

Recognition of Stratum Lithology of Seismic Facies Based on Deep Belief Network

Guohe Li^{1,2,3}, Yang Zheng^{1,2}, Ying Li^{1,2}, Weijiang Wu^{1,2,3}, Yunfeng Hong³ and Xiaoming Zhou³

¹College of Geophysics and Information Engineering, China University of Petroleum, Beijing, 102249, China

²Beijing Key Lab of Data Mining for Petroleum Data, Beijing, 102249, China

³PanPass Institute of Digital Identification Management and Internet of Things, Beijing, 100029, China

Abstract—The Deep Belief Network (DBN) is one of the major algorithms of deep learning. It simulates human brain to extract the features efficiently, so that the model has much strong learning ability. Because it is difficult to extract features from a variety of seismic data effectively, multiple sampling points of seismic data are used as inputs. Then we use DBN to extract the features from seismic data, which can be stacked by RBMs layer-by-layer. The model of lithological recognition can be constructed from previous step, further to recognize stratum lithology. By experiments and practical application, it is proved that partial strata information can be utilized effectively when multiple sampling points of seismic data are used as inputs. In this way, we can effectively recognize the stratum lithology based on DBN.

Keywords—Restricted Boltzmann Machine (RBM); Deep Belief Network (DBN); seismic facies; lithologic recognition

I. INTRODUCTION

Recently, the exploration and development of terrestrial oil and gas basins in China has matured, and the general trend of oil and gas exploration is changing from structural reservoirs to lithology reservoirs [1]. Therefore, effective recognition of lithology becomes the focus of applied research. Seismic inversion has always been an important research topic in geophysics [2], which is beneficial to remaining oil potential tapping. There is a complex nonlinear relationship between layer velocity and lithology, and the stratigraphic velocity of any kind of lithology is not unique. Therefore, the traditional inversion method of recognizing lithology based on velocity has certain limitations, and reservoir recognition is relatively difficult. Artificial intelligence technology has been applied in reservoir recognition. For example, Cui realized the lithology recognition based on the neural network [3]. Wang realized lithology recognition by random forests [4], etc. These intelligent technologies have achieved good results in lithology recognition, but there are some shortcomings in existing intelligent technologies. For example, the neural network with different initial weights is instable, so it is easy to fall into the local optimum in training, and random forest is easy to over fit with large noise. Due to the advantages of DBN, such as fast convergence and high stability, we propose lithology recognition method based on DBN.

II. DEEP BELIEF NETS

The Deep Belief Network (DBN) was proposed by Hinton et al. in 2006 [5, 6], which can be composed of unsupervised

Restricted Boltzmann Machines (RBMs) [7, 8] and a supervised classifier. During the unsupervised training procedure, we can train RBMs to extract features step-by-step, and obtain near-global optimal initial weights. After unsupervised training step, the DBN can be further trained in a supervised way to do classification.

A. Restricted Boltzmann Machines

A RBM is represented by an undirected graph structure consisting of visible units and hidden units, with connection between layers but no between within each layer (shown in Figure 1). When the number of hidden units is enough, a RBM can be fitted to any discrete distribution [9].

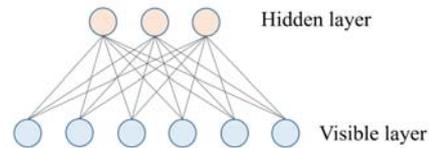


FIGURE 1. ILLUSTRATION OF AN RBM

A RBM is based on the energy model [7], which is expressed as follows:

$$E(v, h | \theta) = -\sum_{i=1}^n a_i v_i - \sum_{j=1}^m b_j h_j - \sum_{i=1}^n \sum_{j=1}^m v_i w_{ij} h_j \quad (1)$$

where θ is the set of parameter $\{w_{ij}, a_i, b_j\}$, w_{ij} is the connection weight between visible unit i and hidden unit j , a_i is the bias of visible unit i , b_j is the bias of hidden unit j , n is the number of visible units and m is the number of hidden units.

