SEISMIC DATA ENHANCEMENT BASED ON BAYESIAN CONVOLUTIONAL NEURAL NETWORK

ZIXUAN QIAO[1,2], XIAOYU CHUAI[3*], ZHENWANG XU[4], NAICHUAN GUO[5], WEI ZHU[6], JINFENG ZHANG[7], WEI CHEN[1,2] and RUI XIA[3]

[1] *School of Geophysics and Petroleum Resources, Yangtze University, Hubei, P.R. China. chenwei2014@yangtzeu.edu.cn*
[2] *Open Research Fund of Key Laboratory of Engineering Geophysical Prospecting and Detection of Chinese Geophysical Society, Beijing, P.R. China.*
[3] *Command Center of Natural Resources Comprehensive Survey, China Geological Survey, Beijing 100055, P.R. China. chuaixiaoyu@126.com*
[4] *Research lnstitute of Petroleum Exploration and Development, Liaohe Oilfield Company, PetroChina, Panjin 124010, P.R. China.*
[5] *Bohai Oilfield Research Institute of CNOOC Ltd., Tianjin Branch, Tianjin, P.R. China.*
[6] *Changqing Industrial Group Co., Ltd, Petro China Changqing Oilfield Branch, Xi'an, P.R. China.*
[7] *No.5 Oil Production Plant, Petro China Changqing Oilfield Branch, Xi'an, P.R. China.*

(Receipt July 23, 2023; accepted September 1, 2023)

ABSTRACT

Qiao, Z.X., Chuai, X.Y., Xu, Z.W., Guo, N.C., Zhu, W., Zhang, J.F., Chen, W. and Xia, R., 2023. Seismic data enhancement based on Bayesian convolutional neural network. *Journal of Seismic Exploration*, 32: 407-425.

The acquisition of high-quality seismic data is an important goal of seismic data processing. Traditional seismic data processing methods are usually used alone to remove noise or improve resolution. They can only improve the quality of seismic data from a certain point of view, and lack the protection of effective detail signals. In order to improve the quality of seismic data from whole angle, protect and highlight the details of geological structures such as faults and fault uplifts, this paper proposes to apply Bayesian Convolutional Neural Network (BCNN) to seismic data processing to enhance seismic data. BCNN is an organic combination of Bayesian theory and neural network, which can avoid network over-fitting and enable the network to learn deeper data features adaptively, with better robustness. In addition, the up-sampling operation at the end of the network model is conducive to preserving the feature information of seismic data in the low-resolution space. In this paper, the seismic data is enhanced based on F3 dataset, and compared with the general full convolutional neural network (FCNN) and construction-oriented filtering methods. The results show that the proposed method can better highlight the structural details, improve the interpretability of seismic data, and is an effective means to enhance fault and uplift structures.

KEY WORDS: seismic data, seismic data enhancement, resolution, Bayes, artificial intelligence.

0963-0651/23/$5.00

INTRODUCTION

Seismic data processing is an important step in structural interpretation and reservoir prediction. Due to the low quality of the initial seismic data, the efficiency and accuracy (Chang et al., 2018; Li et al., 2023; Zhao et al., 2022) of the work will be reduced if the data is directly used in the subsequent interpretation and inversion. For this reason, many methods have been proposed to obtain high quality seismic data, including singular value decomposition algorithm (Claerbout, 1971), median filter (Bednar, 1983), f-x deconvolution (Canales, 1984; Liu and Duan, 2019), wavelet transform (Mallat, 1991; Chen and Song, 2018; Chen and Jin, 2015); Singh and Mittal, 2014) and curve-wave transform (Hennenfent and Herrmann, 2006), empirical mode decomposition (Chen et al., 2017), dictionary learning (Lan et al., 2023). In 1971, Claerbout was the first to apply singular value analysis to the removal of background noise from seismic data, which opened a new direction for the development of seismic data processing and imaging methods. In 1983, Bednar applied the median filtering technique to seismic data enhancement, which improved the quality and reliability of seismic data. f-x deconvolution was proposed by Canales in 1984, and modified by Gulunay (1986), it has become a more common seismic data processing method. Mallat and Hennenfent combined wavelet transform and curvewave transform with seismic data processing in 1991 and 2006, respectively. The above methods usually have certain requirements for the prior knowledge and empirical rules of underground media, and are more sensitive to the initial model and parameter setting of the data. This limits its application in complex geological situations and also increases the subjectivity of processing. To this end, researchers have introduced image enhancement techniques into seismic data processing. These related techniques in the field of image processing provide a more advanced and automated method of seismic data enhancement by introducing new algorithms. And adapt to the processing needs of different geological conditions. Fuzzy set theory, histogram equalization and filtering technology are common image enhancement methods, but due to the limitations of the former two, which cause image distortion and high calculation cost, the filtering technology is more widely used (Singh and Mittal, 2014; Song-Tao and Gang, 2010). In 2010, Yang Peijie applied directional filtering technology and holding filtering technology to seismic data to realize the enhancement of fault images (Yang et al., 2010). Chang-Qing (2012) improved the threshold function used in the traditional wavelet threshold algorithm by combining the characteristics of seismic data. Compared with the improved enhancement method, the seismic image contrast was higher and the edge was clearer and richer. In order to further improve the enhancement effect of seismic data, Yan et al. (2013) proposed a seismic image enhancement processing method based on anisotropic diffusion filtering, taking improving the signal-to-noise ratio of seismic data and protecting the fault structure information as the starting point. In 2018, by combining deconvolution and wavelet scaling, Alaei et al. (2018) designed a scale transform filter by dividing the obtained wavelet magnitude by the original value along the frequency axis of the deconvolution magnitude spectrum, and the resolution of the seismic trace was effectively improved after processing by the filter.

With the development of artificial intelligence, data-driven seismic data processing methods emerge in an endless stream (Deng et al., 2017; Gómez, et al., 2020) and show better results than traditional methods (Kuang et al., 2021). In 2018, Song et al. proposed a deep convolutional self-coding neural network for removing random noise from seismic data, and compared it with classical wavelet and f-x deconvolution algorithms, the results show that this method has stronger denoising ability. In 2019, Wang and Nealon applied the 3D convolutional neural network to the enhancement and noise attenuation of seismic images, and constructed a training set of different noise levels using the difference in lens density. The network trained based on the training set highlighted the geological structure and was easy to interpret. In 2019, Wang et al. built a CNN denoising framework based on data generation and augmentation to solve the problem of insufficient seismic label data, and realized the denoising of CNN seismic data based on small samples. In 2020, Dong and Li proposed a convolutional countermeasure noise reduction network to solve the problem of low signal-to-noise ratio in distributed optical fiber acoustic wave sensing and detection technology. The network improved on the basis of the generated countermeasure network, replaced the original generator in the countermeasure network with a noise canceller, and optimized the noise canceller. Effectively suppress the new noise caused by the bad coupling and recover most of the effective signal. In 2023, in order to make up for the limitations of the technology of improving the resolution of seismic data based on sparse peak inversion, Gao et al. (2023) combined the longitudinal reflectance information of field data and the geological structure characteristics, and trained a U-shaped network integrating residual blocks and attention mechanism to achieve a more consistent thin layer resolution effect. The application of artificial intelligence has greatly improved the efficiency of seismic data processing and interpretation under the premise of ensuring the accuracy (Wang and Nealon, 2019; Zhang et al., 2022; Li et al., 2021; Oliveira et al., 2019; Qiu et al., 2021; Jin et al.. 2018). However, learning algorithms based on neural networks generally have the problem of overfitting, which causes the model to over-rely on the training data, thus reducing its generalization ability. The reason for this is that general neural networks perform a given task by learning examples and estimating the optimal value of each node weight without prior knowledge. However, networks that use point estimates as weights perform well on large datasets, while fail to express uncertainty in areas with little or no data, leading to overconfident or inferior decisions.

To solve this problem, this paper uses a type of BCNN based on variational inference. The combination of Bayesian concept and neural network can treat the weights in the neural network as random variables, and provide a clear prior probability for training by learning the distribution of each parameter. The addition of a prior is equivalent to providing a constraint and regularization for the network, thus preventing overfitting. Considering the robustness of Bayesian neural network to overfitting, this paper creates a dataset of small samples to train BCNN, and gives the parameter description of the network structure and the training results. In addition, the proposed method is compared with structure-oriented filtering

and general full-convolutional neural network (FCNN), and the results show that BCNN has more advantages in seismic data enhancement.

