



# Research on Convolutional Neural Network in the Field of Oil and Gas Exploration

Ting Li<sup>1,2</sup>

<sup>1</sup>College of Earth Science and Engineering, Xi'an Shiyou University, Xi'an, China

<sup>2</sup>Shaanxi Key Laboratory of Petroleum Accumulation Geology, Xi'an Shiyou University, Xi'an, China

Email: 15382200216@163.com

**How to cite this paper:** Li, T. (2023) Research on Convolutional Neural Network in the Field of Oil and Gas Exploration. *Open Access Library Journal*, **10**: e9738. <https://doi.org/10.4236/oalib.1109738>

**Received:** January 6, 2023

**Accepted:** March 11, 2023

**Published:** March 14, 2023

Copyright © 2023 by author(s) and Open Access Library Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

In recent years, Convolutional Neural Network (CNN), as a deep learning algorithm, has been widely used in many fields such as computer vision and speech recognition. In the field of geological exploration, CNN has also made great progress. This paper reviews the latest research progress of CNN in the field of geological exploration, and focuses on the application of CNN in logging and seismic exploration. We first introduce the basic principles of CNN, then introduce the specific applications of CNN in logging and seismic exploration, and analyze the advantages and limitations of CNN with specific examples.

## Subject Areas

Artificial Intelligence, Petrochemistry

## Keywords

CNN, Artificial Intelligence, Deep Learning, Logging Interpretation

## 1. Introduction

Artificial intelligence is a branch of computer science, which studies the computer simulation of human thinking process and intelligent behavior. In 1950, Alan Matheson Turing first proposed the idea of artificial intelligence. The Dartmouth Conference in 1956 was considered to be the origin of artificial intelligence. The core idea of artificial intelligence is to enable machines to judge independently, completely or partially replace human decisions, and maximize the efficiency and benefits of solving problems [1]. Since 2006, due to the emergence of a large amount of data on the Internet, the breakthrough of artificial intelligence algorithms such as GPU computing power and deep learning neural net-

---

works, humans have seen the hope of machines catching up with humans. In 2017, after 40 days of learning, the AI go program developed by Google defeated Ke Jie, the current world's first go player, at 3:0, and AI achieved explosive development.

The large-scale application of artificial intelligence has brought disruptive technological changes to medical, transportation, communication, agriculture and other industries. The same is true for the oil and gas industry. At present, oil and gas exploration targets are becoming more and more refined, which puts forward higher requirements for the efficiency and accuracy of geophysical data processing and interpretation. With the increasing amount of data, it is impossible for humans to quickly identify the characteristics of data and obtain high-quality and high-resolution fine interpretation results. Therefore, the combination of logging, seismic interpretation and artificial intelligence has become an inevitable trend. Artificial intelligence is the most fundamental technical support in the "geological big data" revolution advocated in the field of oil and gas geology. There will be more and more discussions on the theme of machine learning or the combination of artificial intelligence and industry, setting off another round of AI boom [2]. Intelligent oil and gas exploration and development has become a hot spot and development trend in the industry. It is expected to greatly improve the efficiency and quality of oil and gas exploration and development, reduce costs and risks, and improve the exploration and development level of complex oil and gas reservoirs.

Although CNN has achieved some remarkable results in the logging field, there are still some problems to be solved. The number of samples is limited, the data set is not large enough, the universality of the trained model is not high enough, and the network needs continuous training to establish a new mapping. This paper extensively investigates the research status of artificial intelligence in the field of oil exploration and development. Combined with the actual needs of oil and gas exploration and development, this paper expounds on the research progress and application of convolutional neural network in the field of logging and seismic exploration.

## **2. Convolutional Neural Network**

### **2.1. CNN Principle**

Convolutional Neural Networks (CNN) is a kind of feedforward neural network with convolution calculation and deep structure. It is one of the representative algorithms of deep learning. Compared with BP neural network, CNN contains at least one convolutional layer for extracting features. The local connection of neurons and weight sharing between convolutional layers greatly reduce the amount of calculation and improve the training efficiency of neural network. Convolutional neural network is essentially an input-to-output mapping network. It can learn a large number of mapping relationships between input and output without any accurate mathematical expression between input and output.

