DESERT SEISMIC DATA DENOISING AND EFFECTIVE SIGNAL RECOVERY BY USING IMPROVED SHEARLET TRANSFORM BASED ON THE DEEP-LEARNING COEFFICIENT SELECTION

XINTONG DONG[1], YUE LI[2*] and BAOJUN YANG[3]

[1]*Jilin University, College of Instrumentation and Electrical Engineering, Jilin, P.R. China. 18186829038@163.com*
[2]*Jilin University, College of Communication Engineering, Jilin, P.R. China. liyue@jlu.edu.cn*
[3]*Jilin University, College of Geo-exploration Science and Technology, Jilin, P.R. China. yangbaojun@jlu.edu.cn*

(Received August 18, 2020; revised version accepted June 10, 2021)

ABSTRACT

Dong, X., Li, Y. and Yang, B., 2021. Desert seismic data denoising and effective signal recovery by using improved shearlet transform based on the deep-learning coefficient selection. *Journal of Seismic Exploration*, 30: 455-479.

Contamination of seismic data by background noise causes difficulties for imaging, reservoir fluid prediction, and stratigraphic interpretation. Desert seismic data poses a particular problem mainly due to two reasons: (1) low signal-to-noise ratio (SNR); (2) serious frequency spectrum overlapping between the effective signals and low-frequency noise (mainly including random noise and surface waves). Therefore, when apply sparse-transform-based methods to denoise desert seismic data, conventional threshold functions fail to distinguish the effective signal coefficients and low-frequency noise coefficients, which is likely to result in residual noise and signal leakage. To solve this problem, we utilize the convolutional neural network (CNN) to act as a threshold function, thereby establishing an optimal non-linear relationship between noisy coefficients and effective signal coefficients. In addition, in order to achieve multi-scale and multi-direction accurate noise suppression, we construct a corresponding training dataset for each sub-band, so as to obtain a CNN-based coefficient selection model suitable for this sub-band. In this paper, we take shearlet transform as an example to verify the effectiveness of the proposed CNN-based threshold function. Synthetic and real examples demonstrate that our method can effectively suppress the desert low-frequency noise and completely recover the effective signals reflected by layers.

KEY WORDS: desert seismic data, shearlet transform, noise suppression, convolutional neural network, low signal-to-noise ratio, spectrum overlapping.

0963-0651/21/$5.00

INTRODUCTION

Modern seismic exploration requires that seismic data fulfills the requirements of high signal-to-noise ratio (SNR), high resolution, and high fidelity; in general, high SNR is the basis of the latter two (Dong et al., 2019a; Ma et al., 2020a). Suppressing seismic background noise and recovering effective signals is one of the most effective and direct way to increase the SNR of seismic data (Fomel and Liu, 2010; Bonar and Sacchi, 2012; Beckouche and Ma, 2014). Recently, experts have proposed numerous methods to complete this task (Bekara and van der Baan, 2007; Chen et al., 2019; Dong et al., 2020a). Frequency filtering methods, such as band-pass filter (Stein and Bartley, 1983), F-X prediction filter (Harris and White, 1997; Wang, 1999; Gülünay, 2017), time frequency peak filtering (Tian et al., 2014; Xiong et al., 2014) and polarization filter (Kulesh et al., 2007; Akram, 2018), can retain certain frequency and suppress unwanted frequency. Mode-decomposition-based methods including empirical mode decomposition (EMD; Bekara and van der Baan, 2009; Han and van der Baan, 2013), variational modal decomposition (VMD; Li et al., 2017; Ma et al., 2020b), and some variations of the two can decompose noisy seismic data into multiple modes, and then we can preserve the modes containing effective signal components and abandon the modes associated with noise. Rank-reduction-based methods including principle component analysis (Wang et al., 2016) and robust principle component analysis (Dong et al., 2020a) assume that the ideal clean seismic data is a low-rank structure, and noise interference will increase its rank. These approaches usually realize the noise suppression by rank reduction. Some other representative methods include dictionary learning (Beckouche and Ma, 2014; Nazari Siahsar, 2017), median filtering (Duncan and Beresford, 1995; Wang et al., 2020), singular value decomposition (Bekara and van der Baan, 2007; Liang et al., 2018), etc. Although these methods have been successfully applied to numerous real-world problems, these methods still have some limitations (Dong et al., 2019a; Ma et al., 2020a). For example, frequency filtering methods do not work well when effective signals and noise share similar bandwidths (Chen et al., 2019; Dong et al., 2020b); due to mode aliasing, some mode-decomposition- based methods degrade when processing low SNR seismic data (Dong et al., 2020a; Ma et al., 2020b); some rank-reduction-based methods have little effect on random noise suppression (Dong et al., 2020a).

In desert seismic data, reflected effective signals are often contaminated by a large amount of low frequency noise mainly including random noise and surface waves (Zhong et al., 2015a; Li et al., 2017; Dong et al., 2019a); this causes lots of difficulties to the imaging, inversion, and interpretation of the data. The desert seismic data is characterized by two distinct features: low SNR and significant spectrum overlapping between effective signals and low-frequency noise (Zhong et al., 2015a). The low SNR is mainly caused by three reasons. Firstly, sand can absorb partial energy of the reflected effective signals (Dong et al., 2020a) and the target layer is usually deep in desert regions; this two points lead to the weak energy of reflected

effective signals. Secondly, due to the large space between sand grains and low adhesion, the sand dunes tend to produce a large number of surface waves with strong energy. Thirdly, the relative emptiness of desert regions lead to the strong intensity of wind, so the energy of random noise is strong. In desert regions, the dominant frequency of effective signals, random noise and surface waves is mainly distributed in 10-25 Hz, 5-20 Hz, 0-15 Hz, respectively (Li et al., 2017; Zhong et al., 2015a). Therefore, there exists significantly spectrum overlapping between the effective signals and the low-frequency, and how to recover the effective signals from shared frequency bands is a relatively challenging task. Moreover, some related study shows the random noise in desert regions is characterized by non-stationary and non-linear (Zhong et al., 2015a,b). Therefore, how to suppress desert low-frequency noise and recover effective signals is a challenging problem that needs to be solved.

In recent years, sparse transform is developing rapidly, such as wavelet transform (Goudarzi and Riahi, 2012; Yang et al., 2020), curvelet transform (Neelamani et al., 2008), Radon transform (Gholami and Zand, 2017), seislet transform (Fomel and Liu, 2010), shearlet transform (Liang et al., 2018; Dong et al., 2019a), etc. Due to the multi-scale, multi-direction, and good sparse representation, these sparse transforms have been widely used in imaging processing and seismic data denoising (Fomel and Liu, 2010; Goudarzi and Riahi, 2012; Dong et al., 2019b). Sparse-transform-based denoising methods mainly include three steps: sparse representation (forward transform), effective signal coefficient selection, and coefficient reconstruction (or called inverse transform). Obviously, the second step directly determines the denoising performance of spare-transform-based denoising methods. In general, we often use some threshold functions to accomplish the second step, such as hard threshold function, soft threshold function, semi-hard and semi-soft threshold functions, and some improved threshold function based on statistics (Tang et al., 2018; Dong et al., 2019b). We have tried to utilize numerous threshold functions to select the effective signal coefficients of desert seismic data. However, the denoising results are not satisfactory; some false selections by threshold functions lead to residual low-frequency noise in denoising results and obvious signal-leakage energy in the removed noise. This phenomenon is mainly because of two reasons: (1) after scale decomposition, each scale coefficient still contains some noise components due to the serious spectrum overlapping of desert seismic data; (2) partial random noise has roughly the same direction as the effective signals, so it is difficult to realize the separation of effective signal coefficients and noise coefficients by directional decomposition. Furthermore, we often need to set the thresholds manually in the sparse-transform-based denoising methods. Therefore, how to realize the accurate and automatic selection of effective signal coefficients becomes the key to the denoising of desert seismic data.

Convolutional neural network (CNN) is a classical artificial neural network, which can combine the low-dimension features to form more abstract high-dimension features (Zhang et al., 2017; Dong et al., 2019a). Due to its excellent characteristics of local perception and weight sharing,

CNN has been successfully used in image denoising (Remez et al., 2018; Zhang et al., 2017), image super-resolution (Tirer and Giryes, 2019), and pattern recognition, etc. Meanwhile, it is also gradually applied in some areas of seismic exploration, such as waveform classification (Yuan et al. 2018), arrival time picking (Yuan et al. 2018), inversion (Wu and McMechan, 2019; Jin et al., 2017) and seismic data denoising (Dong et al., 2019a; Wang et al., 2019), etc. By extracting inner features of dataset, CNN can establish an end-to-end non-linear mapping relationship between input and output (Zhang et al., 2017; Dong et al., 2020b). In this work, we propose to utilize CNN to establish an optimal relationship between the noisy coefficients and the effective signal coefficients. Specifically, some artificial seismic wavelets and real desert low-frequency noise are used to construct the training dataset, and then we can obtain the optimal CNN-based coefficient selection model by the guidance of training dataset and the minimization of the corresponding mean square error (MSE) loss function.

In this paper, we adopt the shearlet transform and utilize multiple trained CNN-based coefficient selection models to predict the effective signal coefficients in shearlet domain, so as to propose a novel denoising method called improved shearlet transform based on the CNN-based coefficient selection (IST-CNNCS). Firstly, the desert seismic data is decomposed into four sub-bands through the multi-scale and multi-direction decomposition of shearlet transform; at the same time, we construct four training datasets to guide the network training, so as to obtain four CNN-based coefficient selection models for the four sub-bands. Secondly, these models are used to predict effective signal coefficients from noisy coefficients. Finally, we can obtain the denoised desert seismic data by coefficient reconstruction. Synthetic and real examples are included to demonstrate the effectiveness of our method. The denoised results illustrate that IST-CNNCS is superior to traditional seismic denoising methods and some classical deep-learning-based denoising applications.

