SciencePG
Science Publishing Group

The BP Neural Network for Improvement of Classification Accuracy in Remote Sensing Image

Ming Yu[1, 2, *], He-Rong Wang[1], Ting Lan[1]

[1]School of Geographical Sciences, Fujian Normal University, Fuzhou, China

[2]Institute of Geography, Fujian Normal University, Fuzhou, China

Email address:

gisyuming@163.com (Ming Yu)

[*]Corresponding author

Abstract: Remote sensing (RS) data classification is one of the core functions of the system of remote sensing image processing. In this study, back propagation (BP) neural network was introduced into the application of remote sensing image with implementation of MATLAB. To improve measurement accuracy, the BP neural network application includes two schemes of different transfer functions; and 3, 5 and 7 bands of RS images of Landsat 8 OLI were used for validate the accuracy of classification. The experimental results proves that this algorithm is better than tradition classification of supervise and non - supervise methods. Classification accuracy increases as more band information is given; scheme 2 has high classification accuracy than scheme 1. The research results have a certain reference value for the rational use of land resources.

Keywords: BP Neural Network, Remote Sensing Image Classification, Network Parameters, Maximum Likelihood Classification Method

1. Introduction

Technology of remote sensing had been well developed and applied as modern science progress. The core purpose of applying remote sensing technology is to determine the characteristics of the electromagnetic radiation from the surface objects, then to identify the type of the ground object and its distribution [1]. This process also becomes more effective through image classification since classification method is very important while monitoring dynamic changes of land, thematic information extraction and production of thematic maps [2, 3]. Traditional classification methods include supervised classification and unsupervised classification: the principles of both methods are based on the spectral characteristics of a certain object. However, there are always mixed pixels due to the limitation of satellite resolution, and it usually causes that the computer cannot determine the types of ground objects. Scholars put in countless effort to try to modify and improve this condition, but it still cannot fully avoid the phenomenon of "different body with same spectrum" or "same body with different

spectrum", and generate the incorrect classification or with missing date. So the accuracy of the classification strongly influences on the accuracy of the future work. In this study, the back propagation (BP) neural network model was implemented on the MATLAB platform. This model does not only improve the speed of image classification, but also avoid the influence of the shadow and other interference information [4, 5]. Overall, this method overcomes the shortcoming of traditional methods. It strongly improves the classification accuracy and well retained the geographic information of the remote sensing image; it is very convenient to help future image processing.

2. Overview of Artificial Neural

Artificial neural network (ANNs) is also known as neural network or connection model. It simulates basic characteristics of the human brain and natural neural network. The artificial neural network could approximate any complex nonlinear relationship, and it also has the advantages of good robustness and fault tolerance to apply to an unfamiliar or

unknown system [1, 6]. So it has been very popular recently, have been applied to more and more remote sensing image processing projects today [13, 14, 15].

BP network was proposed by a group of researchers led by Rinehart and McClelland in 1986. It is a multilayer feed forward network which originally runs error back propagation algorithm for training purposes. It is also one of the most widely used neural network models right now, and also was widely used in the field of remote sensing image classification. There are three types of layers in a typical BP multilayer perceptron network: input layer, middle layer (hidden layer) and output layer (see figure 1). In figure 1, we can see that all neurons are fully connected to adjacent layers while layers are not connected to each other at all. The BP algorithm has two processes that perform forward computing (forward propagation) of data stream and backward propagation of error signals. The forward propagation first starts in input layer, the move to hidden layer and finally reaches output layer; the status of each layer of neurons only affects the neurons in next layer. Back propagation process will be triggered if the output layer does not reach the expecting value level. This error function of network helps to reach the minimum value of error via the alternation of two propagation processes, and to complete the information extraction process and data storage in next step.
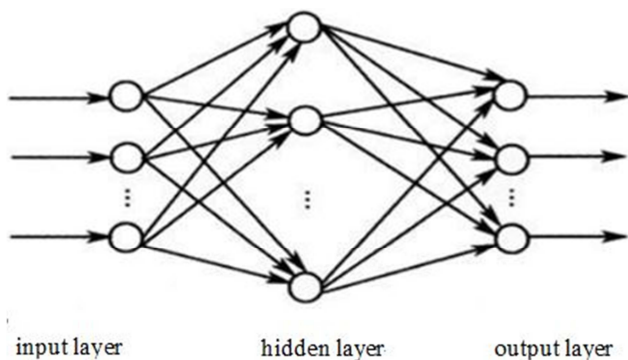

Figure 1. BP neural network.

