Must we have labels for denoising seismic data based on deep learning?

Dawei Liu[1], Zheyuan Deng[1], Xiaokai Wang[1], Wei Wang[1], Zhensheng Shi[1], Cheng Wang[2], Wenchao Chen[1]*
[1]Xi'an Jiaotong University, National Engineering Laboratory for Offshore Oil Exploration, Xi'an, China
[2]Daqing Oilfield Company Ltd., Daqing, China

Summary

Deep convolutional networks (DCN) have been successfully applied to seismic data denoising by training on some large datasets. The excellent denoising performance of convolutional networks is mainly imputed to their powerful ability to learn seismic data priors from some training samples. However, "clean" seismic datasets are very difficult to obtain, which greatly limits the widespread use of convolutional networks. In this abstract, we propose an unsupervised denoising method based on DCNs. We show that the structure of a generator network itself can be used as the prior knowledge for seismic data denoising. Both pre-stack field seismic data example and post-stack field seismic data example demonstrate the validity of our methods.

Introduction

Suppressing random noise has always been a hot topic in seismic signal processing because it will improve results of seismic inversion and other processes. Pre-stack denoising and post-stack denoising are two typical methods for suppressing random noise. Pre-stack denoising is performed on the original seismic records or various gatherings, such as common-reflection-point (CRP) gathers, and it provides a good foundation for the following pre-stack processing methods such as velocity analysis and deconvolution processing. Post-stack denoising can further suppress random noise and improve the signal-to-noise ratio.

In order to suppress random noise in seismic data, many advanced denoising techniques have been proposed. Generally speaking, it can be divided into learning-free methods and learning-based methods.

Learning-free methods are general to handle different seismic data denoising problems and have clear physical meaning. Therefore, they have been successfully applied to denoise seismic data. Principle component analysis (Hagen, 1982), singular value decomposition (Bekara and Van der Baan, 2007) and the Cadzow filter (Trickett, 2008) are some typical learning-free random noise denoising methods which use the correlation between seismic data. Sparse regularization in the transform domain, such as wavelet domain (Jian et al., 2006) and curvelet domain (Neelamani et al., 2008), was also proposed to suppress random noise and achieved good results in practical application. However, they have some common shortcomings. On the one hand, the above methods need prior knowledge to construct models. The hand-crafted prior may not be strong enough so that they cannot sufficiently capture some complicated structures of field seismic data, which reduced their denoising performance. On the other hand, they are time-consuming in the process of model optimization.

Learning-based methods can automatically learn the knowledge needed to suppress random noise from datasets and achieve great denoising performance. As one typical representative of the learning-based methods, DCNs have attracted the attention of many researchers. Liu et al. (2018) constructed some training sample datasets by a complex workflow and trained a 3D-DnCNN to denoise post-stack seismic data. More arc-like random noise can be suppressed by 3D-DnCNN. Li et al. (2018) used the LTFD method to obtain labels in advance and trained a deep residual learning network. The residual network can successfully denoise the scattered ground-roll noise. Although DCNs can achieve satisfying denoising result for seismic data, there are several issues that may limit its further promotion. The most important problem is that it requires a lot of training samples, which is very difficult to satisfy in many cases. Secondly, the generality of well-trained networks is limited, which means that its denoising function for large-scale seismic data is likely to fail.

In this abstract, we propose a new denoising method for seismic data which combines the merits of the two aforementioned learning-free and learning-based methods. Inspired by Ulyanov et al. (2018), we use an untrained generator network with some randomly initialized inputs to learn noisy seismic data. The generator network offers high impedance to the noise and low impedance to the valid signal. Although the final output of the network is the same noisy seismic data, the generator network can obtain different intermediate results in the process of learning. Therefore, we can select the specific number of iterations in the optimization process as denoising results by using early stop to avoid overfitting. We found that the network has some satisfying denoising effects on pre-stack and post-stack datasets, especially for post-stack arc-like noise.

Method

In this section, we present a novel denoising method based on the deep generator network. Unlike previous learning-based methods, the method we proposed does not require the label dataset. Firstly, we investigate the network architecture used in this abstract. Then, our network is trained to fit noisy dataset. With the help of the early stop, the network can output one denoised seismic dataset.

