IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 5, 2020, accepted December 25, 2020, date of publication January 6, 2021, date of current version January 21, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3049479

Residual Learning of Cycle-GAN for Seismic Data Denoising

WENDA LI[1,2,3] AND JIAN WANG[1,2], (Member, IEEE)

[1]Key Laboratory of Petroleum Resources Research, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China
[2]Institute of Earth Science, Chinese Academy of Science, Beijing 100029, China
[3]University of Chinese Academy of Sciences, Beijing 100049, China

Corresponding author: Wenda Li (geophysicslwd@163.com)

This research is supported by National Natural Science Foundation of China (Grant No.U20B200166) and Major State Research Development Program of China (Grant No. 2016YFC0601101).

ABSTRACT Random noise attenuation has always been an indispensable step in the seismic exploration workflow. The quality of the results directly affects the results of subsequent inversion and migration imaging. This paper proposes a cycle-GAN denoising framework based on the data augmentation strategy. We introduced residual learning into the cycle-GAN to improve the training efficiency of the network. We proposed a method for generating labeled datasets directly from unlabeled real noisy data. Then we significantly improve the diversity of the training samples through an augmentation strategy. Through RCGAN, we can realize intelligent seismic data denoising work, which dramatically reduces the manual selection and intervention of denoising parameters. Finally, numerical experiments prove that our method has a remarkably good random noise suppression ability and a minimally damaging effect on useful seismic signals. The experiment tests on synthetic and real data also show the effectiveness and superiority of the proposed method RCGAN compared to the state-of-the-art denoising methods.

INDEX TERMS Geophysical data, geophysics computing, generative adversarial networks, noise reduction, residual learning.

I. INTRODUCTION

In petroleum exploration, the precise processing of seismic data can directly affect subsequent inversion and migration imaging accuracy. Seismic data denoising is an indispensable step to improve the signal-to-noise ratio(SNR) of seismic data, and the result directly affects the quality of subsequent data processing. High SNR is essential for many seismic exploration techniques such as AVO analysis, seismic attribute analysis, and micro-seismic monitoring. The suppression of random noise is essential to improving the SNR. For seismic data acquisition, we usually use broadband to obtain more abundant reflected wave information. Therefore, while acquiring significant waves, various noise interferences are inevitably recorded, reducing the SNR of seismic data. This will seriously affect the SNR and resolution of the seismic migration profile (especially the deep layers), which will cause significant difficulties in data interpretation.

Researchers have proposed many effective algorithms for random noise suppression, mainly including widely used transform-domain denoising algorithms, spatial denoising algorithms, and comprehensive denoising algorithms. The transform domain denoising algorithms are the most widely used and most advanced at this stage. Canales [5] proposed an f-x deconvolution denoising algorithm according to the seismic events' predictability, which has become a standard seismic signal denoising algorithm. Then, Gulunay [7] and Hornbostel (1991) [36] further optimized the f-x deconvolution algorithm to improve random noise suppression. Fieire and Ulrych [26] proposed a denoising algorithm based on singular value decomposition according to the Low-rank hypothesis. At the same time, there are a series of transformation algorithms, such as curvelets transformation [6], seislet transformation [25], and wavelet transformation algorithms [22], all of which have good results. Many researchers further proposed a lot of improved wavelet transform algorithms to better random noise suppression. In recent years, many researchers have developed an adaptive transformation algorithm based on dictionary learning [9], [11], [13], [17], which made full use of the characteristics of the seismic data itself to denoise. The second is the spatial denoising algorithms. Rudin *et al.* [18] proposed an adaptive denoising

The associate editor coordinating the review of this manuscript and approving it for publication was Zihuai Lin.

algorithm based on a partial differential equation. Bonar and Sacchi (2012) [37] proposed a non-local mean value denoising method, which can find similar structure regions in seismic data, and then carry out partition denoising. The third is the comprehensive denoising algorithms, which combined the characteristics of spatial domain and transform domain. This algorithm usually performs structural similarity clustering in the spatial domain and then applies transform algorithms based on the clustering results to achieve a better low-rank effect [28].

Although conventional denoising methods have been gradually developed perfectly, there are still two bottlenecks. One is that the assumptions of the algorithm are inaccurate [23]. The f-x deconvolution [5] method assumed that the random noise is unpredictable, and the coherence signal is predictable, which is difficult to satisfy in the complex geological structure areas. The other is that the denoising process parameters are improper due to manual intervention [23]. In the actual denoising work, selecting denoising parameters usually requires engineers to have a deep understanding of the work area, which significantly increases labor costs and reduces efficiency. For seismic data denoising, can we form a uniform framework without unnecessary assumptions and excessive manual intervention?

Deep learning(DL) networks can help us break these bottlenecks. The secret relationship between seismic data and noise data can be extracted through many training datasets, which does not require too much human intervention and accuracy assumptions.

Since AlexNet [1] won the ImageNet Large-Scale Visual Recognition Challenge(ILSVRC) competition, neural network research has gradually become the mainstream research direction of computer science. In this booming environment, various efficient network structures have been proposed, such as VGGNet (Simonyan and Zisserman, 2014 [43]), U-Net (Ronneberger, 2015 [44]), FCN (Shelhamer *et al.*, 2014 [45]), and CNN [16], etc. These network structures have been widely used in image denoising, face recognition, speech recognition, etc. Hinton *et al.* [6] proposed a deep belief network, and Vincent *et al.* [21] presented a stacked auto-encoder, among others. Lecun *et al.* [16] demonstrated that CNN has fewer parameters and provides superior classification results on the MNIST. Zhang *et al.* [14] proposed a CNN with 17 layers, named DnCNN, for noise attenuation of images. Zhang *et al.* [15] developed a fast and flexible denoised convolutional neural network, FFDNet, to further improve the denoising capability and flexibility of DL networks.

Generative Adversarial Networks (GANs) have achieved impressive image generation results, image editing, and representation learning. GANs have the characteristics of strong parallel processing ability, intense adaptive energy, and good fault tolerance. Generating adversarial networks (GANs) have been called the most exciting idea in machine learning over the past decade by Yann LeCun. GANs are state-of-the-art deep learning neural networks at this stage, so it is meaningful to explore their seismic exploration applications. GANs do not ignore the structure of the input datasets. Due to the seismic data's strong local structure, the application of GANs in seismic exploration has more advantages.

In recent years, neural networks have also been widely used in seismic signal research, such as seismic inversion (Das *et al.*, 2018 [38] and Yang and Ma, 2019 [39]), fault detection (Xiong *et al.*, 2018 [40]), and interpolation (Wang *et al.*, 2018 [41]). In the field of seismic data denoising, neural networks also have good applicability results. Liu *et al.* [17] proposed a data augmentation and U-Net based seismic random noise suppression method using a small amount of synthetic seismic data to perform augmentation and generate labeled datasets to train the network. Yu *et al.* [23] introduced DnCNN in image denoising into seismic data denoising and achieved good results. Zhang *et al.* (2020) [42] used the trained neural network learning from natural image denoising to realize seismic data interpolation.

This paper proposed a cycle-GAN based on residual learning (RCGAN) and achieved noise suppression work for real and synthetic seismic data. Firstly, we introduce the cycle-GAN into the seismic denoising domain and greatly improved the network's training efficiency by employing residual learning. Besides, this paper applied an effective data generation and augmentation method so that the trained network can be better adapted to real seismic data denoising work. Finally, the effectiveness of the method is proved by numerical tests.

II. METHODOLOGY

From seismic data processing, the raw seismic data can be regarded as the sum of signal and noise, which can be expressed as :

$$Y = X + N. \tag{1}$$

It can be expressed as: Where X is the clean data, Y is the collected raw seismic data, and N is the noise data.

A. RCGAN DENOISING WORKFLOW

In this section, we first introduce the workflow of this paper so that the reader can get a better understanding of our methods and the denoising process based on neural networks. Fig.1 shows the diagram of denoising workflow. It is apparent from FIGURE 1. that the raw training and test datasets in this paper are composed of synthetic and real seismic data.

