



A deep learning network for estimation of seismic local slopes

Wei-Lin Huang¹ · Fei Gao¹ · Jian-Ping Liao^{2,3} · Xiao-Yu Chuai⁴

Received: 18 June 2020 / Accepted: 14 September 2020 / Published online: 30 November 2020
© The Author(s) 2020

Abstract

The local slopes contain rich information of the reflection geometry, which can be used to facilitate many subsequent procedures such as seismic velocities picking, normal move out correction, time-domain imaging and structural interpretation. Generally the slope estimation is achieved by manually picking or scanning the seismic profile along various slopes. We present here a deep learning-based technique to automatically estimate the local slope map from the seismic data. In the presented technique, three convolution layers are used to extract structural features in a local window and three fully connected layers serve as a classifier to predict the slope of the central point of the local window based on the extracted features. The deep learning network is trained using only synthetic seismic data, it can however accurately estimate local slopes within real seismic data. We examine its feasibility using simulated and real-seismic data. The estimated local slope maps demonstrate the successful performance of the synthetically-trained network.

Keywords Deep learning · Neural network · Seismic data · Local slopes

1 Introduction

As one of the geometric attributes of seismic signal, the local slopes contain complete information of the reflection geometry, which is of great significance to the analysis of seismic data. An accurate estimation of the slope information can benefit many subsequent procedures such as horizons interpretation (Fomel 2010; Wu and Hale 2015), structure enhancement (Hale 2009; Liu et al. 2010), incoherent/

coherent noise attenuation (Liu et al. 2015; Huang et al. 2017), deblending (Huang et al. 2018), seismic interpolation/reconstruction (Gan et al. 2016; Huang and Liu 2020), inversion (Yao et al. 2020; Liu et al. 2020a) and imaging (Fomel 2007; Zhang et al. 2019b). The introduction of local slopes into seismic data processing can be traced back to the work of Rieber (1936), in this basic, a method of controlled direction reception is developed, which can achieve excellent results in seismic processing and interpretation (Riabinkin 1957). After then, a number of researchers have used several techniques to estimate the local slopes. Ottolini (1983) proposed a picking method of local slopes by local slant stacks. Lambaré et al. (2004) extended Ottolini's method further to make it simpler and more automatic. But it is extremely expensive to cast about the best direction by slant slacking with different slopes. Claerbout (1992) first delved into a plane-wave destruction filter to estimate seismic event slopes in local windows under an assumption of local plane-wave model. Such filter is constructed with the help of an implicit finite-difference scheme for the local plane-wave equation (Fomel 2002). Andersson and Duchkov (2013) described a structure tensor-based local slopes estimation technique and also provided a summary of previous studies in detail. Huang et al. (2017a) proposed an algorithm to track the largest values of the upper envelope in local windows to obtain the event slope.

Edited by Jie Hao and Xiu-Qiu Peng

✉ Wei-Lin Huang
cup_hwl@126.com

✉ Xiao-Yu Chuai
chuaixiaoyu@126.com

¹ State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum-Beijing, Beijing 102249, China

² State Key Laboratory of Coal Resources and Safe Mining, China University of Mining and Technology-Beijing, Beijing 100083, China

³ Hunan Provincial Key Laboratory of Shale Gas Resource Utilization, Hunan University of Science and Technology, Xiangtan 411201, China

⁴ Hebei Coal Research Institute Co., Ltd, Jizhong Energy Group, Xingtai 054000, China

The computer-assisted seismic data processing and interpretation become more popular and have demonstrated outstanding performance. Machine learning refers to the process that perceives the data environment and discovers as well as learns information from the data, and further makes strategic decisions that maximize the chance of successfully achieving specific goals (Murphy 2012; Valentine and Kalnins 2016; Chen 2018;). Compared with traditional technique, machine learning may achieve higher efficiency and precision (Chen et al. 2016; Wu et al. 2019b; Wu et al. 2020b; Zhang et al. 2019b). The machine learning-based technique has drawn a lot of attention from various kinds of fields, such as medical science, biomedical engineering, text processing, (electronic) commerce, internet engineering (Chan et al. 2002; Vafeiadis et al. 2015). In seismological community, machine learning has been successfully applied to noise attenuation (Chen et al. 2019; Zhu et al. 2019; Saad and Chen 2020), signal recognition (Huang 2019), earthquake detection (Jia et al. 2019; Mousavi et al. 2019b), arrival picking (Yu et al. 2018; Zhu and Beroza 2018; Zhao et al. 2019; Jiang and Ning 2019; Zhang et al. 2020), fault detection (Ping et al. 2018), geophysical inversion (Chen et al. 2018; Li et al. 2020; Wu et al. 2020a), traveltime parameters estimation (Liu et al. 2020b) and reservoir porosity prediction (Chen et al. 2020). As one of the machine learning methods, deep learning is composed of multiple processing layers to intelligently extract data features, which can dramatically improve the state-of-the-art in various fields (Lecun et al. 2015). It has become a powerful tool in the seismic processing and interpretation (Waldeland et al. 2018; Mousavi et al. 2019a). Lewis and Vigh (2017) adapted deep learning technique to full-waveform inversion to overcome challenges caused by complex geological stratification. Ross et al. (2018) trained a convolutional neural network using the hand-labeled data archives of the Southern California Seismic Network and used the trained model to detect seismic body-wave phases. Mousavi and Beroza (2019) designed a regressor (MagNet) composed of convolutional and recurrent neural networks to

predict earthquake magnitude from raw waveforms recorded at single stations. Araya-Polo et al. (2018) obtained an accurate gridding or layered velocity model from shot gathers using deep neural networks.

To estimate the local slope information of a seismic profile, we use a deep neural network of six main layers to predict the local slope of the central pixel in each small data patches. To train and validate the neural network, we create 5,562,120 data patches by randomly convoluting 121 different wavelets with different reflectivity models. The neural network parameters are optimized by minimizing a cross-entropy function. The stochastic gradient descent with momentum method is used to achieve solve the optimization problem. Application of the trained neural network on both synthetic and field example shows the effectiveness.

2 Model introduction

2.1 Deep neural network architecture

Our designed network consists of six main layers including three 2D convolution layers and three fully connected layers, as shown in Fig. 1. The three convolution layers serve as a deep feature extractor which reduces the input images into smaller sets of attributes that hold relevant information in local structure. A single convolution layer can be formulated as:

$$Y_j = g\left(f\left(\sum_i W_j * X_i + b_j\right)\right), \tag{1}$$

where Y_j and X_i stand for the j th extracted features map and the i th input data, respectively. W_j and b_j denote the weighting matrix and basis parameters which control the convoluting and horizontal shifting process, which slides over the input map with a specified stride to reduce the spatial dimensionality. f is a nonlinear activation function which

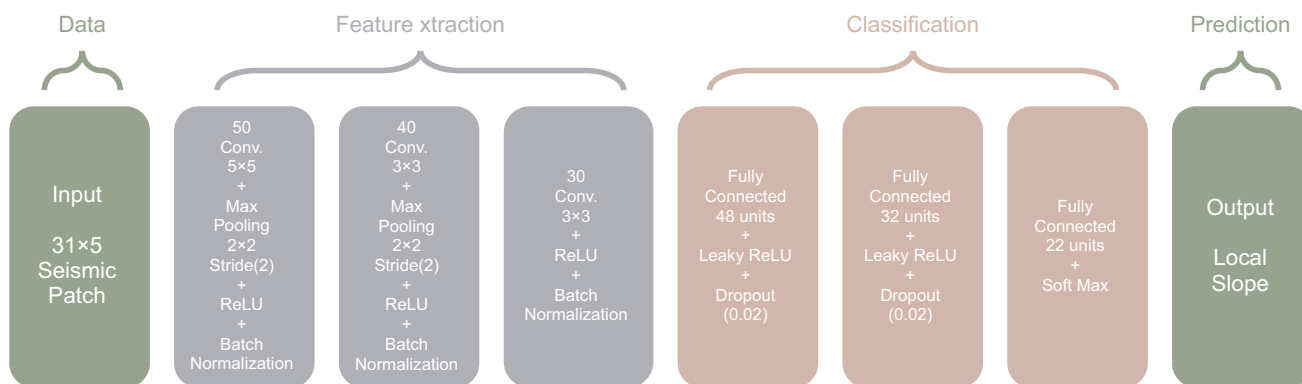


Fig. 1 A demonstration of the neural network architecture

can improve the universality of the deep neural network and its simulation performance of nonlinear process. For the seismic local slope estimation, we choose the state-of-the-art rectified linear unit (ReLU) (Nair and Hinton 2010) as the activation function $f = \max(0, x)$, which returns its argument x when it is greater than zero and returns zero otherwise. The ReLU activation maintains the nonlinearity and off-on characteristic which is similar to a biological neuron. g is the batch normalization operation which accelerates the training process and also reduce the risk of over fitting (Ioffe and Szegedy 2015). For each convolution layer, we pad the data with zeros prior to the convolution operation to keep the sizes of input and output same. To further boost the convergence rate and capacity of resisting disturbance and noise, we insert the max-pooling operation into the first two convolution layers as

$$Y_j = g\left(f\left(p\left(\sum_i W_j * X_i + b_j\right)\right)\right), \quad (2)$$

where p represents the max-pooling operation taking the maximum of the neighborhood.