THEORY

Convolutional Neural Networks

Convolutional neural networks are feedforward neural networks with local connections, weight sharing and other characteristics. A typical convolutional neural network is generally composed (Gal and Ghahramani, 2015) of a convolutional layer, a convergence layer and a fully connected layer. The function of the convolution layer is to extract local features, and the function of the convergence layer is to reduce the size of the output feature map of the convolution layer, so as to reduce the number of parameters and calculation amount of the model, and improve the robustness and generalization ability of the model. The nodes of the fully connected layer integrate the features extracted from the previous layer and map them to the sample label space by connecting with all the nodes of the previous layer. The core operation is the convolution operation, which extracts the high-level features in the input data by sliding the convolution kernel on the input data to realize the abstraction and classification of the data. After the convolution operation, the bias term is usually added, and the nonlinear transformation is carried out by the activation function to further improve the expressivity of the model. The convolution operation can be expressed as the dot product operation (Gal and Ghahramani, 2015) of the input matrix and the convolution kernel:

$$Y_{i,j} = \sum_{k=1}^{K} \sum_{l=1}^{L} W_{k,l} g X_{i+k-1,j+l-1} \tag{1}$$

where Y is the output matrix and i and j represent the row index and column index of the output matrix, respectively. k and l are the row index and column index of the convolution kernel W, where W represents the weight of the convolution kernel at the corresponding position. X is the input matrix, and i+k-1 and j+l-1 represent the position of the center of the current convolution kernel W on the input matrix X. Also, in order for the convolution kernel to slide over the input matrix, the height and width of the convolution kernel will usually be smaller than the size of the input matrix.

Bayes' Fundamental Theorem

$$P(H \mid D) = \frac{P(D \mid H)P(H)}{P(D)} = \frac{P(D \mid H)P(H)}{\int_H P\left(D \mid H^{'}\right)P\left(H^{'}\right)dH^{'}} = \frac{P(D \mid H)}{\int_H P\left(D \mid H^{'}\right)dH^{'}} \tag{2}$$

The core idea of Bayes' fundamental theorem is to calculate the unknown probability according to the known probability, that is, to introduce new information to dynamically update the prior probability on the basis of the original prior probability, so as to approach the actual probability of the event step by step. Bayes theory emphasizes that the data distribution affecting the result is not only related to the collected data, but also related to the prior probability, which affects the estimated result distribution by affecting the posterior probability. Interpreting a neural network model with Bayes' fundamental theorem can be expressed as (Shridhar et al., 2019):

$$p(\theta \mid D) = \frac{p\left(D_y \mid D_x, \theta\right) p(\theta)}{\int_\theta p\left(D_y \mid D_x, \theta'\right) p\left(\theta'\right) d\theta'} \propto p\left(D_y \mid D_x, \theta\right) p(\theta) \tag{3}$$

where θ represents the training parameters of the neural network, D represents the training data, D_x represents the feature data, D_y represents the label data. A general neural network can be seen as a special case of a Bayesian neural network, where each weight is represented as a point estimated using maximum likelihood when the prior probability is constant.

Bayesian backpropagation algorithm

The basic idea of Bayesian backpropagation algorithms is to approximate the true posterior probability by variational inference (Cornfield, 1967). Since the posterior distribution of weights $p(w \mid D)$ is difficult to calculate, a parameterized distribution, the variational distribution $q_\theta(w \mid D)$ is used to approximate $p(w \mid D)$. In order to make the approximate distribution as close as possible to the real posterior distribution, the concept of KL divergence is introduced to measure the distance between the two distributions, and the KL divergence (Shridhar et al., 2019) is minimized by selecting the optimal parameter θ:

$$\theta^{opt} = \arg\min_\theta KL\left[q_\theta(w \mid D)\, Pp(w \mid D)\right] \tag{4}$$

Formula (4) combined with Bayes' theorem can be expanded as:

$$\begin{aligned} & KL\left[q_\theta(w \mid D)\, Pp(w \mid D)\right] \\ & = \int q_\theta(w \mid D) \log \frac{q_\theta(w \mid D) p(D)}{p(w \mid D)} d\theta \\ & = \log p(D) - \int q_\theta(w \mid D) \log \frac{q_\theta(w \mid D)}{p(w \mid D)} d\theta \end{aligned} \tag{5}$$

Since KL is the distance between the two distributions, $KL \geq 0$, can be

obtained from formula (5) $\log p(D) \geq \int q_\theta(w \mid D) \log \frac{q_\theta(w \mid D)}{p(w \mid D)} d\theta$, where $\log p(D)$ is the data likelihood recorded as evidence, then $\int q_\theta(w \mid D) \log \frac{q_\theta(w \mid D)}{p(w \mid D)} d\theta$ is the evidence lower bound ELBO. the assumption remains $\log p(D)$ unchanged, minimizing KL divergence is converted to maximizing ELBO:

$$\begin{aligned} \text{ELBO} &= \int q_\theta(w \mid D) \log \frac{p(w \mid D)}{q_\theta(w \mid D)} d\theta \\ &= \int q_\theta(w \mid D) \log p(D \mid w) d\theta + \int q_\theta(w \mid D) \log \frac{p(w)}{q_\theta(w \mid D)} d\theta \\ &= E_{q_\theta(w|D)} \log p(D \mid w) - KL\left[q_\theta(w \mid D) P_\theta(w)\right] \end{aligned} \tag{6}$$

In eq. (6), $E_{q_\theta(w|D)}$ is the expectation of the parameter θ and $E_{q_\theta(w|D)} \log p(D \mid w)$ is the maximum likelihood of the data. Maximizing ELBO means finding the difference between maximizing the maximum likelihood of the data and minimizing the KL divergence. Since KL is difficult to calculate precisely, it is estimated using Monte Carlo MC sampling, MC sampling from Gaussian distribution $q_\theta\left(w^{(i)} \mid D\right)$ which is based a parameter θ:

$$F(D,\theta) = \sum_{i=1}^{n} -\log p\left(D \mid w^{(i)}\right) + \log q_\theta\left(w^{(i)} \mid D\right) - \log p\left(w^{(i)}\right) \tag{7}$$

In formula (7), n is the sampling times and $w^{(i)}$ is the weights extracted at the i-th data point. The sampled weights are used for the forward propagation of the network to obtain the likelihood loss at the output layer, and the distance $w^{(i)}$ from the prior is calculated at the same time. Finally, the gradient is estimated by the backpropagation algorithm based on reparameterization.

Based on the Bayesian backpropagation algorithm, BCNN considers the feature graph as a random variable in the convolutional layer and operates with two convolutional check data to obtain two output feature graphs, which serve as the mean and variance of the Gaussian distribution respectively, and then samples the two feature graphs to obtain the activation value b_j of this layer as the input of the next layer, as shown in Formula (8) (Shridhar et al.,, 2019):

$$b_j = A_i * \mu_i + \acute{U}_j \, e \, \sqrt{A_i^2 * \left(\alpha_i \, e \, \mu_i^2\right)} \tag{8}$$

where A_i is the receptive field, $*$ indicates the coiling operation, $A_i * \mu_i$ is the mean obtained from the first convolution operation, is the standard deviation obtained from the second $\sqrt{A_i^2 * (\alpha_i \ e \ \mu_i^2)}$ convolution operation. After the mean and standard deviation are obtained, the standard normal sample $\acute{U}_j$ is multiplied by the standard deviation, and the mean is added to obtain the sampled activation value b_j.