As long as the convolutional neural network is trained with a known pattern, the neural network has the mapping ability between input and output.

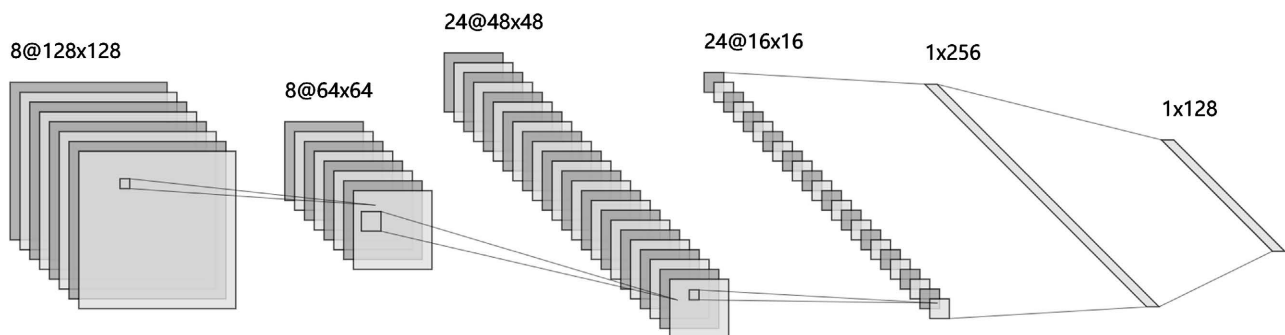
## 2.2. CNN Structure

The basic structure of CNN consists of input layer, convolution layer, pooling layer, fully connected layer and output layer. The network structure is shown in **Figure 1**. Convolution layer and pooling layer are the core modules to realize feature extraction. The network model adjusts the weight parameters in the network layer by layer by using the gradient descent method to minimize the loss function, and improves the accuracy of the network through frequent iterative training.

The convolutional layer consists of multiple feature maps. Each feature map consists of multiple neurons, and each neuron is connected to the local area of the previous feature map. The operation of the connection is carried out through the weight matrix, which is also called the convolution kernel [3]. The convolution kernel performs convolution operations by constantly moving on the feature map of the previous layer, thereby completing feature extraction. During the movement of the convolution kernel, its weight parameters are shared on the same layer. After convolution operation, the expression ability of the network is enhanced by the action of nonlinear activation function. The corresponding calculation formula of convolution layer (1) is as follows:

$$Y_{MN} = f\left(\sum_{j=0}^{q-1} \sum_{i=0}^{p-1} X_{m+i,n+j} * w_{ij}\right) \quad (1)$$

In Formula (1),  $X$  represents the input two-dimensional data;  $w$  represents the convolution filter kernel of size  $p * q$ ;  $b$  denotes bias;  $y$  represents by convolution operation. After convolution, bias term is usually added and nonlinear activation function is introduced. The activation function is usually a nonlinear mapping function, which is introduced to increase the expression ability of the entire network. The Relu function is currently the most popular activation function in neural networks. The pooling layer is usually connected to the local area of the previous feature map after the convolution layer. Its purpose is mainly to reduce the size of the feature map, reduce the number of parameters, improve the robustness of the model, and prevent overfitting [4]. The commonly used pooling



**Figure 1.** CNN structure.

operations are maximum pooling and average pooling. The maximum pooling is to select the maximum value in the pixel neighborhood as the output, and the average pooling is to use the average value of the adjacent area as the output. The pooling layer has no parameters such as weights, but only an operation rule. The pooling method selected in this paper is maximum pooling. The schematic diagram of maximum pooling is shown in **Figure 2**