In the process of neural network training (the flow indicated by the orange arrows in FIGURE 1.), we first added random noise to the clean datasets to obtain the synthetic datasets. Next, we performed data augmentation by a data augmentation strategy proposed in this paper and finally formed the augmentation dataset. We perform a patch sampling operation on the augmentation dataset to obtain the training dataset, which can be expressed as :

$$y^* = P[Y] \tag{2}$$

where $P[\cdot]$ is the patch sampling operator.

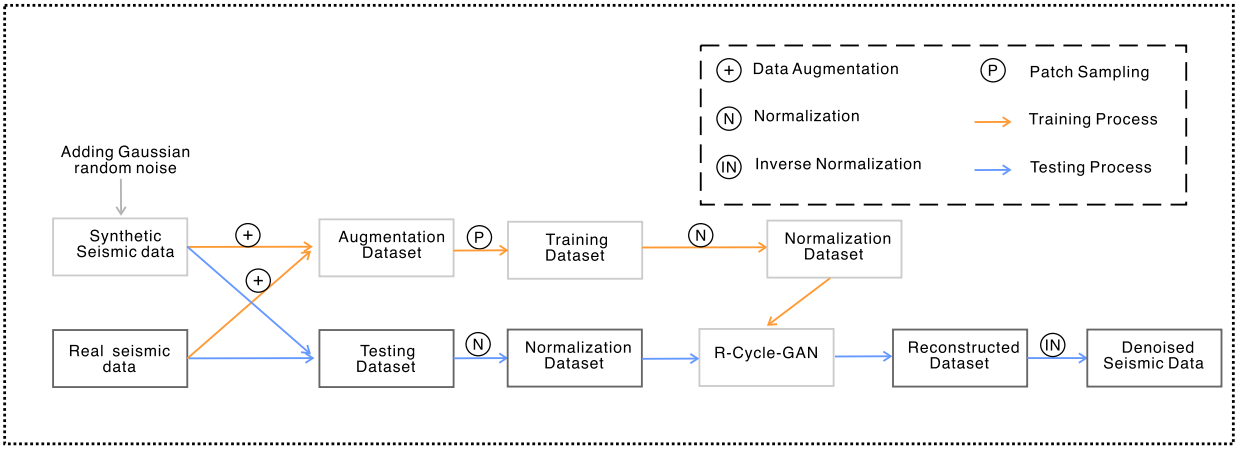

FIGURE 1. The diagram of denoising workflow.

We can split the input into small patches for patch size choice instead of directly using the entire section as the patch size. The most significant advantage of this choice is that it can save memory, and it is feasible because there is no coherence between seismic data and random noise. However, this does not mean that the network testing process's input is still the same patch size. On the contrary, the input size can be selected arbitrarily. Then we performed a normalization operation to make the dataset better adapt to the RCGAN, which can be expressed as :

$$y = \frac{y^* - min}{max - min} \tag{3}$$

The process indicated by the blue arrow in FIGURE 1. is the network testing process. The test dataset also requires normalization operation compared to the training process, but it does not need patch sampling. After the network results are obtained from RCGAN, an inverse normalization operation is required to get the final denoising seismic data.

B. NETWORK ARCHITECTURE AND THEORY

Generative Adversarial Networks (GANs) have achieved impressive image generation results, image editing, and representation learning. Besides, GAN is also very well-suited for predictions between unpaired data. We usually consider the seismic signal and random noise to be typically unpaired data. Our objective is to find a mapping to connect the raw seismic data to the useful seismic signal through a neural network (Cycle-GAN proposed by Zhu *et al.* [11]) for denoising. Besides, we improved the training efficiency by introducing residual learning (Kai *et al.*, 2017 [15]) for the slow training of GAN networks. So the mapping established through our method Residual-Cycle-GAN(RCGAN) is from raw seismic data to random noise.

FIGURE 2. shows the internal architecture of the RCGAN, which contains two mapping functions. We took the raw seismic data as the inputs and the random noise data as the network outputs. The first mapping function is $INPUT \sim FG(INPUT)$, which is based on the forward generator. There must be a reverse mapping function based on backward generator can translate $FG(INPUT)$ to $BG(FG(INPUT))$. RCGAN's cycle consistency means that if we can find a forward mapping, there must also be a reverse mapping. Just like we can translate a sentence from English to French and then translate it back [11]. In addition to the training of two generators in the network, two discriminators are also needed to be trained, as shown in FIGURE 2.

RCGAN contains two generators and two discriminators, respectively. The internal structure of two generators (as shown in FIGURE 3.) is consistent with that proposed by Johnson *et al.* [12], which has four convolution layers and nine residual blocks. For residual blocks, it includes convolution layers, Batch normalization, and ReLU. Similar to Johnson *et al.* [12], we ignore the pooling layer to hold the data size. For discriminators, we adopt the PatchGANs (Li *et al.*, 2016 [46]) with a size of 70×70, which are trained to distinguish between generative data and training data.

C. LOSS FUNCTION AND SOME TRAINING DETAILS

Our goal is to learn mapping functions between two domains $INPUT$ and $OUTPUT$ given training samples $\{x_i\}_{i=1}^N$ and $\{y_i\}_{i=1}^N$ where N is the sampling numbers. We generated the data distribution as $INPUT \sim \{y_i\}_{i=1}^N \in I$ and $OUTPUT \sim \{r_i\}_{i=1}^N = \{y_i - x_i\}_{i=1}^N \in O$ where y_i, x_i and r_i denoted the raw seismic data, clean seismic data and noise data, respectively. For RCGAN of this paper which includes two mappings $FG : I \rightarrow O$ and $BG : O \rightarrow I$ will be introduced two adversarial discriminators D_f and D_b.

Our objective loss function contains two types of terms: adversarial losses for matching the I to O; and cycle consistency loss (Zhu and Alexel, 2017 [11]) to prevent the two learned mappings FG and BG from contradicting each other. Then we will introduce each of the two losses functions.

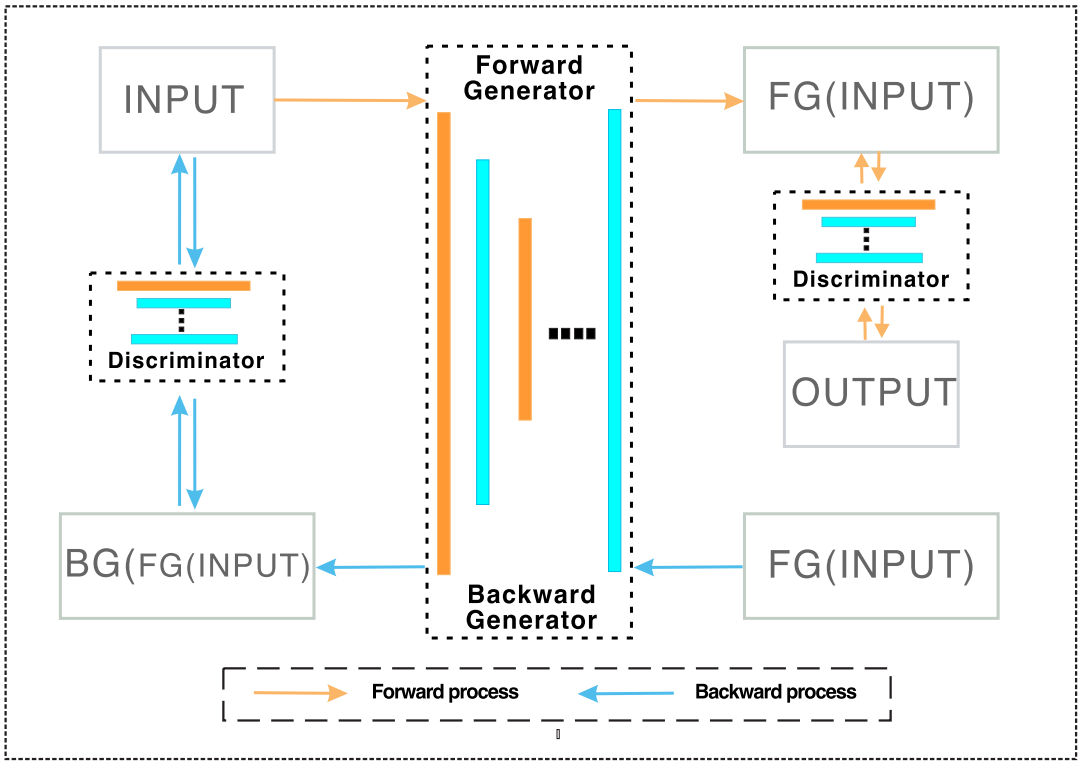

FIGURE 2. The architecture of the RCGAN.

