Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Environmental Sciences 10 (2011) 2360 – 2366

Procedia Environmental Sciences

2011 3rd International Conference on Environmental Science and Information Application Technology (ESIAT 2011)

Wetland Landscape Classification Based on the BP Neural Network in DaLinor Lake Area

YuHai Bao *[1,2,a], Jianbin Ren[1,2,b]

[1] *Inner Mongolia key Laboratory of Remote Sensing and Geography Information Systems,Hohhot 010022, China*

[2] *Inner Mongolia Normal University Geographical Science, Hohhot 010022, China*

[a] *bao18376@sina.com*

Abstract

With the applications of remote sensing in the study of wetland landscape is more and more extensive,computer automatic classification method appears especially important. In this paper, using TM remote sensing data in August 2008 year in Dalinor,Research to application of BP neural network and Traditional classification methods - maximum likelihood.Through the selection of training samples,Choice of training algorithm and Optimal number of hidden neurons to determine,the precision of BP neural network classification and the accuracy of maximum likelihood classification are compared,The results showed that BP neural network based on LM algorithm is better than the maximum likelihood method for classify to remote sensing image Higher accuracy.

Key words: BP neural network;Wetland landscape;Dalinor lake;method of maximum likelihood

Introduction

In recent years, the nerve network,, as a data analysis method, is gaining increasing concern and attention from researchers in a variety of areas and it has become a new branch of mathematics. With a strong Nonlinear fitting ability, the nerve network can map complex nonlinear relationships. It is very easy to learn and be implemented by computer. Its great memory capacity, nonlinear mapping and strong self learning ability, offer the nerve work a big market application prospect.

In remote sensing images classifications, with a high speed and precision of classification , the nerve network can effectively distinguish the details of the landscape,. Therefore, this paper applies the nerve network to Dalinor lake natural reserves, to attain a quick and accurate classification on the landscape types.

Dalinor natural reserves is located in west of ChiFeng city of KeShenKeTengQi, Inner Mongolia. The north-western of the reserves is the Inner Mongolia Xilingol pasture natural reserve, the south is the eastern Tenger Desert, the west is the border of Xilingol league and ChiFeng city, the east is about 80 km away from the BaiyinAobao national natural protection areas. Its geographical coordinates are longitude

doi:10.1016/j.proenv.2011.09.368

116°22′—117°00′E, latitude 43°11′—43°27″N, with a total area of 119413.55 ha. Dalinor natural reserve is a integrated natural reserve features protecting rare birds and their various living ecosystems of lakes and wetlands, grassland and forests land. It is a important migrate channel and gathering place for the migrant birds in the north of China[1].

Research methods

Artificial Neural Network.

"Artificial neural network" (ANN) is a complex network system composed by a large amount of widely interconnected short and simple processing units (neuron),is an engineering system mimicking human brain's structure and mental behaviors on the understanding of its organization and operation mechanism. In fact artificial neural network does not fully reflect the brain's functions. It is only an abstract, simplification and simulation of the living nerve network. The nerve network deals with the information through the interaction of neurons. The knowledge and information are stored in a network of the physical interconnect distributed components. The study and recognition of the nerve network depend on connection weights of each neuron's dynamic evolution[2].

Traditional classification method.

Traditional classification maximum likelihood:It is assumed that category statistics of different bands of the wetlands landscape obey normal distribution. The calculated certain pixel will be distributed to the class of maximum probability. The landscape spectrum characteristics in the research areas are supposed to be similar with most phenomenons in the natural, that is following the normal distribution. Characteristic parameters, such as mean value, variance and covariance calculated from the study area are used to further work out the overall prior probability density function. When it doesn't meet the normal distribution, the accuracy of the wetland landscape category will decline, and cannot use the traditional classification maximum likelihood method. When use this method, we need to build a set of discriminant function with statistics knowledge, according to this set, and according to the set of discriminant function calculate the ownership probability of the each pixel of the unclassified images (For pixel K, its possibility to be category W of the wetland landscape)[3].

Data processing

Data processing includes band selection, characteristic extraction, geometric correction, and the cutting, mosaic and radiometric calibration of the image.

Sample choosing. Six large classes of forest land, water, residential area, sandy land, cultivated land and grassland are categorized according to the geographical features in the study area. Within research areas, we choose several areas with the most obvious characteristics and cut six wavelengths from TM data. Selecting samples has great influence in the accuracy of the classification. So the selective of sample must be typical. To eliminate the influence of same object with different spectra and same spectra with different objects, each kind of land class should have several typical areas.

Limitations and improvement of traditional nerve network.

In the study history of the nerve network, the standard BP network is more mature both in theory and in performance, with the comparison of that of former nerve network. However, it is not completely perfect and it has many constraints to improve.

There are so many limitations in the standard BP algorithm, so, in practical applications, it is rarely used, and we often use an improved BP algorithm. In this paper, we try to use LM training algorithm to overcome the shortages of the BM algorithm, which has slow network training speed and easily getting into local minimum value[4].

