A New Neural Network Model for Rock Porosity Prediction

Youxiang Duan, Yu Li, Gentian Li, Qifeng Sun
College of Computer and Communication Engineering,
China University of Petroleum (East China)
Qingdao, China
Email: liyu34@vip.qq.com, LiGentian@163.com

***Abstract*—Artificial neural network has brought a new way for prediction of geological reservoir physical parameters (e.g. porosity, permeability and saturation). However, it becomes strong pertinence and bad universal in parameters prediction. According to the thought of committee machine, the paper presents a new neural network model, which is based on BP neural network, radial basis function (RBF) neural network and support vector regression (SVR) model. And then, a single layer perceptron (SLP) combines different individual neural network to adjust of network structure and reap beneficial advantages of all model. Eventually, a committee neural network (CNN) was constructed. It eliminated the defects of individual neural network in porosity prediction and improved the accuracy of the prediction. Three well logs are applied for experiment. One was used to establish the CNN model, and the other two were employed to assess the reliability of constructed CNN model. Results show that the CNN model performed better than individual neural network model.**

Keywords—committee machine; committee neural network; BP neural network; radial basis function; support vector regression; porosity prediction

I. INTRODUCTION

Artificial neural network (ANN) is an artificial intelligence technique that simulates the process of human brain, which is derived from the simulation of neural cells. It has characteristics of the massively parallel processing, strong fault tolerance ability and high adaptability. Moreover, it becomes competent for solving some complex nonlinear prediction problem. All of it has been widely used in prediction field. There are many artificial neural network models. They describe and simulate the different levels of biological nervous system from different items, for example, BP neural network, RBF neural network, Hopfiled neural network, generalized regression neural network, support vector regression and so on.

Geological reservoir physical parameters are crucial for building geological model. [1] They are often obtained by the logging data. However in many cases, drilling and coring were confined to certain zones, which results to insufficient analysis of physical data and cannot fully make a reservoir evaluation objectively. In conventional logging interpretation, reservoir physical parameters are calculated by empirical formula or simplified geological conditions. There are many limitations and defects for solving complicated geological problems. Artificial intelligence, especially the development of artificial neural network technology, has brought a new prediction way of geological reservoir physical parameters. At present, much of artificial neural network and improved neural network, which combine with GA algorithm or PSO algorithm specially, are applied extensively for reservoir parameters prediction in international researches. Chen Rong [2] carried on the reservoir porosity and permeability prediction using BP neural network in MATLAB platform. Hamidi H [3] predicted oil reservoir rock porosity based on BP neural network. All they have got expected results. Improved BP neural network combined with particle swarm optimization (PSO) algorithm was used for reservoir parameter dynamic prediction by Zhang Liumei [4] and Pan Shaowei [5]. In the other fields, Ma Lingling [6] predicted water quality based on BP network, and Li Yu [7] carried out complex water quality prediction with improved QGA-BP model. Practical application of improved BP model was better. Optimized spread RBF network was used to predict porosity using well logs by Baneshi M [8]. Wu Xiongjun [9] found the performance of trainable radial basis function was better than simple radial basis function for RBF network. Chen Jing [10] predicted the hydrocarbon reservoir parameter using a GA-RBF network, and Liu Xiaobo [11] predicted chlorophyll-A in Yuqiao reservoir using PCA-RBF model in the Haihe River Basin, China. Baneshi M [12] estimated of reservoir indexes by RBF, ANFIS, and MLP. Support vector regression is more recent prediction model. Na'imi S R [13], Al-Anazi A F (2010) [14], Mollajan A [15] and Al-Anazi A F (2012) [16] performed SVR model for reservoir prediction. And Improved support vector regression model combined with particle swarm optimization algorithm worked by Zhang Jun [17] as well.

However, individual neural network becomes strong pertinence and bad universal for prediction of geological reservoir physical parameters. So, the motivation of this study becomes the quest for better universal, more robust and higher prediction accuracy. Reference [18] has shown that, when individual artificial neural network has a certain prediction accuracy and diversity, by combining them, it could obtain a better model that has a higher and more robust prediction accuracy compared to individual neural network. This model is a parallel network model composed of a variety of artificial neural networks. The experiment shows it performed better than individual neural network model.