There are three distinct advantages of the proposed IST-CNNCS. (1) Compared with the traditional sparse-transform-based denoising methods, IST-CNNCS does not need to set the thresholds manually, and the CNN-based coefficient selection models can automatically and accurately select the effective signal coefficients. (2) Compared with the direct use of CNN in seismic data denoising, IST-CNNCS makes full use of the good sparse representation of shearlet transform; this enables CNN to more easily and accurately extract the internal characteristics of data and thus avoids the appearance of false events. (3) since we construct different CNN-based coefficient selection models for different sub-bands, IST-CNNCS is an accurate multi-scale and multi-direction denoising method (this is similar to setting different thresholds for different sub-bands).

METHOD

Denoising principle of IST-CNNCS

As shown in eq. (1), desert seismic data can be regarded as a linear superposition of effective signals and low-frequency noise (Beckouche and Ma, 2014; Ma et al., 2020a).

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

where y, x, n represent desert seismic data, effective signals, and low-frequency noise, respectively. The desert seismic data is decomposed into a low-frequency scale and a high-frequency scale by the scale decomposition of shearlet transform, and then we utilize the direction decomposition of shearlet transform to decompose the high-frequency scale into three direction components. The low-frequency scale and the three direction components decomposed from the high-frequency scale are the same size as the desert seismic data. For the convenience of expression, the four components are named sub-band 1,2,3,4, respectively.

$$\begin{Bmatrix} C_1(y) = C_1(x) + C_1(n) \\ C_2(y) = C_2(x) + C_2(n) \\ C_3(y) = C_3(x) + C_3(n) \\ C_4(y) = C_4(x) + C_4(n) \end{Bmatrix}, \tag{2}$$

where $C_1(y), C_2(y), C_3(y), C_4(y)$ represent the noisy coefficients of sub-bands 1, 2, 3, 4, respectively; $C_1(x), C_2(x), C_3(x), C_4(x)$ represent the effective signal coefficients of sub-bands 1, 2, 3, 4, respectively; $C_1(n), C_2(n), C_3(n), C_4(n)$ represent the low-frequency noise coefficients of sub-bands 1, 2, 3, 4, respectively. Obviously, how to separate effective signal coefficients from noisy coefficients directly determine the quality of denoising result. Next, we take the sub-band 1 (i.e., low-frequency scale) as an example to illustrate the extraction process of effective signal coefficients in detail.

In IST-CNNCS, we adopt the basic architecture of feed-forward denoising convolutional neural network (DnCNNs; Zhang et al., 2017; Dong et al., 2019a) whose basic structure is described in Appendix B. It is worth mentioning that the residual learning is adopted in DnCNNs, i.e, the noise coefficients are taken as the learning target of DnCNNs by introducing a bypass connection between the input and output. This kind of residual learning not only solves the problem that the deep network model is difficult to train, but also overcomes the network degradation problem caused by the increasing of network depth (Yang et al., 2017; Mou et al., 2018). The basic strategy of CNN-based coefficient selection is to establish an nonlinear mapping relationship between noisy coefficients and low-frequency noise coefficients through residual learning, and then obtain the corresponding

effective signal coefficients through subtraction. The specific process is shown in eqs. (3) and (4):

$$\hat{C}_1(n) = M\left[C_1(y);\theta\right], \tag{3}$$

$$\hat{C}_1(x) = C_1(y) - \hat{C}_1(n), \tag{4}$$

where M represents the mapping relationship between noisy coefficients and low-frequency noise coefficients established by CNN; θ is the network parameter including weight and bias, and its optimization is realized by the supervised learning of training dataset and loss function; $\hat{C}_1(n)$represents the predicted low-frequency noise coefficients, and $\hat{C}_1(x)$ is the corresponding effective signal coefficients. In order to optimize the network parameters and thus obtain an optimal CNN-based coefficient selection model for sub-band 1, we utilize the mean square error (MSE) loss function in eq. (5) to optimize the network.

$$L_{MSE}(\theta) = \frac{1}{2I}\sum_{i=1}^{I}\left\|M\left[C_1^i(y);\theta\right] - C_1^i(n)\right\|_2^F, \tag{5}$$

where $\|\ \|_F$ stands for Frobenious function; I is the batch size; $\left[C_1^i(y), C_1^i(x)\right]$ represents the training dataset for sub-band 1. Training dataset directly determines the denoising effect of IST-CNNCS on desert seismic data, and its specific construction will be introduced in the next part. The smaller the loss function $L_{MSE}(\theta)$, the better the optimization of network parameters θ. In similar way, we can obtain three CNN-based coefficient selection models for the three remaining sub-bands, and then utilize the three CNN-based models to predict effective signal coefficients of the sub-bands 2, 3, 4, respectively. Finally, the denoised desert seismic data can be obtained by coefficient reconstruction.

$$x_{denoised} = CR\left[\hat{C}_1(x), \hat{C}_2(x), \hat{C}_2(x), \hat{C}_3(x)\right], \tag{6}$$

where CR represents the coefficient reconstruction, i.e., inverse transform; $x_{denoised}$ is the denoised desert seismic data; $\hat{C}_2(x), \hat{C}_3(x), \hat{C}_4(x)$ denotes the predicted effective signal coefficients of sub-bands 2, 3, 4, respectively. To be more intuitive, the workflow of IST-CNNCS is shown in Fig. 1.

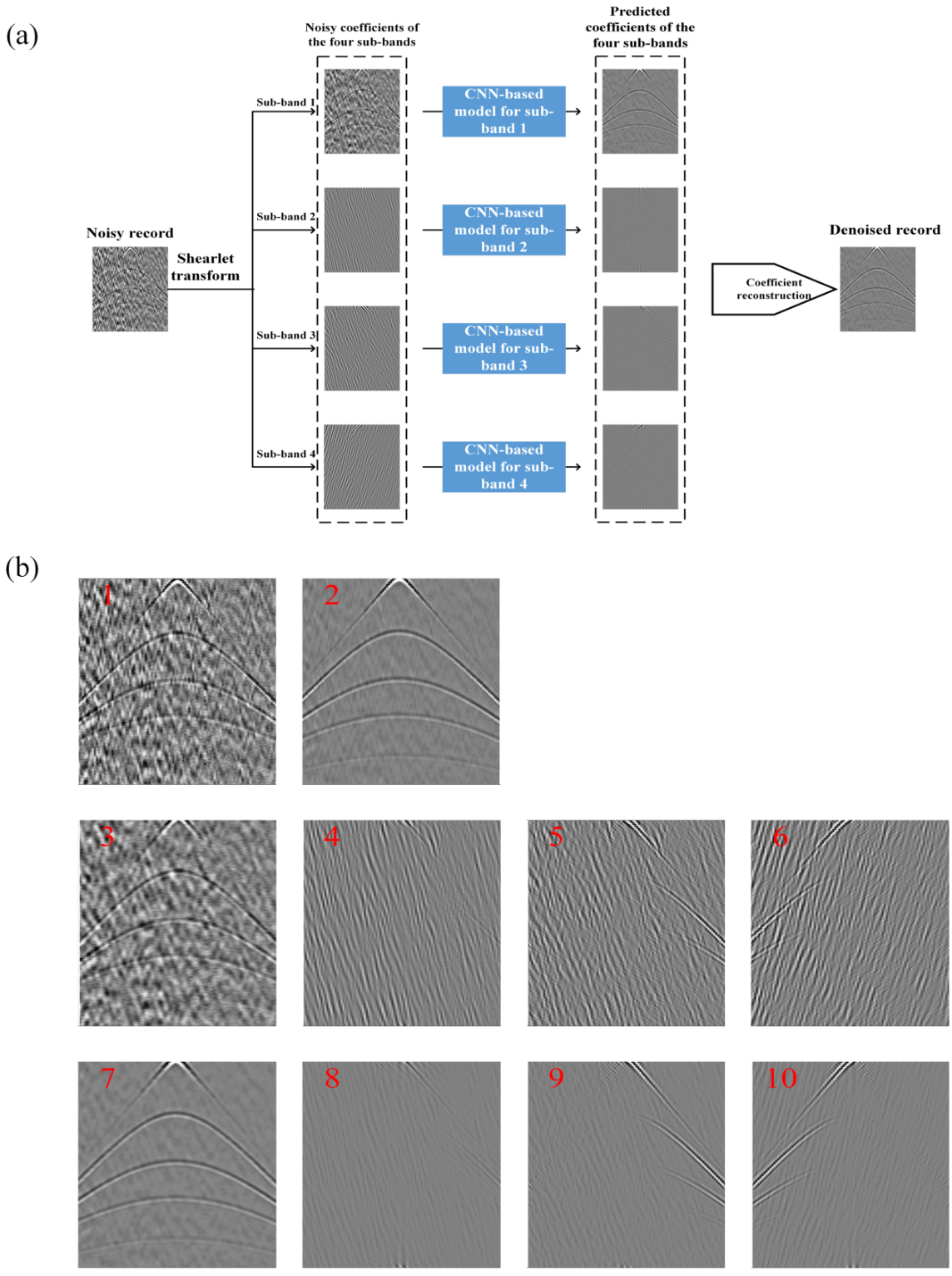

Fig. 1. The summary of IST-CNNCS. (a) The workflow of IST-CNNCS. (b) An enlarged view of the data in Fig. 1(a). In Fig 1(b), the figures numbered by 1 and 2 (first row) are the noisy record and denoised record in Fig. 1(a); the figures numbered by 3,4,5,6 (second row) are the noisy coefficients of the four sub-bands in Fig. 1(a); the figures numbered by 7,8,9,10 (third row) are the predicted coefficients of the four sub-bands in Fig. 1(a).

