Journal of Management & Entrepreneurship
ISSN 2229-5348

UGC Care Group I Journal
Vol-09 Issue-02 Oct 2020

Deep Learning for UnsupervisedSeismic Data

Debarchana Jena
College of Engineering Bhubaneswar

***Abstract*Denoising is a crucial step in the processing of seismic data. The exceptional performance of deep-learning-based seismic data denoising has drawn notice recently. We examine the deep Convolutional Networks (ConvNets) architecture for seismic data denoising in this letter. A generative network with Gaussian noise is applied to a single seismic data profile using the untrained ConvNets. Generative networks with varied handmade designs, starting with random initialized parameters, differ in their capacity to map the seismic data at iterations and separate the Gaussian noise as residuals. The depth, width, and skip connection—the three primary building blocks of a generative network—are created as different architectures to fit noisy, clean, and Gaussian noise seismic data, respectively, in order to investigate the capacity of Gaussian noise separation. Then, as a previous model to the seismic data denoising task, the advantageous network architecture with high and low impedance (a capability to impede data reconstruction) to noise and seismic data is used. In addition, a stopping criterion is created for the data fitting procedure in order to automatically extract the latent clean seismic data. The suggested approach uses network architecture as before and does not require data sets for training. The efficiency of the chosen ConvNet is demonstrated by extensive experiments on both synthetic and field data, and the benefits are assessed by contrasting the denoising outcomes with f-x multi-channel singular spectrum analysis (MSSA) and the most advanced unsupervised neural network (NN)-based technique..**

***Index Terms*— Generative network, network architecture prior, seismic data denoising, unsupervised learning.**

I. Introduction

SEISMIC data denoising is a classical yet still active topic, since the presence of noise will affect subsequent imaging, inversion, interpretation, and other operations [1]–[3]. The seismic data noise is divided into coherent and incoherent noise [4], [5]. For the purpose of this letter, the removal of incoherent noise is considered.

The existing traditional methods for seismic data denoising can be mainly classified into three categories. The first category generally builds a predictive filter for seismic data denoising, such as f-x deconvolution [6], nonstationary predictive filtering [7], and t-x predictive filtering [8]. Considering that the seismic data and noise can be distinguishable in some sparse transform domains, the second category including Fourier transform method [9], wavelet transform [10], [11], curvelet transform [12], dreamlet transform [13], [14], and dictionary learning-based sparse transform [15], [16] are used to enhance the sparsity and then attenuate seismic noise [17]. The third category based on rank reduction has developed rapidly in recent years. It uses the low-rank characteristic of seismic data to remove noise during the rank reduction process, such as multi-channel singular spectrum analysis (MSSA) [18] and Cadzow filtering [19]. Besides the mentioned methods, some hybrid methods are also designed to fulfill the seismic data denoising task [5], [20], [21].

With the success of deep learning in solving computer vision problems, its popularity in seismic data denoising field spread rapidly. The deep learning methods for seismic data denoising can be divided into the supervised methods [22]–[25] and unsupervised methods [26]. The supervised methods usually aim to obtain a nonlinear mapping from noisy seismic data to clean seismic data by high-quality training set and handcrafted regularizer. However, the difficulty of acquiring a large number of label data sets hinders its application to field data. In contrast, the unsupervised methods only use the raw seismic data and can also attenuate the random noise effectively. Due to the promising effects and convenience, the unsupervised methods have been studied intensively recently.

In the field of image processing, Ulyanov *et al.* [27] proposed an unsupervised method called deep image prior (DIP) and designed a general network framework (Fig. 1) for image processing tasks such as super-resolution, inpainting, and denoising. The main components of the framework include the number of filters at depth i for up-sampling ($n_u[i]$), down-sampling ($n_d[i]$), and skip connections ($n_s[i]$). This unsupervised method uses the deep Convolutional Networks (ConvNets) as generative networks to obtain the target image during iterations starting from a single degraded image. They also show that the network architecture serving as prior affects the result of image processing. Wavefield or seismic data are special data set. They are band-limited and the time–space distribution must observe causality which is dictated by wave equations [28]. It decides that the suitable network architecture for unsupervised seismic data processing is different from image processing. Inspired by the outstanding performance of DIP achieved in image processing, Liu *et al.* [29] have designed a special network architecture (deep seismic prior) to seismic missing trace interpolation.