Bayesian convolutional neural network

The BCNN method proposed in this paper fully considers the features of the original seismic data collected in low resolution space, abandons the previous convolution operation in high resolution space, and focuses more on extracting seismic data features in low resolution space. The BCNN structure constructed in this paper is shown in Fig. 1. The first part of the topology is the combination of convolution layer and activation layer. In the first layer, the convolution kernel size is 5×5, the step size is 1, and the zero filling is 2. 64 feature maps are obtained after operation. To extract the nonlinear features of the model, the second layer of convolution kernel size is 3×3, step size is 1, zero fill is 1, input channel is 64, output feature channel is 32, and Tanh activation function is connected. The third layer convolution kernel has the size of 3×3, the step size is 1, the zero fill is 1, the number of input channels is 32 and the number of output channels is 3, connecting the Tanh activation layer. The second part is the form of the upper sampling layer combined with the activation function. The activation function is Sigmoid (as shown in Fig. 2), in which the upper sampling layer uses subpixel convolution, the up-sampling rate is 3, and the low-resolution spatial feature mapping is aggregated, which is an efficient, fast and parameter-free up-sampling method of pixel rearrangement. The way of pixel rearrangement can be expressed as (Shi et al., 2016):

$$PS(T)_{x,y,c} = T_{\lfloor x/r \rfloor, \lfloor y/r \rfloor, c \times r \times \mathrm{mod}(y,r) + c \times \mathrm{mod}(x,r)} \tag{9}$$

where r represents the multiple of the up-sampled graph and c is the final number of channels. In the case of multiple channels, the feature map rearranges the contiguous c channels as a whole to obtain the multichannel up-sampled graph. The subpixel convolution operation at the end of the network model fully preserves more texture features of the seismic image in the low-resolution space, which is conducive to the further establishment of high-resolution seismic images.

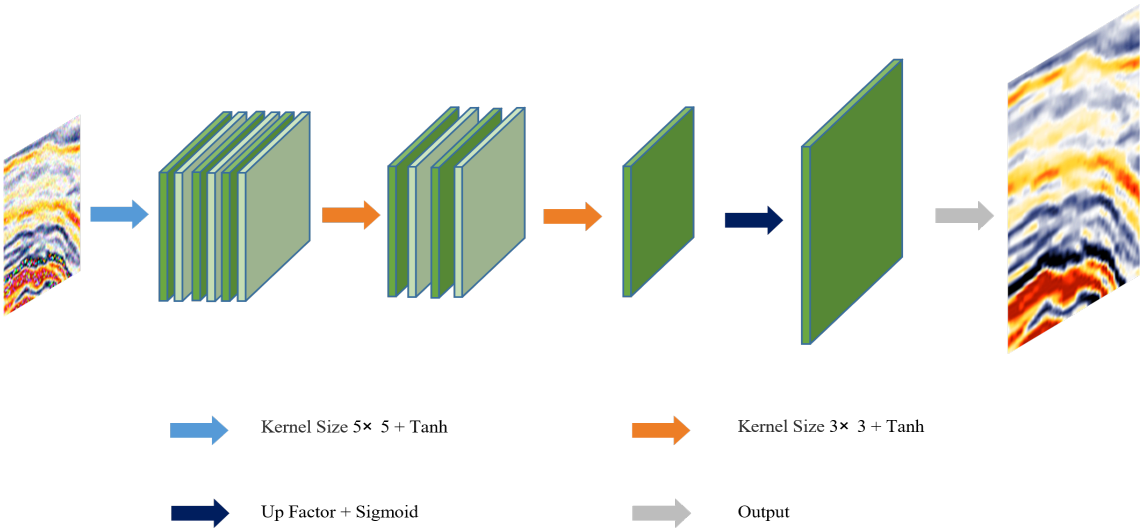


Fig. 1. Schematic diagram of BCNN.

In the traditional sense, neural networks generally need to be trained under artificial fine regulation, and rely on weight control measures such as weight regularization and similar techniques to avoid the impact of overfitting problems. In this paper, we innovatively use a Bayesian convolutional neural network based on variational inference to solve the overfitting problem. BCNN carries out network optimization through random sampling weight distribution. Since the weights obtained from each collection are different, it is equivalent to fusing multiple network models with different parameter combinations of good performance to achieve the effect of model average, so that BCNN can learn more abundant representation and prediction.

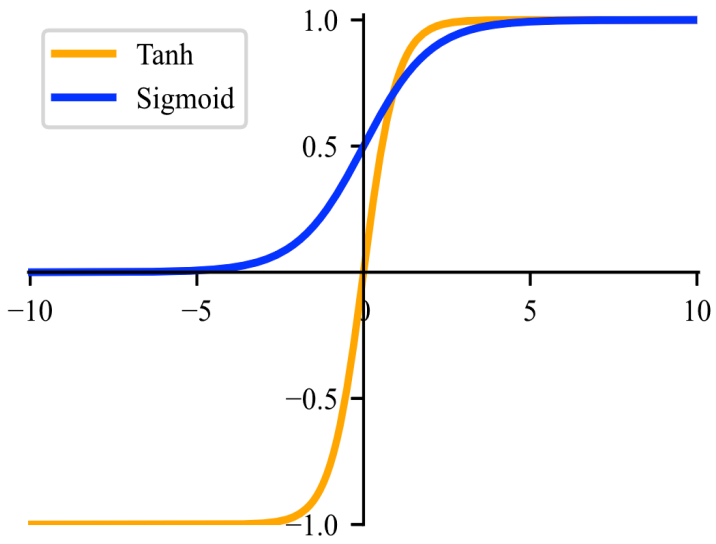


Fig. 2. Activation functions used in the network.

DATASET CREATION AND NETWORK TRAINING

Dataset creation

The data set used in this article is based on the F3 seismic data set published by SEG. The F3 works in the northern part of the Netherlands are well defined and are now well recognized in geology and geophysics. The data set is 122.9 km², with 951 Inline lines and 651 Crosslines, each with 634 sampling points at a sampling rate of 4 ms.

The training data set consists of both pure wave data and enhanced data used as labels. The enhanced data for the label does not really exist. In this paper, the seismic data after manual filtering is selected as the label data to ensure the authenticity of the training data. In order to make the network learn rich data features and give full play to the performance of the network, the data set is made by selecting seismic data containing special geological structures. The specific operation is as follows: Slice the F3 seismic data Inline profile with a sliding window of 85×85, and randomly add Gaussian noise with variance ranging from 0.01 to 0.18 on the basis of manual filtering data to conduct deep damage to geological structures, so as to simulate the pure wave data suitable for training the network. On the basis of the slicing data, the data scale is adjusted to 255×255 as the label data. The training set, verification set and test set are divided into approximately 8:1:1 ratio. Finally, 1995 sets of 85×85 pure wave data, 255×255 labeled data are used as the training set, 230 sets of 85×85 pure wave data, 255×255 labeled data are used as the verification set, 230 sets of 85×85 pure wave data, 255×255 labeled data as the test set. Part of the training data is shown in Fig. 3 for speeding up network training, using GPU for training. The graphics card model is NVIDIA Quadro RTX 4000Ti with 8 GB of separate video memory. The deep learning framework uses PyTorch. This configuration provides enough computing power and memory capacity to efficiently train neural network models.

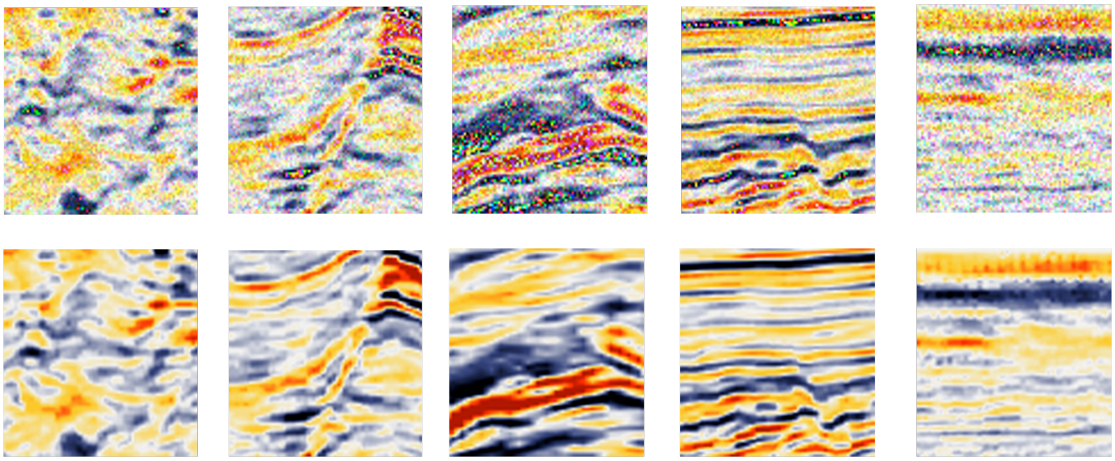
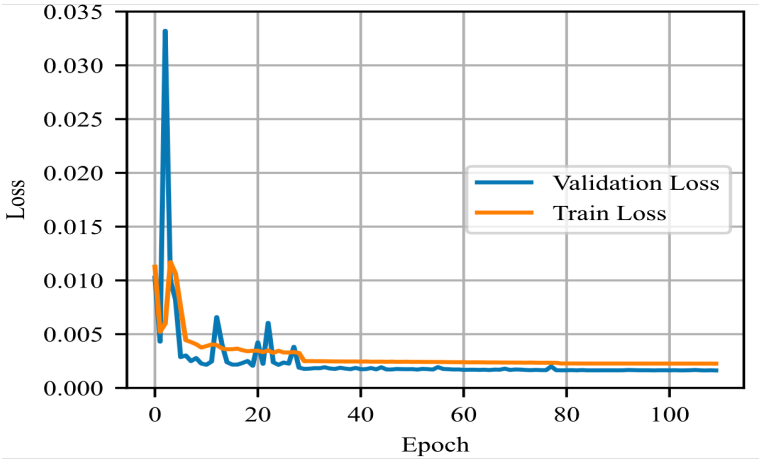