The construction of CNN-based coefficient selection models

Training dataset

It can be discovered from the loss function in eq. (5) that CNN needs a training dataset containing sufficient and representative data, so as to achieve the optimization of network parameters. The differences between the CNN-based coefficient selection model in IST-CNNCS and the CNN model in image processing mainly includes two aspects: (1) the CNN model in image processing is mainly used to process time-domain signals, while IST-CNNCS aims at frequency-domain coefficients; (2) the CNN model in image processing is mainly used to suppress the White Gaussian noise (WGN), so its training set consists of WGN whose characteristics are largely different from the desert low-frequency noise. Therefore, we should construct suitable training dataset to guide the network training of IST-CNNCS. In this paper, we utilize the coefficients of Ricker wavelet, zero-phase wavelet, mixed-phase wavelet and real desert low-frequency noise to construct the training dataset for IST-CNNCS. Furthermore, the coefficient features in different sub-bands are still variable, so we build four different training datasets for sub-bands 1 to 4, respectively. Taking the training set of sub-band 1 as an example, the construction process is described in detail below.

The training dataset consists of an effective signal dataset and a noise dataset, and we firstly introduce the former. In seismic exploration, it is difficult to acquire real effective signals due to the continuous existence of background noise. In order to obtain sufficient data of the effective signals which is similar to real effective signals in desert regions, we select three commonly-used seismic wavelets (Ricker wavelet, zero-phase wavelet and mixed-phase wavelet) to simulate the real effective signals in desert regions. To be more close to real conditions, the parameters of the adopted seismic wavelets are variable. The concrete expression and parameter ranges of the three seismic wavelets are shown in eqs. (7) to (9) and Table 1, respectively.

$$f(t) = V\left[1 - 2\pi^2 f_d{}^2 t^2\right]\exp\left(-\pi^2 f_d{}^2 t^2\right), \tag{7}$$

$$f(t) = V\cos(2\pi f_d t)\exp\left[-\left(\frac{2\pi f_d t}{r_1}\right)^2\right], \tag{8}$$

$$f(t) = V\sin(2\pi f_d t)\exp\left[-\left(\frac{2\pi f_d t}{r_2}\right)^2\right], \tag{9}$$

In eqs. (7) to (9), V represents amplitude; f_d is on behalf of dominant frequency; r_1 and r_2 denote the waveform parameters of zero-phase wavelet and mixed-phase wavelet, respectively.

Table 1. The parameter setting for the three seismic wavelets.

Name	dominant frequency(Hz)	velocity (m/s)	trace interval(m)	waveform parameter
Ricker	5-30	200-4500	10-20	
zero-phase	5-30	200-4500	10-20	1-8
mixed phase	5-30	200-4500	10-20	1-8

Firstly, these seismic wavelets with different parameters are leveraged to construct 20 synthetic clean seismic data with size of 1024×256 (sampling point×trace number), and the 20 clean data contains a variety of events mainly including broken, crossed, hyperbolic events, etc. Secondly, through the decomposition of shearlet transform, we can obtain the corresponding 20 effective signal coefficient data of sub-band 1 with size of 1024×256. Thirdly, a sliding window with size of 60×60 is utilized to randomly intercept 12000 effective signal coefficient patches from the 20 effective signal coefficient data of sub-band 1. Finally, these effective signal coefficient patches after amplitude normalization together form the effective signal dataset for sub-band 1.

The noise set directly determines the denoising target and capability of IST-CNNCS. Thus, the noise set of IST-CNNCS should be composed of random noise and surface waves. For the former, we select a pure desert random noise record acquired from the Tarim Basin of China (mainly desert regions), and this record is obtained under the condition of no source; for the latter, we extract some surface waves from some real desert seismic data. Similarly, after the decomposition of shearlet transform and retention of low-frequency scale, 12000 noise coefficient patches of sub-band 1 can be obtained through the interception of the 60×60 sliding window; 7000 of these noise coefficient patches are from the random noise record, and the remaining 5000 patches are from the extracted surface waves. After the amplitude normalization, we can obtain the noise dataset for sub-band 1.

Training process

Due to the large size of training dataset, the small-batch gradient descent is adopted for the network training. In each iteration, only a small part of samples are used to calculate the gradient. In this way, the global gradient is replaced by local gradient, which not only improves the computational efficiency but also reduces the complexity of network. Furthermore, to enhance the generalization ability of IST-CNNCS to different noise levels, we use a variable factor matrix to adjust the energy ratio of the effective signal coefficient patch and the noise coefficient patch. In summary, the training process is as follows:

(1) We randomly select 64 (batch size) effective signal coefficient patches $\{e_1, e_2, e_3 \ldots\ldots, e_{64}\}$ and 64 noise coefficient patches $\{s_1, s_2, s_3 \ldots\ldots, s_{64}\}$from the training dataset of sub-band 1.

(2) To simulate different noise levels, the $\{s_1, s_2, s_3 \ldots\ldots, s_{64}\}$ is multiplied by a variable factor matrix $[v_1, v_2, v_3 \ldots\ldots, v_{64}]$ where $v_1, v_2 \ldots\ldots, v_{64}$ are random numbers between 1 to 20.

(3) $\{e_1 + s_1 \times v_1, e_2 + s_2 \times v_2, \ldots\ldots, e_{64} + s_{64} \times v_{64}\}$ and $\{s_1 \times v_1, s_2 \times v_2, \ldots\ldots, s_{64} \times v_{64}\}$ are utilized as the input and the label of CNN, respectively. Then, CNN is optimized by minimizing the MSE loss function in eq. (5).

(4) Repeat the steps (1) to (3). As the loss function decreases and finally tends to be stable, we can obtain the CNN-based coefficient selection model for sub-band 1 after 60 epochs.

(5) All network training and the subsequent denoising experiments are completed in MATLAB environment running on a server with an E5-2600 v4 processor, Windows 10 system, and two NVIDA GeForce GTX 1080 GPUs. Also, the network parameters of IST-CNNCS are listed in Table 2.

Table 2. The network parameters of IST-CNNCS.

Parameters	Specification
Network depth (the number of convolution layers)	17
Learning rate	$\lfloor 10^{-3}, 10^{-5} \rfloor$
Batch size	64
Convolution kernel size	3×3
The number of filters	128
The number of epoch	60
Patch size	60×60

SYNTHETIC EXAMPLE

Fig. 2(a) shows a theoretical pure seismic record generated by acoustic equation with source function given by Ricker wavelet (Dong et al., 2019a) with dominant frequency 15 Hz, and its sampling interval, trace interval and receiver number are 0.002 s, 20 m, and 100, respectively. Then, we add the real desert low-frequency noise shown in Fig. 2(b) to Fig. 2(a), so as to obtain the synthetic noisy desert seismic record shown in Fig. 2(c). As shown in Fig. 2(c), most effective signals are completely submerged by desert low-frequency noise and hard to be recognized, especially the weak

effective signals in deep part (1.2 s-1.5 s). The SNR and root mean square error (RMSE) of this noisy seismic record are -3.6861 dB and 3.5774, respectively. In this paper, we utilize SNR and RMSE to measure the quality of denoising results. The definitions of the two measurements are as follows:

$$SNR = 10\log_{10}[\frac{\sum_{n_1=1}^{N_1} x(n_1)^2}{\sum_{n_1=1}^{N_1}[x_d(n_1)-x(n_1)]^2}], \tag{10}$$

$$RMSE = \sqrt{\frac{1}{N_1}\sum_{n_1=1}^{N_1}[x_d(n_1)-x(n_1)]^2}, \tag{11}$$

where $x(n_1)$ and $x_d(n_1)$ represent the pure record and denoised record, respectively; N_1 is the number of sampling points. Larger SNR suggests more thorough suppression for noise, and smaller RMSE indicates more complete recovery for effective signals.

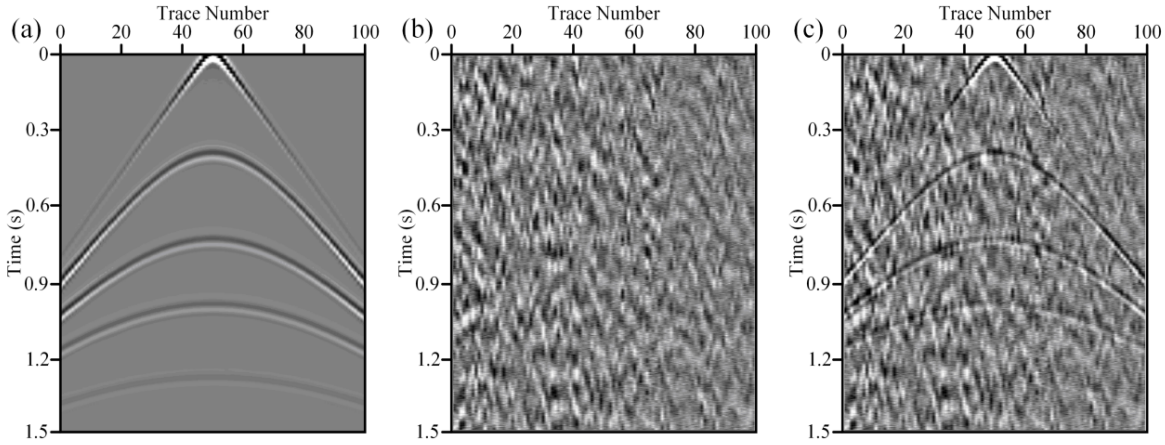

Fig. 2. The synthetic example. (a) The synthetic theoretical pure record. (b) The real desert low-frequency random noise . (c) The synthetic noisy desert seismic record.