In this letter, we explore the ability of generative network architectures to separate seismic data and random noise. In addition to the network depth and width, we also test the impedance to noise and seismic data with or without skip connection. In this way, a generative network architecture beneficial to attenuate Gaussian noise in seismic data is determined and served as prior model for seismic dat

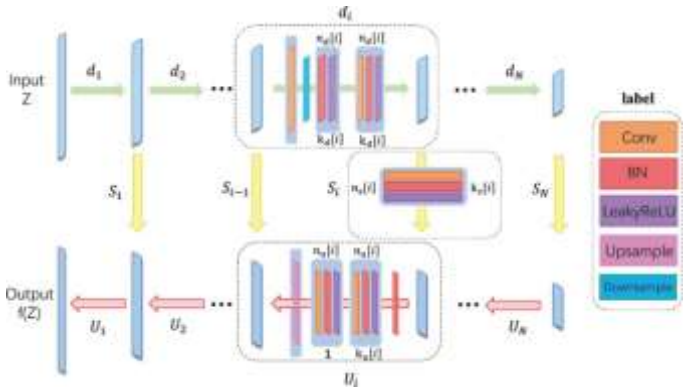

Fig. 1. Illustration of the general DIP architecture for image processing [27] and for image denoising, $n_d[i] = n_u[i] = 128; n_s[i] = 4$, and $k_d[i] = k_u[i] = 3; k_s[i] = 1$ correspond to the respective kernel sizes.

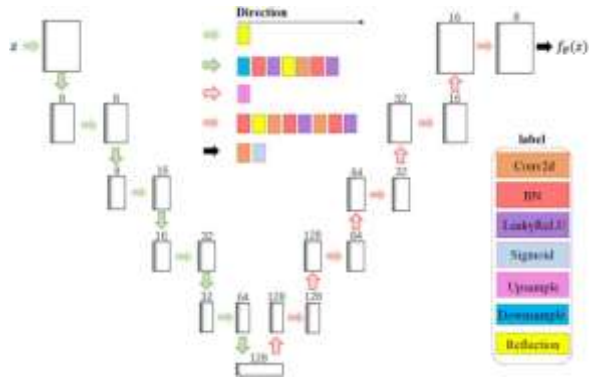

Fig. 2. Illustration of the selected ConvNet architecture for seismic random noise attenuation.

denoising. Meanwhile, we design a stopping criterion for the fitting process to obtain the latent clean seismic data. We apply the favorable generative network to both noisy synthetic and field data sets to demonstrate the denoising performance.

II. Prior Model for Seismic Data Denoising

A. Objective Function

For seismic data denoising, we need to recover clean seismic data x from the contaminated observation y

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

where n denotes the random noise. To recover the clean data, conventional deep learning methods use the optimization function which is defined as

$$\min_x E(f(y), x) + R(x) \tag{2}$$

where f denotes the nonlinear mapping of ConvNet and $R(x)$ stands for the regularizer. In addition to the network architecture, the regularizer based on the prior of the target data or noise also plays an essential role, that is, the performance relies not only on the network architecture but also on the handcrafted regularizer. Here, the explicit regularizer can be dropped and substituted by the implicit prior of data captured by the network parameters [27], as follows:

$$\theta^* = \arg\min_\theta E(f_\theta(Z), y) \tag{3}$$

where Z and θ denote the randomly initialized input and network parameter, respectively. Equation (3) indicates that we can only use the contaminated observation y to obtain a local minimizer θ^* and the optimization problem can be effectively solved by the existing optimizer such as Adam.