978-1-5090-5952-2/16 $31.00 © 2016 IEEE
DOI 10.1109/IIKI.2016.44

II. Committee Neural Network

The paper presents a new neural network model, which is based on BP neural network, radial basis function neural network and support vector regression model. And then, a single layer perceptron combines different individual neural network to adjust of network structure and reap beneficial advantages of all model. Eventually, a committee neural network (CNN) model was constructed in Fig.1. As for CNN model, two major steps are following. The first step, porosity was predicted from the principal components of logging data using BP network, RBF network and SVR model. They are called the basic neural network unit. The basic neural network unit is not only limited to these network, it can be a combination of any variety of neural network. Based neural network unit works alone. Each of it has the same input samples. In the second step, the output porosity of each basic unit is obtained through a certain combination or decision method according to the thought of committee machine. So, single layer perceptron is used to combine the results of each basic unit to one result. This methodology reaps the benefit of all work and enhances the accuracy of final prediction.

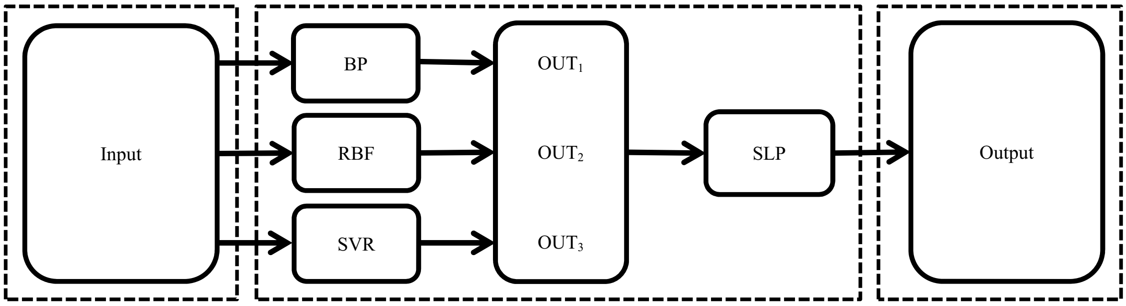

Fig. 1. A schematic diagram shows committee neural network (CNN) with BP neural network, radial basis function (RBF) neural network, support vector regression (SVR). A single layer perceptron (SPL) combines different individual neural network.

A. Neural Network

Neural network (NN) is a storage and analytical system that simulates the process of human brain, which is derived from the simulation of neural cells. Many neurons are arranged in parallel structure named layers, including input layer for gathering input, output layer for producing output, and hidden layer for extracting implicit dependency between input/output data. Neurons of each layer are connected to all the neurons of the next layer by connection weights. Extraction of these weights during the training process of neural network makes the neural network clever. So, after training, the neural network would be precisely able to predict rock porosity for each set of input data. Different archetypes of neural networks have own characteristics. Advantages and disadvantages of BP network, RBF network and SVR model is following in TABLE 1.

TABLE I. Advantages and disadvantages of BP network, RBF network and SVR model

Neural Network	Advantages	Disadvantages
BP	1. Good nonlinear mapping ability. 2. Good generalization ability. 3. Massively parallel processing. 4. Strong fault tolerance ability.	1. Get into local extremum easily. 2. No unified and complete theoretical guidance for the selection of network structure, especially the number hidden layer and neurons.
RBF	1. Strong nonlinear mapping ability (Mapping any complex nonlinear relationship). 2. Massively parallel processing. 3. No local extremum. 4. Faster convergence speed of learning process.	1. Weak fault tolerance ability. 2. Too much training samples lead to larger network size.
SVR	1. Avoid the problem of dimensionality curse. 2. Avoid local extremum theoretically. 3. Good generalization ability and strong robustness.	1. Pay for much machine memory and computing time. 2. Cause the waste of training samples when training accuracy is lower.