The three fully connected layers can be treated as a classifier to predict the class based on the extracted features. Instead of convoluting a local sub-image with a convolutional kernel, the neurons in a fully connected layer are completely interconnected with the next layer neurons, but not connected with neurons in the same layer or cross-layer. A single fully connected layer can be formulated as:

$$y_j = \mathbf{w}_j^T \mathbf{x} + b_j, \quad (3)$$

where y_j denotes j th output neuron. \mathbf{x} is the reshaped 1D column vector from the previous layer neurons. \mathbf{w}_j represents the weighting vector and b_j is the basis. To reduce the risk of gradient vanishing and over fitting, the first two fully connected layers are followed by a leaky ReLU activation l and a dropout operation d as

$$y_j = d\left(l\left(\mathbf{w}_j^T \mathbf{x} + b_j\right)\right), \quad (4)$$

The local slope labels are introduced in the last layer. For the last layer, we insert a soft-max operation to obtain the final prediction result (i.e., the local slope of the central point of the input seismic patch). The final prediction results are classified into 22 groups including a non-signal group and 21 groups of various slopes. One can of course define more output categories, but generally need a more complex or deeper neural network. The detailed parameters we use refer to Fig. 1.

2.2 Training data and labels

The main limitation of applying deep learning in seismic image processing remains in preparing rich training datasets and the corresponding labels (Wu et al. 2019a). Manually labeling a mass of seismic images is highly labor-intensive and tedious. Moreover, the manually labeling is highly related to human experience and the result often goes to be different with different interpreters. The incomplete or inaccurate labeling may mislead the training process and trained neural network will make unreliable predictions (Wu et al. 2019a). For this reason, we train the neural network using synthetically-created seismic patches, in which the local slope information is pre-given. The local slope of a seismic section is defined as the local slope of the corresponding reflection coefficients. In order to have realistic and varied seismic data, we generate 11 Ricker wavelets with equally increasing domain frequencies from 10 to 70 Hz (6 Hz interval). For each Ricker wavelet, we diversify it into 11 wavelets with different phases (0–180°, 18° interval), which leads to $11 \times 11 = 121$ different wavelets in total, as shown in Fig. 2.

We estimate the local slope information of each data point within its neighborhood. The size of the neighborhood patch can be varying for the local slope estimation depending on the user requirement. However, it is to be understood that a small size of neighborhood patch may provide insufficient details and result in an incorrect estimation, while a big one introduces some interference information (e.g., nonlocal information) which may misguide the training and predicting processes. In addition, the bigger the size, the higher demand on the memory and computation performance of the workstation. Take these into account, we set a 31×5 local window to estimate the slope of the central point. We randomly convolute the synthetically created wavelets with different reflectivity models to obtain 1,854,040 small seismic patches including 299,430 (16.15%) non-signal patches and 1,554,610 (83.85%) signal patches with different slopes. To make the synthetic seismic images even more realistic, we add different levels of random noise. We also augment the data sets by randomly decimating and dithering operations to make the trained neural network applicable to the seismic data with relatively complicated structure, which produces 5,562,120 patches (898,052 non-signal and 4,664,068 signal patches) in total. Figure 3 shows the statistical analysis of the training data. 1,112,424 (20%) of the patches are randomly selected as the validation data sets.

2.3 Deep neural network training

The whole training process of the deep neural network can be considered as solving a complex nonlinear inverse problem using interactive forward propagation and back

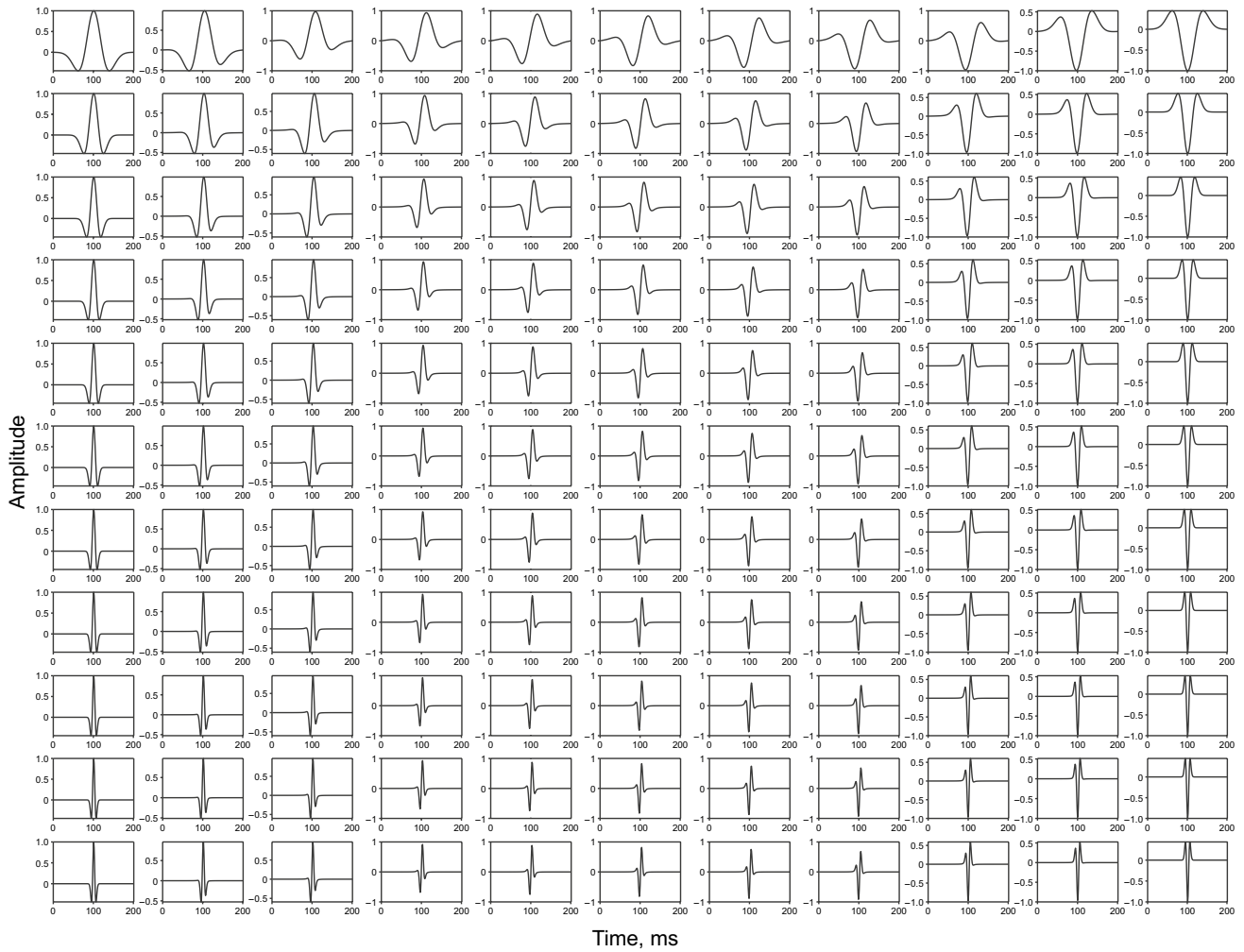


Fig. 2 121 synthetic wavelets for the generation of training data

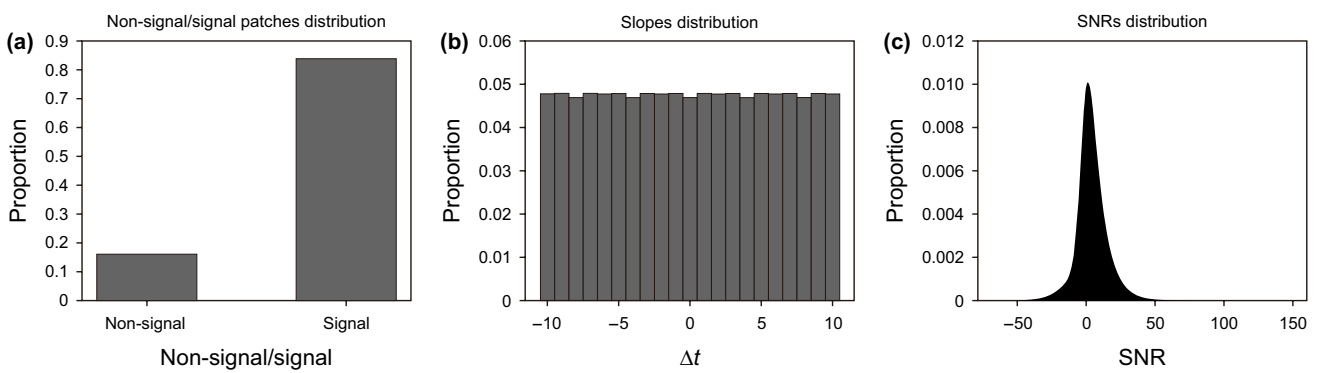


Fig. 3 Statistical analysis of the training data. **a** Non-signal and signal patches distribution. **b** Slopes distribution. **c** SNRs distribution