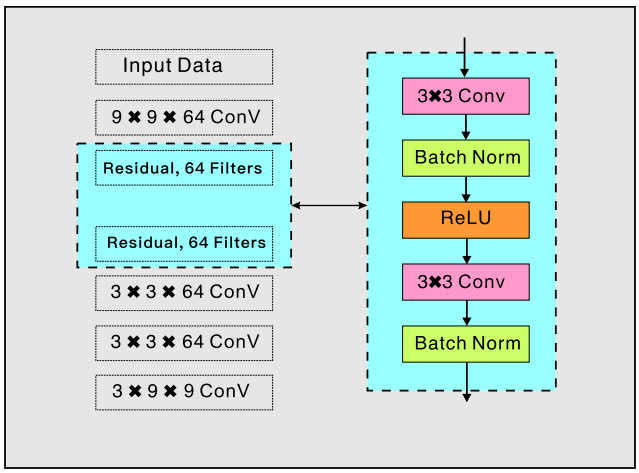

FIGURE 3. The architecture of generators.

1) ADVERSARIAL LOSS

The adversarial losses, which are composed of two parts, are applied to both mapping functions. The first part is the mapping function $FG : I \rightarrow O$ which objective loss function can be expressed as:

$$\mathcal{L}_{RCGAN}\left(FG, D_f, I, O\right) = \mathbb{E}_{r_i \in O}\left[\log D_f(r_i)\right] + \mathbb{E}_{y_i \in I}\left[\log(1 - D_f(FG(y_i)))\right] \quad (4)$$

where FG denoting the forward generator can translated $\{y_i\}_{i=1}^{N}$ to the similar data from domain O, while D_f can distinguish between generated data $FG\left(y_i\right)$ and real samples r_i. The generator FG aims to minimize the objective against while discriminator D_f are the inverse process, which means $\min_{FG}\max_{D_f}\mathcal{L}_{\mathrm{RCGAN}}\left(FG, D_f, I, O\right)$.

The second part is the mapping $BG : O \rightarrow I$, which objective loss function is similar to the $FG : I \rightarrow O$, we express it as:

$$\mathcal{L}_{RCGAN}\left(BG, D_b, O, I\right) = \mathbb{E}_{y_i \in I}\left[\log D_b(y_i)\right] + \mathbb{E}_{r_i \in O}\left[\log(1 - D_b(BG(FG(y_i))))\right] \quad (5)$$

where BG is a backward process to FG, it can translate the previously generated result $FG\left(y_i\right)$ to the similar data from domain I. The adversarial process of backward generator BG with discriminator D_b means $\min_{BG}\max_{D_b}\mathcal{L}_{\mathrm{RCGAN}}\left(BG, D_b, O, I\right)$.

2) CYCLE CONSISTENCY LOSS

In the network training process, we cannot guarantee that the learned function can map individual input y_i to the desired output r_i by only using the adversarial losses alone. We introduced the cycle consistency loss using L_1 norm (Zhu and Alexel, 2017 [11]) to prevent the two learned mappings FG and BG from contradicting each other. We modified the cycle consistency loss to equation 6. As shown in FIGURE 4, for each data y_i from domain I, the cycle-consistency should be able to bring it back to the similarly original state.

$$\mathcal{L}_{cyc}\left(FG, BG\right) = \mathbb{E}_{y_i \in I}\left[\left\|BG\left(FG\left(y_i\right)\right) - y_i\right\|_1\right] \quad (6)$$

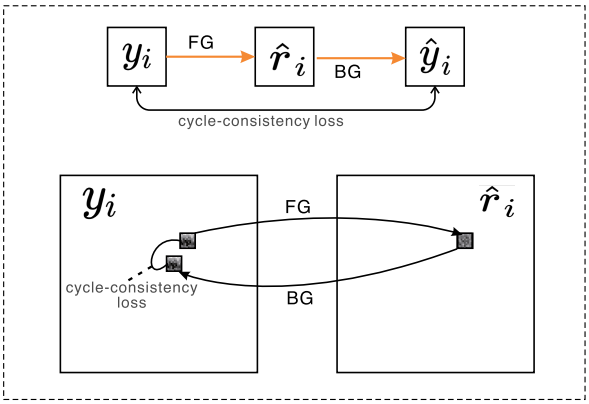

FIGURE 4. The diagram of cycle-consistency loss.

3) FULL OBJECTIVE LOSS FUNCTION

Our full objective loss function is expressed as:

$$\mathcal{L}\left(FG, BG, D_f, D_b\right) = \mathcal{L}_{RCGAN}\left(FG, D_f, I, O\right) + \mathcal{L}_{RCGAN}\left(BG, D_b, O, I\right) + \mathcal{L}_{cyc}\left(FG, BG\right) \quad (7)$$

where Equation 7 is the combination of Equation 4, 5 and Equation 6.

4) SOME TRAINING DETAILS

The RCGAN for seismic data denoising in this paper is based on the open-source machine learning framework PyTorch (https://pytorch.org). Based on the Ubuntu operating system, PyTorch 1.0 version, and Python 3.7 version. The network in this paper is trained on workstations with two GTX960 GPU, Inter(R) i7-4790k @4GHZ CPU,64GB RAM, and uses GPU (graphics processor units) as an acceleration means. The GPU model is GTX960, which has 4GB VRAM and 192 computing cores.

In this paper, we applied two techniques to make our model training process more stable. The first technique [30] is aiming at $\mathcal{L}_{RCGAN}$(Equation 4, 5) by replacing the negative log-likelihood objective by a least-squares loss. Second, we applied Shrivastava *et al.*'s strategy (2016) to reduce the model oscillation.

This section will also introduce the selection of some essential parameters in the network training process. We set the patch size to 70×70 in that we can split the input into small patches instead of directly using the entire section as the patch size, which can greatly save the computing memory. For the choice of batch size, we set it to 128 according to Bengio's (2012) study that the sweet region of the batch size is between 1 and a few hundred. At last, We set the learning rate to 10^{-3}.

D. DATA AUGMENTATION STRATEGY

A qualified deep learning training data set should contain the following three characteristicslarge of numbers, diverse and well-labeled. For the preparation of the dataset for the denoising work of GAN, it should not only contain synthetic seismic data with diversity. Still, it should also include a training dataset of real seismic data, which can significantly improve real data denoising work adaptability.

1) AUGMENTATION STRATEGY OF SYNTHETIC DATASET

For synthetic data, we can add random noise generated by computer simulations to the clean data. To make the network more tolerant to noise variance, we add Gaussian noise with different levels of variance to the training set. Since the signal intensity varies in different regions on the same data, the local SNR of the synthetic data will be various. Therefore, the Gaussian noise added in this paper is additive Gaussian noise with the corresponding variance level added according to the signal intensity in the local region. If we take the mean value of zero, the standard deviation is calculated as follows.

$$\sigma = \gamma \cdot mean(abs(x_{part})), \gamma \in [0.1, 4] \quad (8)$$

where σ is the noise standard deviation, $abs(\cdot)$ means to obtain the absolute value of area x_{part} point-by-point, and $mean(\cdot)$ is to obtain the average value. The value range of γ is obtained based on experience.

The synthetic datasets are downloaded from SEG (Society of Exploration Geophysicists) open datasets (https://wiki.seg.org). In this paper, the training sets are obtained from the following URL. http://s3.amazonaws.com /open.source.geoscience/opendata. If readers need to train the network, they can download the following data for testing.