3. Implementation of BP Neural Networks on Remote Sensing Image Clarification

3.1. Data Source and Data Preprocessing

Remote sensing data in this study was selected from the remote sensing image of Landsat 8 OLI in year 2014. It was downloaded from the international scientific data mirror site of computer network information center directed by Chinese Academy of Sciences (*http://www.gscloud.cn*). The preparation process of the remote sensing images include: image cropping, geometric correction, atmospheric correction and image enhancement. The final size of the experimental image is 800m by 800m, with resolution of 30m by 30m. This image was captured at longitude between 25°43′ to 25°30′N and latitude between 116°8′ to 116°22′E.

3.2. Preparation Stage

The traditional classification method combines three band images into one color image, and then it selects the samples based on colors of the object, and then differentiates the categories of each object through a certain criterion. [6] It is very similar to maximum likelihood classification method in ENVI software, since the number of neurons in input layer of BP neural network are not limited in this study, therefore, they can be used as input layer neurons in 3, 5 and 7 bands, respectively, in order to test classification results and its accuracy of BP neural network under different conditions. The experimental area locates in the western region of Fujian Province (Minxi). According to the information obtained from past studies, the woodland in this region accounts for 80% of the total area, while the grazing grassland only covers 0.02%; the distribution of grass land shatters into fragments and hard to interpret. Types of land usage in Minxi can be categorized into farmland, woodland, water, garden plot, construction land and unused land.

3.2.1. Extermination of Training Sample

Spectral characteristics of Minxi include the following: resultant image consists of 654 bands (RGB) of OLI. The color of image is bright, each layer can easily be identified, and types of surface objects can also easily be categorized. So we used this image of combination band to select the representing samples for each object on the ENVI software. Since this region has a highly complex topography, and the woodland area occupies a large portion of the area, which causes a large portion of shadows (Figure 2). In order to avoid this problem causing potential confusions in future interpretation, shadow is categorized into the class along with woodland for convenient purposes. Also, there is a part of water region shows a different color (Figure 3). In response, we categorized this part into a special group, and combined it with group of water at the end. Sample areas and images are saved after samples of objects are chosen, and then information is extracted for other four bands; finally, we exported the results into TXT files, and saved them as input files to be called in MATLAB.
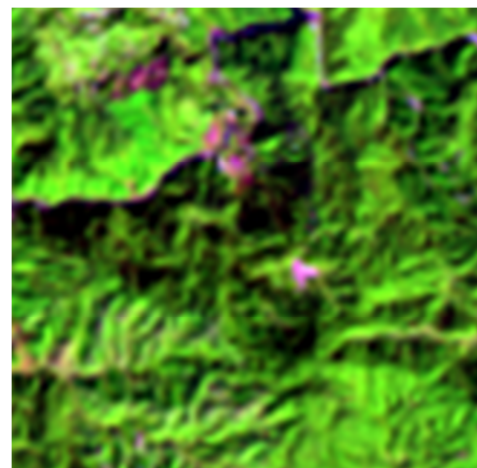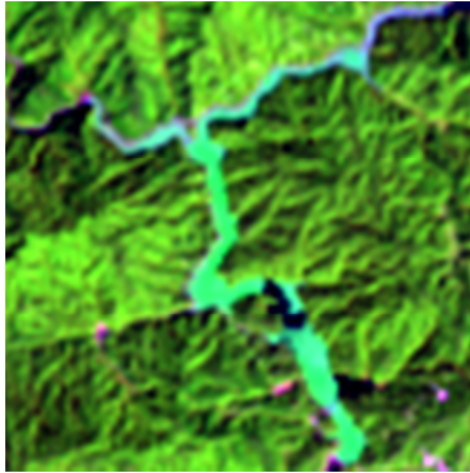

Figure 2. The shadow parts.