propagation, where the former calculates the output (prediction), and the latter updates the parameters (weights and biases). The process of updating weights and biases can be viewed as an optimization problem where a cost

function is minimized. We choose the widely-used cross-entropy as the cost function to train our deep neural network:

$$F(\mathbf{p}, \mathbf{q}) = \sum -\mathbf{p} \log(\mathbf{q}), \quad (5)$$

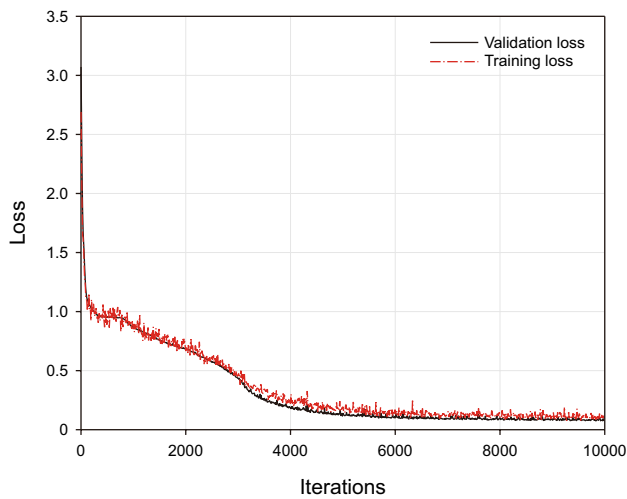


Fig. 4 Loss curves on training set and validation set

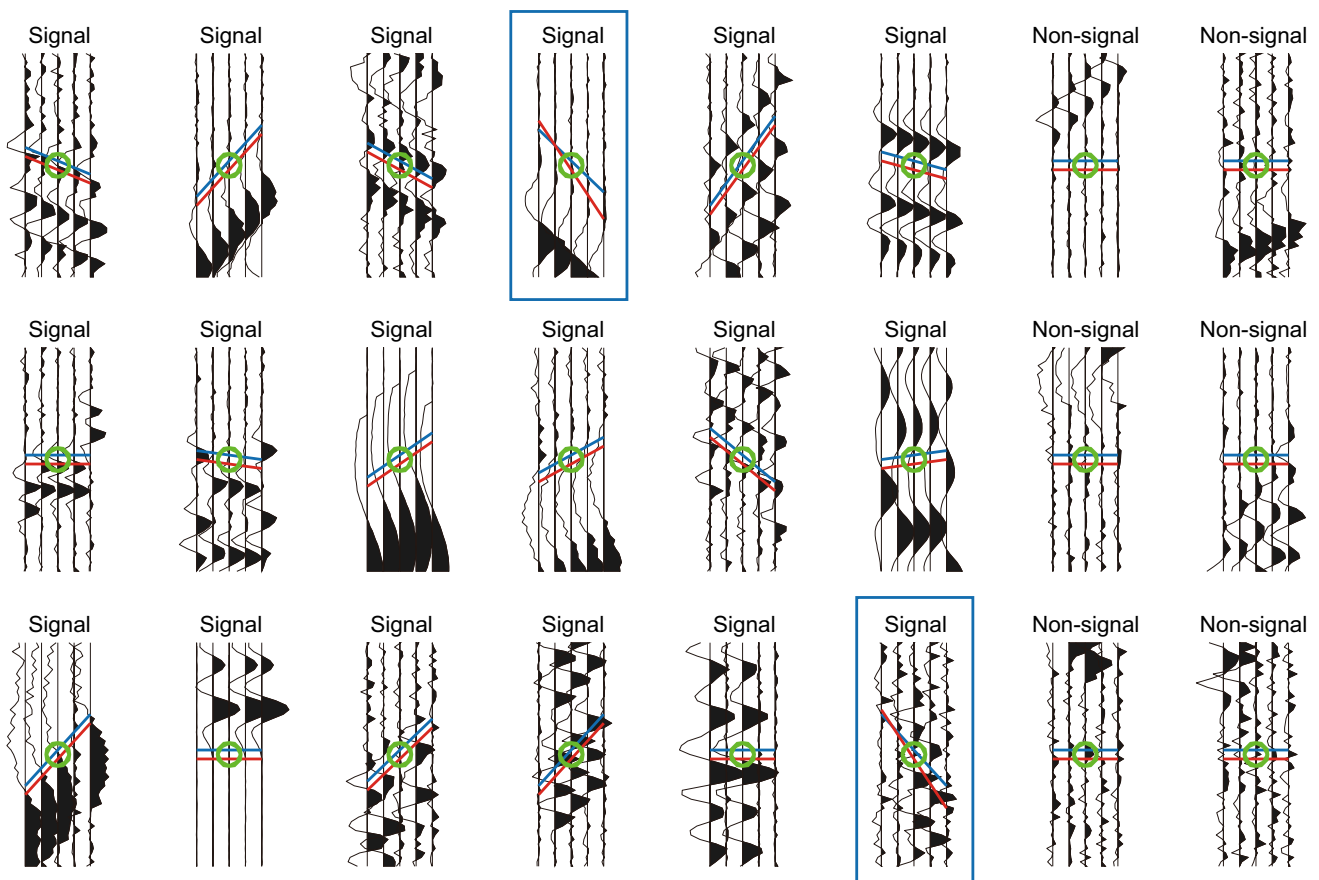


Fig. 5 A demonstration of some testing seismic patches displayed with the true (blue line) and predicted (red line) slopes of the central point (green dot). The blue frameworks mark two wrong predictions

where \mathbf{p} and \mathbf{q} are the true and predicted classification vectors, respectively. We use the stochastic gradient descent with momentum (momentum=0.65) and set a 0.045 initial learning rate and apply a 0.02 dropping factor to it in each epoch to optimize the network parameters. We feed the deep neural network in batches with the size of 1024 patches. Figure 4 presents the loss curves of the first ten thousand iterations on training and validation sets. The final training and validation accuracies are 97.1% and 96.67%, respectively.

Figure 5 shows some randomly-chosen examples of the testing seismic patches (not included in the training data sets) displayed with the true (blue line) and predicted (red line) slopes of the central point (green dot). The last two columns belong to the non-signal category because the signals do not pass the central point. It can be seen that the synthetically-trained network can predict the local slope precisely with only minor error for two patches as marked by the blue frame boxes.

3 Results and discussion

3.1 Synthetic example

We then test the synthetically-trained network on three synthetic seismic data sets which are more realistic, as shown in Fig. 6a–c. The clean signal in each data consists of twelve hyperbolic events with different domain frequencies (20–50 Hz) and initial phases (0°–90°). To test the prediction performance of our network, we added some Gaussian noise to the last two data (Fig. 6b, c). The rate of maximum amplitudes between noise and signal of Fig. 6b and c are 0.75 and 1.5, respectively. Figure 6e and f demonstrates the estimated local slopes maps using our synthetically-trained network. It can be observed that, when the input data are clean (Fig. 6a) or moderately corrupted by noise (Fig. 6b) the network can captures information from the seismic image and predicts the local slopes precisely. Although the performance goes downhill as the

increase of noise (Fig. 6c), the local slope map (Fig. 6f) shows the slope variation tendency of seismic events and the main structure. A by-product i.e., the signal recognition, of each local slope estimations is presented in Fig. 6g–i, where the blue and white areas indicate the spatial–temporal locations of signal and non-signal components of the seismic data. It can be observed that although not very perfect for the last data, all the recognition results graphs the shape of seismic structure.

We need to mention here that the proposed technique is a patch-based network. In other words, for each pixel a neighboring section will be involved in calculation. Therefore, it takes a large amount of time in the network training process. For example, it takes approximate 2 hours for the first ten thousand iterations in the training process. Compared with training process, the prediction step is much faster. For the synthetic data, it takes about 40 s to get the predication result.