In this section, we use IST-CNNCS, shearlet transform with hard threshold function, and two widely-used seismic denoising methods: F-X deconvolution and EMD to process the noisy record in Fig. 2(c). The parameter settings of the three comparative examples are described in the caption of Fig. 3, the denoised results of the four methods are displayed in Figs. 3(a)-(d), and the corresponding removed noise is shown in Figs. 3(e)-(h). As shown in the denoised result after applying IST-CNNCS, the SNR enhancement is visible and most effective signals including the deep weak effective signals become clear and continuous. Also, we can hardly observe signal-leakage energy in the removed noise by IST-CNNCS (Fig. 3e); this demonstrates that the removed noise is extremely close to the pure noise and that IST-CNNCS does little to damage the effective signals. In the

denoised result of shearlet transform with hard threshold function (Fig. 3b), there is obvious residual noise and the effective signals are still too fuzzy to be observed. Also, some false coefficient selections by the hard threshold function lead to the noticeable signal-leakage energy in the removed noise by the shearlet transform with hard threshold function (Fig. 3f). As shown in Figs. 3(c) and 3(d), F-X deconvolution and EMD can suppress a large amount of noise and recover the effective signals to some extent. However, these two approaches have two drawbacks: (1) they have limited ability to recover the weak effective signals in the deeper part, where some residual noise still makes these effective signals unclear and unrecognized. (2) the processing frequency range and the mode abandonment respectively cause some signal-leakage energy in the removed noise by F-X deconvolution and EMD.

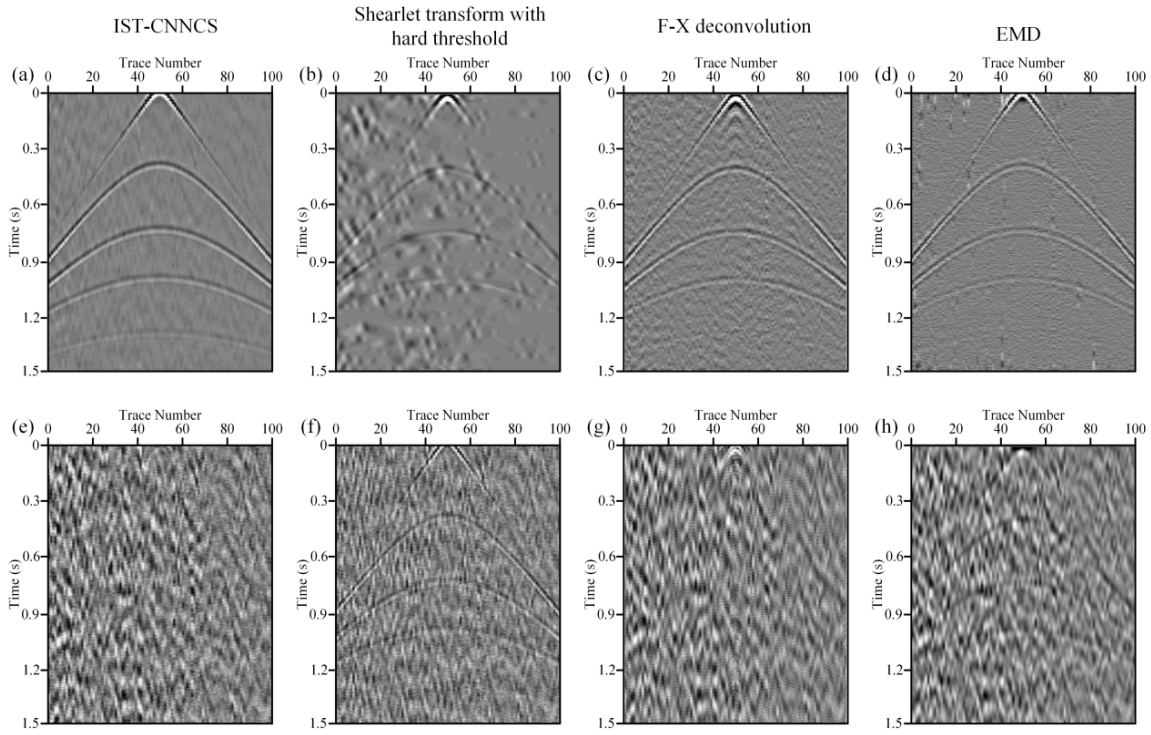

Fig. 3. The comparison of synthetic denoised results. (a)-(d) are denoised results after applying IST-CNNCS, shearlet transform with hard threshold function, F-X deconvolution, and EMD, successively. (e)-(h) are the removed noise by IST-CNNCS, shearlet transform with hard threshold, F-X deconvolution, and EMD, successively. In this synthetic example, for shearlet transform with hard threshold, the scale number is 4, the direction numbers of these 4 scales are 1, 6, 6, 10, respectively and the thresholds of these 4 scales are 2, 1, 1, 10, respectively; for F-X deconvolution, the length of operator is 50, the minimum and maximum frequency is 12 Hz and 200 Hz; for EMD, it decomposes the noisy record into 5 modes, and the sum of 2th, 3th, 4th modes is regarded as the denoised result, i.e., Fig. 3(d).

Next, as shown in Fig. 4, we plot the f-k spectrum of the theoretical pure record in Fig. 2(a), the synthetic noisy record in Fig. 2(c), and the four denoising results in Figs. 3(a)-(d). As can be observed from Figs. 4a and 4b, the effective signals and real desert low-frequency noise share some low-frequency bands (0-10 Hz); this spectrum overlapping brings

difficulties to the separation of noise and effective signals. From Fig. 4(c), we can discover that IST-CNNCS can suppress most desert low-frequency noise. Also, among the f-k spectrum of the four denoising results (Figs. 4c-f), Fig. 4(c) is the most similar to Fig. 4(a), which demonstrates that IST-CNNCS can achieve a good trade-off between noise suppression and effective signal preservation. On the contrary, Figs. 4(d) and 4(e) illustrate that the signal-leakage energy caused by shearlet transform with hard threshold functions and F-X deconvolution; Fig. 4(e) demonstrates the incomplete noise suppression by using EMD.

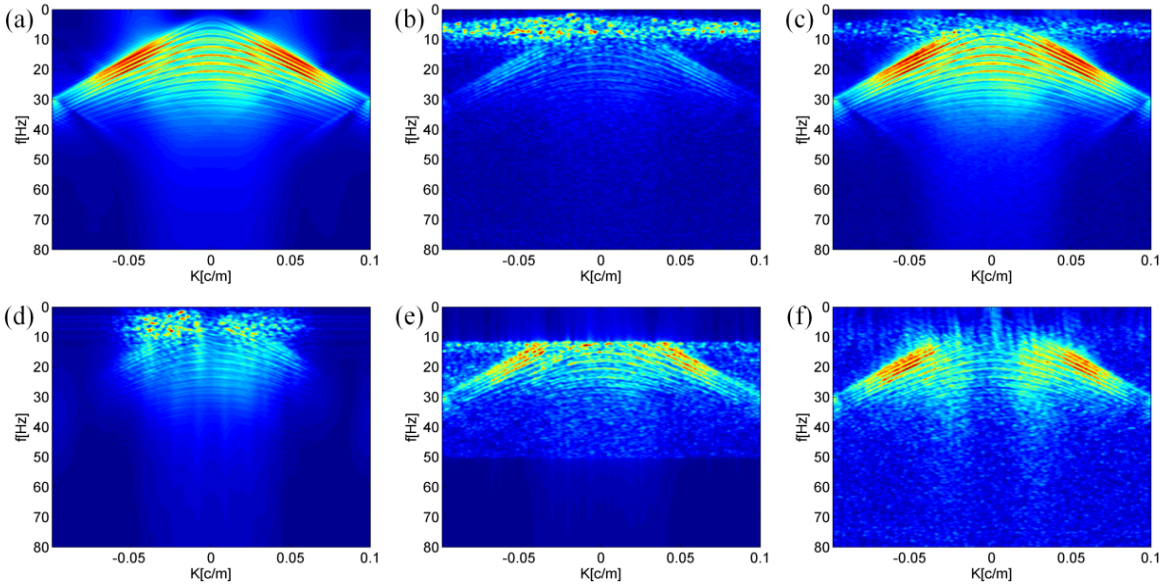

Fig. 4. The comparison of f-k spectrum. (a) The spectrum of the theoretical pure record in Fig. 2(a). (b) The spectrum of the noisy record in Fig. 2(c). (c)-(f) The spectrum of the four denoising results in Figs. 3(a)-(d).

To quantify the denoising performance, we calculate the corresponding SNRs and RMSEs of the four denoised results in Fig. 3 (the bold numbers in Table 3). Also, to demonstrate the reliability to different SNRs, we utilize the above four denoising methods to process four more synthetic noisy records with different SNRs, and the SNRs and RMSEs after denoising are listed in Table 3. It is shown in Table 3 that IST-CNNCS always corresponds to the largest SNR and smallest RMSE.