Figure 3. *The color change of water.*

3.2.2. Network Structure and Parameter Determination

In order to apply BP neuron network in remote sensing image classification, the number of neurons in the input layer is used as the number of the bands to be classified. The number of neurons in the output layer is the number of classes to be classified. The output layer node number is 8 due to the impact of shadow and discoloration of the water body. The number of hidden layer neurons was not uniform. The reference calculation formula of the optimal number for hidden neurons is

$$L=(m+n)/2+c.$$

In this equation, "L" is the number of hidden neurons, "m" is the number of input neurons, "n" is the number of output neurons and "C" is a constant between 1-10 [7]. Also, the number of neurons in input layer and the hidden layer should be set according to different experiments.

Selection of network parameters: Network learning rate is set at 0.01, expected error value level is set at 0.01, and network training number is set for 8500 times.

Establishment of BP network: The range of vector matrices of input layer is from 0 to 1 after normalization, and output value is also controlled within this range, which force the transfer function must contain the logsig function. Two experimental schemes were applied in this training stage for this study [8].

Scheme 1: Hidden layer neurons used logsig function of logarithmic function, output layer neurons.

Used purelin linear transfer function, the training function used traingdx function which was gradient descent method with momentum and adaptive LR. BP neural network was established by newff function. Experiment codes are shown as the following:

net = newff(minmax(input) [1, 9, 10], {'logsig' 'purelin'}, 'traingdx');

Scheme 2: Hidden layer neurons used tansig function, this is a transfer function of the output neuron which uses logsig function, the training function uses Levenberg-Marquardt method by trainlm to optimize algorithm, and BP neural network was established by newff function. Experimental

codes are shown below:

net = newff(minmax(input) [1, 9, 10], {'tansig' 'logsig'}, 'trainlm');

3.3. Learning Phase

After the preparation stage is completed, we used the training samples to calibrate the network. According to the input samples, the output values and errors of both hidden layer and output layer were calculated, so that the network weights and threshold were calibrated until it met the requirements. If the maximum number of pre-stage cannot met tolerance requirements, we would stop calibration process and to reset the network structure and the control parameters. After repeatedly adjusted training time and other control parameters, the learning error of scheme 1 did not reach the value of 0.01, while the number of training reached 8500 times, learning error became relatively small. In this process, scheme 2 converges very quickly, error numbers are below the mark with less time spent and also higher efficiency. The results of learning stage are shown in table 1.

Table 1. *The results of BP neural network learning.*

		Scheme 1		Scheme 2	
		Training times	Error value	Training times	Error value
Band Number	3	8500	0.019	27	0.0099
	5	8500	0.012	10	0.0093
	7	8500	0.011	13	0.0085

3.4. Classification Phase

After the calibration stage is completed, the weight matrices met the expectations of training, and then we move onto the next stage of classification phase. In order to have classification processed properly, 3, 5, 7 bands all be included and put into the MATLAB. And the entire geographic information of the remote sensing images can be called by using the following function:

[RGB, E] = geotiffread('2014_test.tif');% read the images with spatial information

info = geotiffinfo('2014_test.tif');

geotiffwrite('2014_classify.tif',Z,MAP,E, 'GeoKeyDirectoryTag', info.GeoTIFFTags.GeoKeyDirectoryTag);% save image and its spatial information

Based on the weight coefficient matrices of network layers which were calculated in learning phase, the pixel of every image are estimated, as well as the probabilities of each surface object in every image; finally to determine the type of surface object by applying the probabilities from earlier step. The classification results of scheme 1, and scheme 2 are shown in figure 4 and figure 5, respectively, (a), (b), (c) in each figure represent the classification results of bands 3, 5, 7, respectively.