B. Committee Machine

By combining multiple NN into a single system, a committee of machines is formed. The result of a committee system is a combination of the results of the various component systems, which are called the basic neural network units in my research. Different combination or decision methods have a quite affect for prediction accuracy of the whole CNN model. In the case, where the basic neural network units have got higher precision individually or the differences between predicted value and true value are small, the linear-average method or the highest-accuracy method is used. The linear-average method is that the final result is calculated by the average results of the basic neural network unit. The highest-accuracy method is that the highest accuracy of the output of the basic neural network unit is selected as the final output. In the case, where the predicted results of basic neural network unit is quite different, it is necessary to get the weight factor of each basic neural network unit contributing to the final result by further training. So, genetic algorithm or single layer perceptron gets work. Single layer perceptron was chose in the research. Because the advantage of single layer perceptron is that it could get a threshold than genetic algorithm besides the weight factor. This threshold can make a slight adjustment for the final result of the CNN model, but the final prediction accuracy is increased greatly.

III. Data Models and Results

A. Source data

In this study, four kinds of logging data that is compressional wave velocity (DT), formation bulk density (ZDEN), compensated neutron (CNCF) and gamma ray (GR) were selected as input parameters. All they are related to core porosity (NPHI), and they had been proved that they were uncorrelated after the principal component analysis (PCA). Experiment was carried out to get training models with the labeled data firstly, and then training models were used to predict the results. Study chose the well logging data in a certain block, and the number of the three oil wells is NB22-1-1, NB22-1-3 and NB22-1-4. TABLE 2 is a detailed description of three oil wells logging data.

TABLE II. Description of three well logs data

Well	Depth (m)	Data	Input	Output	Training	Predicting	Description
NB22-1-1	2500~4000	1500	DT、ZDEN、CNCF、GR	NPHI	Yes	No	No error data. 1500 sample data that depth range are between 2500m and 4000m are used to establish neural network model.
NB22-1-3	2500~4000	1500	DT、ZDEN、CNCF、GR	NPHI	No	Yes	Error data between 2750m and 2800m are dealt to 0. 1500 data that depth range are between 2500m and 4000m are used to predict porosity using training model, and comparison.
NB22-1-4	2800~4300	1500	DT、ZDEN、CNCF、GR	NPHI	No	Yes	Error data between 3880m and 3920m are dealt to 0. 1500 data that depth range are between 2800m and 4300m are used to predict porosity using training model, and comparison.

B. Models

MATLAB platform was chose for simulation. Based on MATLAB neural network toolbox, study achieved BP network model, RBF network model, SVR model and CNN model through adjustment of neural network structure and size.

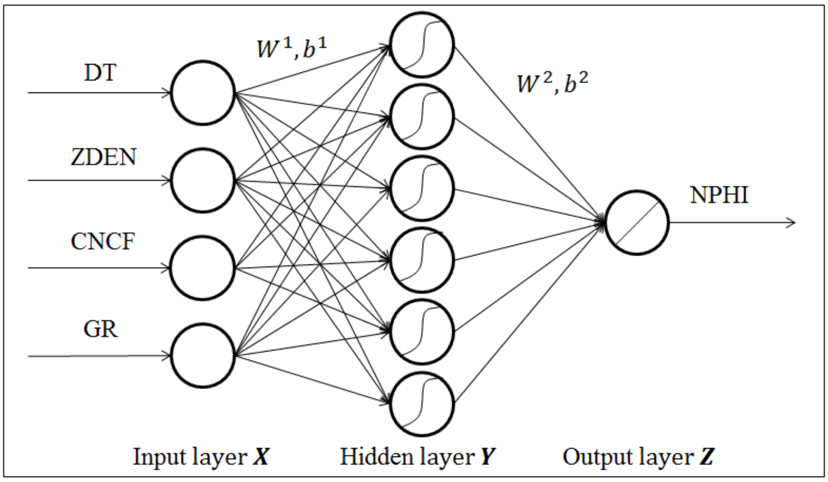

Fig. 2. A structure shows the model of BP neural network including input layer, hidden layer and output layer.