The fully connected layer refers to the connection between each neuron in the output layer and all neurons in the input layer, which is generally located at the end of the whole convolutional neural network and plays the role of classifier. Generally, according to the task requirements, the appropriate number of neurons in the fully connected layer and the output layer is set to complete the corresponding classification or regression tasks. The calculation formula (2) is as follows:

$$X_j^i = f(W_{ij}^l X_i^{l-1} + b_j^l) \quad (2)$$

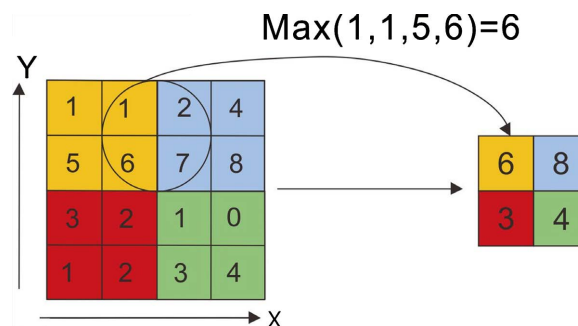
In the formula:  $X_i^{l-1}$  is the output of the  $i^{\text{th}}$  neuron in the  $l-1$  layer,  $X_j^l$  is the output of the  $j^{\text{th}}$  neuron in the  $l$  layer after the fully connected layer,  $W_{ij}^l$  is the weight matrix,  $b_j^l$  is the bias,  $f$  is the activation function. In order to avoid overfitting, the fully connected layer usually adopts Dropout. Dropout is a regularization technique that sets the value of Dropout so that some nodes in the network do not participate in the training process of CNN, thereby reducing the complexity of the model and improving the adaptability of the model.

### 3. Application of CNN in Seismic Interpretation and Processing

Throughout the application of artificial intelligence in the field of oil exploration and development, convolutional neural networks have made great progress in many aspects. Some research results have achieved phased results in actual production, but most of them are still being explored. The main research results and important progress are summarized as follows.

#### 3.1. First Picking

First break picking is a necessary preliminary work for estimating the location of the seismic world and other workflows such as tomography or moment bright



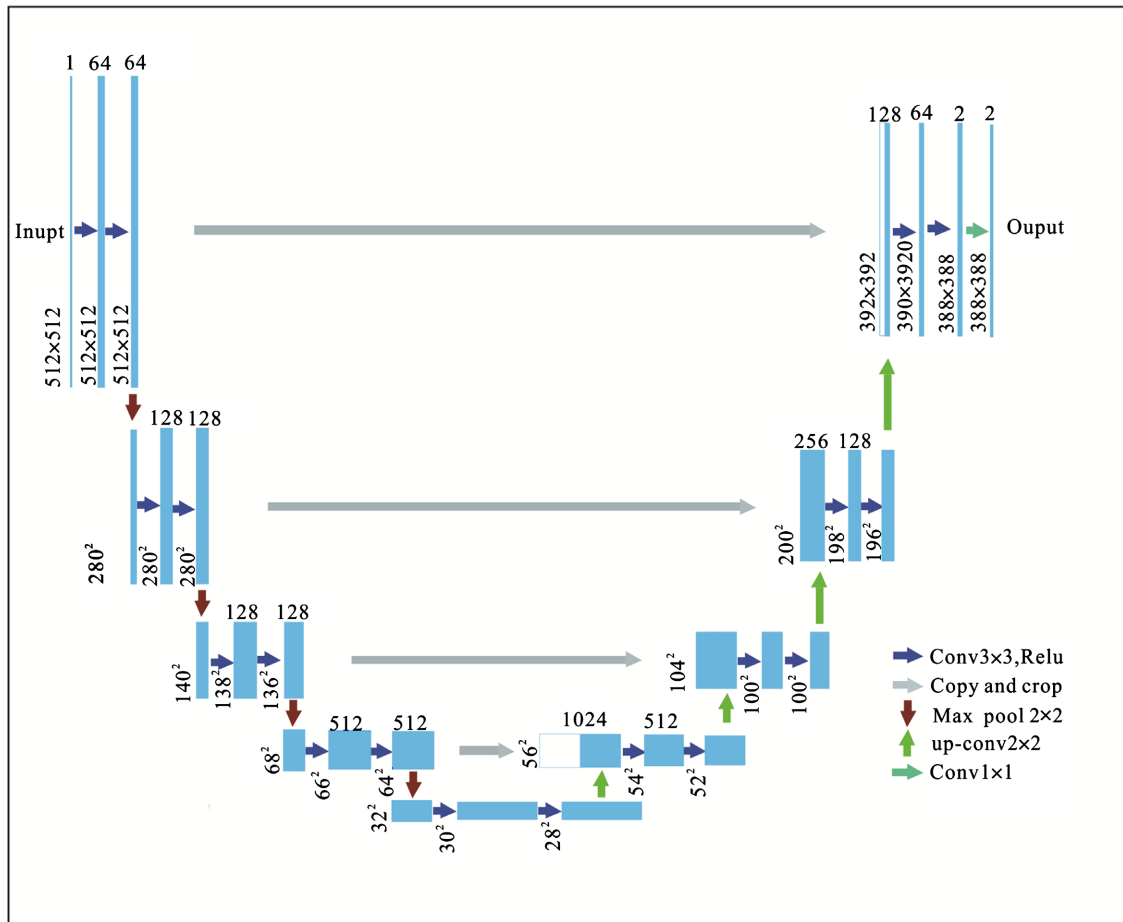
**Figure 2.** The schematic diagram of maximum pooling.