In order to compare the performances of the proposed and traditional methods, we also apply the plane-wave

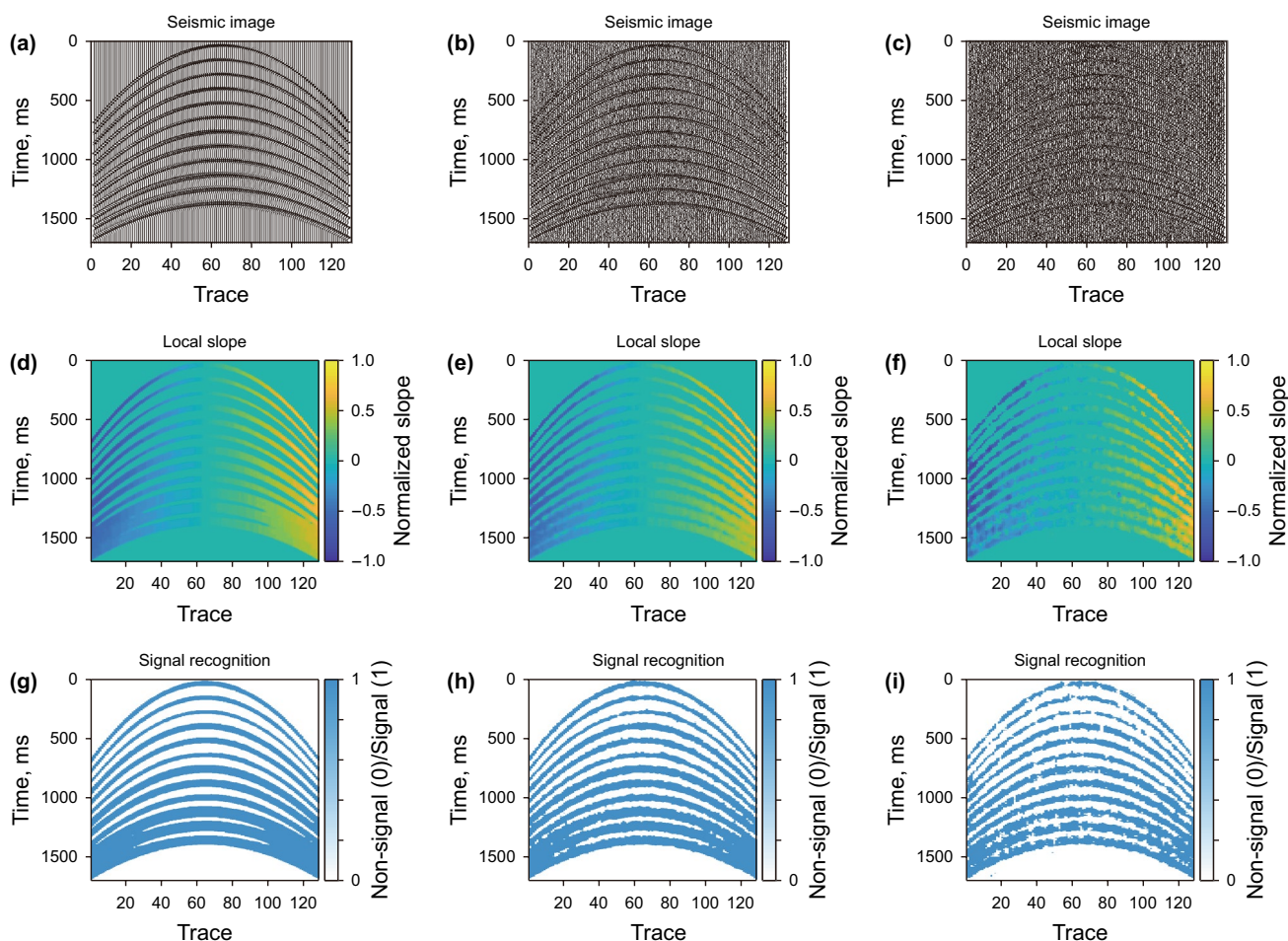


Fig. 6 Synthetic data example. **a–c** Three synthetic seismic images with different levels of added Gaussian noise. **d–f** The estimated local slope maps from Fig. 6a, b and c. **g–i** The corresponding signal recognition results

destruction method (Claerbout 1992) to estimate the local slope map of the data. The plane-wave destructor originates from a local plane-wave model for characterizing seismic data (Fomel 2002). It is constructed as finite-difference stencils for the plane-wave differential equation and acts as a spatial–temporal analog of the spatial-frequency prediction-error filters. Figure 7 presents three groups of estimated local slope maps of the three synthetic seismic images (Fig. 6a–c) using plane-wave destructor with different local windows. From the top to bottom, the sizes of processing windows are 7×3 , 25×5 and 73×7 , respectively. In overall, the results are good. As the size increases, the local slope map become smoother but is with lower resolution to distinguish those events. The small size of windows also depresses the performance of plane-wave destructor in low SNR case (Fig. 7c). We can also observe that plane-wave destructor has difficulty in estimating steep slopes as indicated by the red arrows. Comparing with the deep neural network, the plane-wave destruction method cannot directly obtain the spatial–temporal locations of signal like our trained network does. Figure 8

demonstrates the statistical analysis of the performance of local slope estimation. Because the events are curve continuously, the reference local slope is determined as the first order central difference quotient of the corresponding reflection coefficients. Although not very accurate, it can statistically evaluate the performance to some extent. Figure 8a–l are the statistical analysis of the three synthetic seismic data sets, respectively. The first column corresponds to the proposed deep-learning method, and the second to the last columns correspond to three plane-wave destructors. It can be observed that the deep-learning method obtains the smallest error.

3.2 Field example

To demonstrate how the neural network works in practice, we apply it to two real-reflection seismic data sets. The first real-data set is post-stack and shown in Fig. 9a. This data set is with a relatively high SNR and complex structure, e.g., a steep zone and a fault zone as highlighted by the green and

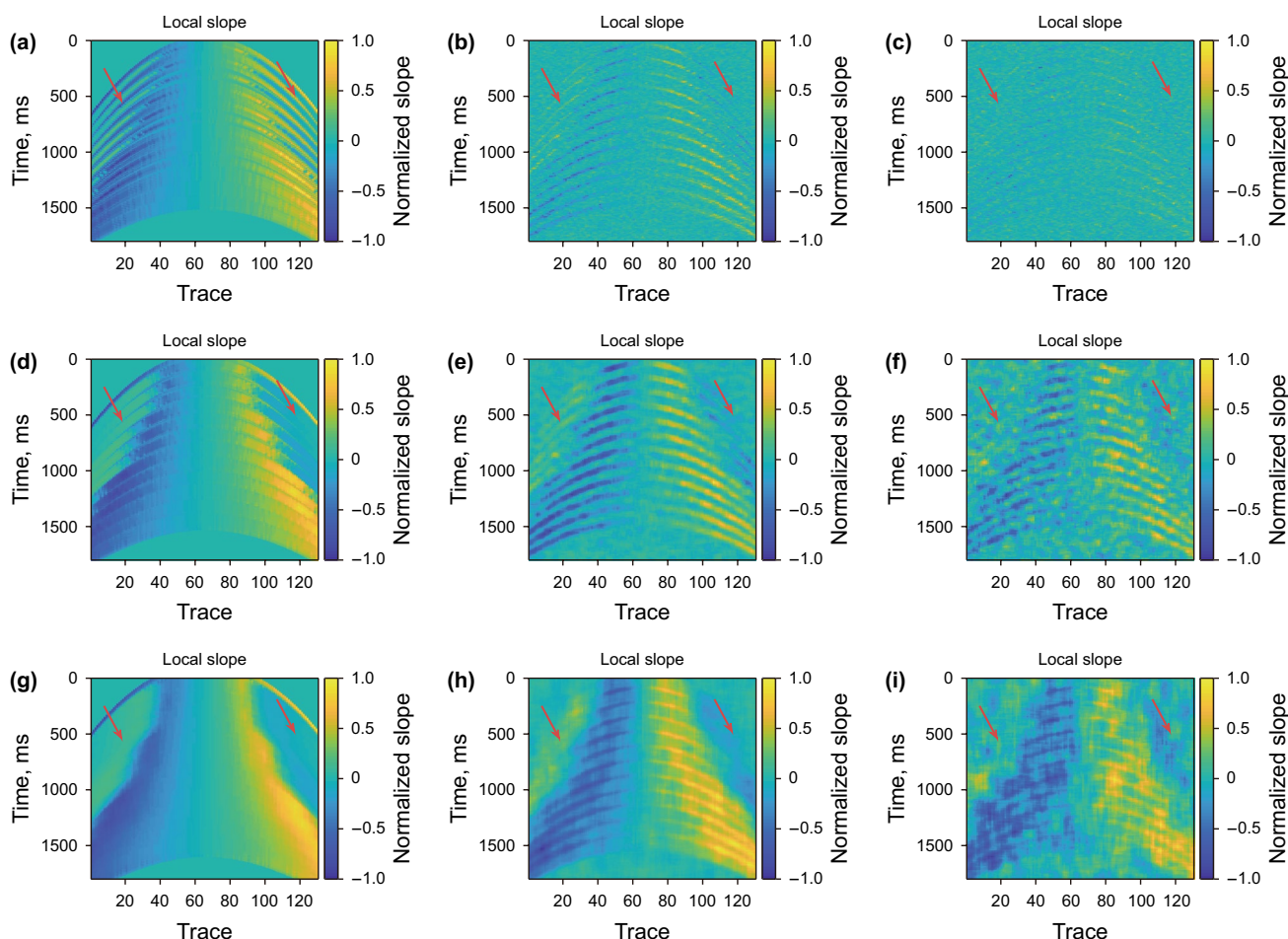


Fig. 7 Local slope estimations of Fig. 6a–c by plane-wave destructor with different sizes of local windows. **a–c** 7×3 windows. **d–f** 25×5 windows. **g–i** 73×7 windows