B. Favorable Network Architecture for Seismic Data Denoising

Benefiting from the high and low impedance to random noise and image, respectively, the architecture displayed in Fig. 1 has good performance on image processing tasks. Due to the difference between seismic data and image, the network architecture must be adjusted to seismic data processing. We divide the network frame (Fig. 1) into three parts: the depth $[l]$(the total numbers of down-sampling or up-sampling layer), the width $[n_d[i], n_u[i]]$ (the numbers of filter at depth i for down-sampling and up-sampling, where $0 < i \leq l$), and the skip connection structure. As shown in the rest of

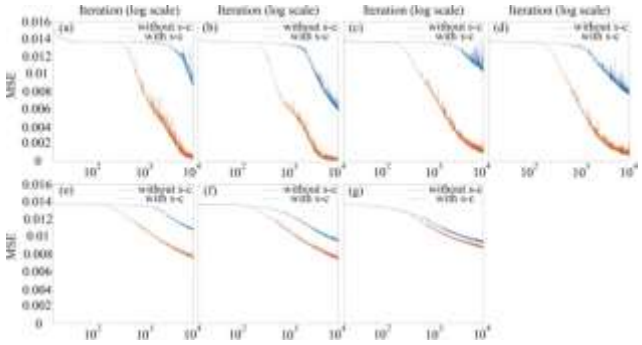

Fig. 3. Noise impedance of various generative networks. (a) $n_d = n_u = [128, 128, 128, 128, 128]$. (b) $n_d = n_u = [128, 128, 128, 128]$. (c) $n_d = n_u = [64, 64, 64, 64, 64]$. (d) $n_d = n_u = [64, 64, 64, 64]$. (e) $n_d = n_u = [8, 16, 32, 64, 128]$. (f) $n_d = n_u = [8, 16, 32, 64]$. (g) $n_d = n_u = [8, 16, 32]$.

this section, ConvNet with $n_d = n_u$ [8, 16, 32, 64, 128] and without skip connection can be a good candidate prior model for seismic data denoising. The detailed network architecture is shown in Fig. 2. It possesses sufficient quality to remove random noise from seismic data effectively. To prove it, Gaussian noise, clean, and noisy seismic data are reconstructed by various generative networks to show the advantages of selected generative network architecture. Here, we use the learning curve (mean square error (MSE) loss) to evaluate the impedance of various generative network architectures for different data.

The high impedance to noise is the most important premise for generative network to noise attenuation. In the following, we first test the impedance of various generative networks to Gaussian noise with $\mu = 0$ and $\sigma^2 = 1$. Fig. 3 shows the learning curves from networks with different depths and widths. The red and blue lines are with skip connection and without skip connection, respectively. First, the blue lines are always above the red lines, which means the generative network has high noise impedance without the skip connection. The skip connection has an important role in recovering the full spatial features of data [30]. However, the network with skip connection in this experiment tends to recover the unstructured features of input Z, so that the random noise is reconstructed earlier, which sharply reduces the impedance to noise. Second, comparing the trend of blue lines in the first and second rows in Fig. 3, the fixed width networks (first row) fit the Gaussian noise better in latter iterations,

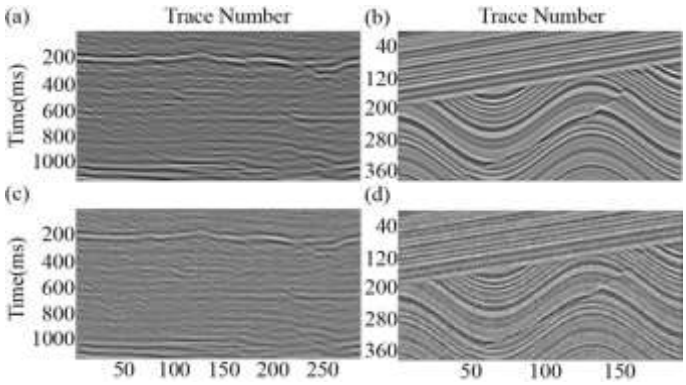

Fig. 4. Data sets for generative networks impedance test. (a) Clean field seismic data obtained by the traditional random noise removal method. (b) Synthetic seismic data. (c) Same data as (a) plus additive Gaussian noise with $\mu = 0$ and $\sigma^2 = 1$. (d) Same data as (b) plus additive Gaussian noise with $\mu = 0$ and $\sigma^2 = 1$.