The inputs can be split into small patches size of 70×70 rather than using the whole sections directly for saving computing memory because random noises are locally incoherent with useful data. But this does not mean that the input size we choose when we use the trained network model for denoising work is still 70×70. Even the input size of the network for testing processing can be any of the sizes. We chose the 2-D and 3-D, pre-stack data, and post-data as the datasets to guarantee the diversity of the training sets. Owing to the data of adjacent shots are very similar, we usually choose the data with a large interval. Finally, we generated a training set contains nearly 25000 samples through the Monte Carlo strategy [24], which can eliminate useless seismic data that are almost all zeros and not helpful for our training.

For the preparation of the dataset, this paper also introduced a data augmentation strategy, which further enhanced the diversity of the dataset by performing the following operations such as rotation transformation, mirror transformation, space-time downsampling, and intensity transformation on synthetic seismic data. As the FIGURE 5 shows, (a) and (e) are the clean synthetic data $\sim x_i$ and contaminated seismic data $\sim y_i$ by random noise, respectively. (b), (c) and (d) are the clean synthetic data augmented from (a) by the rotation + mirror transformation, rotation transformation, and mirror transformation, respectively. (f), (g) and (h) are the corresponding contaminated seismic data.

With the data augmentation method proposed in this paper, the final synthetic data training set of 100,000 samples is

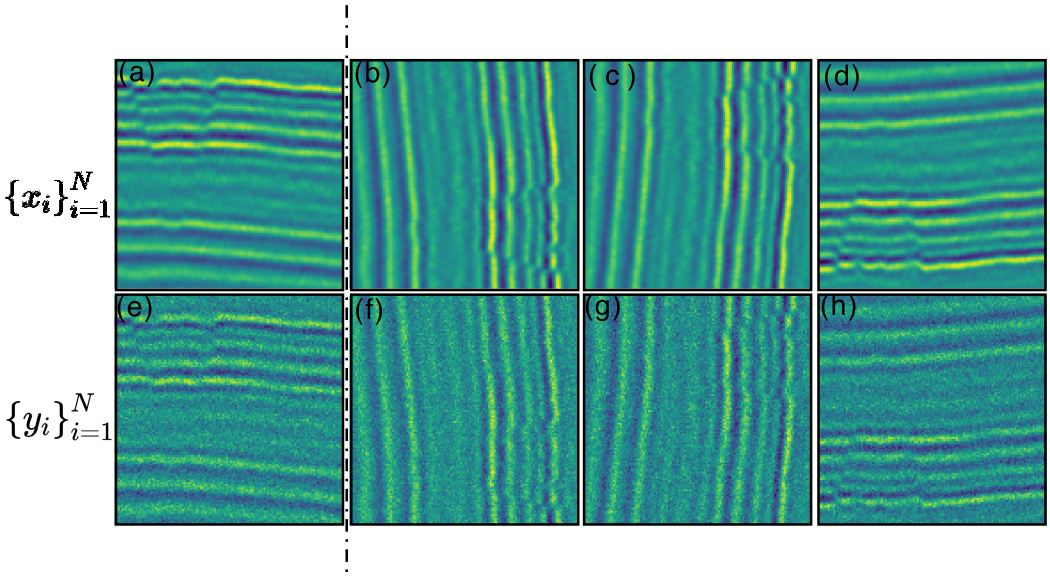

FIGURE 5. The diagram of data augmentation strategy. (a) and (e) are the original synthetic data. (b), (c), (d) and (f), (g), (h) are the augmentation synthetic data.

formed. It can be seen that the data augmentation strategy can effectively enhance the diversity of the training dataset.

2) AUGMENTATION STRATEGY OF REAL DATASET

If the trained denoising neural network does not have good adaptability to real seismic data, it will be useless for actual exploration work. So it is essential to add real seismic data samples to the training dataset to improve its adaptability. However, a big problem here is that there is no clean data in the real seismic data for us to label. Therefore, we applied an existing state-of-the-art denoising algorithm to the real data to get clean data for labeling.

FIGURE 6 shows the denoising results of real seismic data by the f-x adaptive prediction filter(APF) method (Liu *et al.*, 2015 [31]). (a), (b) and (c) are the real seismic data, denoised signal data, and removed noise data, respectively. In this way, we get the data available for labeling.

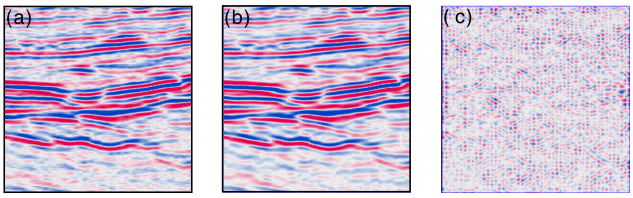

FIGURE 6. The diagram of the real seismic data pre-processing. (a) Real seismic data. (b) Denoised Signal data. (c) Removed noise data.

The augmentation strategy of the synthetic dataset can also be used to increase the diversity of the sample for the clean real seismic data obtained above, such as the rotation transformation, mirror transformation, space-time downsampling, and intensity transformation.

III. NUMERICAL RESULTS

In this section, we will test the denoising performance of the RCGAN on synthetic data and real seismic data. The traditional industrial algorithm f-x deconvolution method [5] (FXDM) and well-known denoising neural network DnCNN are used for comparative experiments. The signal-to-noise ratio (SNR) is used to qualitatively measure the quality of the different algorithms denoised results, which can be expressed as equation 9.

$$SNR = 10\log_{10}\frac{\|x\|_2^2}{\|x-\hat{x}\|_2^2} \tag{9}$$

where x and $\hat{x}$ stand for the clean seismic data and denoised data, respectively, the SNR unit is Decibel(dB).

A. SYNTHETIC DATA TESTING

In this section, we first evaluate the denoising performance of RCGAN in synthetic data and compare it with the results of both DnCNN and FXDM algorithms. Yu *et al.* [23] introduced DnCNN in image denoising into seismic data denoising and achieved good results. To compare our method with DnCNN, we proposed training this model consistent with the RCGAN training dataset with 100,000 augmentation training samples. The training parameters of DnCNN refer to the paper [23] for the best denoising results, which are patch size of 70 × 70, batch size of 128, the convolutional kernel of 70 × 70 and epoch of 50, respectively. For the FXDM, we tested different denoising parameters and finally chose the best denoising results as a comparison. After the network training is completed, we do not need to adjust the network parameters for the testing data-sets to be competent for denoising work. This means that compared to traditional industrial algorithms, AI-based algorithms can

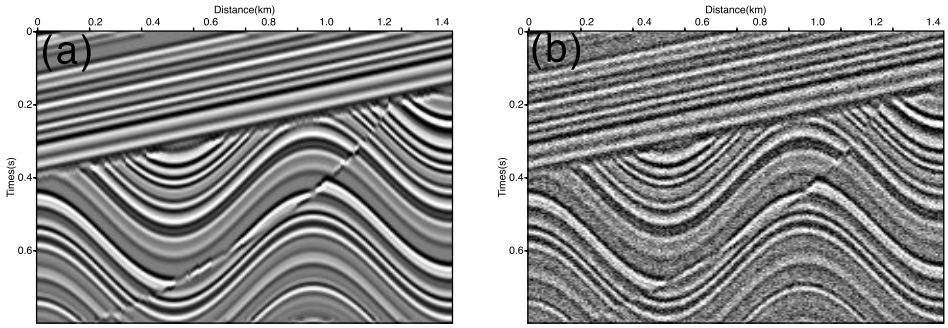

FIGURE 7. Complex synthetic example. (a) Clean seismic data. (b) Contaminated seismic data.

