.769	12.751	12.094	11.550	12.768	12.845
Input SNR(dB)			−2.273			
Wavelet decomposition level Wavelet name	haar	db1	sym5	coif1	bio1.1	rbio3.1
Level 1 output SNR(dB)	9.575	9.573	9.308	8.027	9.630	8.038
Level 2 output SNR(dB)	8.592	8.977	9.812	9.462	8.942	10.076
Level 3 output SNR(dB)	9.504	9.373	8.772	8.648	9.424	9.401
Level 4 output SNR(dB)	9.373	9.307	8.675	8.801	9.091	9.490
Level 5 output SNR(dB)	9.204	9.207	8.776	8.078	9.251	9.370
Input SNR(dB)			−5.203			
Wavelet decomposition level Wavelet name	haar	db1	sym7	coif2	bior1.1	rbio3.1
Level 1 output SNR(dB)	8.581	8.520	6.750	8.717	8.440	7.448
Level 2 output SNR(dB)	8.094	8.140	8.160	8.483	8.073	8.722
Level 3 output SNR(dB)	8.056	8.015	7.807	7.703	8.048	8.402
Level 4 output SNR(dB)	7.974	7.900	7.764	7.620	7.872	8.361
Level 5 output SNR(dB)	7.874	7.912	7.691	8.035	7.943	8.367
Input SNR(dB)			−7.116			
Wavelet decomposition level Wavelet name	haar	db1	sym5	coif2	bior1.1	rbio3.1
Level 1 output SNR(dB)	5.469	5.603	5.270	6.081	5.670	7.573
Level 2 output SNR(dB)	6.444	6.417	6.516	6.547	6.304	6.791
Level 3 output SNR(dB)	5.920	6.015	5.004	5.410	6.032	6.05
Level 4 output SNR(dB)	5.793	5.836	4.884	5.067	5.75	5.926
Level 5 output SNR(dB)	5.726	5.780	4.419	4.503	5.821	5.877

Figure 10 shows the comparison of the denoising results under the optimal performance of Resnet18, DWT-Resnet18, and DWT-Resnet26. From the comparison, it is clear that the DWT-Resnet method performs well, and DWT-Resnet18 performs best at higher SNRs, while the DWT-Resnet26 performs optimally when the SNR is low.

Figure 11 shows the comparison of single-channel seismic data with clean seismic data after denoising by four denoising methods at different noise levels. As the noise level rises, even the denoised data using DWT-Resnet26 is gradually failing to fit the original seismic data well. Therefore, for high noise level data, further improvement of the denoising performance of this method is necessary.

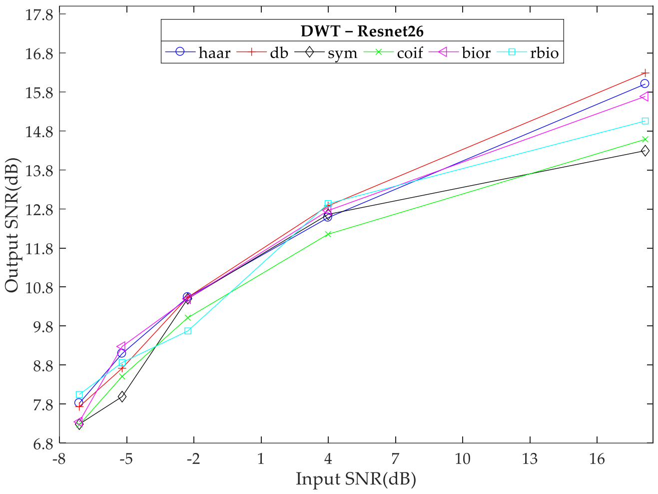


Figure 9. Trends of the best denoising results of DWT−Resnet26 with different SNR real seismic data under different wavelet bases.