Table 3. The comparison of denoising measurement parameters (SNR(dB)/RMSE).

Noisy record	IST-CNNCS	Shearlet transform	F-X	EMD
-3.6861/3.5774	**11.5453/0.6194**	**2.7787/1.6995**	**4.0095/1.4750**	**2.6476/1.7254**
-6.1264/4.8214	9.9045/0.8273	0.2631/2.4024	2.0029/1.9073	0.1482/2.5016
-4.2086/3.8129	10.8725/0.6491	2.0425/1.6243	3.7912/1.4074	1.9527/1.7723
-2.0623/3.1724	12.9013/0.4037	4.1489/1.3791	4.4912/1.2841	4.0128/1.4191
-0.1523/2.7357	14.0726/0.3153	5.9012/1.2162	5.8024/1.2539	4.7385/1.3271

REAL EXAMPLE

In this section, we select some real desert seismic records to testify the effectiveness of IST-CNNCS, and compare our method with the above three traditional methods (shearlet transform with hard threshold function, F-X deconvolution, EMD) and two deep-learning-based methods: DnCNNs and trainable nonlinear reaction diffusion (TNRD).

Example 1

To verify the practical application value of IST-CNNCS, we select a real desert common-shot-point record (Fig. 5a) with sampling interval 0.002 s, receiver number 180 and trace interval 20 m. This record is extracted from a 3D desert seismic dataset and received by a certain receiver-line. As shown in Fig. 5(a), the reflected hyperbolic events are simultaneously contaminated by both low-frequency random noise and surface waves, which seriously destroys the continuity of events. We utilize IST-CNNCS and the three comparative methods adopted in the above synthetic example to process this real common-shot-point record, and the corresponding denoised records are shown in Figs. 5(b)-(e), successively. In this section, we compare the four methods from multiple aspects of random noise suppression, surface wave suppression, preservation ability for effective signals, denoising time.

Random noise suppression

As shown in the red rectangles of Fig. 5(a), the contamination by random noise leads to the poor continuity of events. In the red rectangles of Fig. 5(b), IST-CNNCS can suppress most of the random noise and significantly enhance the continuity of events. On the contrary, in the red rectangles of Figs. 5(c) and 5(d), shearlet transform with hard threshold function and F-X deconvolution can only suppress partial random noise, and the hyperbolic events are still difficult to be recognized; in the red rectangles of Fig. 5(e), EMD can effectively suppress most of the random noise, whereas the events are discontinuous.

Surface wave suppression

As shown in the yellow rectangle of Fig. 5(a), the effective signals are seriously damaged by the strong energy surface waves. In the yellow rectangle of Fig. 5(b), IST-CNNCS can suppress almost all surface waves and recover the events to a great extent, even the deeper events which are completely damaged by the strong surface waves (i.e., 2000 ms to 2200 ms of the yellow rectangle in Fig. 5b). Conversely, the shearlet transform can hardly suppress the surface waves which still interfere the effective signals; in the yellow rectangles of Figs. 5(d) and 5(e), although F-X deconvolution and EMD can remove partial surface waves, they can not fully recover the effective signals and the deep events are still discontinuous.

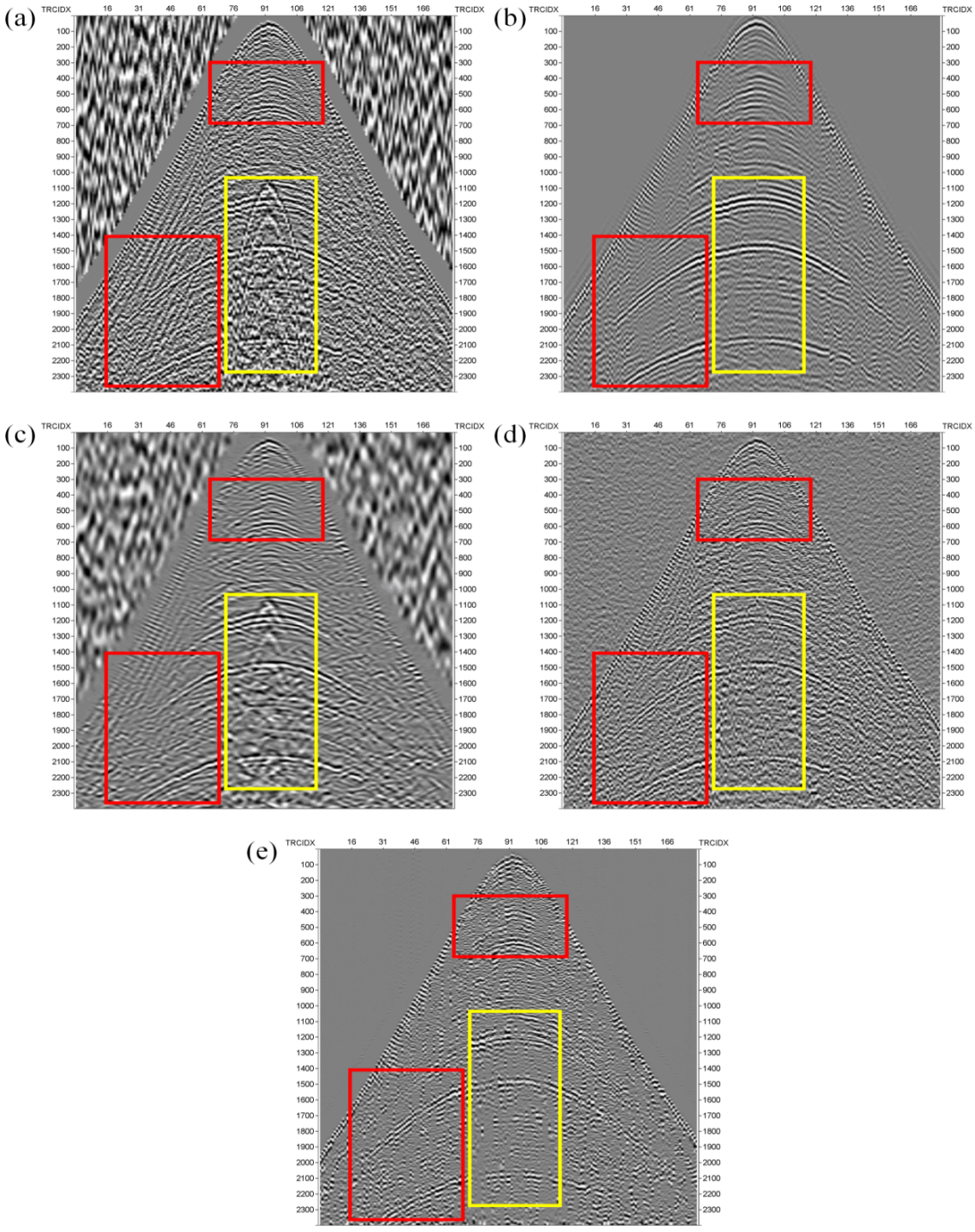

Fig. 5. The real desert common-shot-point record and its denoised results. (a) The real desert common-shot-point record. (b)-(e) are the denoised results of Fig. 5(a) by using IST-CNNCS, shearlet transform with hard threshold function, F-X deconvolution, and EMD, successively. In this real example, for shearlet transform with hard threshold function, the scale number is 5, the direction numbers of these 5 scales are 1, 6, 6, 10, 10, respectively, and the thresholds of these 5 scales are 2, 1, 1, 1, 15, respectively; for F-X deconvolution, the length of operator is 80, the minimum and maximum frequency is 12 Hz and 200 Hz; for EMD, it decomposes the noisy record into 5 modes, and the sum of 2th, 3th modes is regarded as the denoised result, i.e., Fig. 5(e).

Preservation ability for effective signals

To compare the preservation ability for effective signals, we display the difference record of the four methods (Fig. 5a minus Figs. 5b-e, respectively) in Fig. 6. It can be discovered from Figs. 5(c) and 5(d) that there is lots of obvious signal-leakage energy in the difference records of F-X deconvolution and EMD; this phenomenon demonstrates that F-X deconvolution and EMD damage partial effective signals when suppressing the random noise and the surface waves. In the difference record of shearlet transform with hard threshold function (Fig. 5b), we can barely observe the effective signals; this does not mean that shearlet transform with hard threshold has good preservation ability for effective signals, but suggests its poor suppression for low-frequency noise. On the contrary, the different record of IST-CNNCS (Fig. 5a) consists almost entirely of low-frequency noise and we can barely observe signal-leakage energy in this difference record, which proves the superior preservation ability for effective signals of IST-CNNCS.

