Seismic noise attenuation by applying a deep learning method without noise-free labels

Han Wang, Jie Zhang, University of Science and Technology of China (USTC)*

Summary

Seismic data are often interfered by random ambient noise. Noise reduction is a significant effort in seismic data processing. Deep learning methods have been introduced in recent years by training a neural network with noisy data for input and noise-free data for output. However, in practice, it is difficult to obtain noise-free data. In this study, we apply a successful deep learning approach established for denoising of graphic images, which trains the network with noisy data for input, and same data with different noise for output. Because of a large number of training samples associated different random noise but same data, the testing with a new dataset should output data with noise attenuated. Our training data for application include combined naturel images and arbitrary synthetic shot gathers. Testing with synthetics and real data show effective capability to remove random noise and certain coherent impulse as well.

Introduction

In seismic exploration, low signal-to-noise ratio (SNR) data can seriously affect the results of imaging and interpretation (Huang et al., 2017). Noise attenuation and improvement of SNR is one of the major efforts in seismic data processing (Gulunay, 2000; Naghizadeh and Sacchi, 2012). In addition, the coherent noise which is different from data, may also affect the seismic data imaging if it is not attenuated in processing.

Many efforts have been made to solve the problem of noise attenuation and data quality improvement. Traditional attenuation algorithms are mainly based on the transform domain in which signal and noise have different characteristics, and the methods have proven effective in many applications. These include bandpass filtering (Gulunay, 1986), f-x prediction (Canales, 1984), and t-x prediction (Abma and Claerbout, 1995). Other sparse transformation algorithms include the seislet transform (Fomel and Liu, 2010), EMD-seislet transform (Chen and Fomel, 2018), wavelet transform (Liu et al., 2016b), and curvelet transform (Candes et al., 2006). On the other hand, seismic noise varies in different environments, and it deserves further efforts for alternative solutions.

Convolutional Neural Network (CNN) has contributed to solving the noise attenuation problems in recent years. CNN has been widely used in many applications in the field of seismic data processing and imaging. The denoising methods based on deep learning do not require *a priori* noise information or selecting parameters on a case-by-case basis, and can be applied for adaptive noise reduction. In general, deep learning based methods require a large number of noisy and noise-free sample pairs for input and output. However, it is often difficult to obtain noise-free or high-quality waveforms from real data for training.

In this study, we apply a graphics image denoising technique, noise2noise, developed by Lehtinen et al. (2018) to solve a seismic denoise problem. The method takes noisy images as input and the same images with different noise for output. Lehtinen et al. (2018) observe that the network can learn to restore images by only looking at corrupted images, sometimes even better than using noise-free images for training. In the seismic denoise problem, we use a dataset that includes naturel images and arbitrary synthetics along with random noise to train the super resolution residual network (SRResNet, Ledig et al., 2016). Synthetic and field examples demonstrate that the approach is robust.

Method

Traditional neural network denoising methods generally use noisy and noise-free sample pairs as the input and output. Neural network is then trained to fit the mapping relationship between the input and output. Take the $(\hat{\boldsymbol{x}}_i, \boldsymbol{y}_i)$ pair in CNN for example, $\hat{\boldsymbol{x}}_i$ represents the corrupted input and $\boldsymbol{y}_i$ represents the noise-free label. The empirical risk minimization task is:

$$\underset{\theta}{argmin} \sum_i L(f_\theta(\hat{\boldsymbol{x}}_i), \boldsymbol{y}_i) \ , \quad (1)$$

where f_θ is a parametric family of mappings under the loss function L. In practice, it is always difficult to get noise-free signals. Therefore, instead of establishing a mapping relationship between noisy images and noise-free labels, we set the noise-free image hidden by noise as a new label. The empirical risk minimization task is changed to (Lehtinen et al., 2018):

$$\underset{\theta}{argmin} \sum_i L(F_\Theta(\hat{\boldsymbol{x}}_i), \hat{\boldsymbol{y}}_i) \ , \quad (2)$$

where $F_\Theta(*)$ represents the network function parameterized by Θ and L represents the loss function. $\hat{\boldsymbol{x}}_i$ and $\hat{\boldsymbol{y}}_i$ represent the input and output, respectively. They share the same underlying clean image but with different noise.