From Equation (1), we can get the joint probability distribution of visible units and hidden units

$$P(v, h | \theta) = \frac{\exp(-E(v, h | \theta))}{Z(\theta)} \quad (2)$$

where $Z(\theta)$ is a normalization constant, which is given by

$$Z(\theta) = \sum_{v, h} \exp(-E(v, h | \theta)) \quad (3)$$

where v is a vector of visible units, h is a vector of hidden units. To obtain $P(v | \theta)$ one can marginalize out h from

Equation 3.

$$P(v|\theta) = \frac{1}{Z(\theta)} \sum_h \exp(-E(v, h|\theta)) \quad (4)$$

We can get the optimal parameters by maximizing the log-likelihood function,

$$\theta^* = \arg \max_{\theta} L(\theta) = \arg \max_{\theta} \sum_{k=1}^K \log P(v^{(k)}|\theta) \quad (5)$$

where K is the number of samples, and θ^* are the optimal parameters.

The method for training RBMs include MCMC, Gibbs, Contrastive Divergence (CD) [10, 11], Persistent Contrastive Divergence (PCD), etc. Among all the methods, CD is the first practical method for training RBMs, an approximation of the log-likelihood gradient, is a successful update rule which can accelerate the training RBMs.

B. Deep Belief Nets

A DBN is a generative graphical model, which is composed of multiple layers [12], which can be viewed as a composition of RBMs and a classifier. Because Contrastive divergence (CD) algorithm in RBMs is efficient enough, we can train RBMs step-by-step with CD to form a DBN (shown in Figure II). Then the DBN can be trained in a supervised way to perform classification. For example, we can choose Logistic Regression as a classifier, and use the back-propagation algorithm to adjust network weights.

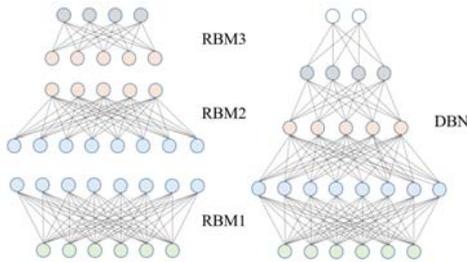


FIGURE II. ILLUSTRATION OF A DBN

We can define the loss function, negative log likelihood function, as follows:

$$J(\theta) = -\sum_{k=1}^K \sum_{j=1}^O I(y^k = j) \log P(y^k = j | x^k, W, b) \quad (6)$$

$$P(y^k = j | x^k, W, b) = \frac{\exp(W_j^T x^k + b_j)}{\sum_{i=1}^O \exp(W_i^T x^k + b_i)} \quad (7)$$

where K is the sample size, O is the number of output units, j is the number of output classes, W is a matrix of pairwise weights between hidden units and visible units, b is the bias for hidden units. Supervised learning in a DBN generally involves doing gradient descent in negative log likelihood

(NLL). The gradient of the NLL for the DBN is given by

$$\frac{\partial J(\theta)}{\partial W_j} = x^k \{P(y^k = j | x^k, W, b) - I(y^k = j)\} \quad (8)$$

$$\frac{\partial J(\theta)}{\partial b_j} = \{P(y^k = j | x^k, W, b) - I(y^k = j)\} \quad (9)$$

where $I(\cdot)$ is the indicator function, so that $I(a \text{ true statement})=1$, and $I(a \text{ false statement})=0$. On each iteration we would perform the update, as follows:

$$W_j = W_j - lr * \frac{\partial J(\theta)}{\partial W_j} \quad (10)$$

$$b_j = b_j - lr * \frac{\partial J(\theta)}{\partial b_j} \quad (11)$$

where lr is the learning rate.

By minimizing $J(\theta)$ with respect to θ , we will have a trained DBN.