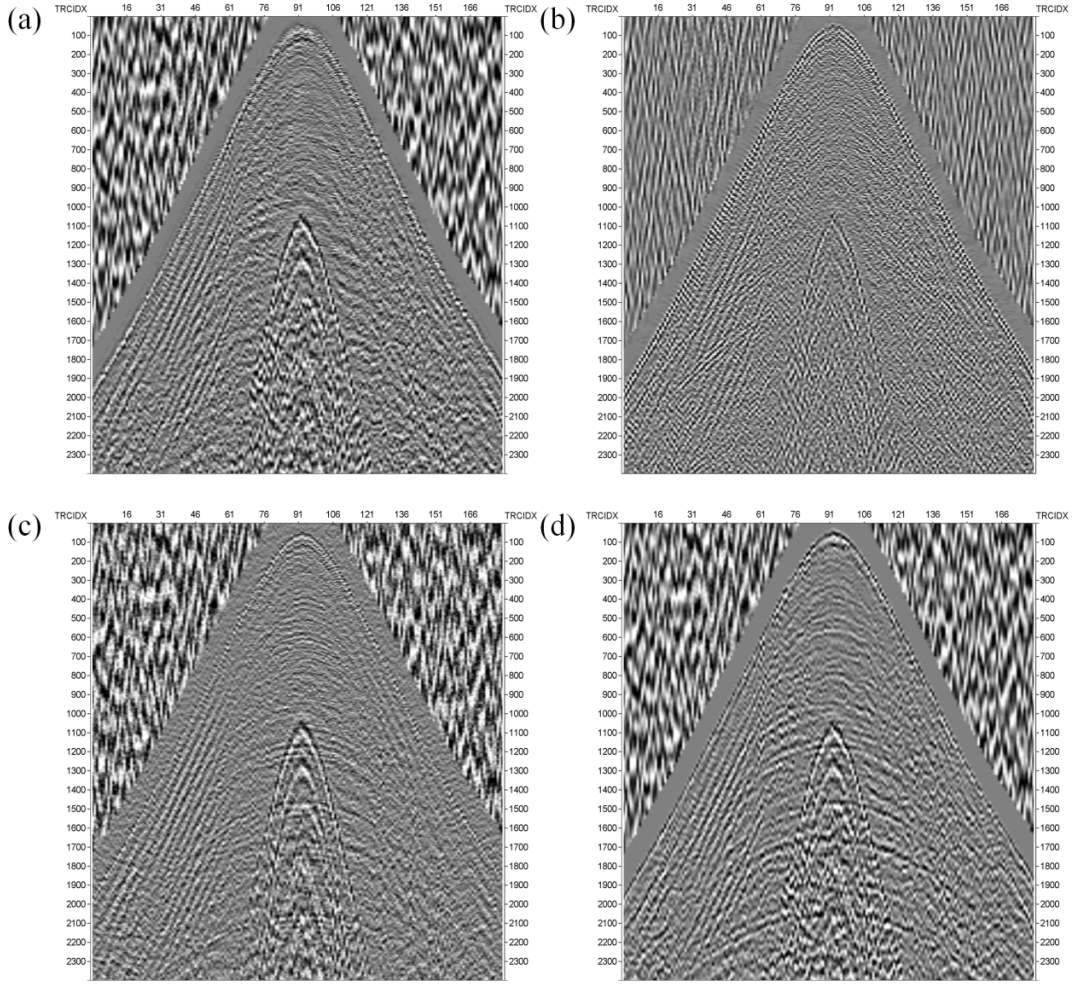

Fig. 6. The comparison of difference records. (a) The difference record of IST-CNNCS (Fig. 5a minus Fig. 5b). (b) The difference record of shearlet transform with hard threshold function (Fig. 5a minus Fig. 5c). (c) The difference record of F-X deconvolution (Fig. 5a minus Fig. 5d). (d) The difference record of EMD (Fig. 5a minus Fig. 5e).

Denoising time

The denoising time-cost of IST-CNNCS mainly includes two parts: training time and denoising time. The former refers to the training time of the four CNN-based coefficient selection models, and the latter refers to the time-cost of multi-scale and multi-direction decomposition by shearlet transform, coefficient prediction by the trained CNN-based models, and coefficient construction. In this paper, the training time of the four CNN-based coefficient selection models is 5.92 hr, 6.13 hr, 6.51 hr, 6.41 hr, respectively, a total of 24.97 hr. However, for the similar desert seismic records, we only need to train the four CNN-based models once; in other words, after completing the network training, the time cost of IST-CNNCS is denoising time rather than the sum of denoising time and training time. Therefore, we just compare the denoising time of the four methods, and not consider the training time of IST-CNNCS.

We utilize IST-CNNCS, shearlet transform with hard threshold, F-X deconvolution, and EMD to process some desert seismic records with different sizes, and the concrete denoising time is listed in Table 4. It can be discovered that, in terms of denoising time, IST-CNNCS is obviously inferior to F-X deconvolution, slightly inferior to shearlet transform with hard threshold function, and superior to EMD. However, considering the denoising performance, compared with the three traditional methods, IST-CNNCS is still more suitable for the denoising of desert seismic records.

Table 4. The comparison of denoising time (s).

Record size	IST-CNNCS	Shearlet transform	F-X	EMD
600×100	4.1324	3.9015	1.2439	37.1073
1000×100	7.9125	7.7139	1.8627	69.1542
1200×200	14.3612	13.9327	2.3614	136.1495
1800×250	22.1592	21.0593	2.8761	241.9372

In general, deep-learning-based algorithms need lots of time to train the network. Decreasing the training time is mainly based on the hardware enhancement and network structure adjustment. The former is not part of our research field, so our next goal is to shorten the training time of CNN by the network structure adjustment. Specifically, we plan to carry out super-parameter optimization and dense-skip connection for CNN, which can significantly reduce the complexity of the network and accelerate the training speed of network.

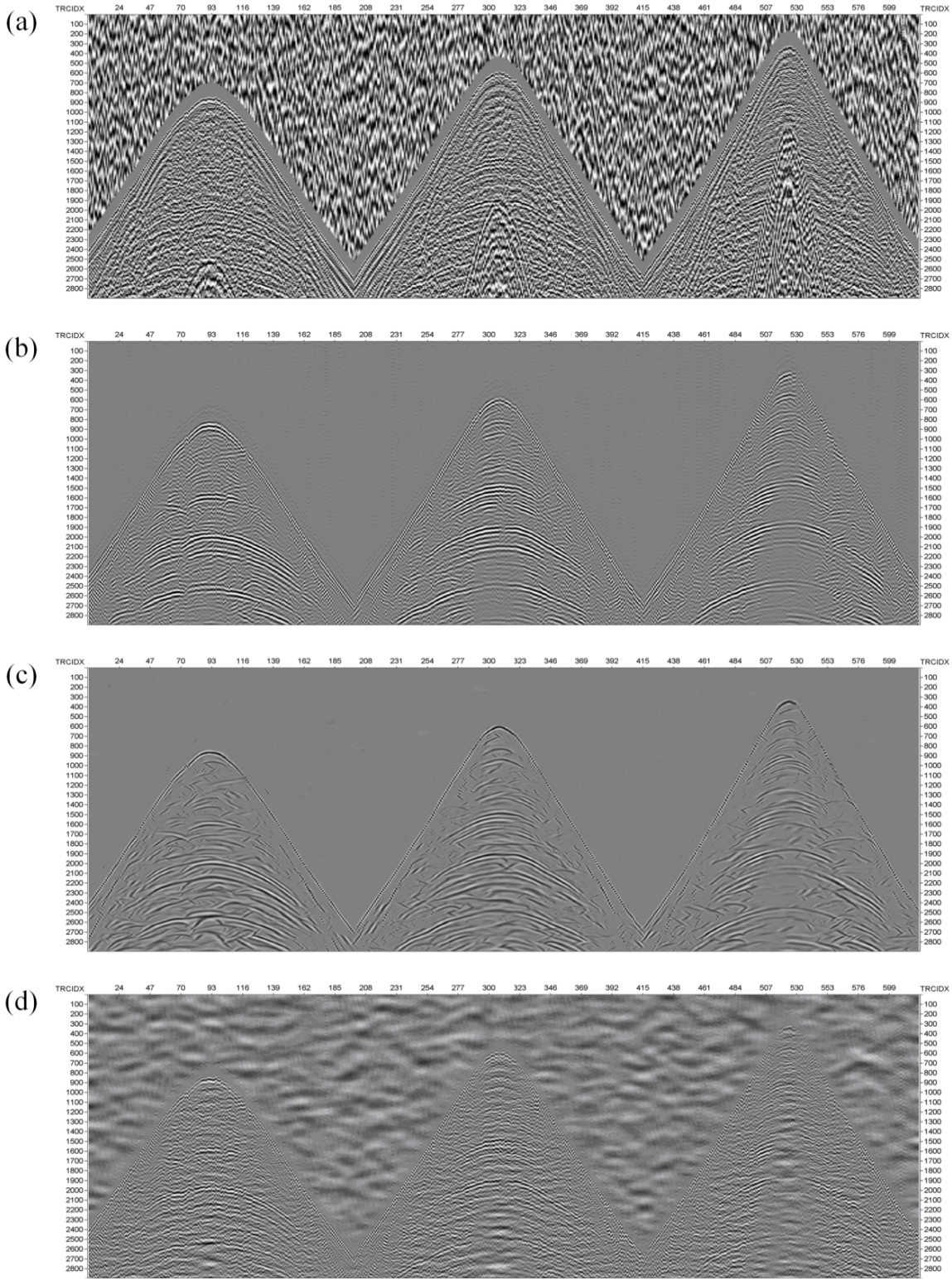

Fig. 7. The denoised results of a shot gather by using deep-learning methods. (a) The shot gather from the desert region. (b) The denoised result by using IST-CNNCS. (c) The denoised result by using DnCNNs. (d) The denoised result by using TNRD. The network parameters of DnCNNs are the same as that of IST-CNNCS. In TNRD, the network depth, filter number in each layer, patch size, and filter size are 8, 24, 100×100 and 5×5. Moreover, we use the 20 synthetic clean seismic data and the real desert low-frequency noise which together make up the training dataset of IST-CNNCS to train DnCNNs and TNRD.