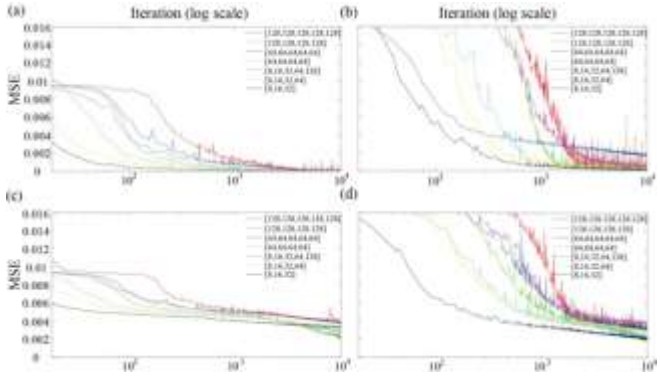

Fig. 5. Learning curve of various generative networks for the test data in Fig. 4. (a) Clean field data [Fig. 4(a)]. (b) Clean synthetic data [Fig. 4(b)]. (c) Noisy field data [Fig. 4(c)]. (d) Noisy synthetic data [Fig. 4(d)].

which means the network with varying width regarding to depth (second row) have high impedance to the Gaussian noise. Third, the networks in Fig. 3(a),(c), (e), Fig. 3(b), (d), (f), and Fig. 3(g) are with depth 5, 4, and 3, respectively. Comparing the blue learning curve between groups, we can conclude that the deeper the network, the higher the impedance to noise. Therefore, for high impedance to Gaussian noise, the selected ConvNet [Fig. 3(e)] with deep depth, varying width, and without skip connections is preferred.

Then, both clean and noisy seismic data are used to verify whether the above networks without skip connection could extract the seismic signal in noisy data. Both the clean field [Fig. 4(a)] and synthetic [Fig. 4(b)] seismic data sets and also with added Gaussian noise with μ 0 and σ^2 1 [Fig. 4(c) and (d)] are used for testing. Fig. 5(a)–(d) shows the corresponding learning curve to the data set in Fig. 4(a)–(d), respectively. The lines of different colors are from networks with different depths and widths. First, comparing the learning curves in Figs. 3 and 5(a) and (b), in general, all networks fit the seismic data at early iterations (approach to 0 after 10^3), which means the networks have low impedance to seismic data. Second, networks with depth from 5 to 3 have decreasing impedance to seismic data. Take also into consideration the learning curve in Fig. 3(b), (d), (f), and (g); the learning curves from the corresponding networks for the noisy seismic data [Fig. 5(c) and (d)] arrive at a too low value of MSE to effectively remove the noise.

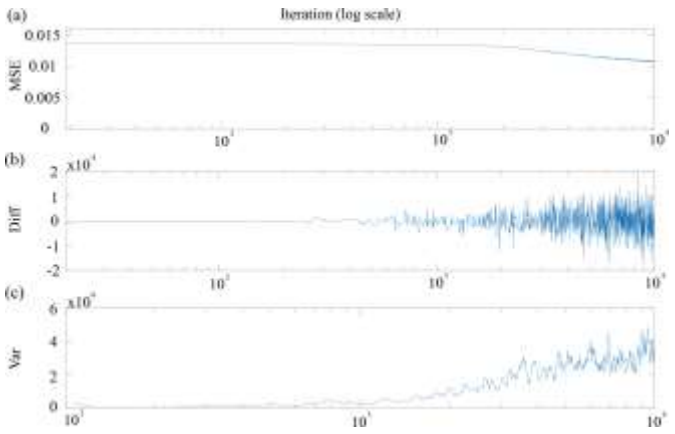

Fig. 6. Analysis of learning curve from the selected ConvNet for random noise. (a) Learning curve (MSE). (b) Difference curve. (c) Variance curve.

The purple lines (selected ConvNet) not only decrease early but also reach higher value of MSE in latter iterations than other lines in Fig. 5(c) and (d). Combining with the result of Fig. 3, the sufficient condition that includes high and low impedance to noise and seismic data, respectively, ensures the selected ConvNet can extract feature of seismic data and attenuate the random noise effectively. Meanwhile, both synthetic and field data are fitted well by the selected ConvNet, which means the corresponding network can be adopted to not only synthetic but also field data denoising task. Due to the computational cost which will increase sharply as the network goes deeper and wider [31], this generative network architecture is used for seismic data denoising without proceeding to deeper depth.