III. LITHOLOGIC RECOGNITION BASED ON THE DBN

A. Seismic Data and Lithological Data

According to the kinematic and dynamic characteristics of seismic waves, seismic attributes include amplitude, frequency, phase, energy, waveform, wave impedance, wave velocity, correlation and ratio. Seismic attribute data are mathematically transformed data relating to geometric, kinematic, kinetic, or statistical characteristics of seismic waves [13, 14].

Lithology is the rock characteristics of formation, such as rock composition, particle size (conglomerate, sandstone, shale, etc.) and so on. Because oil and gas are stored in the stratum with porosity, lithology recognition has important significance for oil and gas exploration and development.

In this paper, we use DBN as a classification and recognition model to achieve the lithological classification based on seismic data.

B. Lithologic Recognition Oriented DBN Model

The stratum shows continuity in the longitudinal direction and any sampling points in the stratum is related to its adjacent points (shown in Figure III). We can use several sampling points of the seismic data to recognize the lithology, which makes use of the information of the stratum in some degree.

Use the multiple sampling points of seismic data as input, lithological data as output. Train a DBN based on RBMs step by step, the specific process is as follows:

Input: Logging lithological data, Borehole side seismic trace

Output: Lithology of the recognized section

Parameters: Neighbor size k

- (1) Choose the neighbor size k , and use the k sampling points of seismic data as the input X
- (2) Take logging lithology data as output Y
- (3) Construct the bottom RBM based on X , use the contrast divergence algorithm to train RBM.
- (4) Take the features extracted by the lower level RBM as the input data of the higher level RBM, and train the higher level RBM
- (5) Repeat (4)
- (6) Build softmax classifier at top-level
- (7) Use back-propagation of error derivatives fine-tune the DBN
- (8) Do iteration (7) until the model converges

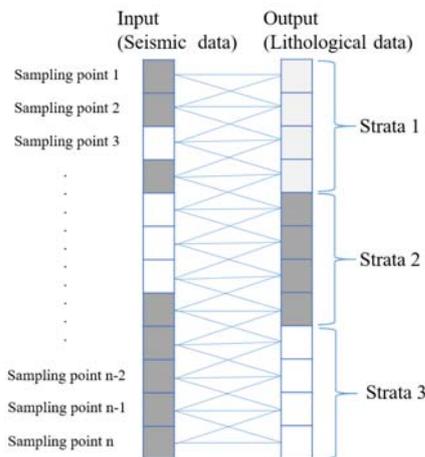


FIGURE III. LITHOLOGICAL RECOGNITION BASED ON MULTIPLE SAMPLING POINTS OF SEISMIC DATA

IV. EXPERIMENT AND ANALYSIS

The post-stack seismic data are from a certain area of north China Oilfield with a total of 1956 traces and 501 sampling points. The sampling rate is 2ms. The amplitude data of different frequencies reflect different stratigraphic information. The low frequency amplitude data reflect the stratum structure distribution better, and the high frequency amplitude data reflect the formation details better.

The region contains a total of 16 wells. As the vertical resolution of seismic data and the lithological data are different, the original lithology curve is coarsened. Because of the different porosities of the rock, the stratigraphic lithology is divided into shale, sandstone, and other lithology, which are denoted by 1, 2 and 3 respectively, as lithology recognition output data.

The experiment runs on Dell workstations, Xeon E5-1607 3.0GHz processor, 8GB DDR3 memory.

A. Model Experiments

Experiments were carried out by using four classification models: DBN, BP, RF and GBDT. The neighbor sampling points of the dataset are taken as 1, and the 5-fold cross validation test is used. The results are shown in Table I. The DBN model is implemented by C++, and different network structures are set. The pre-learning rate is set to 0.02, pre-learning iteration is 10 times, and learning rate is 0.1. The neural network is also implemented by C++, learning rate is set to 0.1. Call RF and GBDT from scikit-learn library and set the different number of trees for experiment.