Example 2

The proposed IST-CNNCS is a deep-learning-based denoising method, so we compare it with two classical deep-learning-based methods: DnCNNs and TNRD; the network parameters and training dataset of these two deep-learning based methods are introduced in the caption of Fig. 7. Fig. 7(a) shows a shot gather including three common-shot-point records, and its sampling interval, trace interval, and receiver number are 0.002 s, 20 m, and 616, respectively. Figs. 7(b), (c), and (d) are the denoised results by using IST-CNNCS, DnCNNs, and TNRD, respectively.

In the denoised result by using TNRD (Fig. 7d), there are still lots of residual random noise and surface waves; the events recovered by TNRD are unclear and hard to be recognized. Comparatively speaking, the denoising performance of TNRD is obviously inferior to that of the other two CNN-based denoising methods. In the denoised results by using IST-CNNCS and DnCNNs (Figs. 7b and 7c), the two CNN-based methods can suppress most of the random noise and surface waves. However, there are some false events in the denoised result by using DnCNNs (Fig. 7c). This phenomenon is mainly due to two reasons: (1) the effective signal dataset is composed of synthetic data which is inevitable different from real data, even if we adopt multiple seismic wavelets and change their parameters; (2) CNN adopts the paired training. On the contrary, as shown in Fig. 7(b), by making use of the good sparse representation of shearlet transform, IST-CNNCS can effectively avoid the appearance of false events. After the processing of IST-CNNCS, most of the random noise and surface waves are reduced and the events become clear and continuous.

CONCLUSION

Lots of traditional seismic denoising methods are constrained by the low SNR and serious spectrum overlapping of the desert seismic data. The denoised results often troubled from two problems: residual low-frequency noise and effective signal attenuation. To solve this problem, we combine shearlet transform and CNN to propose a novel denoising algorithm called IST-CNNCS. In IST-CNNCS, the desert seismic data is decomposed into four sub-bands, and we construct four CNN-based coefficient selection models to achieve multi-scale and multi-direction precise denoising. After the coefficient predictions by the constructed CNN-based models and then coefficient reconstruction, we can obtain the denoised desert seismic data. Both synthetic and real examples demonstrate that IST-CNNCS is superior to the traditional seismic denoising methods and that it can suppress the desert low-frequency noise thoroughly and recover the reflected effective signals completely.

In summary, the proposed IST-CNNCS has three significant advantages. (1) IST-CNNCS is a multi-scale and multi-direction precising denoising algorithm, this is similar to setting different thresholds to different sub-bands according to the coefficient features. (2) we use the trained

CNN-based models to achieve the automatic and accurate coefficient selections, which avoids to set the threshold manually. Therefore, IST-CNNCS can be called intelligent denoising algorithm. (3) through the good sparse representation of shearlet transform, IST-CNNCS can effectively avoid the appearance of false events and significantly enhance the continuity of events. It can be concluded from (2) and (3) that IST-CNNCS is not a simple combination of CNN and shearlet transform, but complement each other.

Currently, almost all deep-learning-based algorithms need clean effective signal data to complete the network training. This greatly limits their application in seismic data denoising due to the unavailability of real seismic effective signals and the inevitable difference between real seismic effective signals and synthetic seismic effective signals. In future studies, we plan to leverage the forward modelling method to minimize this difference, but this strategy cannot solve this problem fundamentally. Therefore, we also plan to adjust the training way and network architecture to realize the unpaired training between the noisy seismic data and the pure noise data. In this way, we can complete the network training only using real data rather than the mixture of real and synthetic data.

ACKNOWLEDGEMENT

We would like to thank the financial support by the Postdoctoral Innovation Talent Support Program of China (BX2021111) and the National Natural Science Foundation of China (41730422), and an anonymous oilfield company for providing the desert seismic dataset.

DATA AVAILABILITY STATEMENT

The synthetic data in this article can be obtained by contacting the corresponding author: Yue Li (liyue@jlu.edu.cn); The real dataset associated with this research is confidential and cannot be released.

REFERENCES

Akram, J., 2018. An application of waveform denoising for microseismic data using polarization-linearity and time-frequency thresholding. Geophys. Prosp., 66: 872-893.

Bekara, M. and van der Baan, M., 2007. Local singular value decomposition for signal enhancement of seismic data. Geophysics, 72(2): V59-V65.

Bekara, M. and van der Baan M., 2009. Random and coherent noise attenuation by empirical mode decomposition. Geophysics, 74(5): V89-V98.

Beckouche, S. and Ma, J.W., 2014. Simultaneous dictionary learning and denoising for seismic data. Geophysics, 79(3): A27-A31.

Bonar, D. and Sacchi, M., 2012. Denoising seismic data using the nonlocal means algorithm. Geophysics, 77(1): A5-A8.

Chen, Y.K., Zhang, M., Bai, M. and Chen, W., 2019. Improving the signal-to-noise ratio of seismological datasets by unsupervised machine learning. Seismol. Res. Lett., 90: 1552-1564.

Dong, X.T., Li, Y. and Yang, B.J., 2019a. Desert low-frequency noise suppression by using adaptive DnCNNs based on the determination of high-order statistic. Geophys. J. Internat., 219: 1281-1299.

Dong, X.T., Jiang, H., Zheng, S., Li, Y. and Yang, B.J., 2019b. Signal-to-noise ratio enhancement for 3C downhole microseismic data based on the 3D shearlet transform and improved back-propagation neural networks, Geophysics 84(4),V245-V254.

Dong X.T., Zhong T. and Li Y., 2020a. New suppression technology for low-frequency noise in desert region: the improved robust principle component analysis based on prediction of neural network. IEEE Transact. Geosci. Remote Sens., in Press.

Dong, X.T., Zhong, T. and Li, Y., 2020b. A deep-learning-based denoising method for multiarea surface seismic data. IEEE Geosci. Remote Sens. Lett., in Press.

Duncan, G. and Beresford, G., 1995. Median filter behaviour with seismic data. Geophys. Prosp., 43: 329-345.

Fomel, S. and Liu, Y., 2010. Seislet transform and seislet frame. Geophysics, 75(3): V25-V38.

Gholami, A. and Zand, T., 2017. Fast L1-regularized radon transforms for seismic data processing. Digit. Sign. Process., 71: 83-94.

Goudarzi, A. and Riahi, M.A., 2012. Seismic coherent and random noise attenuation using the undecimated discrete wavelet transform method with WDGA technique. J. Geophys. Engineer., 9: 619-631.

Guo, K. and Labate, D., 2007. Optimally sparse multidimensional representation using shearlets. SIAM J. Mathemat. Anal., 39: 298-318.

Gülünay, N., 2017. Signal leakage in f -x deconvolution algorithms. Geophysics, 82(5): W31-W45.

Han, J.J. and van der Baan, M., 2013. Empirical mode decomposition for seismic time-frequency analysis. Geophysics, 78(2). O9-O19.

Harris, P.E. and White, R.E., 1997. Improving the performance of f-x prediction filtering at low signal-to-noise ratios. Geophys. Prosp., 45: 269-302.

Jin, K.H., McCann, M.T., Froustey, E. and Unser, M., 2017. Deep convolutional neural network for inverse problems in imaging. IEEE Transact. Image Process., 26: 4509-4522.

Kulesh, M., Diallo, M., Holschneider, M., Kurennaya, K., Kruger, F., Ohrnberger, M. and Scherbaum, F., 2007. Polarization analysis in the wavelet domain based on adaptive covariance method. Geophys. J. Internat., 170: 667-678.

Kutyniok, G. and Lim, W., 2011. Compactly supported shearlets are optimally sparse. J. Approxim. Theory, 163: 1564-1589.

Li, F., Zhang, B., Verma, S. and Marfurt, K.J., 2017. Seismic signal denoising using thresholded variational mode decomposition. Explor. Geophys., 49: 450-461.

Li, G.H., Li, Y. and Yang, B.J., 2017. Seismic exploration random noise on land: Modeling and application to noise suppression. IEEE Transact. Geosci. Remote Sens., 55: 4668-4681.

Liang, X.Q., Li, Y. and Zhang, C., 2018. Noise suppression for microseismic data by non-subsampled shearlet transform based on singular value decomposition. Geophys. Prosp., 66: 894-903.

Lim, W., 2010. The discrete shearlet transform: A new directional transform and compactly supported shearlet frames. IEEE Transact. Image Process., 19, 1166-1180.

Loffe, S. and Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. Proc., 32nd Internat. Conf. Mach. Learning, 448-456.

Ma, H.T, Yao, H.Y., Li, Y. and Wang, H.Z., 2020a. Deep residual encoder-decoder networks for desert seismic noise suppression. IEEE Geosci. Remote Sens. Lett., 17: 529-533.

Ma, H.T, Yan, J. and Li, Y., 2020b. Low-frequency noise suppression of desert seismic data based on variational mode decomposition and low-rank component extraction. IEEE Geosci. Remote Sens. Lett., 17: 337-341.

Mou, L.C., Ghamisi, P. and Zhu, X.X., 2018. Unsupervised spectral-spatial feature learning via deep residual conv-deconv network for hyperspectral image classification. IEEE Geosci. Remote Sens. Lett., 56: 391-406.

Nazari Siahsar, M.A., Gholtashi, S., Abolghasemi, V. and Chen, Y., 2017. Simultaneous denoising and interpolation of 2D seismic data using data-driven non-negative dictionary learning. Sign. Process., 141: 309-321.

Neelamani, R., Baumstein, A.I., Gillard, D.G., Hadidi, M.T. and Soroka, W.L., 2008. Coherent and random noise attenuation using the curvelet transform. The Leading Edge, 27: 240-248.