To construct BP network model, hyperbolic tangent function was selected as activation function from input layer to hidden layer. As well, pure linear function was selected as activation function from hidden layer to output layer. Fig.2 shows the structure of BP model. So the state of the hidden layer is $Y = tansig(W^1 X + b^1)$. The state of output layer is $Z = purelin(W^2 Y + b^2)$. Two formulas where *X, Y, Z* are the state of input layer, hidden layer and output layer, W is the network weight factor, and b is the network threshold. Through many simulation experiments and comparison, the parameters of BP network were determined. Input layer was set 4 nodes. Hidden layer was set 6 nodes. When the node of hidden layer is more than 6, prediction accuracy has a little change but training speed is very slow. When the node of hidden layer is less than 6, training speed is fast but prediction accuracy is poor. Output layer was set 1 node. Moreover, weight initialization algorithm was set *Nguyen-Widrow* algorithm. Back propagation learning algorithm was set *Levenberg-Marquardt* technology. Learning rate was set 0.01 and maximum number of training was set 1000 and training precision was set 0.001. Fig.4 (a) and Fig 7 (a) show the comparison between core porosity and predicted porosity and Fig.4 (e) and Fig.7 (e) show the correlation coefficient between core porosity and predicted porosity using BP model.

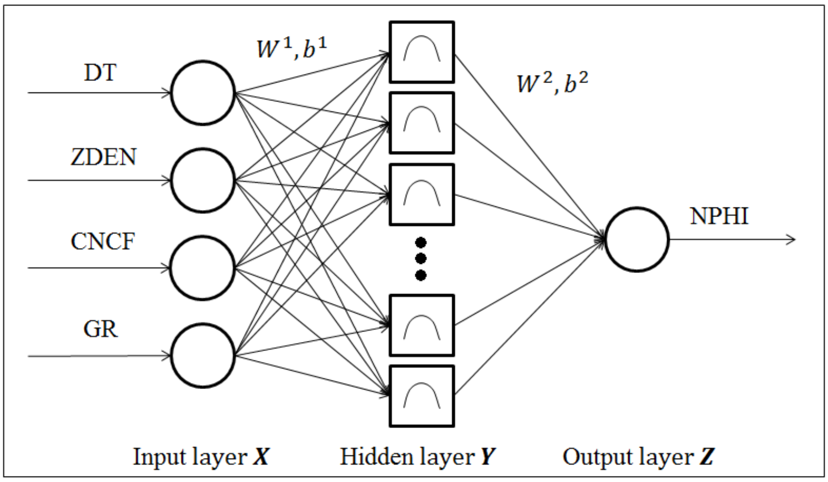

Fig. 3. A structure shows the model of RBF neural network including input layer, hidden layer and output layer.

Radial basis neurons of RBF network model was set the distance function (Euclidean distance) multiplied by threshold, and the Gaussian radial basis function was selected as the activation function. [19] Fig.3 shows the structure of RBF model. So the state of the hidden layer is $Y = radbas(\| W^1 - X \| \bullet b^1)$. The state of output layer is $Z = purelin(W^2 Y + b^2)$. Two formulas where *X, Y, Z* are the state of input layer, hidden layer and output layer, W is the network weight factor, and b is the network threshold. Radial basis neurons were added successively to the hidden layer until the desired performance has been achieved in processing of establishing RBF model. As well, we found fault tolerance ability of the model became lower when the training precision was higher, that result in unsatisfactory prediction results. That means that training precision must be set reasonable value. Experimental results show that final prediction accuracy was higher when the training precision was set 0.02, and at this case, hide the layer had 17 neurons. Gaussian radial basis parameter was selected as the default value because the sample data is unified in range of -1 to 1. Fig.4 (b) and Fig.7 (b) show the comparison between core porosity and predicted porosity and Fig.4 (f) and Fig.7 (f) show the correlation coefficient between core porosity and predicted porosity using RBF model.