C. Stopping Criterion and Workflow of Seismic Data Denoising

For the generative networks, they reconstruct the input data (including the valid signal and noise) if the fitting process runs long enough [27]. To obtain the latent denoising result automatically, we design a heuristic algorithm with stopping criterion for iterations. We first use the difference operation to eliminate the linear trend of the learning curve [Fig. 6(a)] for random noise. Then, by sliding a window with length 100 on the difference curve, the variance curve can be calculated. In Fig. 6(c), after 10^3 iterations, the difference curve [Fig. 6(b)] starts to fluctuate significantly and its variance [Fig. 6(c)] increases sharply. Hence, the optimization function in the proposed heuristic algorithm is defined as

$$\arg\min_{t>10^3} \mathrm{Var}_{100}(\mathrm{Diff}(\mathrm{loss}(t))) \tag{4}$$

where t denotes the epoch of iterations; $\mathrm{loss}(t)$ denotes the value of MSE at t_{th} epoch; Diff denotes the differential operation; Var denotes the variance operation in the sliding window.

After setting a tolerance (default:1000) to determine the local minimizer of variance curve, the stopping criterion can find a suitable t to stop iterations and the latent clean seismic data are then obtained. The proposed seismic data random noise attenuation workflow is shown in Algorithm 1.

III. Experimental Results

In this section, we apply the selected ConvNet (Fig. 2) to the synthetic and field seismic data for denoising. The selected

Algorithm 1 Workflow of Seismic Random Noise Attenuation With the Selected ConvNet

Input: Noisy seismic data $\boldsymbol{y}$; Maximum number of iterations $\boldsymbol{N} = 1000$; learning rate $\boldsymbol{lr} = 0.01$; Uniform distribution Parameters $\boldsymbol{a} = \mathbf{0}, \boldsymbol{b} = \mathbf{0.1}$.

1: Normalize $\boldsymbol{y}$ to the range $\mathbf{0}$ —$\mathbf{1}$;

2: Initialization: Network parameter $\boldsymbol{\theta^{(0)}}$; input $\boldsymbol{z} \sim \boldsymbol{U(a,b)}$. $\boldsymbol{n} = \mathbf{0}$.

3: While $\boldsymbol{n} < \boldsymbol{N}$ **do**

$\boldsymbol{loss^{(i)}} = \boldsymbol{E}\ \boldsymbol{f}_{\boldsymbol{\theta^{(i)}}}(\boldsymbol{z}), \boldsymbol{y}$;

$\boldsymbol{\theta^{(i+1)}} = \boldsymbol{\theta^{(i)}} + \boldsymbol{lr} * \frac{\partial \boldsymbol{loss}^{(i)}}{\partial \boldsymbol{\theta}^{(i)}}$;

if the stopping criterion is met:

$\boldsymbol{\theta}^{*} = \boldsymbol{\theta}^{(n-1000)}$;

break;

end if

$\boldsymbol{n} = \boldsymbol{n+1}$;

end while

Output: Clean seismic data $\boldsymbol{x} = \boldsymbol{f}_{\boldsymbol{\theta}^{*}}(\boldsymbol{z})$.

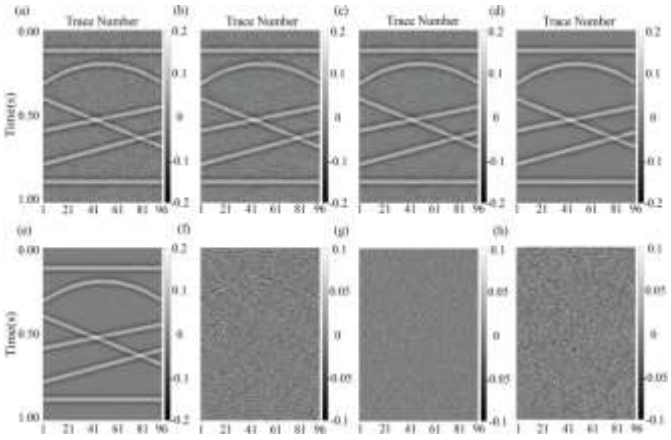

Fig. 7. Denoising results for synthetic data set. (a) Noisy synthetic seismic data (0 dB). (b) f-x MSSA (25.07 dB). (c) Zhang *et al.* [26] (26.74 dB). (d) Selected ConvNet (**31.27 dB**). The figures below (f)–(h) are the corresponding residuals. The ground truth is listed in (e) for comparison.