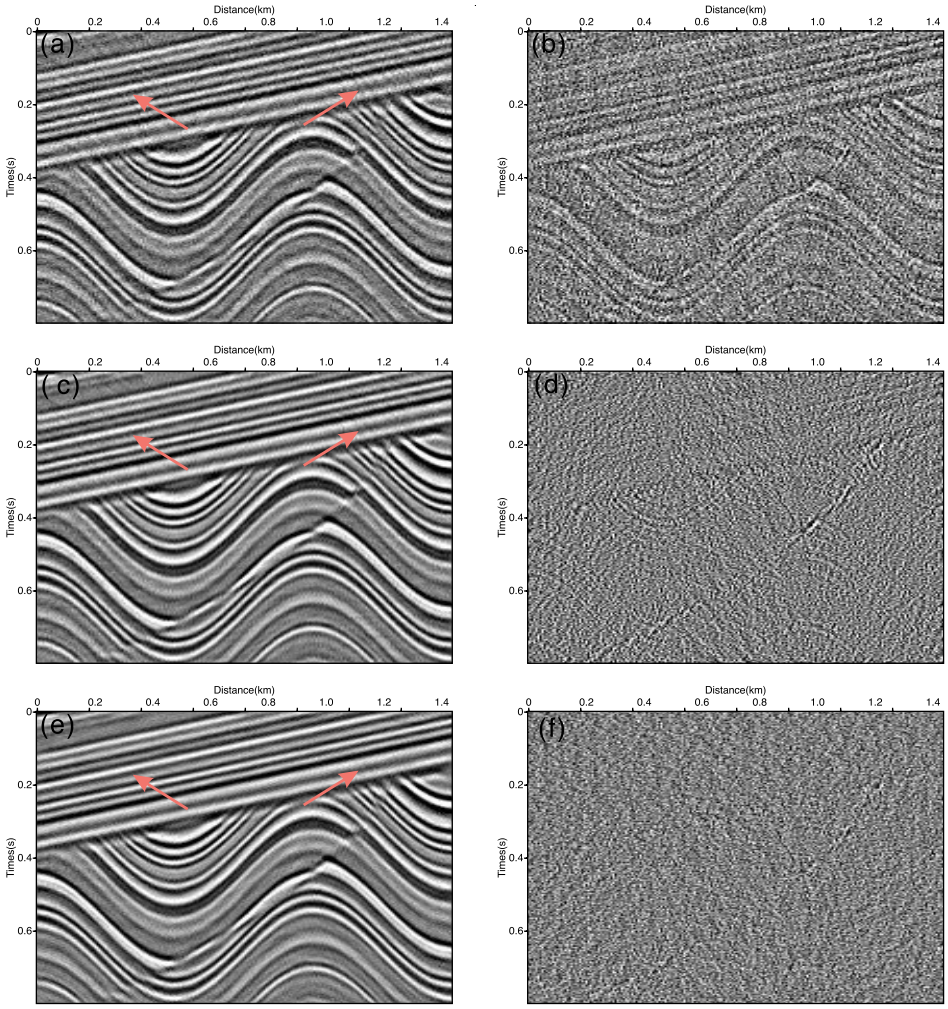

FIGURE 8. The comparison of three methods denoised results and removed noise. (a) and (b) FXDM denoised result and removed noise. (c) and (d) DnCNN denoised result and removed noise. (e) and (f) RCGAN denoised result and removed noise.

eliminate the tedious work of denoising parameter selection and significantly reduce manual intervention in denoising work.

FIGURE 7(a) shows the synthetic data model borrowing from Claerbout (2009) [47]: A synthetic seismic image containing dipping beds, an unconformity, and a fault. We added incoherent noise to the clean raw data and obtained FIGURE 7(b).

FIGURE 8 shows the comparison of the denoising results of RCGAN, DnCNN, and FXDM. FIGURE 8(a), (c), and (e) are the denoised results by FXMD, DnCNN method, and RCGAN method, respectively. We can easily find that the

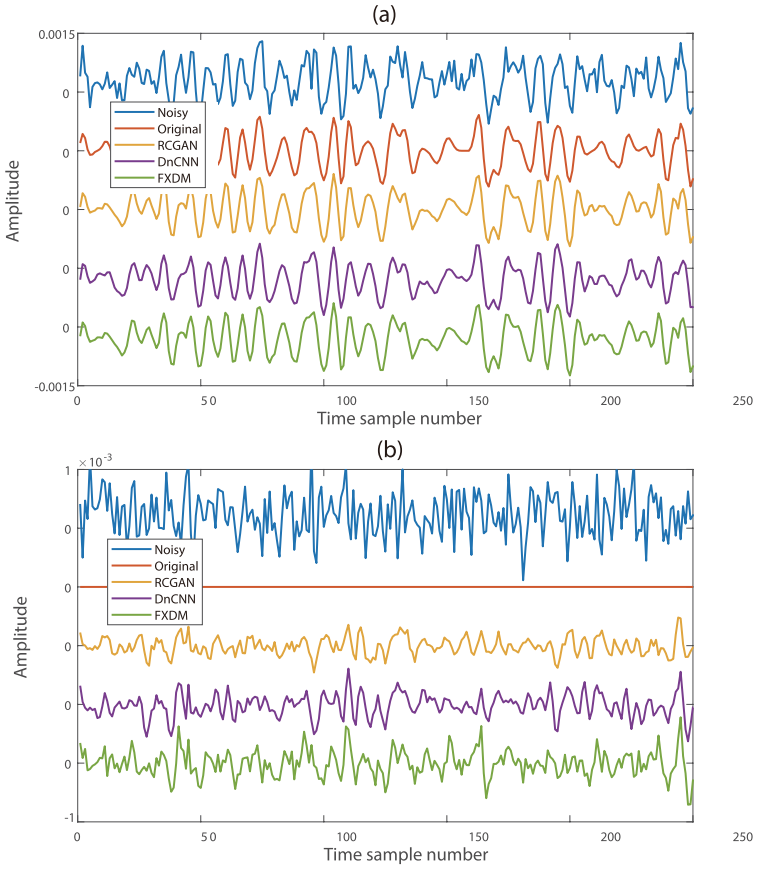

FIGURE 9. The denoised 50th trace (a) and the difference (b) with respect to the original trace.

denoising effect of the f-x deconvolution method commonly used in the industry is not very satisfactory, and there is still much noise remaining. In addition, it even cannot escape too much manual intervention and excellent denoising parameter settings. The DnCNN denoising method implemented by Yu and Ma (2018) [23] can achieve a good intelligent denoising effect without a manual set of denoising parameters, as shown in FIGURE 8 (c). Nevertheless, we can still notice some residual noise. The RCGAN method in this paper can also realize the intelligent denoising of seismic data and further improve the ability to suppress random noise, as shown in FIGURE 8(e). We can find that through RCGAN, the slight noise pointed by the red arrows in FIGURE 7(b) is well removed, which demonstrated that our RCGAN has better denoising performance than previous CNN-based denoising methods.

To better compare the differences in the three methods' denoising performance, we give the three removed noise data images. FIGURE 8 (b) is the noise data removed by the classical FXDM, which contains too many useful seismic data, such as geologic folds, faults, and unconformities. For the DnCNN method, as shown in FIGURE 8 (d), the noised data removed contains a small amount of useful seismic data such as faults information, folds, and unconformities information. FIGURE 8 (f) is the noise data that hardly contains useful geological data. This means that our method has a strong ability to retain details.

For better quantitative observation of the denoised results, we extracted the traces of several methods to compare the denoised ability further. FIGURE 9(a) shows the 50th traces from the denoised results and (b) shows the corresponding difference concerning the original trace. From top to bottom are noisy data, original data, RCGAN, DnCNN, and FXDM. We can see that the trace of RCGAN is not only the closest to the original data, with the smallest differences from the original data.

Then, we increased the random noise level to more comprehensively test the RCGAN's denoising ability. FIGURE 10 shows the denoising results of FXDM (b), DnCNN (c), and RCGAN (d). We can see that our method still obtained a very good denoising result. However, the denoising ability of the DnCNN method decreases significantly as the noise level increases.