4. Results and Analysis

In order to validate the accuracy of classification results more

effectively, this study applied the classification overall accuracy and the Kappa coefficient of BP neural network. The overall classification accuracy was calculated by taking sum of correctly classified pixels, divided by total number of pixels. Kappa coefficient is a multivariate statistical method for accuracy evaluation of the classification; it represents the ratio of the error reduction of the total random classification. The Kappa coefficient calculation result is called KHAT statistics [11]
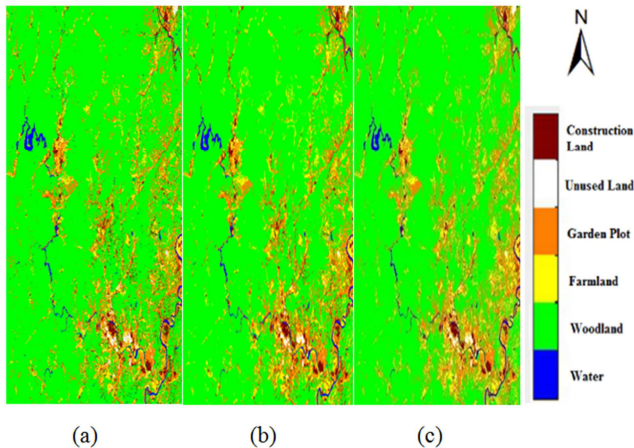

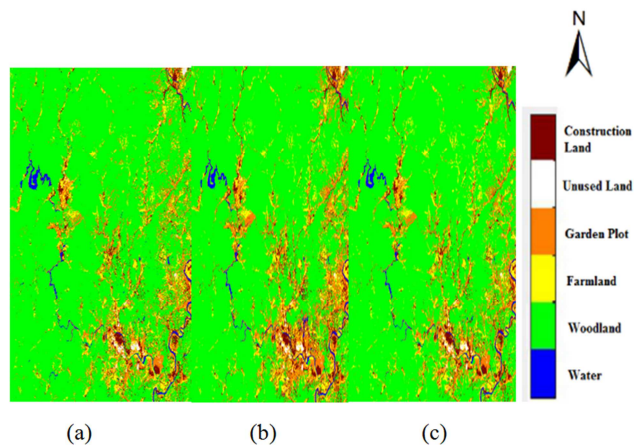*Figure 4. The classification results of scheme 1 of BP neural network.*

Figure 5. The classification results of scheme 2 of BP neural network.

By applying visual analysis, the classification results of the two schemes both reflect the actual situation of the surface. But we also found more patches of classification results exist with fewer bands are given; more shadow areas are also categorized as water under this condition. For testing BP neural network classification results, only 100 pixel points were randomly selected in each sample since there are too many samples for each kind of surface features. Then these pixel points were applied in network and been re-calculated which class each one of them belong to. The results of this classification accuracy test on two schemes are shown in table 2. Experimental results show that the total accuracy of BP neural network classification method is good, accuracy increases as more bands information are given. With the same number of bands, the classification accuracy of scheme 2 (see figure 5) is higher than that of scheme 1 (see figure 4); scheme 2 also converges faster than

scheme 1. The classification method of 7 bands has the best results, its classification accuracy is 97.67%, and the Kappa coefficient is 0.972. Table 3 shows the result of classification error statistics of scheme 2 [12].

Table 2. The classification accuracy of BP neural network.

		Scheme 1		Scheme 2	
		Total accuracy	Kappa coefficient	Total accuracy	Kappa coefficient
Band Number	3	93.83%	0.926	95.83%	0.95
	5	96.67%	0.96	96.83%	0.962
	7	96.83%	0.962	97.67%	0.972

Table 3. Classification error statistics of scheme 2.