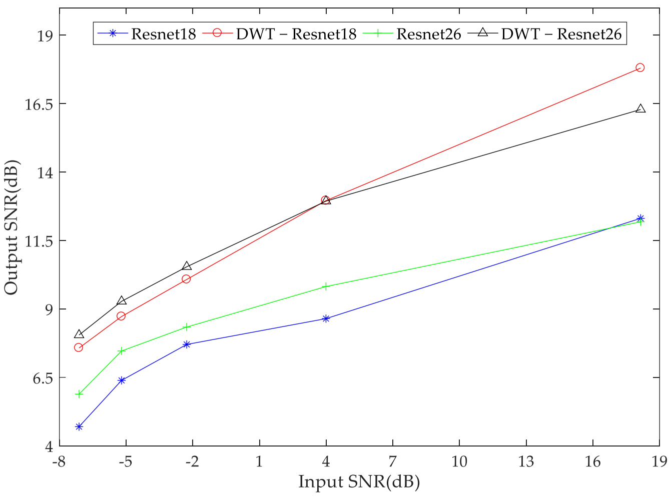
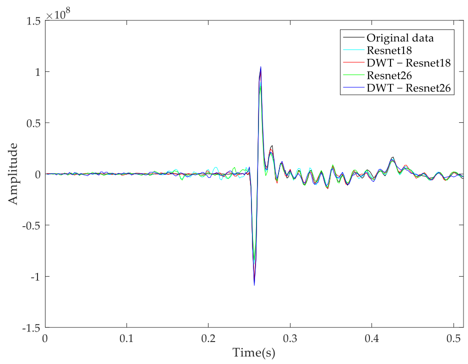
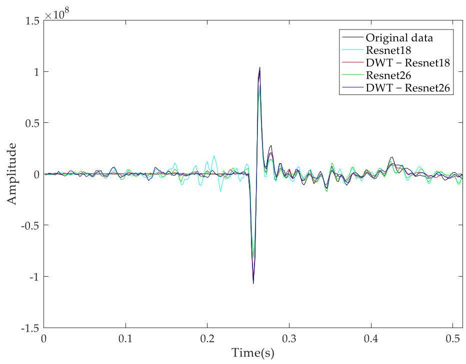


Figure 10. Comparison of denoising results using Resnet18, Resnet26, DWT−Resnet18, DWT−Resnet26 with different input SNRs.

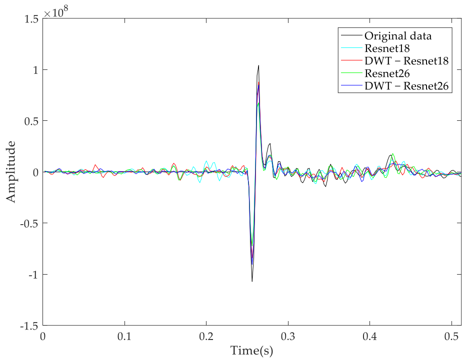


(**a**)

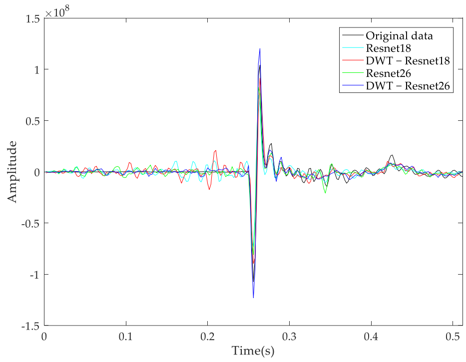


(**b**)

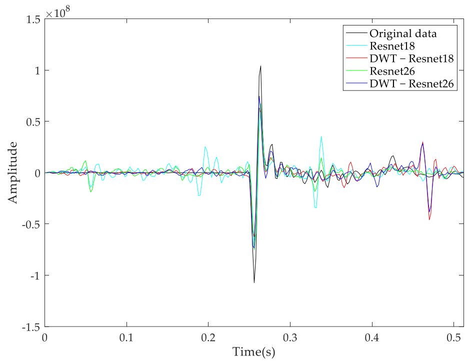
Figure 11. *Cont*.



(**c**)



(**d**)



(**e**)

Figure 11. (**a**–**e**) Illustrates the single−channel real data compared with the original single−channel real data after denoising with four denoising methods at SNRs ranked from high to low.

5. Conclusions and Discussion

In this paper, combining traditional methods and deep learning methods, we propose an improved deep learning denoising method for random noise denoising of seismic data, that is DWT-Resnet. It is reflected in (1) the preliminary processing of noisy seismic data by utilizing the ability of wavelet transform separating the effective signals from noise in traditional denoising methods, and then feeding the processed signal into the residual neural network. With this step, the capability of network denoising is improved. (2) By varying the number of layers of the neural network, a deep network applied to the wavelet domain noisy data is proposed, further improving the denoising effect of the network.