ConvNet is compared with f-x MSSA and Zhang *et al.* [26] to demonstrate the advantages.

A. Synthetic Data Set

The SNR is used to measure the synthetic data denoising performance, which is defined as follows:

$$\mathrm{SNR} = 10\log_{10}\ \frac{||X||_F^2}{||X - X^*||_F^2} \tag{5}$$

where X denotes the clean signals; X^* denotes the reconstructed data; $||\ ||_F^2$ denotes the F norm.

The synthetic seismic data in Fig. 7 contain 96 traces, each with 500 time-samples and 2-ms time interval. Fig. 7(a) shows the noisy data which added Gaussian white noise with $\mu = 0$ and $\sigma^2 = 1$. The value of SNR in noisy data is 0 dB. Fig. 7(b)–(d) shows the denoising result by f-x MSSA (25.07 dB), Zhang *et al.* [26] (26.74 dB), and the selected ConvNet (**31.27 dB**), respectively. The corresponding residuals are shown in Fig. 7(f)–(h) and the ground truth is listed in Fig. 7(e) for comparison. As it can be seen clearly in Fig. 7, the selected ConvNet can extract and reconstruct the reflections of seismic signal effectively. Meanwhile, benefiting from the high impedance to noise, there is no significant residual noise in the denoising result of the selected ConvNet like other methods.

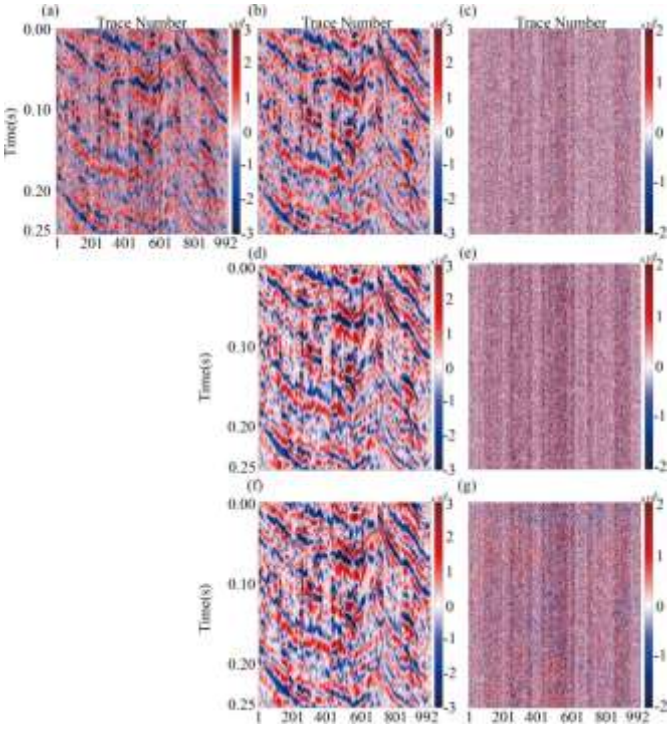

Fig. 8. (a) Noisy field seismic data. The denoising results by (b) f-x MSSA, (d) Zhang *et al.* [26] and (f) Selected ConvNet. Figures in (c), (e), and (g) are the corresponding residuals for (b), (d), and (f).

B. Field Data Set

The field data in Fig. 8(a) contain 992 traces, each with 128 time-samples and 2-ms time interval. The second column in Fig. 8 shows the denoising results by f-x MSSA [Fig. 8(b)], Zhang *et al.* [26] [Fig. 8(d)], and the selected ConvNet [Fig. 8(f)], respectively, and the corresponding residuals are shown in the third column. In the corresponding residuals shown in Fig. 8(b), (d), and (f), we can hardly find obvious seismic reflection events except the residual of f-x MSSA, which indicates that the signals are not seriously damaged during the denoising process of Zhang *et al.* [26] and the selected ConvNet. Also, it can be seen that although the selected ConvNet and Zhang *et al.* [26] have good performance on this data set, the residual of the selected ConvNet indicates that the selected ConvNet can remove more noise than Zhang *et al.* [26] while preserving valid signals.