To qualitatively measure the quality of the different algorithms denoised results, we compared the SNR variations with regard to the noise level in FIGURE 11. The green, Sky blue, and magenta lines represent the SNR values of denoised results via RCGAN proposed in this paper, DnCNN, and

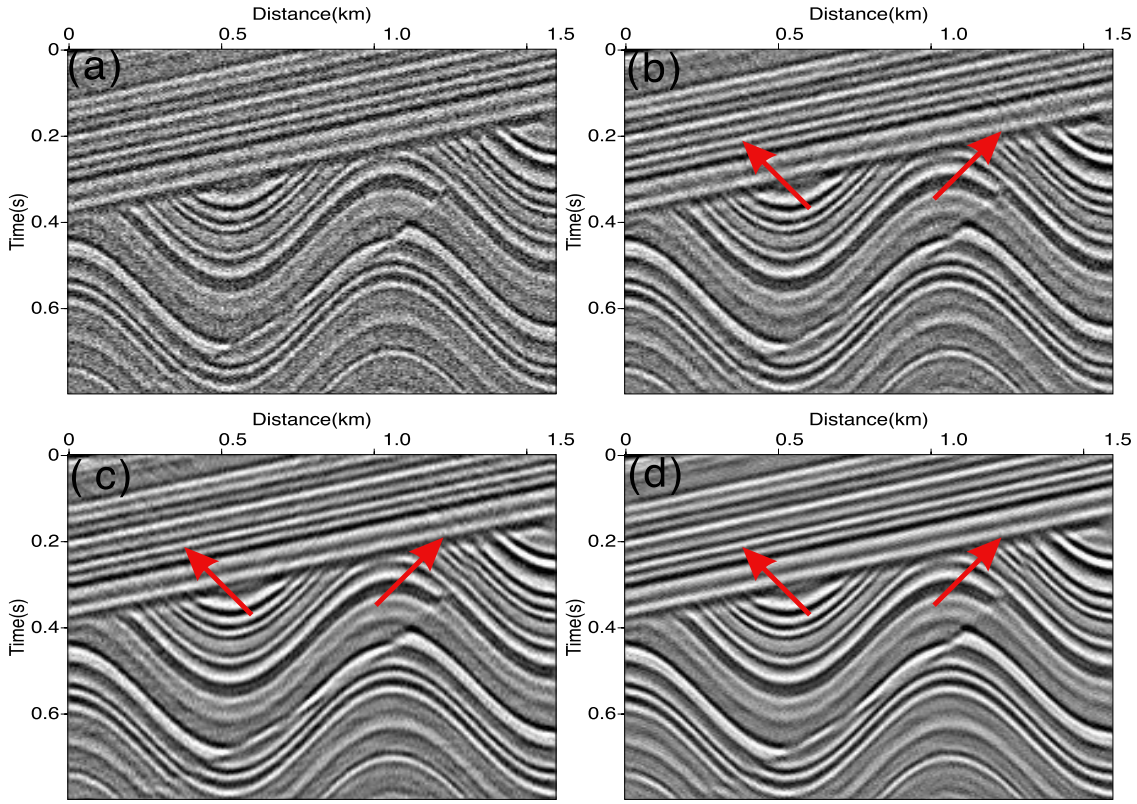

FIGURE 10. The comparison of three methods denoised results. (a) Contaminated seismic data. (b) FXDM denoised result. (c) DnCNN denoised result. (d) RCGAN denoised result.

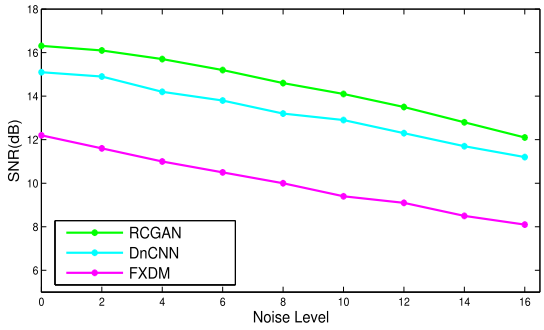

FIGURE 11. The SNR comparison of three methods with regard to the noise level.

FXDM, respectively. For each method, the SNR will decrease as the noise level increases. We can significantly find that the SNR of our method is remarkably greater than the other two methods for the same noise level, which proves the remarkable denoising performance of our method RCGAN.

B. REAL DATA TESTING

In this section, We applied the RCGAN of this paper to test its applicability in real seismic data denoising work. We have found that our method's denoising results are much better than FXDM in testing synthetic data, so we will no longer compare it with the traditional denoising algorithm(FXDM) in real data testing. We applied the APF method for denoising when constructing the real data training set. So, when testing the trained neural network's denoising ability, we will compare it with the APF method.

FIGURE 12 (a) shows the real seismic data. FIGURE 12 (b), (c), and (d) display the denoised results of DnCNN, APF method, and RCGAN, respectively. We can find that the DnCNN method has a good random noise suppression ability, but there are still several random noise remaining in the denoised result. The APF method and RCGAN will be significantly better at suppressing random noise than DnCNN. The APF method contains very slight noise in the denoised result, while random noise is almost invisible in the RCGAN denoised result. As indicated by the red arrows in the FIGURE 12, we can see that the APF method and the RCGAN method can preserve the continuity of the axis significantly better.

FIGURE 13 (a), (b), (c), and (d) are the results of a partial enlargement of the red box section in FIGURE 12. We can find more clearly that the latter two methods will have much better reflection signal continuity than the DnCNN method, which indicates that the signals are not seriously damaged during the denoising process. Since the denoised results are zoomed-in partially, the differences between the three methods can be observed more clearly. In FIGURE 13 (c), we can still find very subtle random noise. The RCGAN method