The comparison of the denoising results of the same denoising method for seismic data with different SNRs shows that excellent denoising results can be obtained for high SNR seismic data. As for DWT-Resnet, it can separate the effective signal from the noise by wavelet decomposition, which initially improves the quality of the data entering the network, and therefore improves the denoising ability of the neural network. Yet, as the noise contaminates the data, this ability of wavelets to separate the effective signal from the noise gradually diminishes, making DWT-Resnet subject to this limitation of wavelets. Furthermore, with deeper neural networks, the training time of DWT-Resnet keeps growing, therefore, it is a direction requiring further research to train the network quickly and well. In this paper, an attempt is made to find a wavelet basis and wavelet decomposition levels that can achieve better results for different data at different SNR levels through experiments. However, synthetic seismic data and real seismic data show that there is no generalized rule that can be obtained on which wavelet to choose for different data, at different neural network depths, to obtain the best denoising effect for noisy seismic data. Iterative wavelet base selection trials may remain necessary for applications of denoising other seismic data, which somewhat dissipates the advantage of just using a neural network approach to denoising, i.e., not requiring a more manual selection of parameters. Moreover, with high noise levels, although the DWT-Resnet26 denoising effect is improved compared to Resnet18, the improvement is limited. Therefore, how to develop a new neural network structure to more effectively combine the advantages of wavelet transform and the neural network with a fixed number of wavelet-based transforms and decomposition levels to obtain a more general DWT-Resnet method with better denoising effects will be the focus of the next research work.

Author Contributions: T.L.: conceptualization, methodology, software; Z.Z.: review and guidance, supervision; L.H.: review and guidance, funding acquisition, supervision; J.Z.: data curation, validation. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (grant no. 42074154).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors are very grateful for all constructive comments that helped us improve the original version of the manuscript.

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

References

1. Liu, G.H.; Chen, X.H.; Jing, D.; Wu, K.L. Nonstationary autoregression in f-x domain for random noise attenuation. In *SEG Technical Program Expanded Abstracts*; Society of Exploration Geophysicists: Tulsa, OK, USA, 2011; pp. 3607–3611.
2. Chen, K.; Mauricio, D.S. Making f-x projection filters robust to erratic noise. In *SEG Technical Program Expanded Abstracts*; Society of Exploration Geophysicists: Tulsa, OK, USA, 2014; pp. 4371–4375.
3. Bekara, M.; Van der Baan, M. Random and coherent noise attenuation by empirical mode decomposition. *Geophysics* **2009**, *74*, V89–V98. [CrossRef]
4. Hooshmand, A.; Jalileh, N.; Hamid, R.S. Seismic Data Denoising Based on the Complete Ensemble Empirical Mode Decomposition. In Proceedings of the International Geophysical Conference and Oil & Gas Exhibition, Istanbul, Turkey, 17–19 September 2012; pp. 1–4.
5. Chao, Z.; Li, Y.; Lin, H.B.; Yang, B.J. Seismic Random Noise Attenuation and Signal-Preserving by Multiple Directional Time-Frequency Peak Filtering. *Comptes Rendus Geosci.* **2015**, *347*, 2–12.
6. Wang, Y.H.; Liu, X.W.; Gao, F.X.; Rao, Y. Robust Vector Median Filtering with a Structure-Adaptive Implementation. *Geophysics* **2020**, *85*, V407–V414. [CrossRef]
7. Alsdorf, D. Noise Reduction in Seismic Data Using Fourier Correction Coefficient Filtering. *Geophysics* **1997**, *62*, 1617–1627. [CrossRef]
8. Zhou, J.X.; Lu, W.K.; He, J.W.; Liu, B.; Ren, T. A Data-Dependent Fourier Filter Based on Image Segmentation for Random Seismic Noise Attenuation. *J. Appl. Geophys.* **2015**, *114*, 224–231. [CrossRef]
9. Mohanalin, J.; Prabavathy, S.; Torrents-Barrena, J.; Puig, D. A Novel Wavelet Seismic Denoising Method Using Type II Fuzzy. *Appl. Soft Comput.* **2016**, *48*, 507–521.
10. Kourouniotis, F.P.; Kubichek, R.F.; Boyd III, N.G.; Majumdar, A.K. Application of the wavelet transform in seismic data processing for the development of new noise reduction techniques. In *Wavelet Applications in Signal and Image Processing IV SPIE, Proceedings of the SPIE's 1996 International Symposium on Optical Science, Engineering, and Instrumentation, Denver, CO, USA, 4–9 August 1996*; SPIE: Denver, CO, USA, 1996; pp. 620–631.
11. Oliveira, M.S.; Henriques, M.V.C.; Leite, F.E.A.; Corso, G.; Lucena, L.S. Seismic Denoising Using Curvelet Analysis. *Phys. A Stat. Mech. Its Appl.* **2012**, *391*, 2106–2110. [CrossRef]
12. Zhao, X.; Li, Y.; Zhuang, G.H.; Zhang, C.; Han, X. 2-D TFPF Based on Contourlet Transform for Seismic Random Noise Attenuation. *J. Appl. Geophys.* **2016**, *129*, 158–166. [CrossRef]
13. Liu, J.; Gu, Y.; Chou, Y.; Gu, J. Seismic random noise reduction using adaptive threshold combined scale and directional characteristics of shearlet transform. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 1637–1641. [CrossRef]
14. Liu, L.; Ma, J.; Plonka, G. Sparse graph-regularized dictionary learning for suppressing random seismic noise. *Geophysics* **2018**, *83*, V215–V231. [CrossRef]
15. Nazari Siahsar, M.A.; Gholtashi, S.; Kahoo, A.R.; Chen, W.; Chen, Y.K. Data-Driven Multitask Sparse Dictionary Learning for Noise Attenuation of 3D Seismic Data. *Geophysics* **2017**, *82*, V385–V396. [CrossRef]