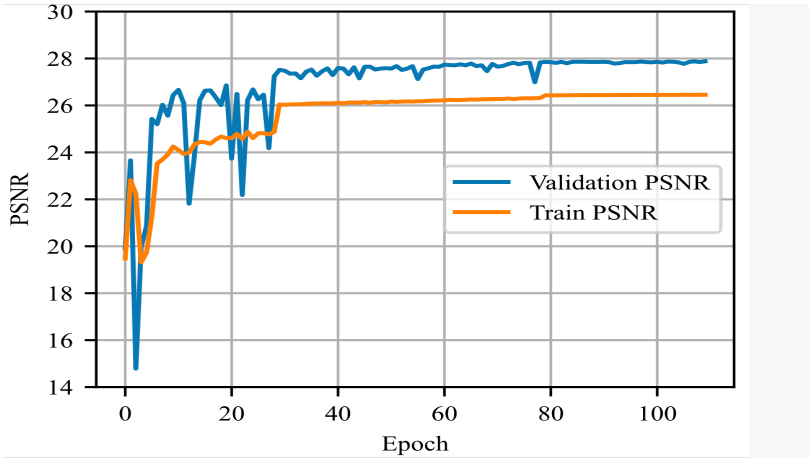
Fig. 3. Part of the training data sample.

Network training and optimization

A total of 200 rounds of network training were conducted with a batch size of 32. The loss function was selected as MSE and the Adam optimization algorithm was used to minimize the loss function. When loading the data, the data of each batch was randomly arranged to introduce rich sample variations for training, helping the optimization algorithm to better explore the parameter space. In addition, during the network training process, set the initial learning rate to 0.01, use MultiStepLR, an easy-to-adjust learning rate adjustment strategy, set parameter Gamma to 0.1, and adjust the learning rate at Epoch 30 and 80, respectively, to improve the convergence speed and prediction accuracy of the model. The results of network training are shown in Fig. 4 (the network starts training to the convergence process).



(a)



(b)

Fig. 4. Network training results. (a) Loss function of network training; (b) PSNR change curve during training.

Fig. 4(a) shows the loss function of training set and verification set during training. The loss function of the network on the training set and the verification set shows the same trend, and with the increase of the network training times, the loss on the verification set is lower than that on the test set. Fig. 4(b) shows the PSNR change curve of training data and verification data in the training process. PSNR is the peak signal-to-noise ratio, whose expression is

$$\mathrm{PSNR} = 10 \times \lg\left(\frac{\mathrm{MAX}^2}{\mathrm{MSE}}\right) \quad , \tag{10}$$

where MAX is the maximum possible value of the signal (e.g., 255 for an 8-bit grayscale image) and MSE is the mean square error, which represents the average of the square of the difference between each pixel of the original image and the enhanced image.

$$\mathrm{MSE}=\frac{\sum[\sum(\mathrm{I(i,j)-K(i,j)})^2]}{\mathrm{M*N}} \quad , \tag{11}$$

where M and N represent the width and height of the image respectively, I(i, j) represents the pixel value of the original image at the pixel position (i, j), and K(i, j) represents the pixel value of the processed image at the pixel position (i, j). The first summation means summing all the rows of the image, and the second summing means summing all the pixels of each row. When the model loss tends to be stable, using PSNR to quantify the performance results of the network on the training set and verification set, it can be found that the PSNR of the final result of the training set is stable below 27 dB, while the PSNR of the final result of the verification set is obviously slightly higher than the training set, stable around 28 dB. The results show that the neural network with Bayes fusion has strong generalization ability and is not easy to overfit.

COMPARATIVE ANALYSIS OF ENHANCED RESULTS

In order to further demonstrate the enhancement effect of the network applied in this paper on seismic data, the residual section, amplitude spectrum and spectrogram before and after the enhancement are analyzed by selecting seismic data containing two special geological structures, fault and uplift respectively.

Fault structure enhancement

Select the original seismic section at Inline 104, where the fault structure is obvious, but the details are blurred due to noise interference. Fig. 5(a) is the original seismic section, and Fig. 5(b) is the enhanced seismic section after the method in this paper. Compared with the pre-

enhanced seismic section, the noise is significantly suppressed, the in-phase axis is more continuous and clear, the difference between layers is significant, and the fault structure is more prominent. Fig. 5(c) shows the seismic section after FCNN enhancement. Compared with the original section, the noise removal is obvious, but compared with the method in this paper, a small amount of noise remains and the detail enhancement of the in-phase axis is poor. The construction-oriented filtering is based on anisotropic diffusion smoothing algorithm, which only smooths the information parallel to the seismic in-phase axis, but does not smooth the information perpendicular to the seismic in-phase axis. Fig. 5(d) shows the enhancement result of the construction-oriented filtering. The in-phase axis processed by the algorithm is more continuous, and the fault structure is also effectively enhanced. However, because the information perpendicular to the in-phase axis is not processed, the smoothness of the formation edge is low, and the overall visual resolution is significantly lower than that of the proposed method.