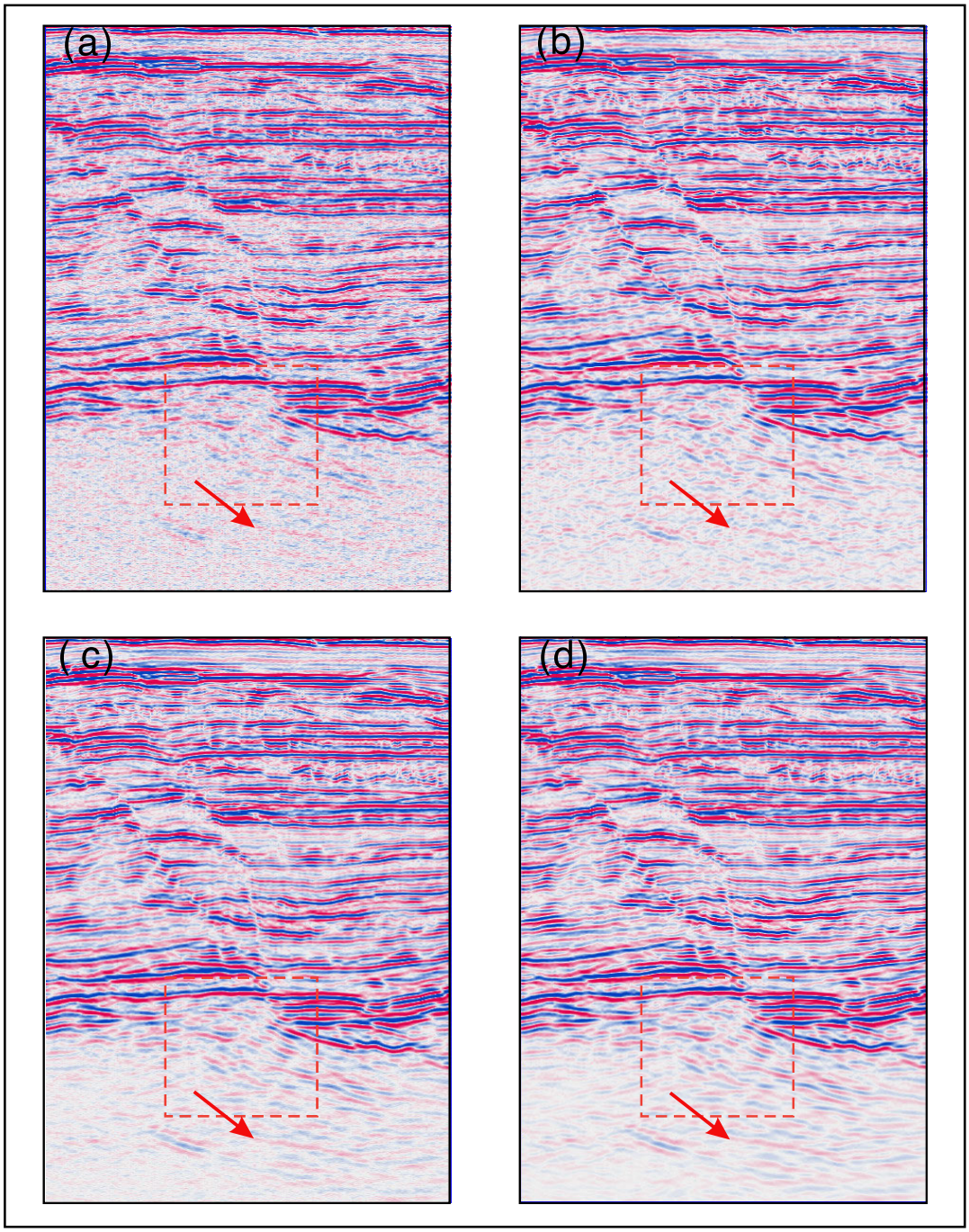

FIGURE 12. The comparison of three methods denoised results in real seismic data. (a) The raw real seismic data. (b) Denoised result using DnCNN. (c) Denoised result using APF. (d) Denoised result using RCGAN.

is perfect for removing random noise and maintaining the continuity of the reflection signal.

We use the APF method for denoising the real seismic data to construct the training set, but why does our RCGAN method have better denoising performance than the APF method? We consider that the main reason is that this paper adopted a real data augmentation strategy, which greatly expanded the training set of real data and thus improved the denoising capability of the network.

FIGURE 13 (a), (b), and (c) shows the removed noise of the above three methods DnCNN, APF method, and RCGAN, respectively. We can see that the continuous reflected signal is apparent in FIGURE 13 (a), while the FIGURE 13 (b) and (c) have very slight damage to the effective signal. Therefore, compared with the DnCNN and

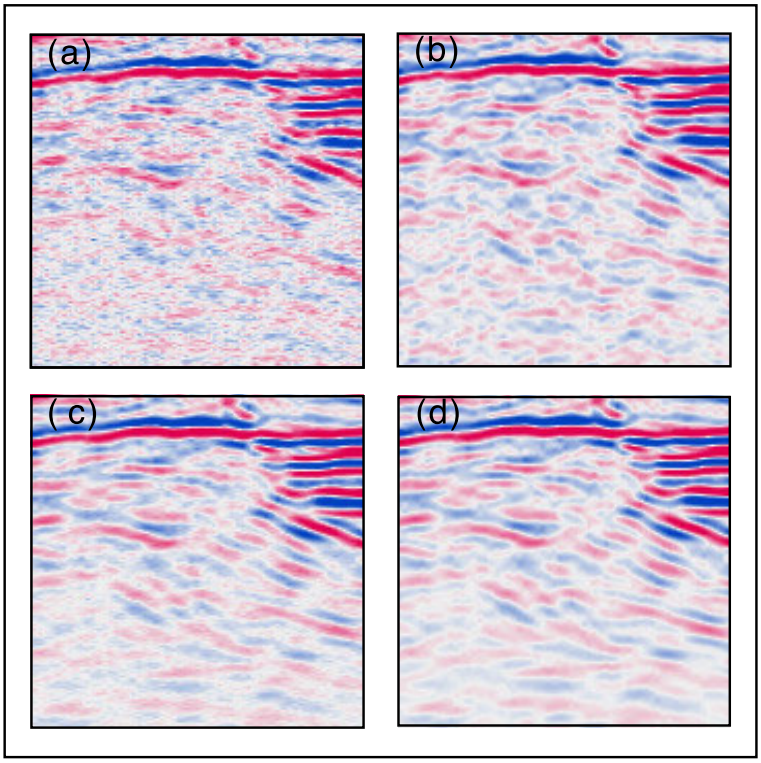

FIGURE 13. Zoomed denoised results test on the real seismic data. (a) The raw real seismic data. (b) Denoised result using DnCNN. (c) Denoised result using APF. (d)Denoised result using RCGAN.

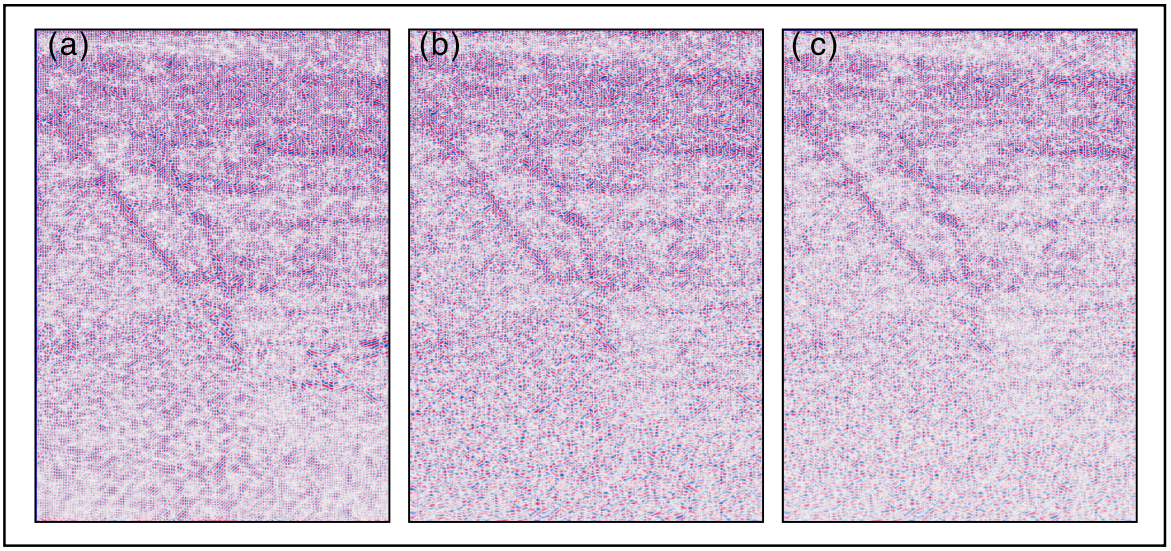

FIGURE 14. The comparison of three methods removed noise in real seismic data. (a) Removed noise using DnCNN. (b) Removed noise using APF. (c) Removed noise using RCGAN.

APF method, the proposed method can effectively attenuate the noise and fully protect the effective signals.

Compared to the APF method, our RCGAN can significantly reduce manual intervention and accuracy assumptions while maintaining noise suppression capability. Moreover, compared to DnCNN, our proposed method can improve denoising performance while retaining the advantages of neural networks. It means that our approach is very advantageous in the work of suppressing random noise.

IV. CONCLUSION

Seismic data denoising has always been an indispensable step in the seismic exploration workflow. The quality of the results directly affects the results of subsequent inversion

and migration imaging. The traditional denoising methods require a lot of complicated denoising parameters selection work. An experienced data processor must make appropriate parameter selection according to the work area's geological conditions. A large number of manual interventions are not only time-consuming and labor-intensive but also add much uncertainty to the denoising results.

This paper proposed an RCGAN method based on a data augmentation strategy for seismic data noise suppression work. Firstly, we introduce the cycle-GAN into the seismic denoising domain and greatly improved the network's training efficiency by employing residual learning. Besides, this paper applied an effective data generation and augmentation method so that the trained network can be better adapted to real seismic data denoising work.

Compared to FXDM and DnCNN, our method has a better ability to remove random noise and retain data details in synthetic data testing. Our RCGAN method will be better adapted to real seismic data denoising work due to introducing a data augmentation strategy. Besides, the lack of real data training samples faced in network training can also be effectively overcome by the real data augmentation strategy. Compared to DnCNN, our proposed method can improve denoising performance while maintaining the advantages of neural networks. Furthermore, compared to the APF method, our RCGAN can significantly reduce manual intervention and accuracy assumptions while maintaining noise suppression capability. Experiments with synthetic and real seismic data confirm that the RCGAN method performs superior to other state-of-the-art denoising algorithms.