16. Li, G.; Müller, M.; Ghanem, B.; Koltun, V. Training graph neural networks with 1000 layers. In Proceedings of the 38th International Conference on Machine Learning. Proceedings of Machine Learning Research. PMLR, Virtual, 18–24 July 2021; pp. 6437–6449.
17. Sarvamangala, D.R.; Kulkarni, R.V. Convolutional neural networks in medical image understanding: A survey. *Evol. Intell.* **2022**, *15*, 1–22. [CrossRef] [PubMed]
18. Dong, Y.; Liu, Q.; Du, B.; Zhang, L. Weighted feature fusion of convolutional neural network and graph attention network for hyperspectral image classification. *IEEE Trans. Image Process.* **2022**, *31*, 1559–1572. [CrossRef] [PubMed]
19. Yuan, Y.; Si, X.; Zheng, Y. Ground-roll attenuation using generative adversarial networks. *Geophysics* **2020**, *85*, WA255–WA267. [CrossRef]
20. Yang, L.; Chen, W.; Liu, W.; Zha, B.; Zhu, L. Random noise attenuation based on residual convolutional neural network in seismic datasets. *IEEE Access* **2020**, *8*, 30271–30286. [CrossRef]
21. Tang, S.; Ding, Y.; Zhou, H.W.; Zhou, H. Reconstruction of sparsely sampled seismic data via residual U-Net. *IEEE Geosci. Remote Sens. Lett.* **2020**, *19*, 1–5. [CrossRef]
22. Zhu, W.Q.; 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]
23. Saad, O.M.; Chen, Y.F.; Savvaidis, A.; Chen, W.; Zhang, F.X.; Chen, Y.K. Unsupervised Deep Learning for Single-Channel Earthquake Data Denoising and Its Applications in Event Detection and Fully Automatic Location. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–10. [CrossRef]
24. Ghodrati, V.; Shao, J.; Bydder, M.; Zhou, Z.; Yin, W.; Nguyen, K.L.; Yang, Y.; Hu, P. MR image reconstruction using deep learning: Evaluation of network structure and loss functions. *Quant Imaging Med. Surg.* **2019**, *9*, 1516–1527. [CrossRef] [PubMed]
25. He, K.; Zhang, X.; Ren, S.; Sun, J. Identity mappings in deep residual networks. In *European Conference on Computer Vision, ECCV 2016*; Springer: Cham, Switzerland, 2016; Volume 9908, pp. 630–645.