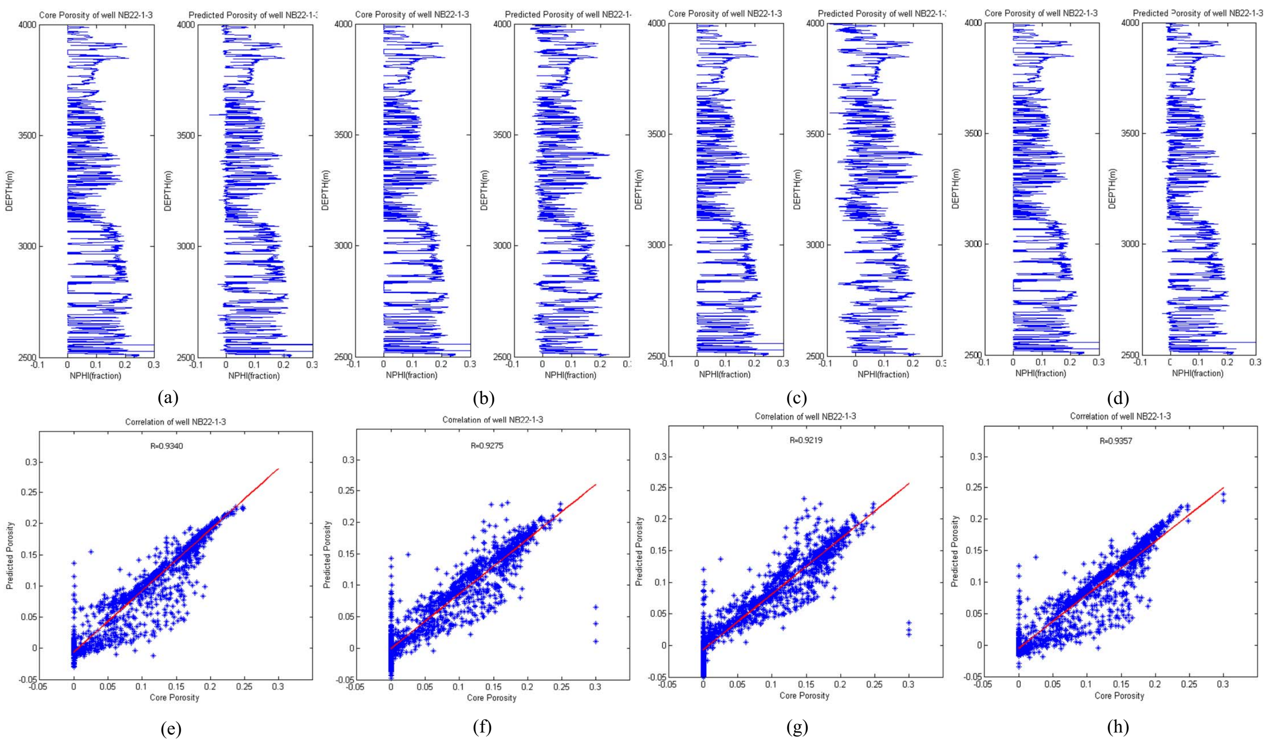

Fig. 4. Plot shows comparison between core porosity and predicted porosity of BP model (a) (e), RBF model (b) (f), SVR model (c) (g) and CNN model (d) (h) for well **NB22-1-3**. Left of (a) (b) (c) (d) are the data of core porosity. Right of (a) (b) (c) (d) are the data of predicted porosity. (e) (f) (g) (h) are the correlation coefficient between core porosity and predicted porosity.

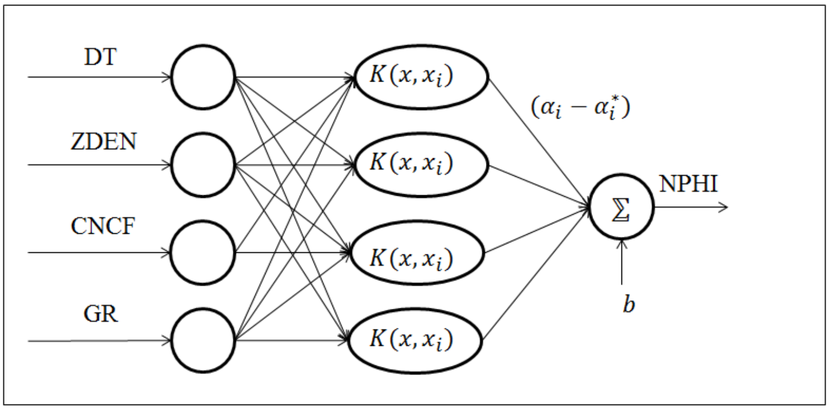

Fig. 5. A structure shows the model of SVR including input layer and output layer.