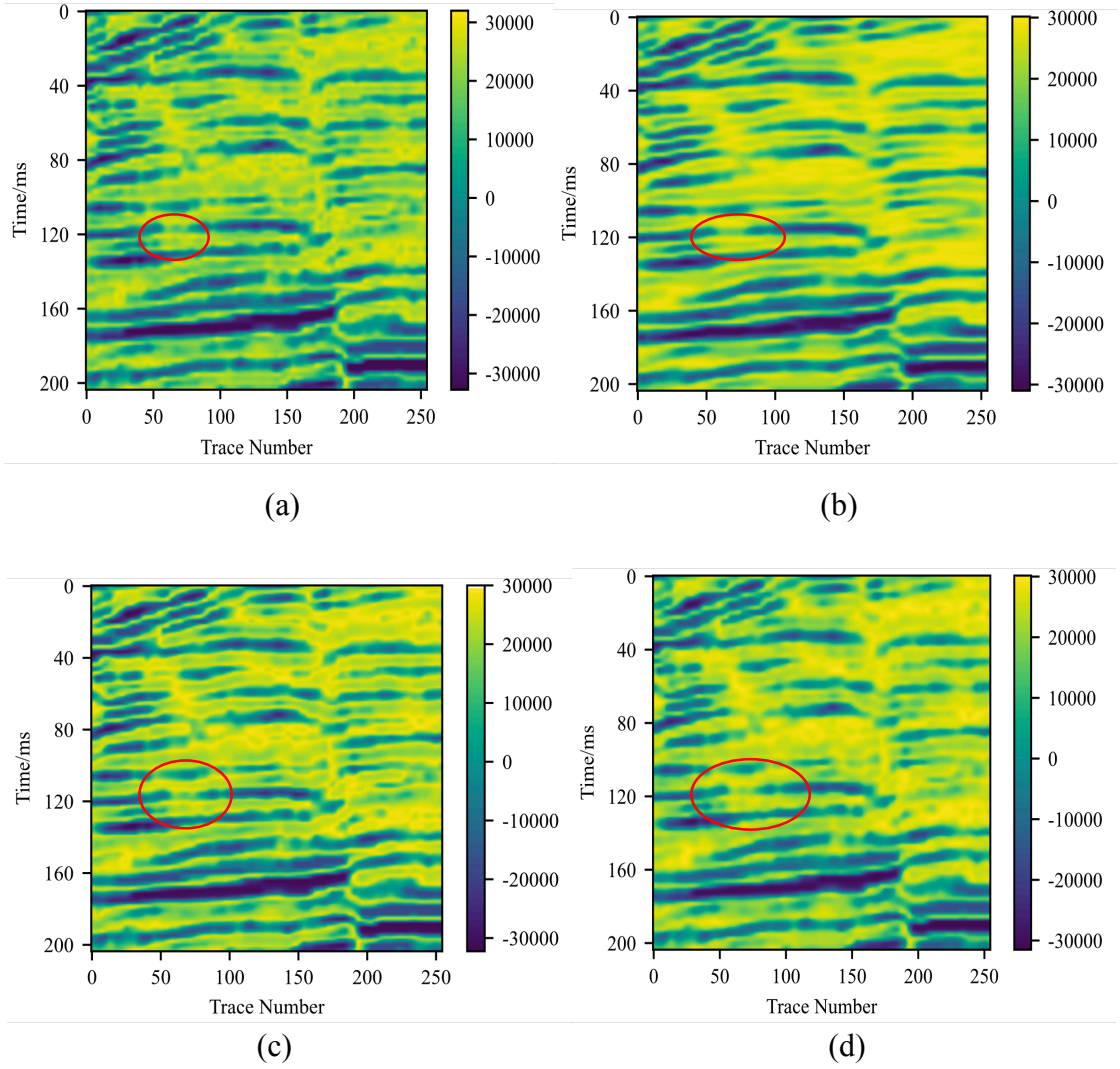


Fig. 5. Comparison of fault enhancement effects of different methods. (a) Original fault seismic data; (b) BCNN enhancement results; (c) FCNN enhanced results; (d) Construct guided filter enhancement results.

Fig. 6 shows the residual profiles corresponding to the three methods. Near the 250th channel, the residual plots of the three methods all contain a small amount of relatively clear in-phase axis information, indicating that some effective signals are over-enhanced. In the residual section corresponding to FCNN, the in-phase axis information is particularly obvious, and even distortion occurs at the break point. Compared with the other two methods, the method presented in this paper can enhance the fault structure and avoid distortion better. Fig. 6(d) shows the average amplitude spectrum of fault seismic data before and after enhancement by the three methods, in which the effective signals are mainly concentrated within 40 Hz. Due to its special geological structure, the amplitude at low frequency is larger. Compared with the three methods, the overall frequency of seismic data enhanced by structure-oriented filtering is lower than that of the original data. The high-frequency part of the seismic data enhanced by FCNN is slightly higher than that of the original data, and the enhanced amplitude spectrum of the proposed method is more consistent with the amplitude distribution of the original data, indicating that the proposed method has good amplitude preservation.

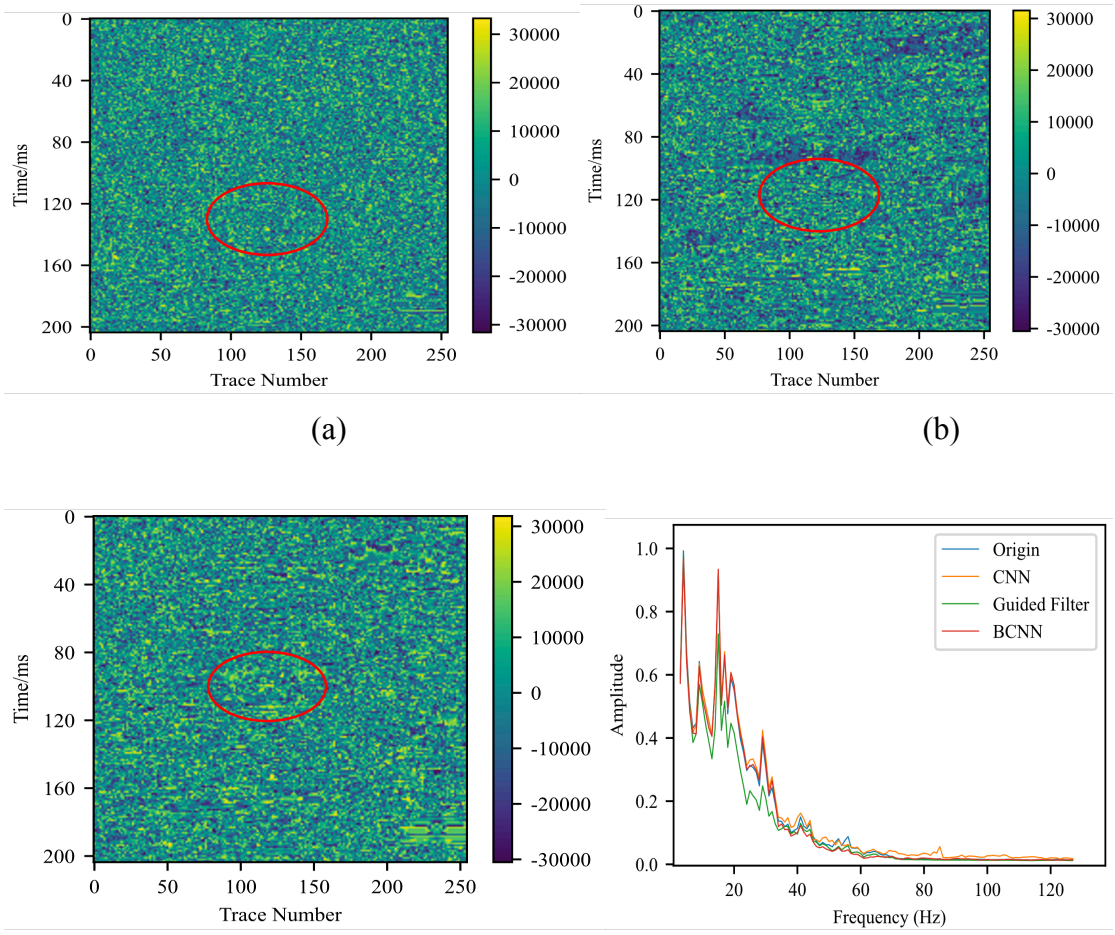


Fig. 6. Residual sections and amplitude spectra of different fault methods: (a) BCNN fault residual profile; (b) FCNN fault residual profile; (c) Construct guided filter fault residual profile; (d) Amplitude spectrum before and after fault tectonic reinforcement.

The f-k spectra corresponding to the original data and different methods are shown in Fig. 7, where the brighter the color indicates the stronger the signal energy of the corresponding frequency. The original acquired data contains a large amount of incoherent noise, and the energy of the f-k spectrum presents a dispersed state. Compared with the original data, the corresponding f-k spectrum energy after processing by the three methods is convergent. In general, the f-k spectrum energy of the method proposed in this paper is more concentrated, indicating that BCNN has more advantages in enhancing the details of fault structure. The enhanced PSNR of the three methods was recorded, and the values of the proposed method, FCNN and construction-oriented filter were 29.93 dB, 28.41 dB and 29.30 dB, respectively.