Seismic data $\hat{\boldsymbol{x}}$ is usually expressed as the superposition of effective reflected signal $\hat{\boldsymbol{s}}$ and the additive noise **n**:

$$\hat{\boldsymbol{x}} = \hat{\boldsymbol{s}} + \boldsymbol{n} \ , \quad (3)$$

To predict the possible clean seismic signal, the SRResNet (Ledig et al., 2016) learns the mapping relationship in equation (2) by taking the noisy data $\hat{\boldsymbol{x}}$ as input. The denoised seismic image **s** is represented by:

$$\boldsymbol{s} = F_{\Theta}(\hat{\boldsymbol{x}}) \,, \tag{4}$$

The mapping function $F_{\Theta}(*)$ is learned by the SRResNet. SRResNet contains multiple residual blocks, which show advantages in learning high frequency details and residual network (He et al., 2016). In the training process, ADAM (Kingma & Ba, 2015) is selected to optimize network parameters. The loss function is mean square error (MSE):

$$L(\Theta) = \frac{1}{N}\sum_{i=1}^{N}(F_{\Theta}(\hat{\boldsymbol{x}}_{\boldsymbol{i}}) - \hat{\boldsymbol{y}}_{\boldsymbol{i}})^2 \,, \tag{5}$$

where Θ represents the trainable parameters. The architecture and parameters of SRResNet are the same as those used in Ledig et al. (2016).

Example

The training dataset includes both natural graphics images and synthetic seismic data with 64*64 size. Graphics images are randomly extracted from the well-known BSD300 dataset (Martin et al., 2001), which contains 300 natural images with complex features, and the arbitrary synthetic 2-D seismic dataset including 1348 shot gathers is downloaded from SEG (Society of Exploration Geophysicists) open datasets. The features in natural images are likely to be more complex than those in real seismic images. The purpose of including synthetic 2-D seismic images in training is to make the network more suitable for our denoising task. For each patch, we add the Gaussian white noise at two different noise levels, and set the standard deviation of the noise between 0 and 50. Figure 1 shows an example of synthetic training patches along with different noise for input and output. In order to deal with coherent impulse noise, we also train the case of adding impulse noise to the dataset. Note that the output image can also be noisier than the input image.

We present two synthetic tests and one field example to demonstrate the performance of our method. In both numerical examples, we use the same geometry with 64 receivers and the same velocity model with two layers. The sample interval is 2 ms. The dominant frequencies of the three events are 38 Hz, 32 Hz, and 30 Hz, respectively.

1) Numerical example 1

In the first synthetic test, we add coherent impulse noise to the clean seismic record, which is generally encountered in the reprocessing of the original seismic data collected in the past. The noise shows the following characteristics: spike;

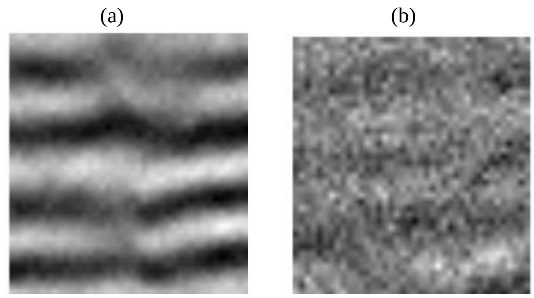

Figure 1: An example of synthetic training datasets. (a) is the noisy image as input and (b) is the same image with different noise for output.

coherence; amplitude variable; the position of noise is uncertain; short duration (Jing, 2001). Since the noise has the characteristics of sharp pulse and infinite frequency band, it may affect the subsequent processing if it is not eliminated completely in the pretreatment.

Figure 2 shows the denoising effect of the proposed method and the bandpass method. Figure 2(b) is the seismic records with coherent pulse interference. Figure 2(c) and Figure 2(d) are the results of the bandpass method (10-80 Hz) and our method, respectively. The bandpass method can attenuate part of the impulse noise, but the remaining noisy part may still affect stacking process. The noise2noise method performs better in removing this type of noise. It does not need to select parameters and has the ability of adaptive denoising.

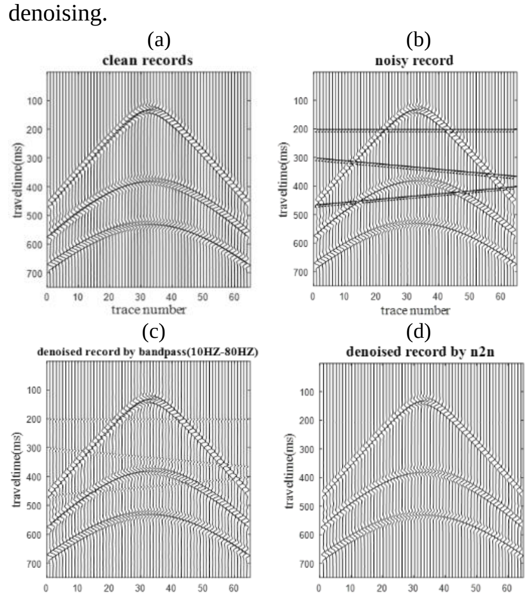

Figure 2: The denoising results of the bandpass (10-80 Hz) (c) and noise2noise (d) method.