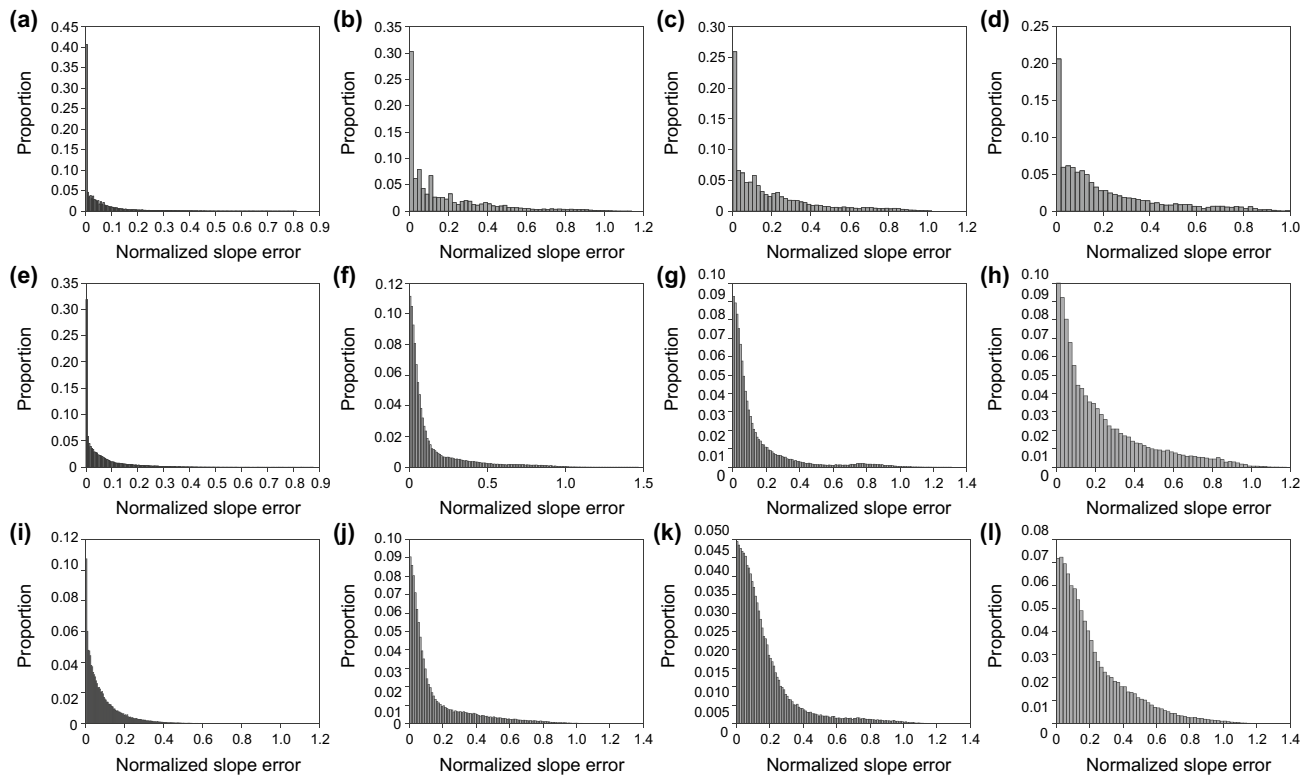


Fig. 8 A statistical analysis of the performance of local slope estimation. From the top to bottom: the statistical analysis of the three synthetic seismic data sets, respectively. From the left to right: the statistical analysis corresponds to the proposed destructors

blue frame boxes, respectively. To obtain the dip information of seismic events, we apply the trained neural network to calculate the local slope map (Fig. 9b). It can be seen that the neural network obtains a satisfactory performance and the change in color suggests the dip variation. Figure 9c–f presents the magnified sections marked by the green and blue frame boxes in Fig. 9a, where a better view is held to observe the performance of slope estimation in detail. In Fig. 9c, e the local slopes are plotted as the red bars to give us a more intuitive presentation. Figure 9c, d correspond to the blue frame box, which show us the details of the fault zone. We can observe that the color conflicts highlight the location of the faults. The data and local slope information corresponding the green frame box are shown in Fig. 9e, f. It can be seen that the almost all the red bars follow the direction of both steep and complanate seismic events spreading indicating a successful performance. Figure 10 demonstrates a comparison of the local slope estimation by the plane-wave destructor technique. We observe that although most slope information is picked accurately, there some disorganized and very likely false picked slopes as indicated by the red ellipse in Fig. 10a and the blue ellipses in Fig. 10b, d.

The second example explores the performance of the synthetically-trained network on a pre-stack data set. The original data set is shown in Fig. 11a. Similarly, we apply

the trained neural network to estimate the local slope map and recognize the position of signals. Figure 11b shows the local slope estimation, which presents a distinct color piece change from blue to green and yellow (from left to right) in the shallow (0–2000 ms), but shows a somehow disorganized image in the deep (2000–4000 ms) due to the poor SNR. The seismic data and corresponding slope information in the blue and green frame boxes in Fig. 11a are presented in Fig. 11c–f. Figure 12 presents the local slope estimation by the plane-wave destructor technique where we can find some inaccurate slopes by visual inspection as highlighted by the blue ellipses in Fig. 12b, d. The slope estimation by the synthetically-trained network can be further used to facilitate the subsequent seismic data processing and underground structure analysis/interpretation.

4 Conclusions

We have designed a deep learning-based neural network to automatically estimate the local slope and spatial–temporal position of seismic events. In the designed network, three convolution and three fully connected layers serve as the feature extractor and classifier, respectively. Our neural network is trained by using only synthetic seismic data

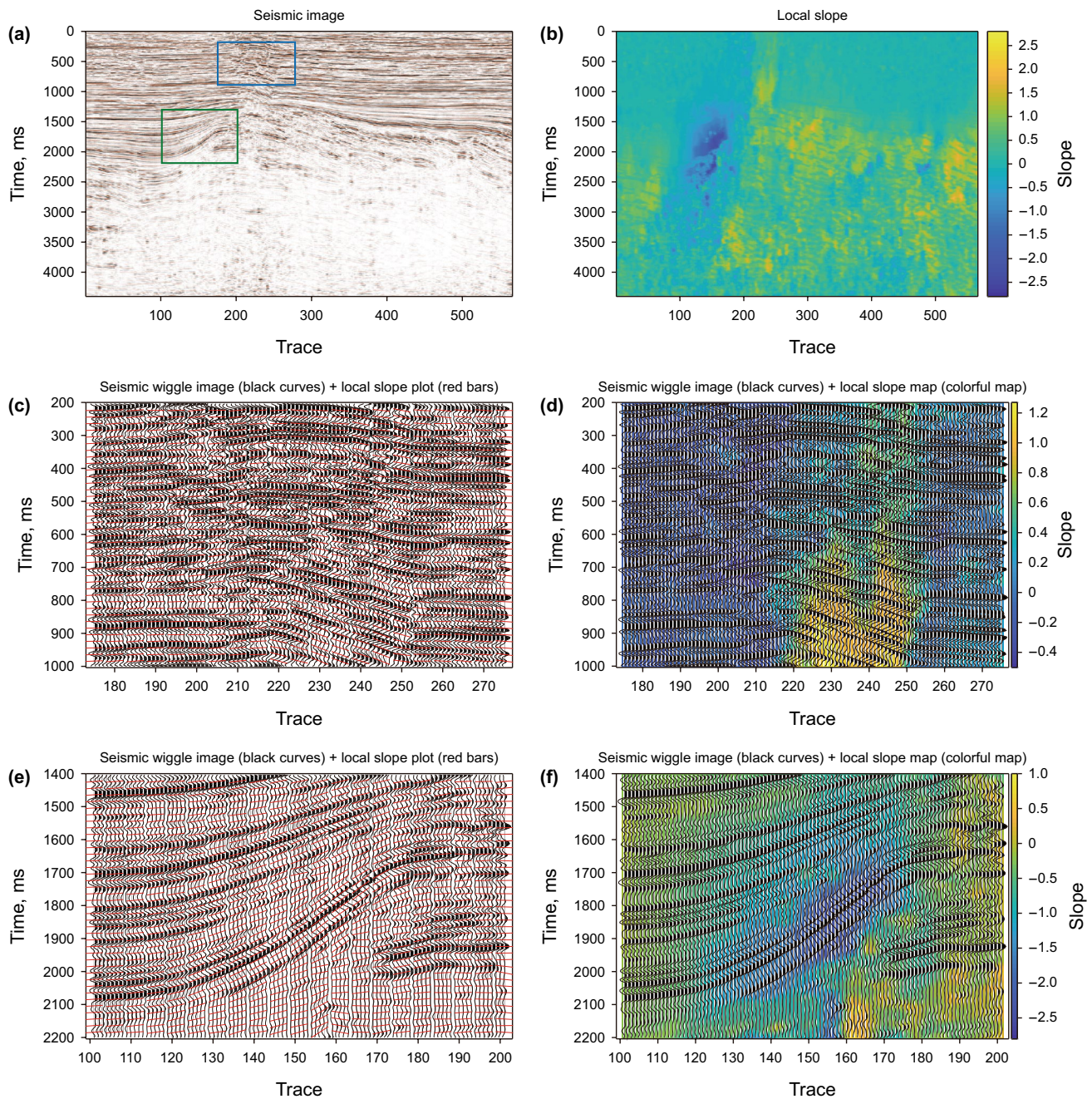


Fig. 9 The field post-stack data example. **a** Input data. **b** Local slope map obtained by the learned network. **c** and **d** Magnified sections correspond to the blue frame box. **e** and **f** Magnified sections correspond to the green frame box