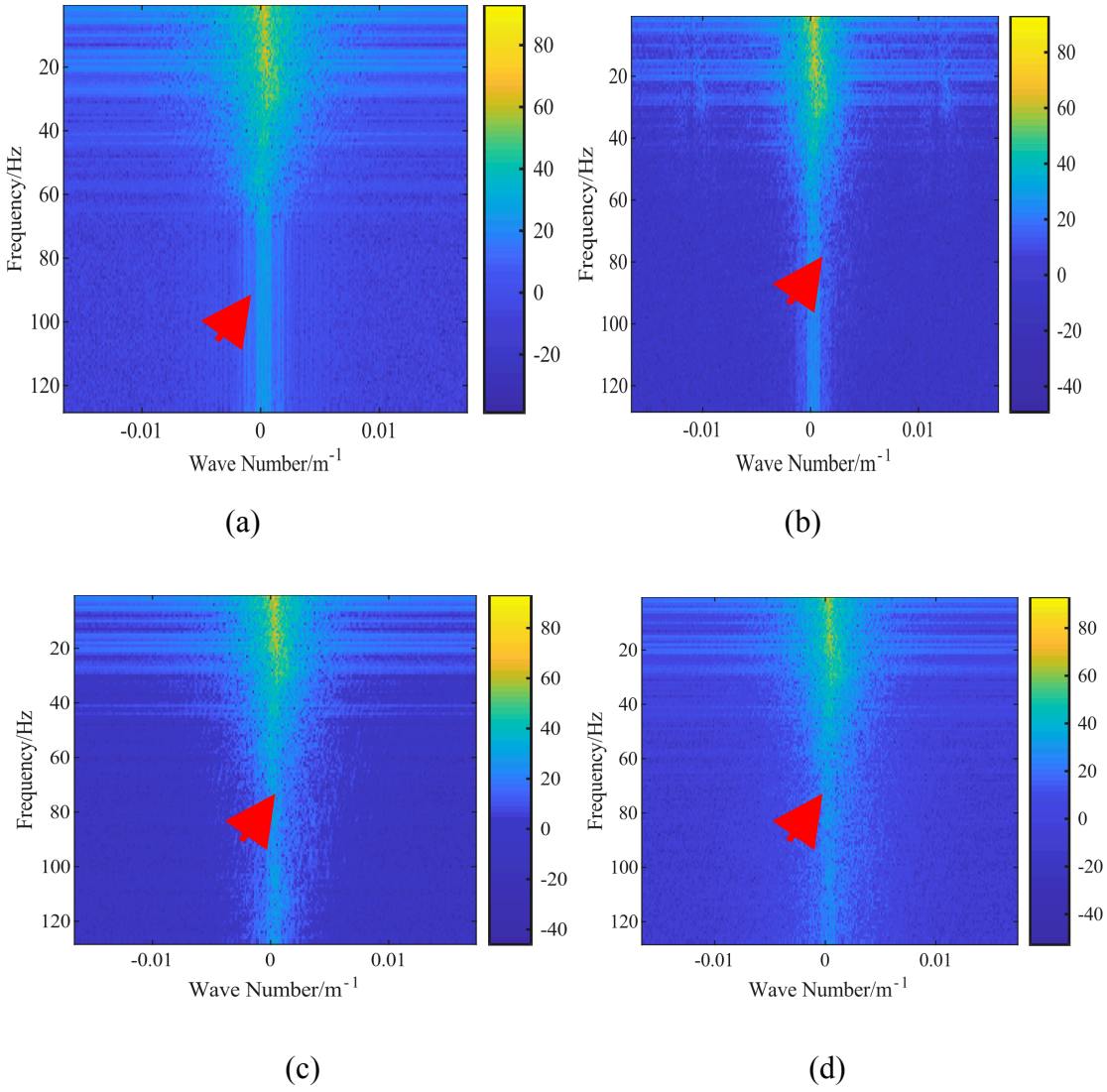


Fig. 7. f-k spectra of different methods. (a) f-k spectrum of original fault seismic data; (b) f-k spectrum of seismic data enhanced by BCNN; (c) f-k spectrum of seismic dataenhanced by FCNN; (d) Construct f-k spectrum of seismic data enhanced by guided filter.

Fault uplift structural enhancement

Select the local section of the original fault rise at Inline 420. The fault rise structure is complex, the in-phase axis is dense, and some of the in-phase axis is difficult to trace. Fig. 8(a) shows the local section of the original fault rise structure, and Fig. 8(b) shows the enhanced effect of the method in this paper. In contrast, the noise of the enhanced seismic data has been effectively removed. Moreover, near the 130th channel, the continuity of the inclined in-phase axis is significantly improved. Fig. 8(c) shows the enhancement effect of FCNN, and the noise is suppressed to a certain extent. Compared with the method in this paper, the enhancement effect of the tilt in-phase axis continuity needs to be improved. Fig. 8(d) shows the enhancement effect of constructing oriented filtering. This method removes a large amount of noise and has a remarkable enhancement effect in the direction parallel to the in-phase axis. However, because it does not smooth the transverse discontinuous in-phase axis, it is slightly inferior in the processing of details compared with BCNN.

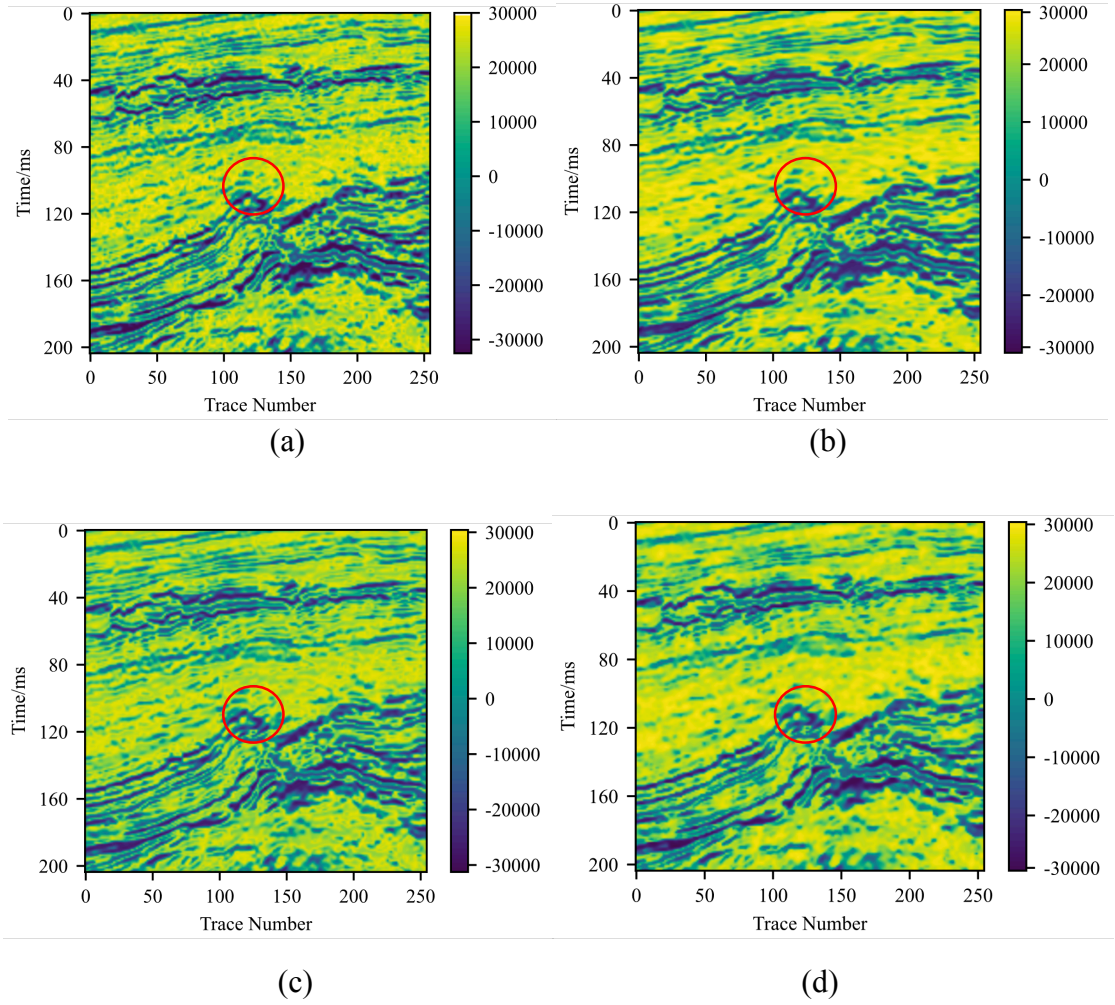


Fig. 8. Comparison of local enhancement results of fault uplift structures by different methods. (a) Original local uplift seismic data; (b) BCNN enhancement results; (c) FCNN enhanced results; (d) Construct guided filter enhancement results.

Fig. 9 shows the residual profiles corresponding to different methods. There are no effective signals observed in the residual plots of the proposed method and the construction-oriented filter, but there are subtle effective signals in the residual plots of FCNN, which indicates that the protection of effective signals by FCNN method is poor. Fig. 9(d) shows the average amplitude spectrum of the section where the fault rise structure is located after enhancement by different methods. The amplitude of the seismic data enhanced by the proposed method is more consistent with the original data than that of the construction-oriented filtering and FCNN, indicating that BCNN also has better amplitude preservation in the enhanced fault rise structure. The f-k spectra enhanced by the original data and different methods are shown in Fig. 10. It can be seen that the energy of the f-k spectra of the proposed method is more concentrated. The PSNR of BCNN, FCNN and construction-oriented filtering are 29.70 dB, 28.67 dB and 29.40 dB, respectively. The proposed method has the highest PSNR, which is superior to FCNN and structure-oriented filtering.