TABLE I. ERROR OF LEARNING ALGORITHMS ON THE LITHOLOGY RECOGNITION TASK

Model	Network structure	Training error	Testing error
DBN	9×50×20×3	0.145	0.286
	9×100×20×3	0.119	0.282
	9×100×50×10×3	0.127	0.290
BP	9×50×20×3	0.149	0.310
	9×100×20×3	0.112	0.268
	9×100×50×10×3	0.163	0.301
Model	Number of trees	Train error	Test error
RF	50	0.139	0.521
	100	0.124	0.538
	200	0.119	0.535
GBDT	50	0.228	0.518
	100	0.108	0.522
	200	0.022	0.532

The training error mainly expresses the credibility of the model, while the testing error mainly expresses the availability of the model (generalization ability). The smaller training error and the larger testing error reflect the model over-fitting, and the larger training error and the smaller test error reflect the model under-fitting. Both over-fitting and under-fitting affect the popularization value of the model. Both smaller training error and smaller testing error can show that the model has good generalization ability. According to the error analysis, we can know from Table I that the training error of RF and GBDT can be as low as 0.15, but the testing error is obviously too high. However, training error and testing error of DBN and BP are relatively small. The results show that the BP and DBN have good performance on lithology recognition in this area, and the generalization ability of DBN is relatively better than that of BP.

B. Sampling Number Experiment

In the seismic data, selecting different number of sampling points will lead to different effects of lithology recognition. Due to odd sampling points have the center, we select different odd sampling points of seismic data to do experiment. The DBN has the following architecture: $9 * k$ inputs (where k is the number of sampling points), 3 outputs, and the two hidden layers (selected layers sizes are 100 and 20 respectively). We use cross validation to assess our algorithms effectiveness, and the results are shown in Table II. From Table II, we can see that the model with multiple sampling points as inputs has smaller error than that with single sampling point. It indicating that there is local correlation in the seismic data, especially in the number of sampling points is 7, the training error and testing error are minimum.

TABLE II. ERROR OF DBNS WITH DIFFERENT SAMPLING NUMBER AS INPUT

Sampling Number	Training Error	Testing Error
1	0.119	0.282
3	0.030	0.137
5	0.029	0.127
7	0.028	0.122
9	0.046	0.126
11	0.085	0.177

The trained DBN is applied to the recognition of lithological profile. The results are shown in Figure IV.

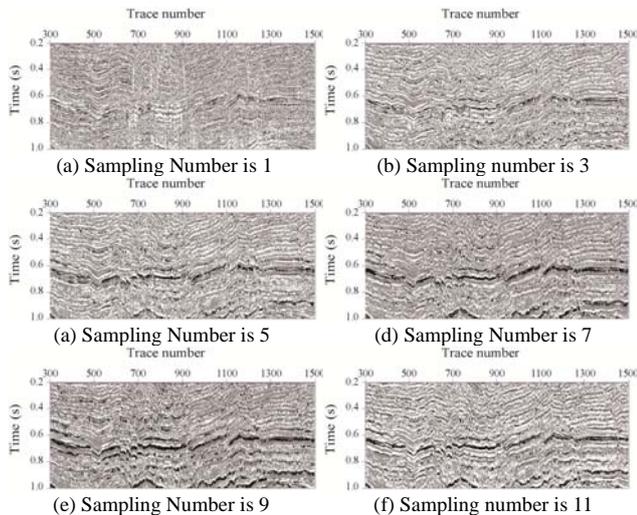


FIGURE IV. COMPARISON OF THE LITHOLOGICAL PROFILES WITH DIFFERENCE SAMPLING NUMBER OF SEISMIC DATA AS INPUT. (A-F) SHOWS THAT THE SECTIONS OF THE LITHOLOGICAL PROFILES, WHERE THE SAMPLING TIME IS FROM 0.2 TO 1.0 SECOND, THE TRACK NUMBER IS FROM 300 TO 1500.