C. Discussion

Compared with the denoising performance of the above methods in both the synthetic and field data, it can be seen that the selected ConvNet can remove the random noise well from not only synthetic but also field data. Different from f-x MSSA based on rank reduction and Zhang *et al.* [26] which is a patch-neural network (NN)-based method, it possesses

TABLE I
COMPARISON OF SEISMIC DATA DENOISING METHODS

Operation / Method	Data Preprocessing	Parameters Selection	Result Reprocessing	Time Cost
F-x MSSA	×	√	×	0.7s
Zhang [26]	√	√	√	152s
The selected ConvNet	×	×	×	195s

extensive adaptability to seismic data and can better capture the global feature of seismic signal. We list some operation requirements (data preprocessing, parameters selection, result reprocessing, and the time cost using single GTX1060 GPU on the synthetic experiment) in Table I for further comparison. f-x MSSA and Zhang *et al.* [26] need to select parameters while denoising different seismic data; Zhang *et al.* [26] also need to process data with patch sampling and recombination. However, the selected ConvNet can be used to denoise the seismic data directly without any pre and postoperations except normalization. Though the cost time for single data set usually needs several minutes, the convenient operation of the selected ConvNet still makes it good option in practice.

IV. CONCLUSION

This letter presents an unsupervised seismic random noise attenuation method based on ConvNet denoising. When fitting noisy seismic data, the generative network may distinguish random noise during iterations by utilizing the high and low impedance of the network to noise and seismic data. In order to automatically extract the latent clean seismic data, a stopping criterion is also established, and the advantageous generative network architecture that served as the previous model for seismic data denoising is investigated.. Examples using both synthetic and field data show how effective it is. In order to obtain the desired outcome, the suggested method typically requires a significant number of iterations—many thousands—and is consequently computationally demanding. Because of its straightforward operation and unsupervised nature, it can be used to create label data for conventional approaches.

REFERENCES

[1] M. Xiong, Y. Li, and N. Wu, "Random-noise attenuation for seismic data by local parallel radial-trace TFPF," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4025–4031, Jul. 2014.

[2] Z. Gao, Z. Pan, and J. Gao, "Multimutation differential evolution algorithm and its application to seismic inversion," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3626–3636, Jun. 2016.

[3] S. Yu, S. Osher, J. Ma, and Z. Shi, "Noise attenuation in a low-dimensional manifold," *Geophysics*, vol. 82, no. 5, pp. V321–V334, Sep. 2017.

[4] D. Bonar and M. Sacchi, "Denoising seismic data using the nonlocal means algorithm," *Geophysics*, vol. 77, no. 1, pp. A5–A8, Jan. 2012.

[5] M. A. N. Siahsar, S. Gholtashi, A. R. Kahoo, H. Marvi, and A. Ahmadifard, "Sparse time-frequency representation for seismic noise reduction using low-rank and sparse decomposition," *Geophysics*, vol. 81, no. 2, pp. V117–V124, Mar. 2016.

[6] L. L. Canales, "Random noise reduction," in *Proc. SEG Tech. Program Expanded Abstr.*, Jan. 1984, pp. 525–527.

[7] G. Liu, X. Chen, J. Du, and K. Wu, "Random noise attenuation using f-x regularized nonstationary autoregression," *Geophysics*, vol. 77, no. 2, pp. V61–V69, Mar. 2012.

[8] R. Abma and J. Claerbout, "Lateral prediction for noise attenuation by t-x and f-x techniques," *Geophysics*, vol. 60, no. 6, pp. 1887–1896, Nov. 1995.

[9] D. Alsdorf, "Noise reduction in seismic data using Fourier correction coefficient filtering," *Geophysics*, vol. 62, no. 5, pp. 1617–1627, Sep. 1997.

[10] S. M. Mousavi and C. A. Langston, "Hybrid seismic denoising using higher-order statistics and improved wavelet block thresholding," *Bull. Seismol. Soc. Amer.*, vol. 106, no. 4, pp. 1380–1393, Aug. 2016.

[11] Z. Yu, R. Abma, J. Etgen, and C. Sullivan, "Attenuation of noise and simultaneous source interference using wavelet denoising," *Geophysics*, vol. 82, no. 3, pp. V179–V190, May 2017.

[12] G. Hennenfent and F. J. Herrmann, "Seismic denoising with nonuniformly sampled curvelets," *Comput. Sci. Eng.*, vol. 8, no. 3, pp. 16–25, May/Jun. 2006, doi: 10.1109/MCSE.2006.49.