 10.1190/iwmg2019_08.1

Network architecture

The seismic data, which is contaminated by random noise, can be expressed as :

$$x_0 = x + n \tag{1}$$

where x is the valid signal, x_0 is the noisy signal, and n is the random noise. The basic idea of the generator network is to map a code vector z to seismic data x as follows:

$$x = f_\theta(z) \tag{2}$$

where z is a fixed randomly initialized vector, θ is the network parameters including the weights and bias of the convolutional filters in the network and $f_\theta(\cdot)$ stands for the network architecture. Then, we build the reconstruction problem model: given seismic data x_0, we want to find the θ^* to rebuild x_0 as follows:

$$\theta^* = \arg\min_\theta E\left(f_\theta(z); x_0\right), \qquad x^* = f_{\theta^*}(z) \tag{3}$$

where $E\left(f_\theta(z); x_0\right)$ is an energy function used to measure the differences between $f_\theta(z)$ and x_0. In this abstract, $E\left(f_\theta(z); x_0\right)$ is formulated as follows:

$$E\left(f_\theta(z); x_0\right) = \left\| f_\theta(z) - x_0 \right\|^2. \tag{4}$$

An U-Net type network including several convolutional layers, downsampling layers, upsampling layers, batch normalization, and skip-connections layers are adopted in this abstract. As shown in Figure 1, the network has 5 depths. The number of filters increases from 8 to 128 with the increase of depth-i, while the output size of each layer decreases by half with the increase of depth-i. This enhances the network's ability to extract some features at different scales and compromises calculation cost. In order to weaken the gradient vanishing and speed up network training, we use some skip connections at depth-4 and depth-5. As for downsampling and upsampling technique, we simply use the strides implemented within the convolution process and bilinear upsampling, respectively. In addition, reflection paddings instead of zero paddings in convolution layers are adopted in order to keep the output size the same as the input size.

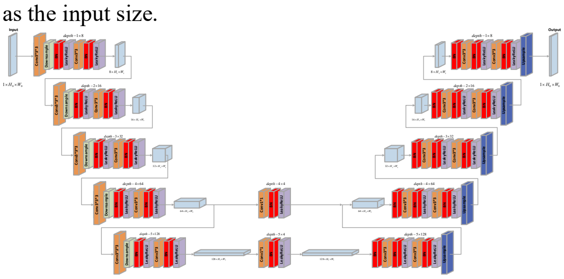

Figure 1: The network structure of our denoising method.

Model training

Seismic data is usually three-dimensional (3D) or higher. However, our network is two-dimensional (2D). We select 2D slices from 3D seismic dataset and denoise them slice by slice. At each denoising process, z is fixed and randomly initialized with the uniform noise between 0 and 0.1.

In order to find θ^* in equation (3), we use ADAM optimizer to iteratively solve the optimization problem. Given almost any seismic data x_0, the network can find optimal θ^* to recovery it after enough number of iterations. However, although the network can finally fit almost any seismic data, the process of iterative solution is different.

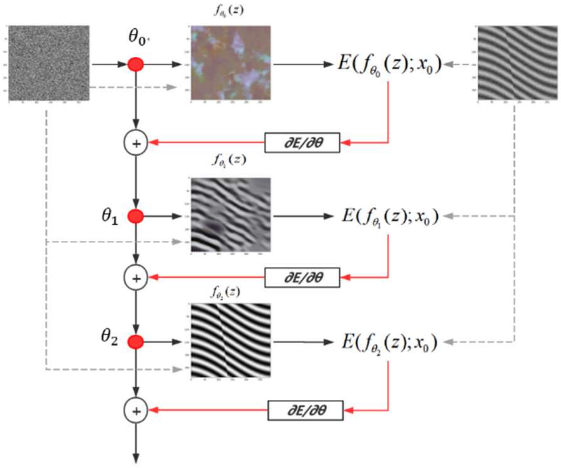

Figure 2: The optimization process of reconstructing noisy synthetic seismic data.