inversion. It is essentially a process of pattern recognition. Artificial first break takes a lot of manpower and time. With the massive generation and accumulation of seismic data, manual picking cannot be completed. Therefore, many scholars continue to explore the method of automatic first break picking.

For the first time, Murat successfully applied artificial neural networks to the extraction of first breaks in noisy background. Since then, the effectiveness of artificial neural networks and even convolutional networks in first breaks picking has been verified by multiple teams. The advantage of neural networks is that once the training is completed, the processing efficiency is greatly improved compared to manual processing [5]. McCormac *et al.* found that the picking efficiency of the trained network 3D dataset was 8 times higher than that of manual picking [6]. After these two scholars, the related research is mainly based on two aspects: improving the training and processing speed and picking accuracy of neural networks. In terms of training speed, the use of logical fuzzy network can ensure that the training speed is improved without losing accuracy. (Multi-layer perceptron) performs well, but RBF (radial basis) neural network has more advantages in training speed. In terms of accuracy, Yuan *et al.* first successfully applied artificial neural networks to the extraction of first-arrival waves in noisy backgrounds, and convolutional neural networks can be used for multi-channel data analysis [7]. Maity *et al.* designed a new hybrid automatic picking network structure and found that the neural network is more stable than other methods in the case of low SNR [8]. Yuan *et al.* pointed out that artificial neural networks rarely use the spatial coherence characteristics of waveforms, so they proposed to use CNN for first break picking, but did not compare the results of previous artificial neural networks. Extracting which seismic attributes are more helpful for first arrival picking is also one of the problems worth studying in the future.

### 3.2. Fault Automatic Recognition

Fault automatic recognition based on deep learning has gradually become a hot direction. Many scholars have used convolutional neural networks to train on synthetic seismic record data sets or actual seismic data sets to construct intelligent fault recognition models and automatically identify parameters such as the probability of fault existence and dip angle. As shown in **Figure 3**, Wu *et al.* developed a convolutional neural network model based on encoding and decoding. Using U-net network training and data sets, the intelligent fault recognition effect of image-to-image is truly realized. The model can simultaneously realize fault detection and slope estimation. In order to train the network, thousands of three-dimensional synthetic noise seismic images and corresponding fault images, clean seismic images and seismic normal vectors are automatically generated. Compared with other common methods, the fault identification based on U-net network can better describe the fault boundary than the same method, and the continuity of the fault is better. The fault detection and reflection slope calculation are obviously better than the traditional method [9].



**Figure 3.** U-net network structure.

### 3.3. Seismic Inversion

Seismic inversion is mainly to convert the conventional interface-type reflection profile into a rock-type logging profile, and the seismic data can be directly compared with the logging data. In recent years, the application of artificial intelligence technology in this field has made great progress, and the algorithms used are mainly CNN, RNN, DNN, and GAN.