(d)

2) Numerical example 2

The signal-to-noise ratio (SNR) is used to partly judge the quality of denoised results. It is defined as:

$$SNR = 10log_{10}\left(\frac{\|\boldsymbol{A}\|_F^2}{\|\boldsymbol{A}-\boldsymbol{B}\|_F^2}\right), \tag{6}$$

Where matrix **A** and **B** represent the clean data and denoised data, respectively.

In the second synthetic test, we add strong Gaussian noise to a clean 2-D synthetic seismic dataset. As shown in Figure 3 (a) and (b). The SNR of noisy seismic data is -4 db. We compare the noise2noise method with f-x deconvolution and Time Variant Spectral Whitening (TVSW) methods. The length of the autoregressive operator of f-x deconvolution is 32 ms. The frequency range of TVSW is 8-48 Hz. The results of three methods are shown in Figure 3. The recovery quality of f-x deconvolution to three seismic events is disturbed by strong noise. The TVSW produces high resolution denoised results, but it loses a lot of power of effective signals. Our method shows great improvement in effective signals reservation, resolution, and high signal-to-noise ratio. Table 1 shows the signal-to-noise ratio of the denoising results of the above three methods.

Table 1: The SNR of f-x deconvolution, TVSW and the proposed method.

Method	noisy	F-x deconvolution	TVSW	Noise2noise
SNR(dB)	-4	-0.4686	0.3147	10.4754

Furthermore, we conduct other tests on the synthetic data example (Figure 2a) with Gaussian white noise at different levels, and compare them with the bandpass method to prove the effectiveness of the proposed method on SNR.

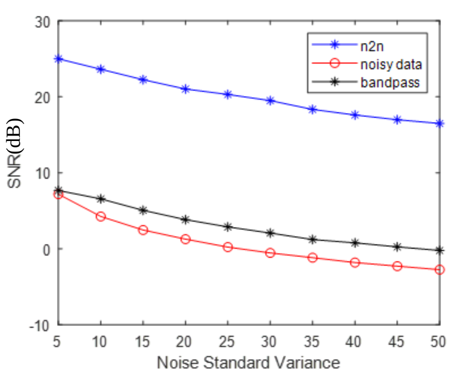

Figure 4: SNR after applying bandpass (8-56 Hz) and the proposed method with different noise levels.