[13] B. Wang, R.-S. Wu, X. Chen, and J. Li, "Simultaneous seismic data interpolation and denoising with a new adaptive method based on dreamlet transform," *Geophys. J. Int.*, vol. 201, no. 2, pp. 1180–1192, 2015.

[14] B. Wu, R.-S. Wu, X. Jiang, W. Sun, and J. Gao, "Dreamlet: A new representation and migration of seismic wavefield in full local domains," *Commun. Comput. Phys.*, vol. 28, no. 1, pp. 111–127, 2020.

[15] L. Liu and J. Ma, "Structured graph dictionary learning and application on the seismic denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 1883–1893, Apr. 2019, doi: 10.1109/TGRS.2018.2870087.

[16] Y. Yang, J. Gao, Z. Wang, G. Zhang, and X. Zhu, "2-D seismic random noise attenuation via self-paced nonnegative dictionary learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 12, pp. 5391–5401, Dec. 2019, doi: 10.1109/JSTARS.2019.2957017.

[17] B. Wu, J. Yu, H. Ren, Y. Lou, and N. Liu, "Seismic traffic noise attenuation using l_p -norm robust PCA," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 11, pp. 1998–2001, Nov. 2020, doi: 10.1109/LGRS.2019.2955737.

[18] V. Oropeza and M. Sacchi, "Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis," *Geophysics*, vol. 76, no. 3, pp. V25–V32, May 2011.

[19] W. Huang, D. Feng, and Y. Chen, "De-aliased and de-noise Cadzow filtering for seismic data reconstruction," *Geophys. Prospecting*, vol. 68, no. 2, pp. 553–571, 2018.

[20] L. Deng, S. Yuan, and S. Wang, "Sparse Bayesian learning-based seismic denoise by using physical wavelet as basis functions," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 11, pp. 1993–1997, Nov. 2017.

[21] R. Anvari, M. A. N. Siahsar, S. Gholtashi, A. R. Kahoo, and M. Mohammadi, "Seismic random noise attenuation using synchrosqueezed wavelet transform and low-rank signal matrix approximation," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6574–6581, Nov. 2017.

[22] X. T. Dong, Y. Li, and B. J. Yang, "Desert low-frequency noise suppression by using adaptive DnCNNs based on the determination of high-order statistic," *Geophys. J. Int.*, vol. 219, no. 2, pp. 1281–1299, Nov. 2019.

[23] Y. Zhang, H. Lin, Y. Li, and H. Ma, "A patch based denoising method using deep convolutional neural network for seismic image," *IEEE Access*, vol. 7, pp. 156883–156894, 2019.

[24] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.

[25] C. Min, G. Wen, B. Li, and F. Fan, "Blind deblurring via a novel recursive deep CNN improved by wavelet transform," *IEEE Access*, vol. 6, pp. 69242–69252, 2018.

[26] M. Zhang, Y. Liu, and Y. Chen, "Unsupervised seismic random noise attenuation based on deep convolutional neural network," *IEEE Access*, vol. 7, pp. 179810–179822, 2019, doi: 10.1109/ACCESS.2019.2959238.

[27] V. Lempitsky, A. Vedaldi, and D. Ulyanov, "Deep image prior," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 9446–9454.

[28] R. Wu, Y. Geng, and B. Wu, "Physical wavelet defined on an observation plane and the dreamlet," in *Proc. SEG Tech. Program Expanded Abstr.*, Jan. 2011, pp. 3835–3839.

[29] Q. Liu, L. Fu, and M. Zhang, "Deep-seismic-prior-based reconstruction of seismic data using convolutional neural networks," 2019, *arXiv:1911.08784*. [Online]. Available: http://arxiv.org/abs/1911.08784

[30] M. Drozcal, E. Vorontsov, G. Chartrand, S. Kadoury, and C. Pal, "The importance of skip connections in biomedical image segmentation," in *Deep Learning and Data Labeling for Medical Applications*. Cham, Switzerland: Springer, 2016, pp. 179–187.

[31] K. He and J. Sun, "Convolutional neural networks at constrained time cost," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 5353–5360, doi: 10.1109/CVPR.2015.7299173.