Classification	1	2	3	4	5	6	Sum
1	100						100
2		100					100
3			98			2	100
4				96	4		100
5			2		98		100
6		1	5			94	100
Sum	100	101	105	96	102	96	600

Notes: 1, 2, 3, 4, 5, 6 stand for Water, Wood land, Farm land, Garden plot, Unused land and Construction land

5. Conclusions

This study applies BP neural network with images of 3, 5 and 7 bands, with application of two types of schemes, different transfer functions were also involved, compared and analyzed accordingly. This study shows that the BP neural network classification has higher overall accuracy with different parameters. The nonlinear transfer function of the scheme 2 does not only speed up convergence rate, but also has better classification results, which greatly improves the classification accuracy. This study also overcame the shortcomings of MATLAB in processing remote sensing images, and retained the integrity of remote sensing images of geographic information, for better post-processing and further application. The methods researches have certain significance to the reasonable protection of land resources.

Acknowledgements

References

[1] Atkinson, P. M. and A. R. L. Tattnall, Neural Networks in Remote Sensing, international journal of Remote Sensing, 1997, 18 (4): 699-709.

[2] JIAO Li-min, Wu Su. Analyzing the characteristics of the Expansion of the Metropolises in china from 1990 to 2010 Using Self-organizing Neural Network [J]. Geomatics and Information science of Wuhan University. 2014. 39 (12): 1435-1440, 1471.

[3]   Ming Yu, Ting-hua AI. Study of RS data classification based on rough sets and C4.5 algorithm. 2009 Proceedings of SPIE, 2009.10.13-2009.10.14.

[4]   LI CHUN-HUA, SHA Jin-Min. Knowledge Based Self-Organizing Neural Network Remote Sensing Image Classification Approach，Remote Sensing Technology and Application, 2006, 21 (6): 507-511.

[5]   Xu Lei, Lin Jian, et al. Classifying Remote Sensing Image Based on BP Neural Network [J]. Geospatial Information, 2012, 10 (4): 83-85.

[6]   Zhang S L, Chang T C. A Study of Image Classification of Remote Sensing Based on Back-Propagation Neural Network with Extended Delta Bar Delta [J]. Mathematical Problems in Engineering, 2015, 501: 178598.

[7]   Y. Du, S. Zhou, and Q. Si. Application and contrast research on remote sensing image classification base on ANN [J]. Journal of Science of Surveying and Mapping, 2010, 35 (4): 121-125.

[8]   Xiao Jincheng, Ou Weixin, et al. Land cover classification of Yancheng Coastal Natural Wetlands based on BP neural network and ETM+ remote sensing data [J]. Acta Ecologica Sinica, 2013, 33 (23): 7496-7504.

[9]   Wu Chuang Ju, Song Shuang Jie, et al. Research of new method of RS Image classification Based on neural network [J]. China Science and Technology Review, 2014 (7): 321-322.

[10]   Li Haiyang, Fan Wenyi. Matlab Realization of Sensing Image Classification Based on Probabilistic Neural Network [J]. Journal of Northeast Forestry University, 2008, 36 (6): 55-56.

[11]   Jensen, J. R., Qiu, F. and K. Patterson, A Neural Network Image Interpretation System to Extract Rural and Urban Land U se and Land Cover Information from Remote Sensor Data. Geocarto International. 2001, 16 (1): 19-28.

[12]   Ke Huaming, CHEN Chaozhen, et al. Application of BP Neural Network Classification with Optimization of Genetic Algorithm for Remote Sensing Imagery Based on Matlab [J]. Journal of Southwest University of Science and Technology, 2010, 25 (3): 55-59.

[13]   Burks T F, Shearer S A, Gates R S, et al. Backpropagation neural network design and evaluation for classifying weed species using color image texture [J]. Transactions of the ASAE, 2000, 43 (4): 1029-1037.

[14]   Sexton R S, Dorsey R E. Reliable classification using neural networks: a genetic algorithm and backpropagation comparison [J]. Decision Support Systems, 2000, 30 (1): 11-22.

[15]   Li Shuang, Ding Shengyan, et al. The Comparative Study of Remote Sensing Image Classification. Journal of Henan University (Natural Science), 2002, 02: 70-73.