We take the reconstruction problem of noisy synthetic seismic data as x_0 to illustrate the phenomenon as shown in Figure 2. At the start stage of the optimization process (corresponds to a very small number of iteration), the output of the network cannot reconstruct the target x_0 well. As the number of iterations increases, the resulted approximation recreates the shape of the objects but still corrupted. As the iteration process continues, there are only the valid signal and very little noise in the output of the network. The above optimization process shows that the network resists “bad" solutions and tends to generate the seismic valid signal. To demonstrate it, we do an experiment by taking random noise and seismic valid signals as different choices for x_0, as shown in Figure 3. Obviously, the valid signal converges much faster than noise, which means the network offers high impedance to noise and low impedance to valid signals. Therefore, there is a range of iterations in which the network fits the effective signal well, but it has not yet begun to fit the noise. By choosing the appropriate number of iterations, the network can realize the function of denoising and output the denoised seismic valid signal.

10.1190/iwmg2019_08.1

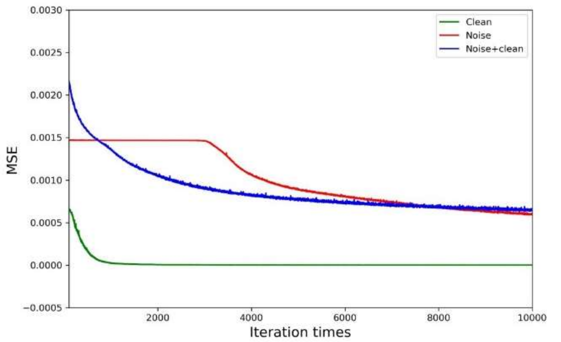

Figure 3: Learning curves for the reconstruction task using: seismic valid signal, random noise, and noisy seismic dataset.

Examples

We tested the denoising performance of our network on one pre-stack and post-stack field dataset, respectively. Figure 4 shows the denoising result of a CRP gather. The Cadzow filter method is applied to the CRP gather as shown in Figure 4b and Figure 4c. It can be seen that the random noise is successfully removed from the original gather. The denoising performance of our method is shown in Figure 4d and Figure 4e. It can be observed that our method is better than the Cadzow filter since it removes more random noise and has greater fidelity.

Figure 5 shows a denoising result for a seismic profile from a post-stack seismic dataset in Daqing Oilfield. It can be seen that random noise has been removed effectively. In addition, arc-like noise that is difficult to handle by the conventional method is suppressed well. The above two experiments prove the validity of our method.

Conclusions

In this abstract, we propose a random noise removal strategy in seismic data based on a generator network. Learning-based methods usually require a large amount of samples. On the contrary, our method only utilizes the noisy seismic data for denoising. We use the generator network to reconstruct the noisy seismic data and obtain the denoised seismic data in the intermediate process of optimization iteration. Two field seismic data examples demonstrate the validity of our method. In addition, our method has a strong ability to suppress arc-like noise of post-stack seismic data.

Acknowledgements

The research was funded by National Natural Science Foundation of China (41774135, 41504092, 41504093), National Key Research and Development Program of China (2017YFB0202902), and the Fundamental Research Funds for the Central Universities. We would like to thank Daqing oilfield for providing the dataset.