Phan *et al.* combined the cascade method with the convolutional neural network to construct a deep learning model by minimizing an energy function similar to the least squares solution of the inverse problem, and then used the network to perform pre-stack seismic inversion. The network learns the nonlinear relationship between rock properties and seismic amplitude to predict impedance. The results show that the algorithm can capture all the features in the training data set, accurately reconstruct the input logging curve of the well point, and generate a geologically reasonable impedance profile [10].

## 4. The Application of CNN in Logging Field

### 4.1. Prediction of Logging Physical Parameters

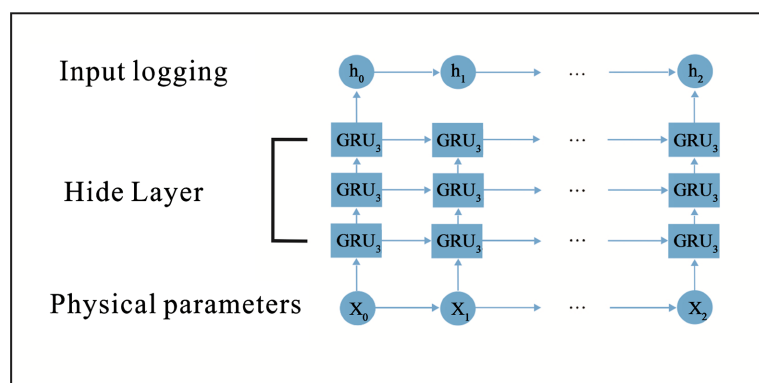
Determining the properties of oil and gas reservoirs is extremely important in

the exploration industry. Among them, porosity and permeability are key physical parameters. Traditional physical parameter prediction methods mainly include core experiments and logging data interpretation. With the development of deep learning theory, the prediction accuracy and recognition ability of deep models are continuously improved. Deep learning strategies are constantly applied to practical problems and breakthrough progress has been made. Wang *et al.* proposed a deep bidirectional recurrent neural network (DBRNN) model, which superimposes multiple bidirectional RNNs to effectively solve the spatial scale problem of porosity prediction and take into account the time scale effect. Compared with BRNN, DNN and other models, its accuracy and effectiveness are outstanding [11].

Zhang *et al.* proposed a multi-logging parameter combination prediction method based on GRU neural network. The structure diagram of the three-layer GRU neural network model is shown in **Figure 4**. Firstly, the logging parameters most related to porosity are selected by using the correlation measurement method based on coupla function. Then GRU neural network is used to identify the nonlinear mapping relationship between logging data and porosity parameters. The application results in an exploration area of Ordos Basin show that this method is superior to multiple regression analysis and recurrent neural network method, indicating that GRU neural network is more effective in predicting a series of reservoir parameters such as porosity. It can not only make full use of the response characteristics of different logging parameters to different formations, but also get rid of the limitations of traditional empirical formula linear prediction [12].

#### 4.2. Logging Lithofacies Category Prediction

With the increasing difficulty of oil and gas resource development, higher requirements are put forward for complex lithofacies prediction technology. The traditional logging interpretation scheme based on expert experience or manual intervention is difficult to apply to complex reservoir geological conditions and diversified logging data. In the actual logging, the cost of obtaining the actual lithofacies type is usually very high, and the related engineering is extremely



**Figure 4.** The structure diagram of the three-layer GRU neural network model.

cumbersome [13]. Since the deep learning model can efficiently and automatically extract features from massive data, and solve complex classification or prediction problems through layer-by-layer feature changes, thus effectively solving multi-dimensional nonlinear problems and big data problems, more and more deep learning technologies have begun to be applied to the field of lithofacies identification.

Gu *et al.* of the three models, Gu *et al.* proposed an improved probabilistic neural network based on continuous restricted Boltzmann machine and particle swarm optimization, namely CRBM-PSO-PNN. Experiments show that this method is effective for the prediction of complex lithology [14].