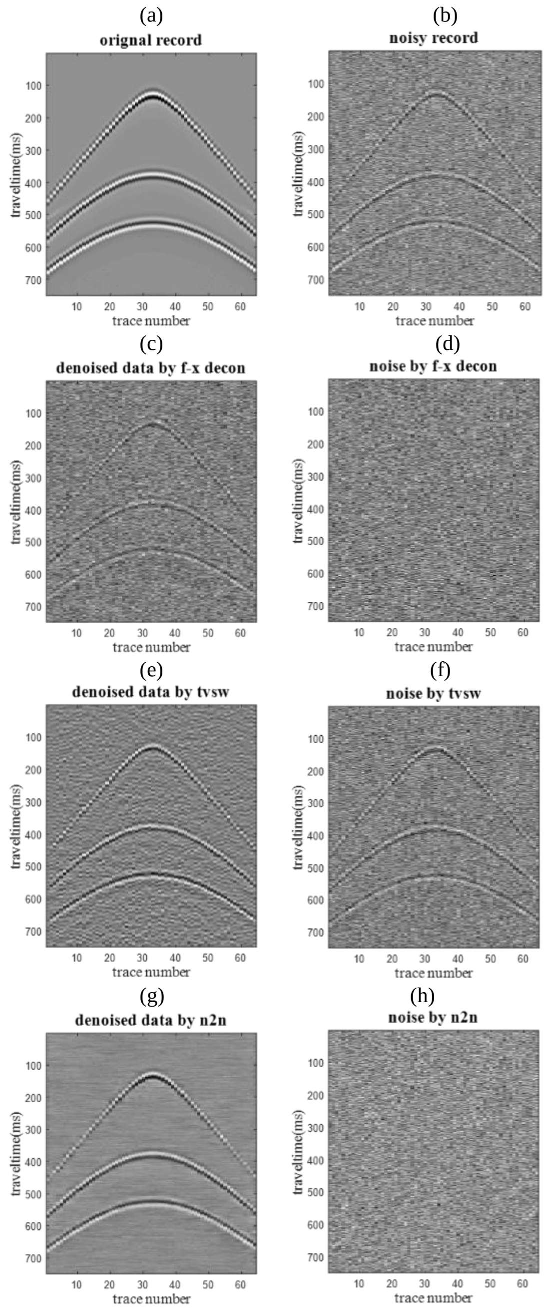

Figure 3: The noise attenuation results and difference for synthetic record. (a) clean 2-D seismic data. (b) noisy data with SNR=-4dB. (c) Result after applying f-x deconvolution. (d) Difference between (c) and (b). (e) Result after applying the TVSW. (f) Difference between (e) and (b). (g) Result after applying our method. (h) Difference between (h) and (b).

3) Field example

We further apply our method to the field data acquired on land as shown in Figure 5(a). Comparing the bandpass (8-56 Hz) with the proposed method, Figure 5(b) and Figure 5(c) show the results of the two methods, respectively. It can be seen from the shot gathers that the two denoising results show improved data quality. In the red rectangular box, the noise2noise approach shows better enhancement for reflections.

Conclusions

We apply a deep learning algorithm to attenuate seismic noise without using noise-free labels, i.e., taking noisy data as input and the same data with different noise as output. The SRResNet neural network is used to extract complex features from both natural images and synthetic seismic data. Strong Gaussian white noise and pulse noise at different noise levels are injected for training. The numerical examples indicate that our approach can remove certain coherent impulse noise and has obvious attenuation effect on random noise with high noise level. Finally, the method is applied to the field data. The comparison with the bandpass method shows that our method can help make improvement. We also point out that this method can only remove the types of noise that have been trained.

Acknowledgments

We would like to appreciate GeoTomo Hefei for the support of our study, allowing us to utilize the TomoPlus software to perform our research.

(a) (b) (c)

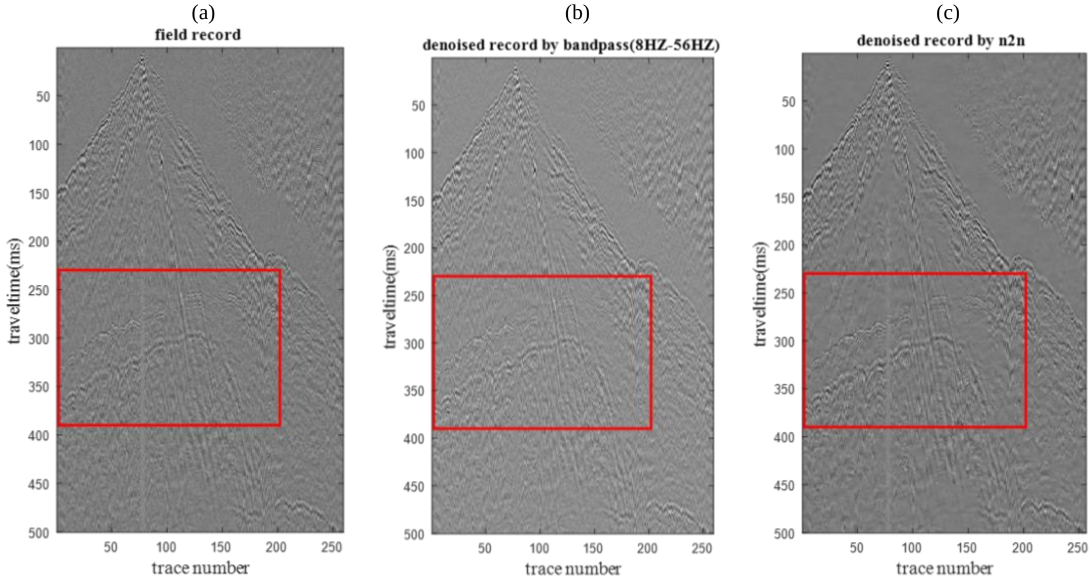

Figure 5: The field seismic data. Denoised result by (b) bandpass (8-56 Hz) and (c) the proposed method.