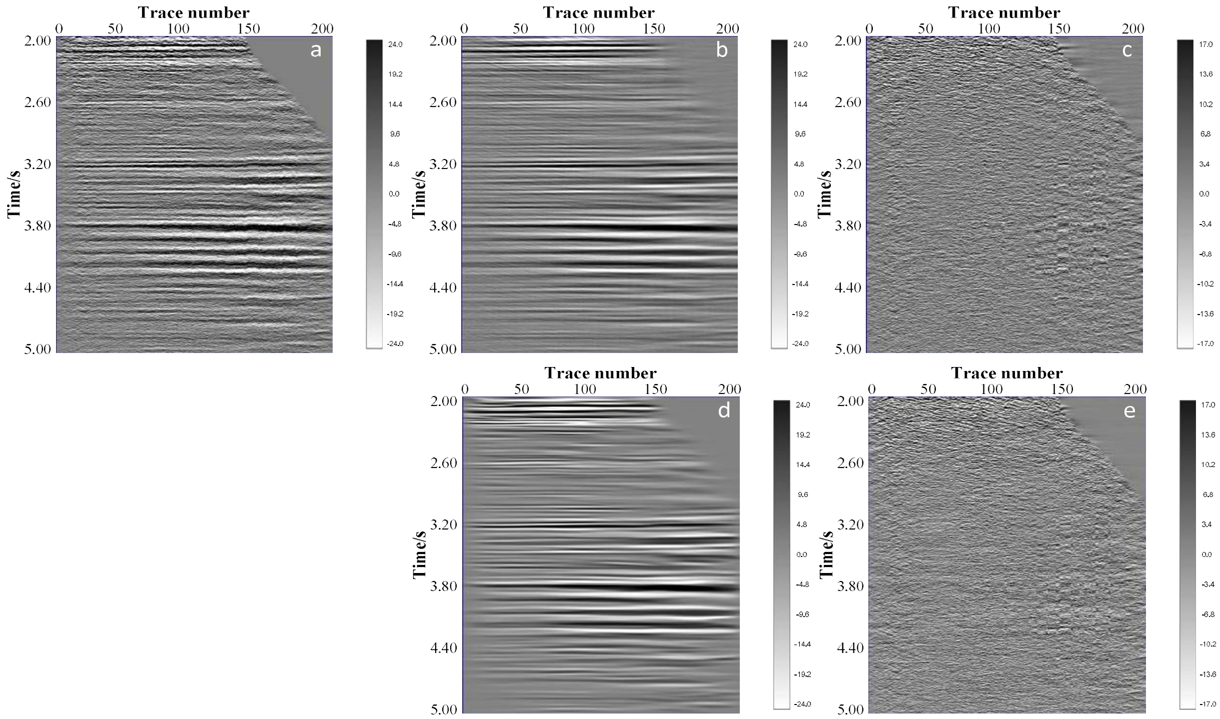

Figure 4: Denoising results for a pre-stack CRP gather. (a) an original gather. (b) denoised by Cadzow filter. (c) the noise removed by Cadzow filter. (d) denoised by our method. (e) the noise removed by our method.