sets which greatly simplify the process of training data collection because no manual labeling is needed. The application of the trained model to multiple synthetic and field seismic data sets recorded at varied surveys demonstrate an encouraging performance. The estimated local slopes and

spatial–temporal position can be further utilized for many subsequent procedures such as signal detection, normal move-out correction, time-domain imaging and structural interpretation. Although this technique is currently tested on 2D seismic data but can be easily extended to 3D cases.

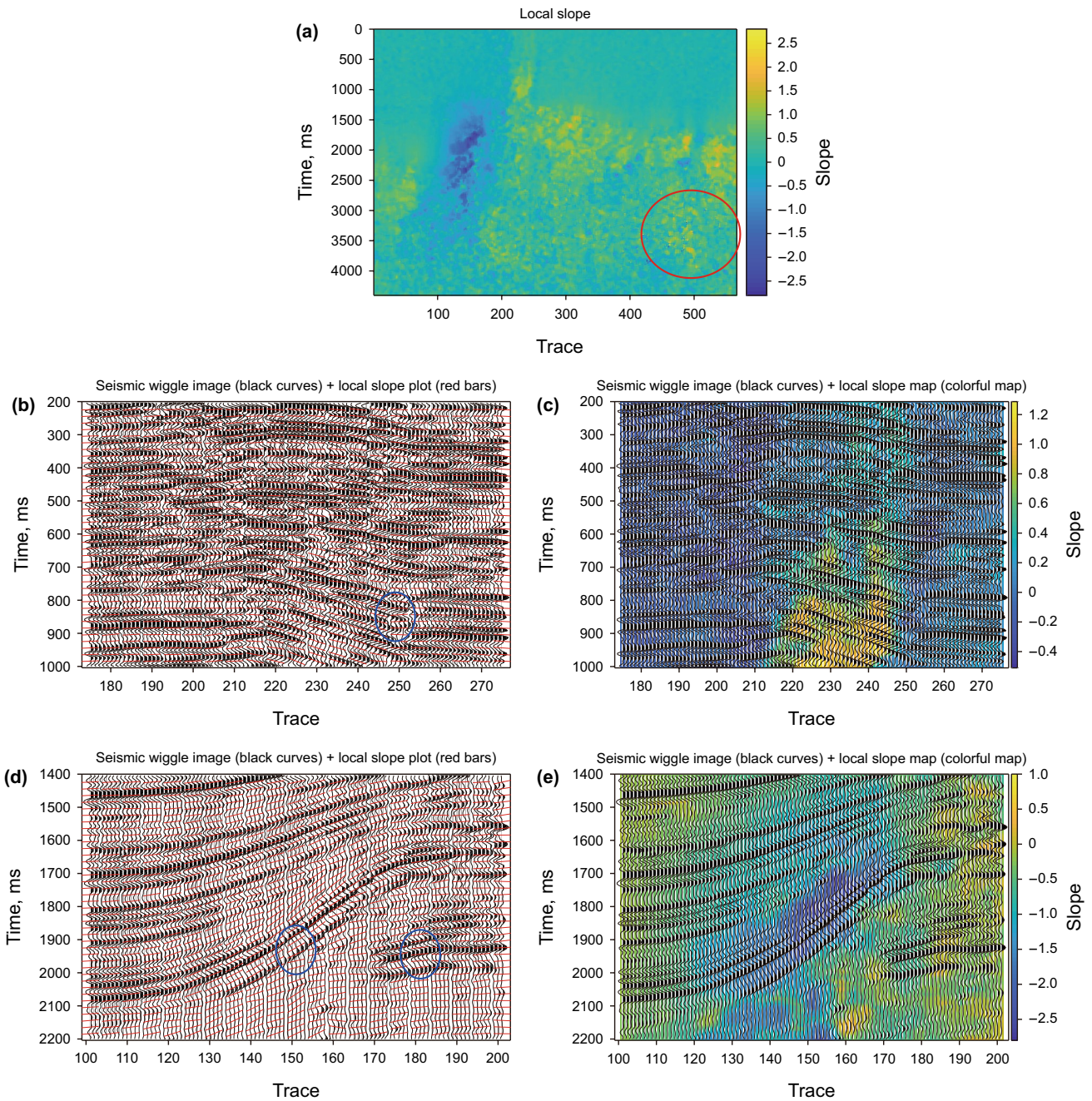


Fig. 10 The field post-stack data example. **a** Local slope map obtained by the plane-wave destructor. **b** and **c** Magnified sections correspond to the blue frame box in Fig. 9a. **d** and **e** Magnified sections correspond to the green frame box in Figure

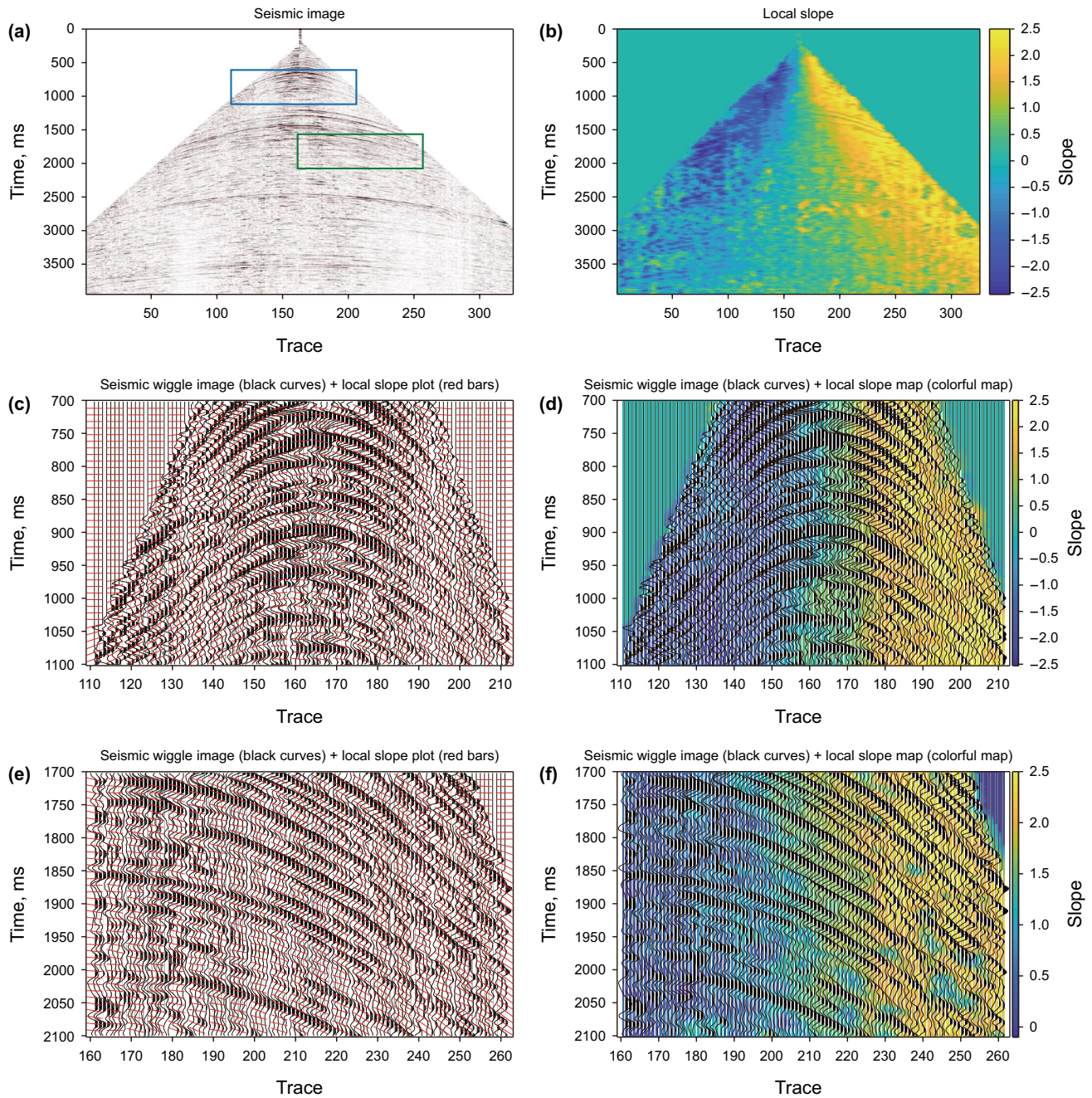


Fig. 11 The field pre-stack data example. **a** Input data. **b** Local slope map obtained by the learned network. **c** and **d** Magnified sections correspond to the blue frame box. **e** and **f** Magnified sections correspond to the green frame box