ACKNOWLEDGMENT

The author is very grateful to Siwei-Yu and Kai-Zhang for their papers, inspiring the author.

REFERENCES

[1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.

[2] A. Shrivastava, T. Pfister, O. Tuzel, J. Susskind, W. Wang, and R. Webb, "Learning from simulated and unsupervised images through adversarial training," 2016, *arXiv:1612.07828*. [Online]. Available: http://arxiv.org/abs/1612.07828

[3] B. Russell, D. Hampson, and J. Chun, "Noise elimination and the Radon transform: Part 1," *Geophysics*, vol. 9, no. 10, pp. 18–23, 1990.

[4] C. P. Ross and D. M. Cole, "A comparison of popular neural network facies-classification schemes," *Lead. Edge*, vol. 36, no. 4, pp. 340–349, Apr. 2017.

[5] L. L. Canales, "Random noise reduction," in *Proc. SEG Tech. Program Expanded Abstr.*, 1984, p. 329.

[6] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006.

[7] N. Gulunay, "FXDECON and complex Wiener prediction filter," in *Proc. SEG Tech. Program Expanded Abstr.*, Jan. 1986, pp. 279–281.

[8] G. Hennenfent and F. J. Herrmann, "Seismic denoising with nonuniformly sampled curvelets," *Comput. Sci. Eng.*, vol. 8, no. 3, pp. 16–25, May 2006.

[9] G. Tang, J.-W. Ma, and H.-Z. Yang, "Seismic data denoising based on learning-type overcomplete dictionaries," *Appl. Geophys.*, vol. 9, no. 1, pp. 27–32, Mar. 2012.

[10] J. Y. Cheong and I. K. Park, "Deep CNN-based super-resolution using external and internal examples," *IEEE Signal Process. Lett.*, vol. 24, no. 8, pp. 1252–1256, Aug. 2017.

[11] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2223–2232.

[12] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proc. ECCV*. Springer, 2016, pp. 694–711.

[13] S. H. A. O. Jie, S. U. N. Chengyu, T. A. N. G. Jie, and W. Dunshi, "Microseismic data denoising based on sparse representations over learned dictionary in the wavelet domain," *Oil Geophys. Prospecting*, vol. 51, no. 2, pp. 254–260, 2016.

[14] 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.

[15] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a fast and flexible solution for CNN-based image denoising," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4608–4622, Sep. 2018.

[16] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[17] L. Zhu, E. Liu, and J. H. McClellan, "Seismic data denoising through multiscale and sparsity-promoting dictionary learning," *Geophysics*, vol. 80, no. 6, pp. WD45–WD57, Nov. 2015.

[18] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, Nov. 1992.

[19] 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.

[20] M. Tassano, J. Delon, and T. Veit, "An analysis and implementation of the FFDNet image denoising method," *Image Process. Line*, vol. 9, pp. 1–25, Jan. 2019.

[21] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *J. Mach. Learn. Res.*, vol. 11, no. 12, pp. 3371–3408, Dec. 2010.

[22] R. Zhang and T. J. Ulrych, "Physical wavelet frame Denoising," *Geophysics*, vol. 68, no. 1, p. 225, 2003.

[23] S. Yu, J. Ma, and W. Wang, "Deep learning for denoising," *Geophysics*, vol. 84, no. 6, pp. V333–V350, 2019.

[24] S. Yu, J. Ma, and S. Osher, "Monte Carlo data-driven tight frame for seismic data recovery," *Geophysics*, vol. 81, no. 4, pp. V327–V340, Jul. 2016.

[25] S. Fomel and Y. Liu, "Seislet transform and seislet frame," *Geophysics*, vol. 75, no. 3, pp. V25–V38, May 2010.

[26] S. L. M. Freire and T. J. Ulrych, "Application of singular value decomposition to vertical seismic profiling," *Geophysics*, vol. 53, no. 6, pp. 778–785, Jun. 1988.

[27] S. Hornbostel, "Spatial prediction filtering in the t-x and f-x domains," *Geophysics*, vol. 56, no. 12, p. 2019, 2012.

[28] W. Dong, G. Shi, and X. Li, "Nonlocal image restoration with bilateral variance estimation: A low-rank approach," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 700–711, Feb. 2013.

[29] X. Wu, Y. Shi, S. Fomel, and L. Liang, *Convolutional Neural Networks for Fault Interpretation in Seismic Images*. New York, NY, USA: Academic, 2018.

[30] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, and S. Paul Smolley, "Least squares generative adversarial networks," 2016, *arXiv:1611.04076*. [Online]. Available: http://arxiv.org/abs/1611.04076

[31] Y. Liu, N. Liu, and C. Liu, "Adaptive prediction filtering in t-x-y domain for random noise attenuation using regularized nonstationary autoregression," *Geophysics*, vol. 80, no. 1, pp. V13–V21, Jan. 2015.

[32] Y. Bengio, "Practical recommendations for gradient-based training of deep architectures," *Neural Netw., Tricks Trade*, vol. 7700, no. 3, pp. 437–478, 2015.

[33] Y. Shi, X. Wu, and S. Fomel, "Automatic salt-body classification using deep-convolutional neural network," in *Proc. SEG Tech. Program Expanded Abstr.*, Aug. 2018, pp. 1971–1975.

[34] Y. Jia and J. Ma, "What can machine learning do for seismic data processing? An interpolation application," *Geophysics*, vol. 82, no. 3, pp. V163–V177, May 2017.

[35] L. Zhu, E. Liu, and J. H. McClellan, "Joint seismic data denoising and interpolation with double-sparsity dictionary learning," *J. Geophys. Eng.*, vol. 14, no. 4, pp. 802–810, Aug. 2017.

[36] S. Hornbostel, "Spatial prediction fifiltering in the t-x and f-x domains," *Geophysics*, vol. 56, no. 12, pp. 2019–2026, 1991.

[37] D. Bonar and M. Sacchi, "Denoising seismic data using the nonlocal means algorithm," *Geophysics*, vol. 77, no. 1, p. 5, 2012.

[38] V. Das *et al.*, "Convolutional neural network for seismic impedance inversion," *Geophysics*, pp. 1–66, 2019.

[39] M. Zhang, Y. Liu, M. Bai, and Y. Chen, "Seismic noise attenuation using unsupervised sparse feature learning," *IEEE Trans. Geosci. Remote Sens.*, to be published.

[40] W. Xiong, X. Ji, Y. Ma, Y. Wang, and Y. Luo, "Seismic fault detection with convolutional neural network," *Geophysics*, vol. 83, no. 5, pp. 1–28, 2018.

[41] W. Benfeng *et al.*, "Deep-learning-based seismic data interpolation: A preliminary resultDeep learning for interpolation," *Geophysic*s, 2018.

[42] H. Zhang, X. Yang, and J. Ma, "Can learning from natural image denoising be used for seismic data interpolation?" *Geophysics*, vol. 85, no. 4, pp. 1–142, 2020.

[43] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *Comput. Sci.*, 2014. [Online]. Available: http://arxiv.org/abs/1409.1556v6

[44] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2015.

[45] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, 2015.

[46] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016.

[47] J. F. Claerbout. (2009). *Basic Earth Imaging: Stanford Exploration Project*. Accessed: Oct. 2014. [Online]. Available: http://sepwww.stanford.edu/sep/prof/

WENDA LI received the bachelor's and master's degrees in geophysics from the China University of Geoscience, Beijing, China, in 2017 and 2019, respectively. He is currently pursuing the Ph.D. degree with the Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing.

He has published three articles in international journals. He has participated in many academic conferences and made academic reports.

JIAN WANG (Member, IEEE) received the bachelor's degree in geophysics from the China University of Geoscience, Beijing, China, in 2014, and the Ph.D. degree, in 2018.

He did postdoctoral research work with the Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing. He is currently an Assistant Researcher with the Institute of Acoustics, Chinese Academy of Sciences.