SVR model was considered as a two layer network model, and the radial basis function was set as the kernel function. Fig.5 shows the structure of SVR model. So the state of output layer is $y = \sum_i^k (\alpha_i - \alpha_i^*) exp(-\gamma \| x - x_i \|^2) + b$, where α_i and α_i^* is Lagrange factor of each sample, γ is the parameter of kernel function, b is the threshold. In the SVR model, the input is the 4 dimension vector, and the training sample is labeled as the 1 dimension vector. Kernel function selected the Gaussian radial basis function [19]. There are three important parameters that are the penalty coefficient, kernel function parameter and fitting accuracy. Through many experiments and comparison, the penalty coefficient was set 1000, kernel function parameter was set 0.01, and the fitting precision was set 0.02. Fig.4 (c) and Fig.7 (c) show the comparison between core porosity and predicted porosity and Fig.4 (g) and Fig.7 (g) show the correlation coefficient between core porosity and predicted porosity using SVR model.

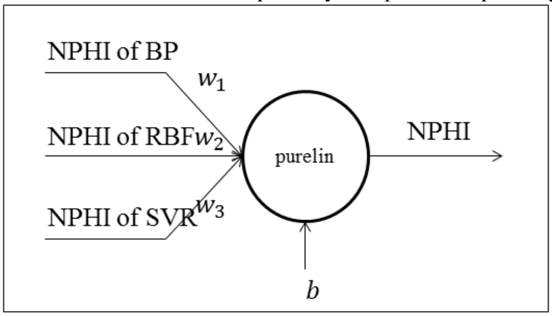

Fig. 6. A structure shows SLP that is set a pure linear activation function combines the results of BP, RBF and SVR.

In the final stage of this study, CNN model was constructed after three basic neural network models were established. A single layer perceptron was embedded to combine the results of basic neural network model to one porosity result. Fig.6 shows the structure of SLP model combining the results of BP, RBF and SVR. It adjusted network structure and reaped beneficial advantages of all models. Pure linear activation function was set in SLP. So the state of output is $y = purelin(p_1 * w_1 + p_2 * W_2 + p_3 * w_3) + b$, where y is final porosity, p is results of basic neural network model (BP model, RBF model and SVR model), w is weight value that each model contributes for final result, and b is threshold. In the experiment, the weight values are 0.7910, 0.1674 and -0.0213 to BP, RBF and SVR model. The threshold value is 0.0052. Fig.4 (d) and Fig.7 (d) show the comparison between core porosity and predicted porosity and Fig.4 (h) and Fig.7 (h) shows the correlation coefficient between core porosity and predicted porosity using CNN model.

C. Results and Analysis

The sample data of NB22-1-1 well was used to train to establish BP model, RBF model, SVR model and CNN model, and then the data of NB22-1-3 well and NB22-1-4 well were employed to predict porosity using each training model, which assessed the reliability of different model. Finally the predicted porosity of the two wells was compared and analyzed with true core porosity. The comparison results were expressed by mean square error (MSE) and correlation coefficient (R). Results were shown in TABLE 3.

TABLE III. MEAN SQUARE ERROR AND CORRELATION COEFFICIENT BY COMPARING PREDICTED POROSITY WITH CORE POROSITY

Model	**NB22-1-3**		**NB22-1-4**	
	MSE	***R***	***MSE***	***R***
BP	0.0007669	0.9340	0.0009377	0.8595
RBF	0.0008761	0.9275	0.0009397	0.8532
SVR	0.0011	0.9219	0.0009395	0.8714
CNN	0.0007451	0.9357	0.0008185	0.8739