Remez, T., Litany, O., Giryes, R. and Bronstein, A.M., 2018. Class-aware fully convolutional Gaussian and Poisson denoising. IEEE Transact. Geosci. Remote Sens., 27: 5707-5722.

Stein, R.A. and Bartley, N.R., 1983. Continuously time-variable recursive digital band-pass filters for seismic signal processing. Geophysics 48: 702-712.

Tang, N., Zhao, X., Li, Y. and Zhu, D., 2018. Adaptive threshold shearlet transform for surface microseismic data denoising. J. Appl. Geophys., 153: 64-74.

Tian, Y.N., Li, Y. and Yang, B.J., 2014. Variable-eccentricity hyperbolic-trace TFPF for seismic random noise attenuation. IEEE Transact. Geosci. Remote Sens., 52: 6449-6458.

Tirer, T. and Giryes, R., 2019. Super-resolution via image-adapted denoising CNNs: Incorporating external and internal learning. IEEE Signal Processing Letters 26(7), 1080-1084.

Wang, F. and Chen, S., 2019. Residual learning of deep convolutional neural network for seismic random noise attenuation. IEEE Geosci. Remote Sens. Lett., 16: 1314-1318.

Wang, Y., 1999. Random noise attenuation using forward-backward linear prediction. J. Seismic Explor., 8: 133-142.

Wang, Y., Liu X., Gao F. and Rao Y., 2020. Robust vector median filtering with a structure-adaptive implementation. Geophysics, 85(5): V407-V414.

Wang, Y., Lu, W., Wang, B. and Liu, L., 2016. A robust polynomial principal component analysis for seismic noise attenuation. J. Geophys. Engineer., 13: 1002-1009.

Wu, S., Li, G.Q., Deng, L., Liu, L., Wu, D., Xie, Y. and Shi, L.P., 2019. L1-norm batch normalization for efficient training of deep neural networks. IEEE Transact. Neural Netw. Learning Syst., 30: 2043-2051.

Wu, Y. and McMechan, G.A., 2019. Parametric convolutional neural network-domain full-waveform inversion. Geophysics, 84(6): R881-R896.

Xiong, M.J., Li, Y. and Wu, N., 2014. Random-noise attenuation for seismic data by local parallel radial-trace TFPF. IEEE Transact Geosci. Remote Sens., 52: 4025-4031.

Yang, W., Feng, J., Yang, J., Zhao, F., Liu, J., Guo, Z. and Yan, S., 2017. Deep edge guided recurrent residual learning for image super-resolution. IEEE Transact. Image Process., 26: 5895-5907.

Yang Y., Liu C. and Langston C.A., 2020. Processing seismic ambient noise data with the continuous wavelet transform to obtain reliable empirical Green's functions. Geophys. J. Internat., 222: 1224-1235.

Yuan, S.Y., Liu, J.W., Wang, S.X., Wang, T.Y. and Shi, P.D., 2018. Seismic waveform classification and first-break picking using convolution neural networks. IEEE Geosci. Remote Sens. Lett., 15: 272-276.

Zhang, K., Zuo, W.M., Chen, Y.J., Meng, D.Y. and Zhang, L., 2017. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. IEEE Transact. Image Process., 26: 3142-3155.

Zhong, T., Li, Y., Wu, N., Nie, P.F. and Yang, B.J., 2015a. Statistical analysis of background noise in seismic prospecting. Geophys. Prosp., 63: 1161-1174.

Zhong, T., Li, Y., Wu, N., Nie, P.F. and Yang, B.J., 2015b. A study on the stationarity and Gaussianity of the background noise in land-seismic prospecting. Geophysics, 80(4): V67-V82.

APPENDIX A

THE BASIC THEORY OF SHEARLET TRANSFORM

Shearlet transform proposed by Guo and Labate (2007) is constructed by an affine system with composite dilations (i.e., the expansion, shear and translation of basic function). Compared with some conventional multi-scale geometric analysis methods including wavelet transform, curvelet transform, and contourlet transform, it has a simpler mathematical structure and requires fewer coefficients when approximating a curve (Liang et al., 2018; Tang et al., 2018; Dong et al., 2019b). Furthermore, shearlet transform has many advantages including anisotropy, sensitive directivity and optimal sparse representation (Guo and Labate, 2007; Lim, 2010), so it can well capture the geometric characteristics of multidimensional signals and provide the optimal sparse representation for signals (Dong et al., 2019b).

For 2D space, the affine systems with composite dilations can be expressed as follows (Kutyniok and Lim, 2011):

$$M_{AB}(\psi)=\left\{\psi_{j,l,k}(x)=\left|\det A\right|^{j/2}\psi\left(B^{l}A^{j}I-k\right)\quad j,l\in Z,k\in Z^{2}\right\},\tag{A-1}$$

where $\psi\in L^{2}(R^{2})$, I represents input signal, A,B are 2×2 invertible matrices and $\left|\det B\right|=1$. Dilation matrices A^{j} control the scale transformations, j、l、k are the scale, location and direction parameters, respectively, and matrices B^{j} are associated with geometrical transformations including rotations and shear. If $M_{AB}(\psi)$ satisfies the Parseval frame (i.e., tight frame), that is for any $f\in L^{2}(R^{2})$,

$$\sum_{j,l,k}\left|\left\langle f,\psi_{j,l,k}\right\rangle\right|^{2}=\left\|f\right\|^{2}.\tag{A-2}$$

The elements of above affine systems with composite dilations are called complex wavelet. Shearlet is a special form of the composite wavelet.

When $A=A_{0}=\begin{bmatrix}4 & 0\\ 0 & 2\end{bmatrix}$, $B=B_{0}=\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}$, shearlet can be expressed as follows:

$$\psi_{j,l,k}=\left|\det A_{0}\right|^{j/2}\psi\left(B_{0}^{l}A_{0}^{j}I-k\right)\quad j,l\in Z,k\in Z^{2},\tag{A-3}$$

where $\psi\in L^{2}(R^{2})$ is the band-limited function; A_0 is anisotropic dilation matrix whose function is scale transformation; B_0 is the shear matrix related to direction transformation. It can be clearly seen from eq. (A-3) that shearlet is obtained by proper dilation, shear and translation of the basis function.

APPENDIX B

THE BASIC ARCHITECTURE OF DnCNNs

In IST-CNNCS, we adopt the basic architecture of DnCNNs which is a classical network for noise suppression. As shown in Fig. B-1, the DnCNNs consists of convolution layers (Conv) , batch normalization layers (BN) and rectified linear units (ReLU). The Conv contains numerous convolutional kernels which can extract the data features by the convolution operation (Yuan et al., 2018; Dong et al., 2020b), and the concrete expression of its function is displayed in eq. (B-1).

$$C = w * I + b, \tag{B-1}$$

where C and I are the output and input of Conv, w and b represent weight and bias. BN is mainly used to control the mean value and variance of the input data in each layer to be approximately 0 and 1, respectively (Loffe and Szegedy, 2015; Wu et al., 2019). The basic function of BN is shown in the eqs. (B-2) to (B-5):

$$\mu_{BN} = \frac{1}{B}\sum_{i}^{B} a_{i,} \tag{B-2}$$

$$\delta_{BN} = \left[\frac{1}{B}\sum_{i=1}^{B}\left(a_i - \mu_{BN}\right)^2\right]^{\frac{1}{2}} \tag{B-3}$$

$$\hat{a}_i = \frac{a_i - \mu_{BN}}{\sqrt{\delta_{BN}^2 + \xi}}, \tag{B-4}$$

$$b_i = BN_{\gamma,\beta}(a_i) = \gamma\hat{a}_i + \beta, \tag{B-5}$$

where B represents batch size; a_i and b_i are the input and output of BN; ξ denotes a constant which can guarantee numerical stability; γ and β are learnable parameters optimized by network training. Some related research demonstrates that BN can effectively avoid the phenomenon of gradient dispersion and significantly accelerate the convergence speed of network (Loffe and Szegedy, 2015; Dong et al., 2020a).

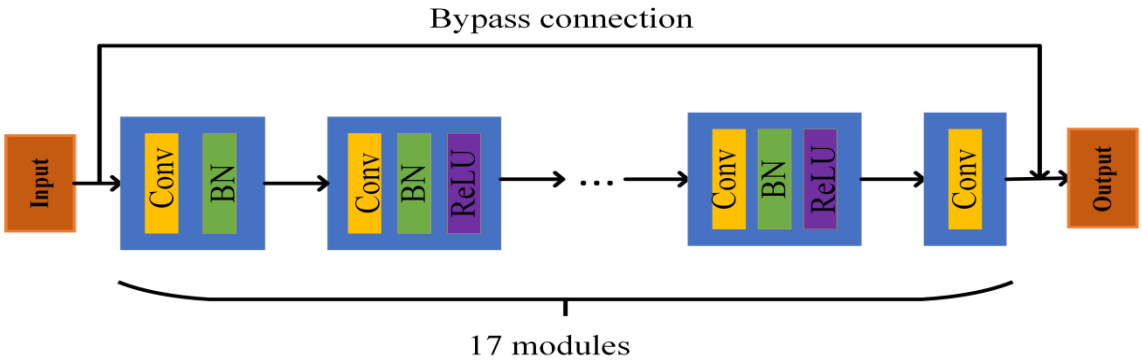

Fig. B-1. The architecture of DnCNNs.

ReLU is a widely-used activation function, whose function can be expressed as $f(x) = \max(0, x)$. Compared with traditional activation functions including Sigmoid and tanh, ReLU can significantly increase the speed of gradient descent and reduce the operation cost (Zhang et al., 2017).