Supervised learning generally requires a large amount of labeled logging data as support. However, most geophysical problems have unbalanced data sets and scarce labels. Using semi-supervised or unsupervised learning algorithms to train classification models can greatly improve the accuracy of various types of predictions, thereby constructing more accurate lithofacies prediction models. For example, Chang *et al.* [15] introduced the unsupervised domain adaptive method into lithology identification under the premise of no lithology mark in new wells, developed a two-stream multi-layer neural network, and trained the network according to the maximum mean difference optimization method. This method was applied to Jiyang depression in Bohai Bay Basin, which effectively alleviated the performance degradation caused by the difference of logging data distribution.

Aiming at the problem of poor universality of the model, Xie *et al.* proposed an intelligent logging lithology identification framework combining outlier detection and extremely random tree based on a multi-class lithology classification model with outlier data sets. An unsupervised learning method is used to detect outliers in the data set, and then a very random tree classifier is used for coarse-to-fine inference. The lithofacies prediction accuracy of the model in the DGF and HGF regions is as high as 89.4% and 91.1%, respectively, which is significantly better than other classifiers, so it has the ability to identify sandstones with high accuracy [16].

### 4.3. Reconstituted Log

Curve reconstruction is to use deep learning, correlation analysis and other algorithms to find the correlation between logging curves, and to reconstruct the data of wrong, inappropriate and missing logging curves. The artificial intelligence algorithms used include neural networks, combinatorial learning algorithms, clustering algorithms, etc. (Yang *et al.*, 2019) selected four logging parameters of compensation density, acoustic time difference, natural gamma and shale content as independent variables of convolutional neural network (CNN) to reconstruct the change curve of dependent variables such as porosity [17]. (Zhang Dongxiao *et al.*, 2018) proposed a method based on recurrent neural network (RNN), namely long short-term memory neural network (LSTM), to reconstruct logging curves. After verification with real logging curves, it was



found that the accuracy was higher than that of traditional methods [18].

## 5. Conclusions and Prospect

This paper reviews the development of convolutional neural networks in the field of oil and gas fields, and focuses on the research progress of convolutional neural networks in logging processing and interpretation (such as lithology identification, physical parameter calculation, curve reconstruction, etc.), seismic processing and interpretation (such as first arrival wave picking, fault identification, etc.). The research results show that big data and artificial intelligence technology can provide efficient analysis methods for complex exploration, development and application requirements, and have great development potential. The integrated development of artificial intelligence and oil and gas industry will lead the disruptive innovation of oil and gas technology and promote the oil and gas industry from digitization to intelligence.

It should be pointed out that there are still three main problems in the application of artificial intelligence in oil and gas exploration: Firstly, the limited number of samples cannot accurately and comprehensively reflect the actual geological conditions; secondly, it focuses on the discussion from the perspective of artificial intelligence, and often ignores the preprocessing process of logging data such as missing value processing, outlier correction and standardization. When different distribution data appear, the network needs to be continuously trained to establish new mapping, which hinders the application of neural network. The robustness is poor, and learning methods such as deep reinforcement and adaptive neural network tuning are not introduced, so the machine cannot gradually improve the analysis and adaptability in the process of building the database. While designing the implementation method, researchers should take into account its economic applicability and scalability as much as possible.

## Conflicts of Interest

The author declares no conflicts of interest.

## References

- [1] Zhang, B. (2019) Artificial Intelligence Is Entering the Post-Deep Learning Era. *Chinese Journal of Intelligent Science and Technology*, 4-6. (in Chinese)
- [2] Li, N., Xu, B.S., Wu, H.L., Feng, Z., Li, Y.S., Wang, K.W. and Liu, P. (2021) Application Status and Prospect of Artificial Intelligence in Logging Formation Evaluation. *Acta Petrolei Sinica*, **42**, 508-522. (in Chinese)
- [3] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2017) Imagenet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, **60**, 84-90. <https://doi.org/10.1145/3065386>
- [4] Li, Z.C., Wang, W.B., Gao, M. and Sheng, L. (2022) Microseismic First Break Picking Based on Convolutional Neural Network. *Progress in Geophysics*, **37**, 1060-1069. (in Chinese)
- [5] Murat, M.E. and Rudman, A.J. (1992) Automated First Arrival Picking: A Neural