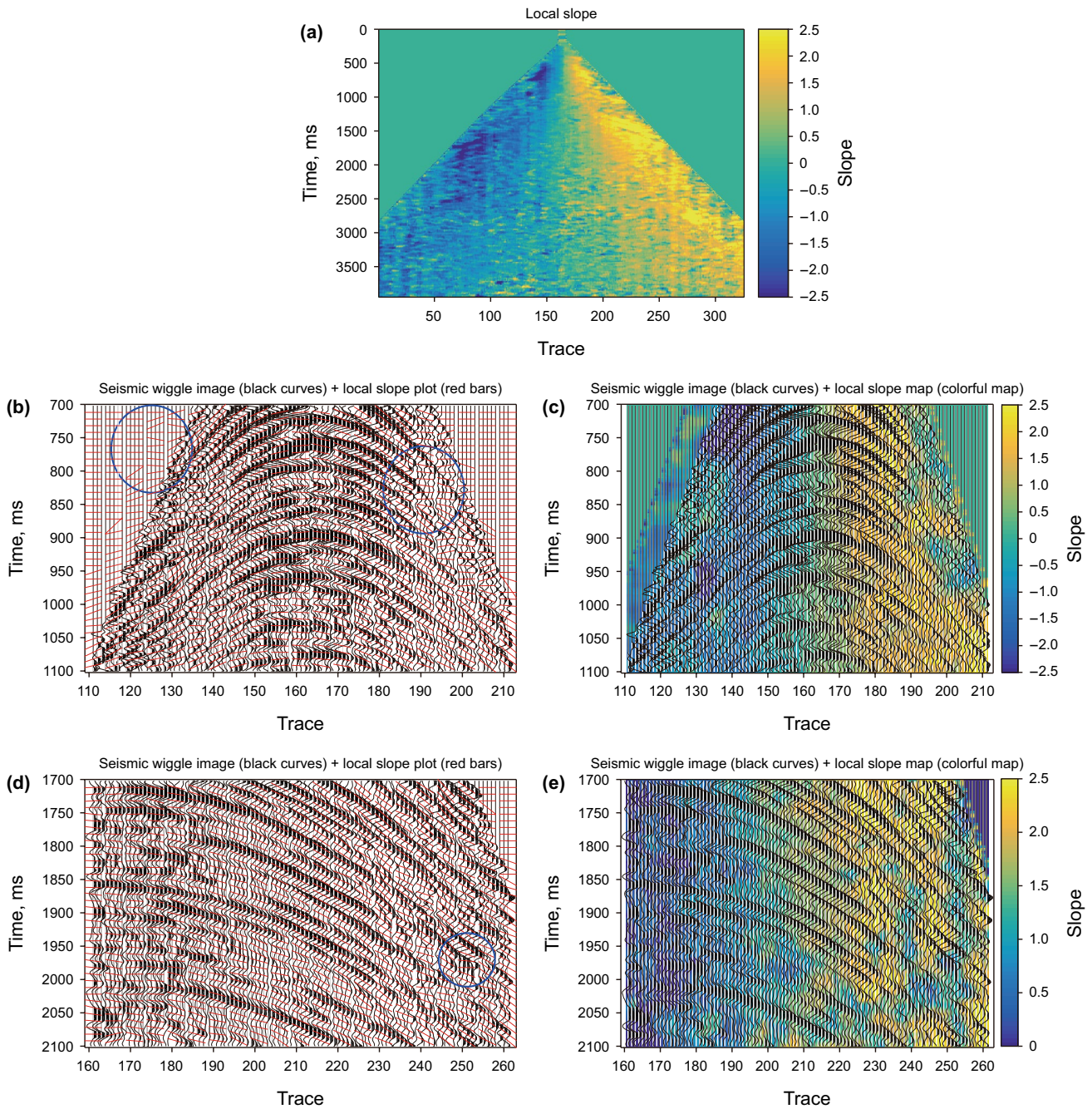


Fig. 12 The field pre-stack data example. **a** Local slope map obtained by the plane-wave destructor. **b** and **c** Magnified sections correspond to the blue frame box in Fig. 11a. **d** and **e** Magnified sections correspond to the green frame box in Fig. 11a

Acknowledgements This work is supported by the Science Foundation of China University of Petroleum, Beijing under Grant No.: 2462018YJRC020 and 2462020YXZZ006, the National Natural Science Foundation of China under grant no.: 41904098, the Young Elite Scientists Sponsorship Program by CAST (YESS) under Grant No.: 2018QNRC001. The work of Jianping Liao is partly supported by the National Natural Science Foundation of China under Grant No.: 41874156 and 42074167. The authors appreciate Binbin Qi for constructive discussions.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not

permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Andersson F, Duchkov AA. Extended structure tensors for multiple directionality estimation. *Geophys Prospect*. 2013;61(6):1135–49. <https://doi.org/10.1111/1365-2478.12067>.
- Araya-Polo M, Jennings J, Adler A, Dahlke T. Deep-learning tomography. *Lead Edge*. 2018;37(1):58–66. <https://doi.org/10.1190/tle37.010058.1>.
- Chan K, Lee TW, Sample PA, Goldbaum MH, Weinreb RN, Sejnowski TJ. Comparison of machine learning and traditional classifiers in glaucoma diagnosis. *IEEE Trans Biomed Eng*. 2002;49(9):963–74. <https://doi.org/10.1109/TBME.2002.802012>.
- Chen YK. Fast waveform detection for microseismic imaging using unsupervised machine learning. *Geophys J Int*. 2018;215:1185–99. <https://doi.org/10.1093/gji/ggy348>.
- Chen YK, Ma JW, Fomel S. Double-sparsity dictionary for seismic noise attenuation. *Geophysics*. 2016;81(2):V193–206. <https://doi.org/10.1190/GEO2014-0525.1>.
- Chen XQ, Wang RQ, Huang WL, Jiang YY, Chen Y. Clustering-based stress inversion from focal mechanisms in microseismic monitoring of hydrofracturing. *Geophys J Int*. 2018;3:1887–99. <https://doi.org/10.1093/gji/ggy388>.
- Chen YK, Zhang M, Bai M, Chen W. Improving the signal-to-noise ratio of seismological datasets by unsupervised machine learning. *Ssmol Res Lett*. 2019. <https://doi.org/10.1785/0220190028>.
- Chen W, Yang LQ, Zhai B, Zhang M, Chen YK. Deep learning reservoir porosity prediction based on multilayer long short-term memory network. *Geophysics*. 2020;85(4):1–69. <https://doi.org/10.1190/geo2019-0261.1>.
- Claerhout JF. Earth soundings analysis: processing versus inversion. London: Blackwell Scientific Publications; 1992.
- Fomel S. Applications of plane-wave destruction filters. *Geophysics*. 2002;67(6):1946–60. <https://doi.org/10.1190/1.1527095>.
- Fomel S. Velocity-independent time-domain seismic imaging using local event slopes. *Geophysics*. 2007;72:S139–47.
- Fomel S. Predictive painting of 3D seismic volumes. *Geophysics*. 2010;75(4):A25–30. <https://doi.org/10.1190/1.3453847>.
- Gan SW, Wang SD, Chen YK, Chen XQ, Huang WL, Chen HM. Compressive sensing for seismic data reconstruction via fast projection onto convex sets based on seislet transform. *J Appl Geophys*. 2016;130:194–208. <https://doi.org/10.1016/j.jappgeo.2016.03.033>.
- Hale D. Structure-oriented smoothing and semblance. *CWP Rep*. 2009;635:261–70.
- Huang WL. Seismic signal recognition by unsupervised machine learning. *Geophys J Int*. 2019;219:1163–80. <https://doi.org/10.1093/gji/ggz366>.
- Huang WL, Liu JX. Robust seismic image interpolation with mathematical morphological constraint. *IEEE Trans Image Process*. 2020;29:819–29. <https://doi.org/10.1109/TIP.2019.2936744>.
- Huang WL, Wang RQ, Yuan YM, Gan SW, Chen YK. Signal extraction using randomized-order multichannel singular spectrum analysis. *Geophysics*. 2017a;82(2):V59–74. <https://doi.org/10.1190/geo2015-0708.1>.
- Huang WL, Wang RQ, Zhang D, Zhou YX, Yang WC, Chen YK. Mathematical morphological filtering for linear noise attenuation of seismic data. *Geophysics*. 2017b;82(6):V369–84. <https://doi.org/10.1190/geo2016-0580.1>.
- Huang WL, Wang RQ, Gong XB, Chen YK. Iterative deblending of simultaneous-source seismic data with structuring median constraint. *IEEE Geosci Remote Sens Lett*. 2018;15(1):58–62. <https://doi.org/10.1109/LGRS.2017.2772857>.
- Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. 2015. arXiv preprint arXiv:1502.03167.
- Jia J, Wang FY, Wu QJ. Review of the application of machine learning in seismic detection and phase identification. *China Earthq Eng J*. 2019;41(6):1419–25. <https://doi.org/10.3969/j.issn.1000-0844.2019.06.1419>. (in Chinese).
- Jiang YR, Ning JY. Automatic detection of seismic body-wave phases and determination of their arrival times based on support vector machine. *Chin J Geophys*. 2019;62(1):361–73. <https://doi.org/10.6038/cjg2019M0422>. (in Chinese).
- Lambaré G, Alerini M, Baina R, Podvin P. Stereotomography: a semi-automatic approach for velocity macromodel estimation. *Geophys Prospect*. 2004. <https://doi.org/10.1111/j.1365-2478.2004.00440.x>.
- Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436–44. <https://doi.org/10.1038/nature14539>.
- Lewis W, Vigh D. Deep learning prior models from seismic images for full-waveform inversion. *SEG Technical Program Expanded Abstracts 2017*. 2017;1512–1517. <https://doi.org/10.1190/segam2017-17627643.1>.
- Li J, Wang XM, Zhang YH, Wang WD, Shang J, Gai L. Research on the seismic phase picking method based on the deep convolution neural network. *Chin J Geophys*. 2020;63(4):1591–606. <https://doi.org/10.6038/cjg2020N0057>. (in Chinese).
- Liu Y, Fomel S, Liu GC. Nonlinear structure-enhancing filtering using plane-wave prediction. *Geophys Prospect*. 2010;58:415–27.
- Liu Y, Fomel S, Liu C. Signal and noise separation in prestack seismic data using velocity-dependent seislet transform. *Geophysics*. 2015;80:WD117–28. <https://doi.org/10.1190/geo2014-0234.1>.
- Liu D, Huang J, Wang Z. Convolution-based multi-scale envelope inversion. *Pet Sci*. 2020;2:352–362.
- Liu SY, Zhang YZ, Li C, Sun WY. Automatic estimation of traveltimes parameters in vti media using similarity-weighted clustering. *Pet Sci*. 2020. <https://doi.org/10.1007/s12182-019-00423-y>.
- Mousavi SM, Beroza GC. A machine-learning approach for earthquake magnitude estimation. *Geophys Res Lett*. 2019. <https://doi.org/10.1029/2019GL085976>.
- Mousavi SM, Zhu WQ, Sheng YX, Beroza GC. Cred: a deep residual network of convolutional and recurrent units for earthquake signal detection. *Sci Rep*. 2019;9:1–4. <https://doi.org/10.1038/s41598-019-45748-1>.
- Mousavi SM, Sheng YX, Zhu WQ, Beroza GC. Stanford earthquake dataset (STEAD): a global data set of seismic signals for AI. *IEEE Access*. 2019;7:179464–179476. <https://doi.org/10.1109/ACCESS.2019.2947848>.
- Murphy KP. Machine learning: a probabilistic. Perspective. 2012. https://doi.org/10.1007/978-94-011-3532-0_2.
- Nair V, Hinton GE. Rectified linear units improve restricted boltzmann machines. In: proceedings of the 27th international conference on machine learning (ICML-10). 2010; 807–14.
- Ottolini R. Velocity independent seismic imaging. *Stanford Exploration Project Report*; 1983, 37.
- Ping L, Matt M, Seth B, Cody C, Yuan X. Using generative adversarial networks to improve deep-learning fault interpretation networks. *Lead Edge*. 2018;37(8):578–83. <https://doi.org/10.1190/tle3708057.8.1>.
- Riabinkin LA. Fundamentals of resolving power of controlled directional reception (CDR) of seismic waves: Slant-stack processing SEG. *Soc Expl Geophys* 1957;P36–P60.
- Rieber F. A new reflection system with controlled directional sensitivity. *Geophysics*. 1936;1(1):97–106. <https://doi.org/10.1190/1.1437082>.
- Ross ZE, Meier MA, Hauksson E, Heaton TH. Generalized seismic phase detection with deep learning. *Bull Seismol Soc Am*. 2018;108(5A):2894–901. <https://doi.org/10.1785/0120180080>.

- Saad OM, Chen YK. Deep denoising autoencoder for seismic random noise attenuation. *Geophysics*. 2020. <https://doi.org/10.1190/geo2019-0468.1>.
- Vafeiadis T, Diamantaras KI, Sarigiannidis G, Chatzisavvas KC. A comparison of machine learning techniques for customer churn prediction. *Simul Model Pract Theory*. 2015;55:1–9. <https://doi.org/10.1016/j.simpat.2015.03.003>.
- Valentine A, Kalnins L. An introduction to learning algorithms and potential applications in geomorphometry and earth surface dynamics. *Earth Surf Dyn*. 2016. <https://doi.org/10.5194/esurf-2016-6>.
- Waldeland AU, Charles JA, Leiv-J G, Anne HS. Convolutional neural networks for automated seismic interpretation. *Lead Edge*. 2018;37(7):529–37. <https://doi.org/10.1190/tle37070529.1>.
- Wu XM, Hale D. Horizon volumes with interpreted constraints. *Geophysics*. 2015;2(80):1942–2156. <https://doi.org/10.1190/geo2014-0212.1>.
- Wu XM, Liang LM, Shi YZ, Geng ZC, Fomel S. Multitask learning for local seismic image processing: fault detection, structure-oriented smoothing with edge-preserving, and seismic normal estimation by using a single convolutional neural network. *Geophys J Int*. 2019;219(3):2097–109. <https://doi.org/10.1093/gji/ggz418>.
- Wu BY, Qiu WR, Jia JX, Liu NH. Landslide susceptibility modeling using bagging-based positive-unlabeled learning. *IEEE Geosci Remote Sens Lett*. 2020. <https://doi.org/10.1109/LGRS.2020.2989497>.
- Wu BY, Meng DL, Wang LL, Liu NH, Wang Y. Seismic impedance inversion using fully convolutional residual network and transfer learning. *IEEE Geosci Remote Sens Lett*. 2020;99:1–5. <https://doi.org/10.1109/LGRS.2019.2963106>.
- Yao G, Wu D, Wang SX. A review on reflection-waveform inversion. *Pet Sci*. 2020;1:334–51. <https://doi.org/10.1007/s12182-020-00431-3>.
- Yu ZY, Chu RS, Sheng MH. Pick onset time of P and S phase by deep neural network. *Chin J Geophys*. 2018;61(12):4873–86. <https://doi.org/10.6038/cjg2018L0725>. (in Chinese).
- Zhang TF, Tilke P, Dupont E, Zhu LC, Liang L, Bailey W. Generating geologically realistic 3D reservoir facies models using deep learning of sedimentary architecture with generative adversarial networks. *Pet Sci*. 2019a;16(03):541–9.
- Zhang R, Huang JP, Zhuang SB, Li ZC. Target-oriented Gaussian beam migration using a modified ray tracing scheme. *Pet Sci*. 2019b;16(6):1301–19. <https://doi.org/10.1007/s12182-019-00388-y>.
- Zhang GY, Lin CY, Chen YK. Convolutional neural networks for micro-seismic waveform classification and arrival picking. *Geophysics*. 2020;85(4):1–75. <https://doi.org/10.1190/geo2019-0267.1>.
- Zhao M, Chen S, Fang LH, Yuen DA. Earthquake phase arrival auto-picking based on U-shaped convolutional neural network. *Chin J Geophys*. 2019;62(8):3034–42. <https://doi.org/10.6038/cjg2019M0945>. (in Chinese).
- Zhu WQ, Beroza GC. PhaseNet: a deep-neural-network-based seismic arrival-time picking method. *Geophys J Int*. 2018;1:1. <https://doi.org/10.1093/gji/ggy423>.
- Zhu WQ, Mousavi SM, Beroza GC. Seismic signal denoising and decomposition using deep neural networks. *IEEE Trans Geosci Remote Sens*. 2019;99:1–3. <https://doi.org/10.1109/TGRS.2019.292677>.