By contrast, the DBN with different sampling number as inputs have different results. When the sampling number is 1, the recognized profile is more detailed, but the error is also larger. As the sampling number increase, the main horizons of the recognized lithological profiles become clear. But when the sampling number is too large, the lithological horizons become unclear, because there is local correlation rather than global correlation in the strata. From the Table II, we can see that the testing error is the smallest when the sampling number is 7, and the recognized lithological profile is also the best. In practical applications, if we have little well data, the experimental results can be used as a reference, cannot be used as the only standard to judge.

V. CONCLUSION

The intelligent technology can solve the nonlinear relationship between seismic data and lithological data, which has been applied in oil-gas exploration field. Machine learning is an effective means to establish nonlinear model, especially with the arrival of large data, deep learning has become a new research and application hotspot. The deep belief network (DBN) is a type of deep learning, which include two stages for training model. In the unsupervised learning step, RBMs are used to extract feature from training data. After this step, the

DBN can be trained in a supervised way to optimize parameters and perform classification. Because the DBN has the characteristics of feature automatic extracted, it has better generalization ability and greater application value. With different number of sampling points as input, the trained DBN can lead to different stratigraphic horizon information. We select suitable network structure and parameters to train a DBN using contrastive divergence and stochastic gradient descent algorithm. Then, the trained model is used to recognize the lithology of seismic profile. In practice, the data quantity of logging lithology is smaller, the choice of the profile need to be judged from the experimental error and the stratigraphic continuity. The next step, we will focus on the automatic determination of deep belief network structure for lithological recognition of seismic facies, and the study on the imbalance of the lithological data distribution.

ACKNOWLEDGEMENT

This work is partly supported by the Nature Science Foundation of China under Grant No. 60473125, the Innovation Foundation of CNPC under Grant No. 05E7013; National Key Project Foundation of Science under Grant No. G5800-08-ZS-WX, Science Foundation of China University of Petroleum-Beijing at Karamay.

REFERENCES

- [1] Tian Yu-kun, Zhou Hui, Yuan San-yi. Lithologic discrimination method based on Markov random-field. *Chinese Journal of Geophysics*, 2013, 56(4):1360-1368.
- [2] Zhang Ying-bo, Zhang Ji-dong. New technology and application of Lithologic Prediction. *China Offshore Oil and Gas(Geology)*, 2000, 14(2):138-144.
- [3] Cui Li-ling, Qi Long-yu, Liu Song-feng, Hu Bei-lai. Lithology Prediction with Multiple Seismic Attributes and Neural Network. *Acta Scientiarum Naturalium Universitatis Nankaiensis*, 2007, 40(2):99-104.
- [4] Wang Zhi-hong, Han Lu, Qi Lei. Random forest classification method in the application of reservoir lithology recognition. *Journal of Liaoning Technical University(Natural Science)*, 2015, 34(9):1083-1088.
- [5] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks. *Science*, 2006, 313(5786):504-507.
- [6] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets. *Neural Computation*, 2006, 18(7):1527-1554.
- [7] Smolensky P. Information processing in dynamical systems: Foundations of harmony theory//*Parallel Distributed Processing: explorations in the Microstructure of Cognition*. 1986:194-281.
- [8] Freund Y, Haussler D. Unsupervised learning of distributions of binary vectors using two layer networks. Santa Cruz: Computer Research Laboratory, University of California, 1994.
- [9] Roux N L, Bengio Y. Representational power of restricted Boltzmann machines and deep belief networks. *Neural Computation*, 2008, 20(6): 1631-1649.
- [10] Hinton G E. Training products of experts by minimizing contrastive divergence. *Neural Computation*, 2002, 14(8): 1771-1800.
- [11] Tieleman. Training restricted boltzmann machines using approximations to the likelihood gradient. *ICML 2008*.
- [12] Hinton G E. Deep belief networks. *Scholarpedia*. 2009, 4 (5): 5947.
- [13] Wu Han-bing. The extraction and application of local frequency and bandwidth in seismic data. Chang'an University, 2009.
- [14] Ni Feng-tian. Reservoir Parameter Predicted Method Research Based on Seismic Attribute Analysis. China University of Petroleum(EastChina), 2008