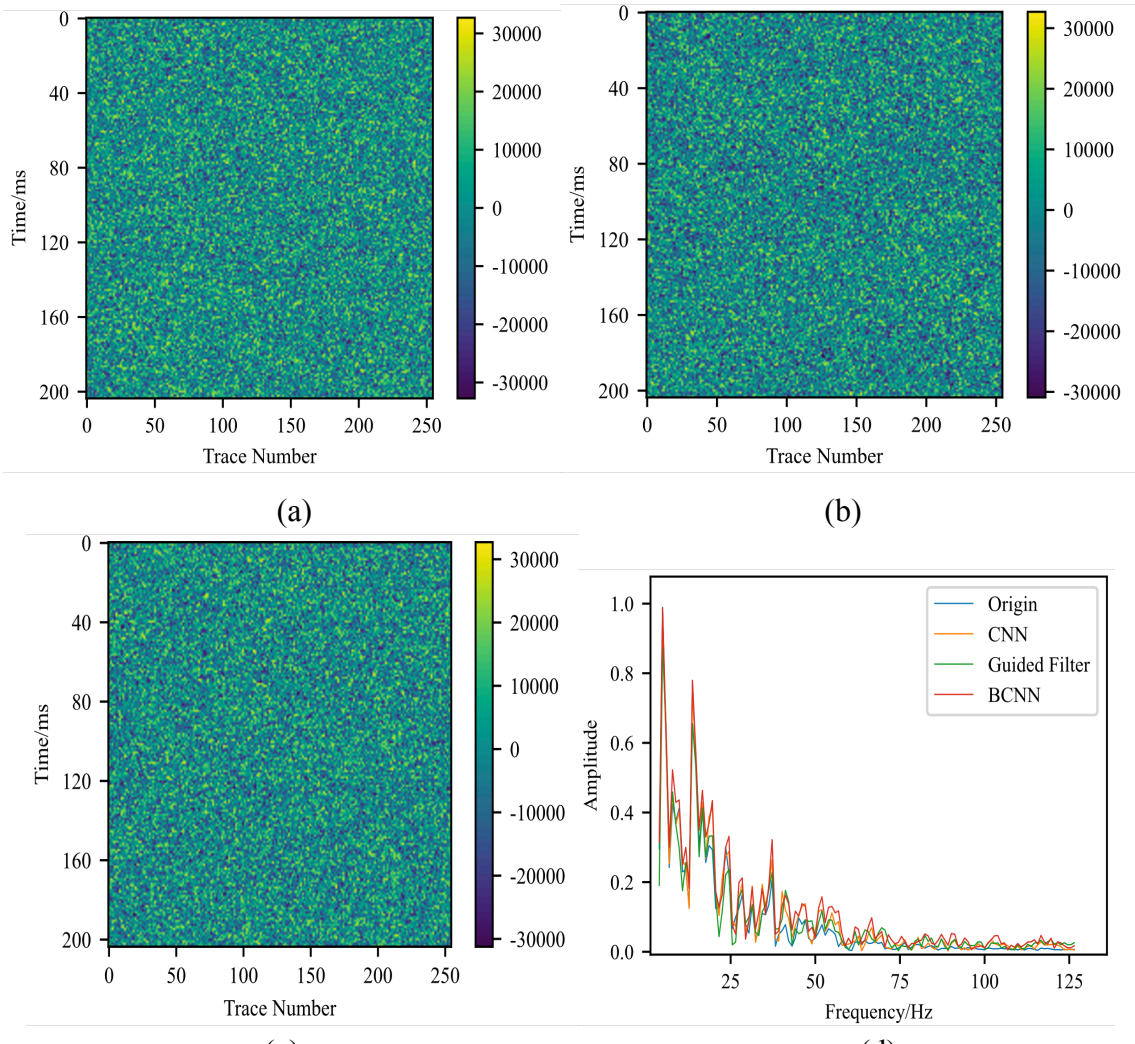


Fig. 9 Residual profiles of fault uplift structures by different methods. (a) BCNN fault rise residual profile; (b) FCNN back rise residual profile; (c) Construct a guided filter rise residual profile; (d) Amplitude spectrum before and after fault uplift enhancement.

CONCLUSION

As a kind of Bayesian modeling network based on variational inference for weights, BCNN has achieved good results in many fields such as image recognition and image super resolution. The author applies BCNN to seismic data enhancement. Compared with FCNN and construction-oriented filtering algorithm, the proposed algorithm can effectively remove random noise, protect and highlight geological structure details, improve the visual resolution of seismic data on the whole, and have more advantages in the retention of effective detail information. This is due to the fact that the random sampling operation of the weight distribution in the training process enables BCNN to have a deeper data representation ability and the up-sampling operation at the end of the network protects the training data in the low resolution spatial feature information. In addition, compared with FCNN trained with the same data set, the training results and test results of BCNN fully demonstrate that the network has stronger generalization ability and can effectively avoid overfitting problems.

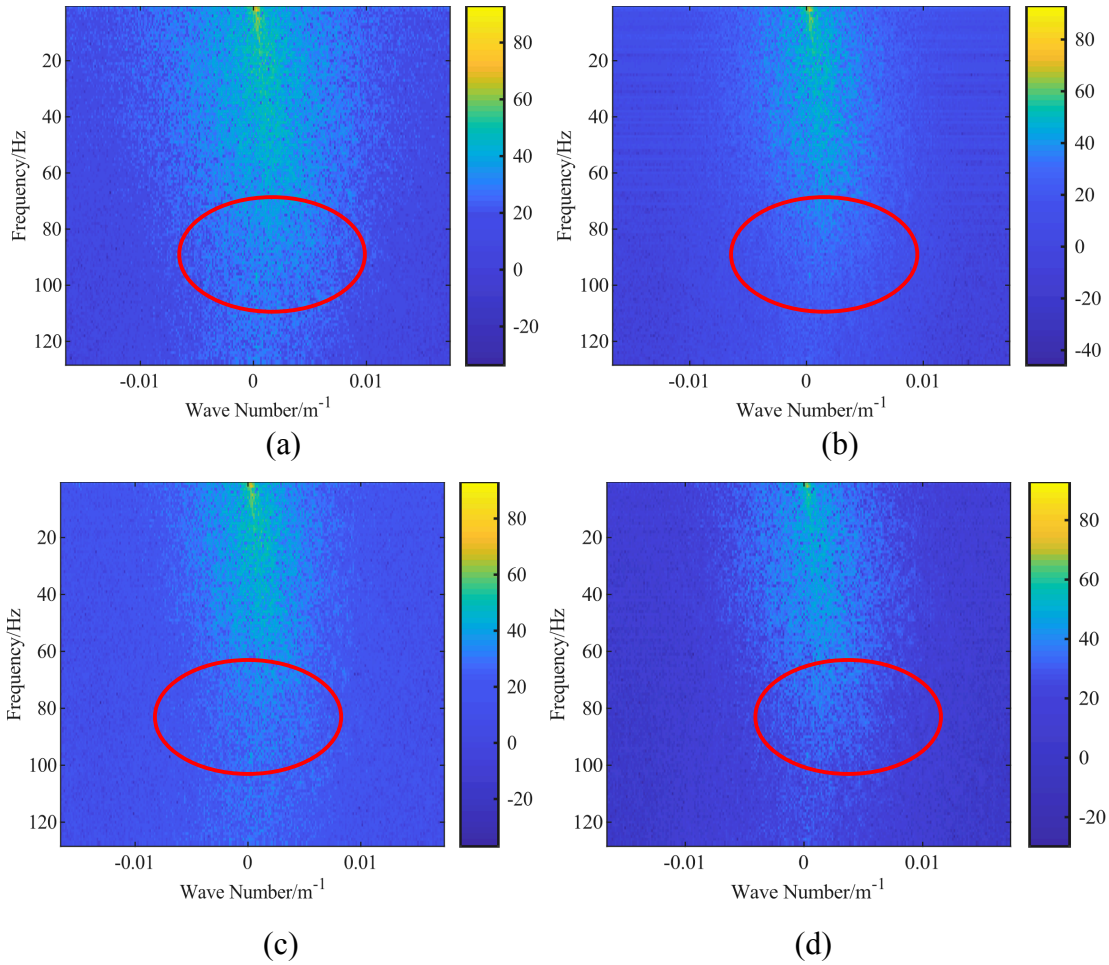


Fig. 10 f-k spectra of different methods. (a) f-k spectrum of original local fault rise seismic data; (b) f-k spectrum of seismic data enhanced by BCNN; (c) f-k spectrum of seismic data enhanced by FCNN; (d) Construct f-k spectrum of seismic data enhanced by guided filtering.