10.1190/iwmg2019_08.1

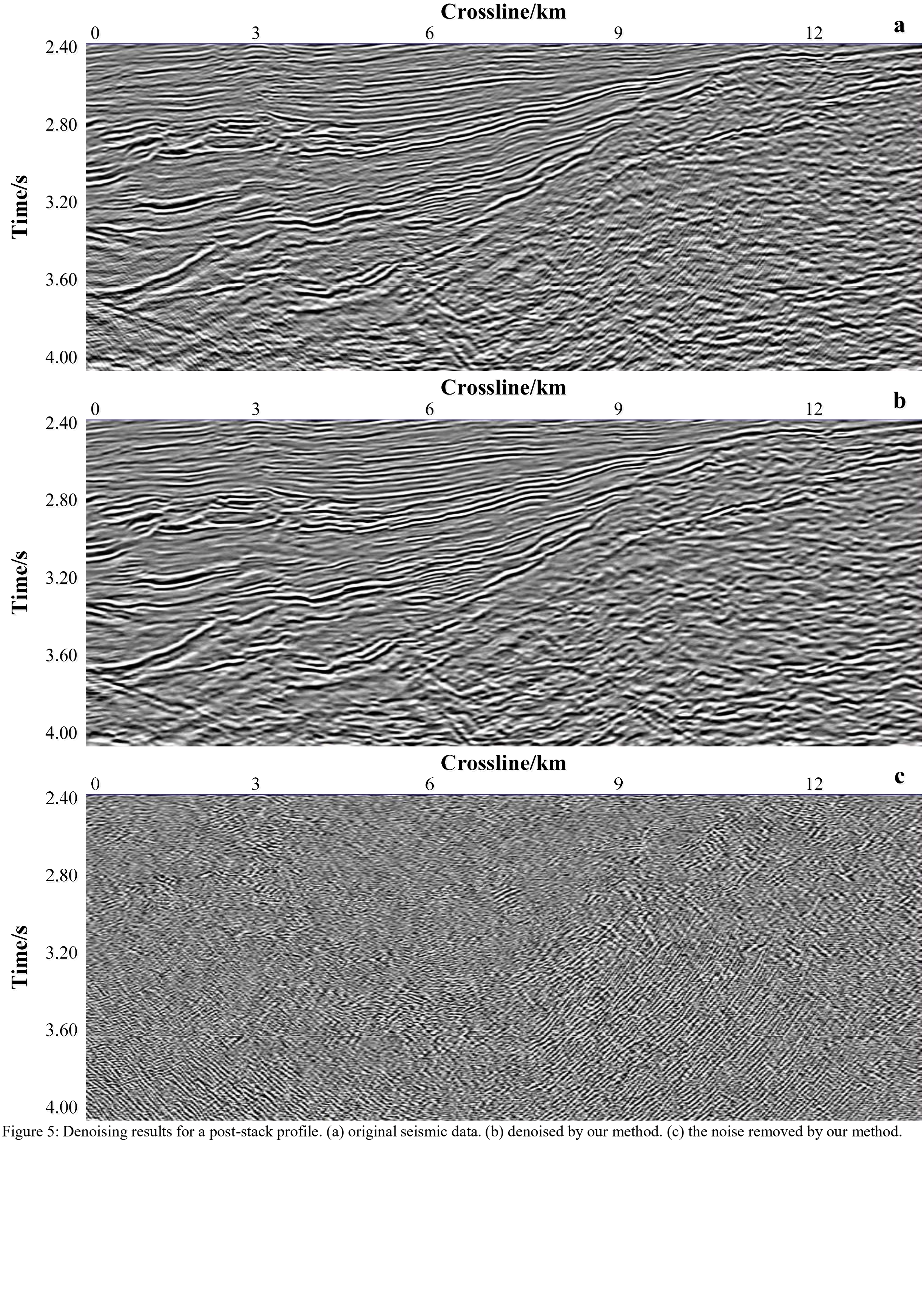

Figure 5: Denoising results for a post-stack profile. (a) original seismic data. (b) denoised by our method. (c) the noise removed by our method.

REFERENCES:

Bekara, M., and M. Van der Baan, 2007, Local singular value decomposition for signal enhancement of seismic data: Geophysics, **72**, no. 2, V59-V65, https://doi.org/10.1190/1.2435967.

Hagen, D. C, 1982, The application of principal components analysis to seismic data sets: Geoexploration, **20**, no. 1, 93-111, https://doi.org/10.1093/gji/ggx325.

Li, H., W. Yang, and X. Yong, 2018, Deep learning for ground-roll noise attenuation: 88th Annual International Meeting, SEG, Expanded Abstracts, 1981-1984, https://doi.org/10.1190/segam2018-2981295.1.

Liu, D., W. Wang, W. Chen, X. Wang, Y. Zhou, and Z. Shi, 2018, Random noise suppression in seismic data: What can deep learning do?: 88th Annual International Meeting, SEG, Expanded Abstracts, 2016-2020, https://doi.org/10.1190/segam2018-2998114.1.

Mao, J., J. Gao, and W. Chen, 2006, On the denoising method of prestack seismic data in wavelet domain: 76th Annual International Meeting, SEG, Expanded Abstracts, 2852-2855, https://doi.org/10.1190/1.2370118.

Neelamani, R., A. I. Baumstein, D. G. Gillard, M. T. Hadidi, and W. L. Soroka, 2008, Coherent and random noise attenuation using the curvelet transform: The Leading Edge, **27**, no. 2, 240-248, https://doi.org/ 10.1190/1.2840373.

Trickett, S, 2008, F-xy Cadzow noise suppression: 78th Annual International Meeting, SEG, Expanded Abstracts, 2586-2590, https://doi.org/10.1190/1.3063880.

Ulyanov, D., A. Vedaldi, and V. Lempitsky, 2017, Deep image prior: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, https://doi.org/10.1109/CVPR.2018.00984.