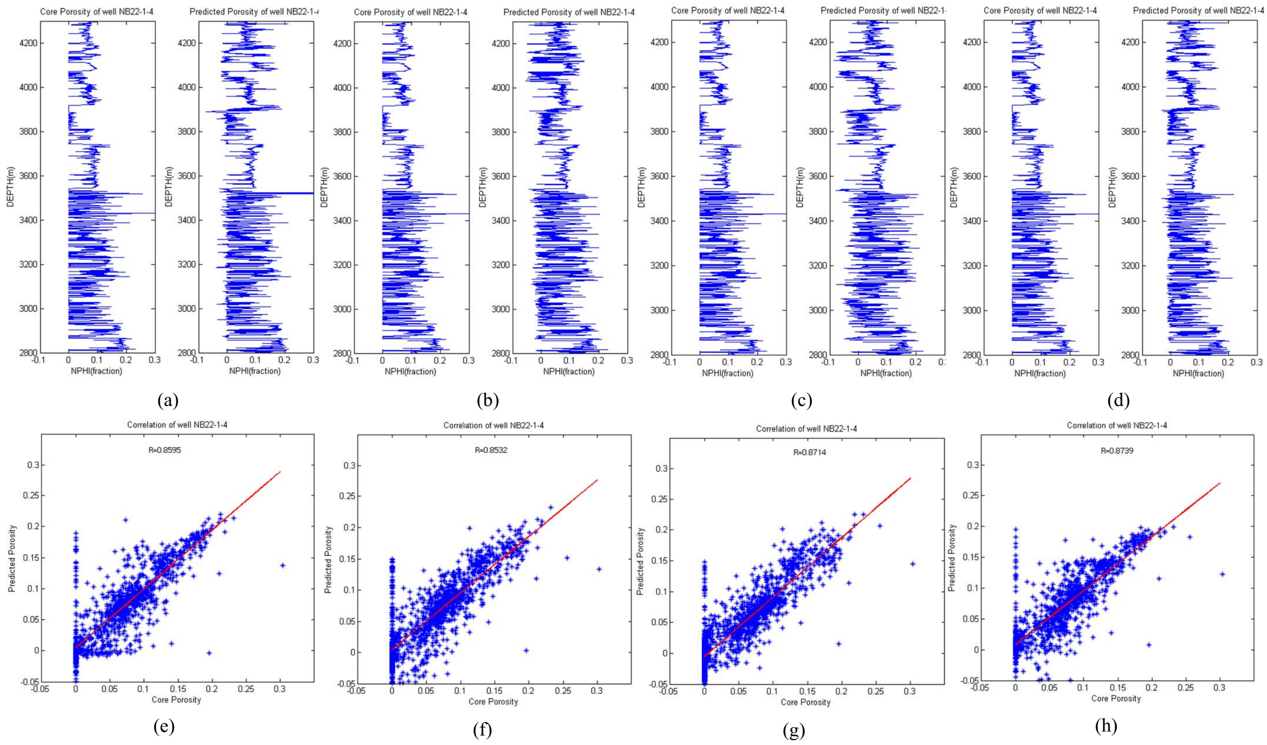

Fig.7. Plot shows comparison between core porosity and predicted porosity of BP model (a) (e), RBF model (b) (f), SVR model (c) (g) and CNN model (d) (h) for well **NB22-1-4**. Left of (a) (b) (c) (d) are the data of core porosity. Right of (a) (b) (c) (d) are the data of predicted porosity. (e) (f) (g) (h) are the correlation coefficient between core porosity and predicted porosity.

From comparison results of four models, it showed that mean square error of CNN model was the least and its correlation coefficient was the highest for NB22-1-3 well and NB22-1-4 well. Therefore, results showed that the CNN model performed better than individual neural network model. In experiments, we found that the predicted result of CNN model tended to the best one of three models. Because it eliminates the defects of individual neural network and reaps beneficial advantages of all model. At the same time, the predicted results of NB22-1-3 was better than NB22-1-4 well. Because the data for training model is derived from 2500m depth to 4000m depth. Training model for predicted performance was better in the same depth of NB22-1-3 well. And the predicted performance is poor for NB22-1-4 well between 4000m to 4300m.

IV. Conclusions

As the branch of artificial intelligence, artificial neural network has solved the many difficult practical problems in pattern recognition and classification prediction field successfully. But it has some disadvantages as well, such as the poor universal ability of structure, difficult analysis of performance, etc. Rock porosity is the most crucial and at the same time complicated to determine the parameter of hydrocarbon reservoirs. This paper presented a new CNN model by a SLP combined with BP network, RBF network and SVR model, which eliminated the defects of individual neural network in porosity prediction and reaped beneficial advantages of all model. It improved the accuracy of the rock porosity prediction. Results show that the CNN model performed better than individual neural network model. In general, the basic neural network units of CNN model can be replaced and changed to other neural network. The better performance basic neural network units perform, the better performance of CNN model will get.

With the development of deep learning, deep neural network will be paid more and more attention. The research in the field of geological reservoir parameter prediction using deep neural network will be valuable and developed.