Although BCNN has achieved a good effect on seismic data enhancement, the whole network training process lacks the combination with the seismic data itself. In the follow-up work, the internal relationship of seismic data should be fully explored, and high-quality training data should be created to further improve the enhancement effect of the network for special geological structures.

ACKNOWLEDGMENT

This research is supported by the Open Research Fund of Key Laboratory of Engineering Geophysical Prospecting and Detection of the Chinese Geophysical Society.

REFERENCES

Alaei, N., Roshandel Kahoo, A. and Kamkar Rouhani, A., 2018. Seismic resolution enhancement using scale transform in the time-frequency domain. Geophysics, 83(6): V305-V314.

Bednar, J.B., 1983. Applications of median filtering to deconvolution, pulse estimation, and statistical editing of seismic data. Geophysics, 48: 1598-1610.

Canales, L., 1984. Random noise reduction. Expanded Abstr., 54th Ann. Internat. SEG Mtg., Atlanta: 525-527.

Chang, D.K., Yang, W.Y., Wang, Y.H., Yang, Q., Wei, X.J. and Feng, X.Y., 2018. Random noise suppression for seismic data using a non-local Bayes algorithm. Appl. Geophys., 15: 91-98.

Chang-Qing, M., 2012. Research and application on seismic image enhancement based on wavelet transformation. Adv. Automat. Robot., 2: 197-202.

Chen, W., Zhang, D. and Chen, Y., 2017. Random noise reduction using a hybrid method based on ensemble empirical mode decomposition. J. Seismic Explor., 26: 227-249.

Chen, W. and Song, H., 2018. Automatic noise attenuation based on clustering and empirical wavelet transform. J. Appl. Geophys., 159: 649-665.

Chen, Y. and Jin, Z., 2015. Simultaneously removing noise and increasing resolution of seismic data using waveform shaping. IEEE Geosci. Remote Sens. Lett., 13: 102-104.

Claerbout, J.F., 1971 Toward a unified theory of reflector mapping. Geophysics, 36: 467-481.

Cornfield, J., 1967. Bayes theorem. Revue de l'Institut International de Statistique, 1967: 34-49.

Deng, L., Yuan, S. and Wang, S., 2017. Sparse Bayesian learning-based seismic denoise by using physical wavelet as basis functions. IEEE Geosci. Remote Sens. Lett., 14: 1993-1997.

Dong, X, and Li, Y., 2020. Denoising the optical fiber seismic data by using convolutional adversarial network based on loss balance. IEEE Transact. Geosci. Remote Sens., 59: 10544-10554.

Gal, Y. and Ghahramani, Z., 2015. Bayesian convolutional neural networks with Bernoulli approximate variational inference. arXiv preprint arXiv:1506.02158.

Gao, Y., Zhao, D.F., Li, T.H., Li, G.F. and Guo, S.W., 2023. Deep learning vertical resolution enhancement considering features of seismic data. IEEE Transact. Geosci. Remote Sens., 61: 1-13.

Gómez, J.L., Velis, D.R. and Sabbione, J.I., 2020. Noise suppression in 2D and 3D seismic data with data-driven sifting algorithms. Geophysics, 85(1): V1-V10

Gulunay, N., 1986. FXDECON and complex Wiener prediction filter. Expanded Abstr., 56th Ann. Internat. SEG. Mtg., Houston: 279-281.

Hennenfent, G. and Herrmann, F.J., 2006. Seismic denoising with nonuniformly sampled curvelets. Comput. Sci. Engineer., 8(3): 16-25.

Jin, Y., Wu, X., Chen, J., Han, Z. and Hu, W., 2018. Seismic data denoising by deep-residual networks. Expanded Abstr., 88th Ann. Internat. SEG Mtg., Anahein: 4593-4597.

Kuang, L., Liu, H., Ren, Y., Luo, K., Shi, M., Su, J. and Li, X., 2021. Application and development trend of artificial intelligence in petroleum exploration and development. Petrol. Explor. Develop., 48: 1-14.

Lan, N.Y., Zhang, F.C., Sang, K.H. and Yin, X.Y., 2023. Simultaneous denoising and resolution enhancement of seismic data based on elastic convolution dictionary learning. Petrol. Sci. doi: 10.1016/j.petsci.2023.02.023.

Li, C., Peng, S., Cui, X., Du, W. and Lin, P., 2023. Seismic data enhancement based on common-reflection-surface-based local slope and trimmed mean filter. IEEE Transact. Geosci. Remote Sens., 61: 1-9.

Li, J., Wu, X. and Hu, Z., 2021. Deep learning for simultaneous seismic image super-resolution and denoising. IEEE Transact. Geosci. Remote Sens., 60: 1-11.

Liu, W. and Duan, Z., 2019. Seismic signal denoising using f-x variational mode decomposition. IEEE Geosci. Remote Sens. Lett., 17: 1313-1317.

Mallat, S., 1991. Zero-crossings of a wavelet transform. IEEE Transact. Informat. Theory, 37: 1019-1033.

Oliveira, D.A.B., Ferreira, R.S., Silva, R., 2019. Improving seismic data resolution with deep generative networks. IEEE Geosci. Remote Sens. Lett.,16: 1929-1933.

Qiu, C., Wu, B., Liu, N., Zhu, X. and Ren, H., 2021. Deep learning prior model for unsupervised seismic data random noise attenuation. IEEE Geosci. Remote Sens. Lett., 19: 1-5.

Shi, W., Caballero, J., Huszar, F., Totz, J., Aitken, A.P., Bishop, R., Rueckert, D. and Wang, Z., 2016. Real-time single image and video super-resolution using and efficient sub-pixel convolutional neural network. Proc. IEEE Conf. Comput. Vis. Patt. Recognit.: 1874-1883.

Shridhar, K., Laumann, F. and Liwicki, M.A., 2019. Comprehensive guide to Bayesian convolutional neural network with variational inference. arXiv preprint arXiv:1901.02731.

Singh, G. and Mittal, A., 2014. Various image enhancement techniques - a critical review. Internat. J. Innovat. Sci. Res., 10: 267-274.

Song, H., Gao, Y., Chen, W., Xue, Y.-J., Zhang, H. and Zhang, X., 2020. Seismic random noise suppression using deep convolutional autoencoder neural network. J. Appl. Geophys., 178: 104071.

Song-Tao, L and Gang, W., 2010. A novel method for image enhancement based on generalized histogram. Optic Control, 3: 12-14.

Wang, E. and Nealon, J., 2019. Applying machine learning to 3D seismic image denoising and enhancement. Interpretation, 7(3): SE131-SE139.

Wang, Y.Q., Lu, W.K., Liu, J.L., Zhang, M. and Miao, Y.K., 2019. Seismic random noise suppression based on data augmentation and CNN. Chin. J. Geophys., 62: 421-433.

Yan, Z., Gu, H.M. and Cai, C.G., 2013. Seismic image enhancement processing based on anisotropic diffusion filtering. Oil Geophys. Prosp., 48: 390-394.

Yang, P., Mu, X. and Zhang, J.T., 2010. Directional boundary preserving fault enhancement technique. Chin. J. Geophys., 53: 2992-2997.

Zhang, H., Alkhalifah, T., Liu, Y., Birnie, C. and Di, X., 2022. Improving the generalization of deep neural networks in seismic resolution enhancement. IEEE Geosci. Remote Sens. Lett. doi: 10.1109/LGRS.2022.3229167.

Zhao, B., Han, L., Zhang, P. and Yuchen, Y., 2022. Weak signal enhancement for passive seismic data reconstruction based on deep learning. Remote Sens., 14: 5318.