- Network Approach. *Geophysical Prospecting*, **40**, 587-604. <https://doi.org/10.1111/j.1365-2478.1992.tb00543.x>
- [6] McCormack, M.D., Zauha, D.E. and Dushek, D.W. (1993) First-Break Refraction Event Picking and Seismic Data Trace Editing Using Neural Networks. *Geophysics*, **58**, 67-78. <https://doi.org/10.1190/1.1443352>
- [7] Yuan, S., Liu, J., Wang, S., *et al.* (2018) Seismic Waveform Classification and First-Break Picking Using Convolution Neural Networks. *IEEE Geoscience and Remote Sensing Letters*, **15**, 272-276. <https://doi.org/10.1109/LGRS.2017.2785834>
- [8] Maity, D., Aminzadeh, F. and Karrenbach, M. (2014) Novel Hybrid Artificial Neural Network Based Autopicking Workflow for Passive Seismic Data. *Geophysical Prospecting*, **62**, 834-847. <https://doi.org/10.1111/1365-2478.12125>
- [9] Wu, X., Liang, L., Shi, Y., *et al.* (2019) Deep Learning for Local Seismic Image Processing: Fault Detection, Structure-Oriented Smoothing with Edge-Preserving, and Slope Estimation by Using a Single Convolutional Neural Network. In: *SEG Technical Program Expanded Abstracts 2019*, Society of Exploration Geophysicists, 2222-2226. <https://doi.org/10.1190/segam2019-3215251.1>
- [10] Phan, S. and Sen, M.K. (2019) Deep Learning with Cross-Shape Deep Boltzmann Machine for Pre-Stack Inversion Problem. In: *SEG Technical Program Expanded Abstracts 2019*, Society of Exploration Geophysicists, 2293-2297. <https://doi.org/10.1190/segam2019-3215696.1>
- [11] Wang, J., Cao, J.X. and Zhou, X. (2022) Reservoir Porosity Prediction Based on Deep Bidirectional Recurrent Neural Network. *Progress in Geophysics*, **37**, 267-274. (in Chinese)
- [12] Zhang, Z., Wang, Y. and Wang, P. (2021) On a Deep Learning method of Estimating Reservoir Porosity. *Mathematical Problems in Engineering*, **2021**, Article ID: 6641678. <https://doi.org/10.1155/2021/6641678>
- [13] Hu, B., Li, F., Cheng, G.J., Wang, L. and Tong, S.Q. (2020) Application of Artificial Neural Network Lithology Identification in Geological Modeling. *Petrochemical Industry Application*, **39**, 94-104. (in Chinese)
- [14] Gu, Y., Bao, Z., Song, X., *et al.* (2019) Complex Lithology Prediction Using Probabilistic Neural Network Improved by Continuous Restricted Boltzmann Machine and Particle Swarm Optimization. *Journal of Petroleum Science and Engineering*, **179**, 966-978. <https://doi.org/10.1016/j.petrol.2019.05.032>
- [15] Chang, J., Li, J., Kang, Y., *et al.* (2021) Unsupervised Domain Adaptation Using Maximum Mean Discrepancy Optimization for Lithology Identification. *Geophysics*, **86**, ID19-ID30. <https://doi.org/10.1190/geo2020-0391.1>
- [16] Xie, Y., Zhu, C., Hu, R., *et al.* (2021) A Coarse-To-Fine Approach for Intelligent Logging Lithology Identification with Extremely Randomized Trees. *Mathematical Geosciences*, **53**, 859-876. <https://doi.org/10.1007/s11004-020-09885-y>
- [17] Yang, L.Q., Chen, W. and Cha, B. (2019) Prediction and Application of Reservoir Porosity by Convolutional Neural Network. *Progress in Geophysics*, **34**, 1548-1555. (in Chinese)
- [18] Zhang, D.X., Chen, Y.T. and Meng, J. (2018) Synthetic Well Logs Generation via Recurrent Neural Networks. *Petroleum Exploration and Development*, **45**, 598-607. (in Chinese) [https://doi.org/10.1016/S1876-3804\(18\)30068-5](https://doi.org/10.1016/S1876-3804(18)30068-5)