References

[1] Xu J, Wang C. Predicting reservoir physical parameters by comprehensively analysing geologic, drill and seismic data[J]. Oil Geophysical Prospecting, 1996

[2] Chen R, Wang F. Application of MATLAB-based of BP Neural Network in Reservoir Parameters Prediction[J]. Well Logging Technology, 2009, 33(1):75-78.

[3] Hamidi H, Rafati R. Prediction of oil reservoir porosity based on BP-ANN[C]// Innovation Management and Technology Research (ICIMTR), 2012 International Conference on. IEEE, 2012:241-246.

[4] Zhang L, Ma J, Wang Y, Pan S. PSO-BP Neural Network in Reservoir Parameter Dynamic Prediction.[C]// International Conference on Computational Intelligence and Security, Cis 2011, Sanya, Hainan, China, December. 2011:123-126..

[5] Pan S, Liang H, Liang L I, Wang J. Dynamic prediction on reservoir parameter by improved PSO-BP neural network[J]. Computer Engineering & Applications, 2014.

[6] Ma L, Zhou L, Wang T. Water Quality Prediction Based on BP Neural Network at Dahuofang Reservoir, China[J]. Nature Environment & Pollution Technology, 2015, 14(3):727-732.

[7] Li Y U, Wang J Q. Application Research on Complex Water Quality Prediction with Improved QGA-BP Model[J]. Pattern Recognition & Artificial Intelligence, 2012, 25(4):705-708.

[8] Baneshi M, Behzadijo M, Rostami M, Schaffie M, Nezamabadipour H. Using Well Logs to Predict a Multimin Porosity Model by Optimized Spread RBF Networks[J]. Energy Sources Part A Recovery Utilization & Environmental Effects, 2015, 37(22):2443-2450.

[9] Wu X J, Jiang G C, Wang X J, Wang X J, Fang N, Zhao L, Ma Y M, Luo S J. Prediction of reservoir sensitivity using RBF neural network with trainable radial basis function[J]. Neural Computing & Applications, 2013, 22(5):947-953.

[10] Chen J, Li Z, Zhao D. Prediction of Hydrocarbon Reservoir Parameter Using a GA-RBF Neural Network[J]. Communications in Computer & Information Science, 2009, 51:379.

[11] Liu X, Dong F, He G, Liu J. Use of PCA-RBF model for prediction of chlorophyll-a in Yuqiao Reservoir in the Haihe River Basin, China[J]. Water Science & Technology Water Supply, 2014, 14(1):73-80.

[12] Baneshi M, Schaffie M, Nezamabadipour H, Behzadijo M, Rostami M. Accurate Estimation of Petrophysical Indexes by RBF, ANFIS, and MLP Networks[J]. Energy Sources Part A Recovery Utilization & Environmental Effects, 2015, 37(17):1874-1882.

[13] Na'imi S R, Shadizadeh S R, Riahi M A, Mirzakhanian M. Estimation of reservoir porosity and water saturation based on seismic attributes using support vector regression approach[J]. Journal of Applied Geophysics, 2014, 107: 93-101.

[14] Al-Anazi A F, Gates I D. Support vector regression for porosity prediction in a heterogeneous reservoir: A comparative study[J]. Computers & Geosciences, 2010, 36(12):1494-1503.

[15] Mollajan A. Estimation of reservoir water saturation using support vector regression in an Iranian carbonate reservoir[C]// American Rock Mechanics Association. 2013.

[16] Al-Anazi A F, Gates I D. Support vector regression to predict porosity and permeability: Effect of sample size[J]. Computers & Geosciences, 2012, 39(2):64-76.

[17] Zhang J, Yin K, Wang J, Huang F. Displacement prediction of baishuihe landslide based on time series and PSO-SVR model[J]. Chinese Journal of Rock Mechanics & Engineering, 2015.

[18] Hansen L K, Salamon P. Neural network ensembles[J]. IEEE transactions on pattern analysis and machine intelligence, 1990, 12: 993-1001.

[19] LOU Jun Gang JIANG Yun Liang SHEN Qing JIANG Jian Hui Department of Computer Science and Technology, College H T, Huzhou. Evaluating the Prediction Performance of Different Kernal Functions in Kernel Based Software Reliability Models[J]. Chinese Journal of Computers, 2013, 36(6):1303-1311.