Sample training.

The number of neurons in hidden layer can be first determined by the initial value formula N = log2 (M)[5]. Through trial and error, we eventually determine the number of hidden neurons to be 10. The faster the function converged, the better the classification.

This paper used a new method of building the training sample —— delete the areas of no interest: this method can be achieved by combining use of the remote sensing image processing software, such as ENVI or ERDAS, with MATLAB software. Its realization steps are as follows:

Firstly, establish a series of interested documents, select many areas with same feature type in one document, and mask the uninterested zone in 0.

Secondly, before training sample for network, sample must be processed. Delete all the pixels whose value is 0 in all bands (unusampled areas) and establish the every kind of pure training samples.

By doing this, it can eliminate the influence of non-sample pixels, and this method can realize the same training set with the maximum likelihood method, make it possible to compare the classification precision of the two algorithms. Also, it can avoid differences in classification accuracy, get massive pure sample pixels, build a large training keys, realize the full training of neural network, and thus realize high precision remote sensing image classification[6].

The MATLAB code realization is as follows:

```
a=[bb1，bb2，bb3，bb4，bb5，bb6]
b(:,find(～any(a)))=[]
```

Bbi is the single band data from a sample image, b is a certain subsets including mask pixels.

The second code is used to found the location of the pixels who is masked to be 0 (i.e. all the uninterested areas), delete them and then get pure samples.

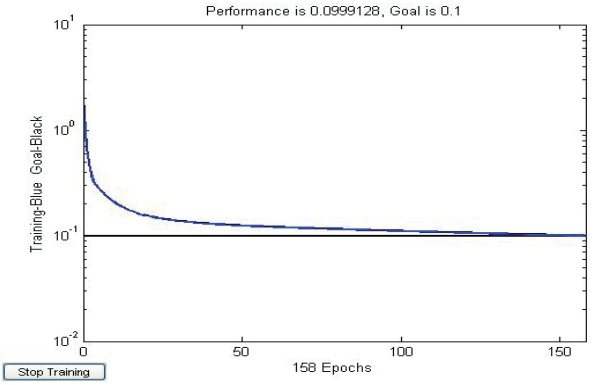

Fig. 1 Function converging condition based on method of deleting the areas of no interest to construct the training sample

When training 158 times, the error is 0.0999128 and less than the setting 0.1.

This method has a high speed of training and high classification. This paper adopts the method of deleting the areas of no interest to construct the training sample and train the BP neural network.

The classification effects of neural network simulation and maximum likelihood.

On the basis of full training with neural network, use the trained network to simulate the original image, and finally get the output.

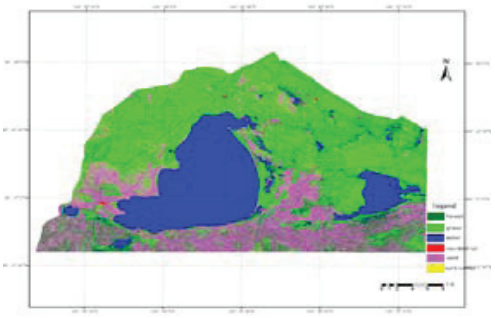

Fig. 2 The classification effects of neural network simulation

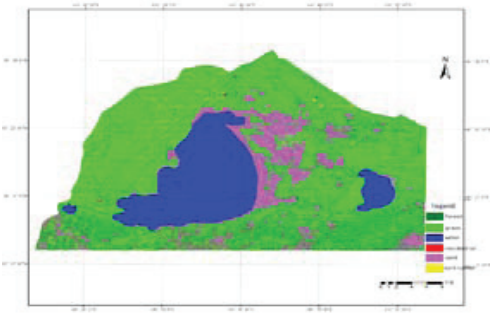

Fig. 3 The classification effects of maximum likelihood

Precision contrast analysis

Using confusion matrix verify the accuracy.

In order to test the classification effect of every classification method, this paper adopts confusion matrix to evaluate the classification accuracy, and confusion matrix can show mix points situation of various categories to a certain extent. Test samples collected by random sampling and distributed evenly in remote sensing images.

Tab 1 Classification confusion matrix for BP neural network

actual / classification	grassland	cultivated land	residential area	forest land	sandy land	waters
grassland	4077	0	24	0	21	0
cultivated land	272	3626	33	10	33	2
Residential area	34	2	186	28	41	0
forest land	0	190	13	989	25	6
sandy land	189	134	35	101	1385	0
waters	0	74	0	0	6	818
total	4538	4024	291	1100	1470	826
Kappa	0.8781		Total precision		90.71%	

Tab 2 Classification confusion matrix for method of maximum likelihood

actual / classification	grassland	cultivated land	residential area	forest land	sandy land	waters
grassland	3676	412	5	2	5	1
cultivated land	668	3553	8	166	4	0
Residential area	66	0	172	12	41	0
forest land	194	59	50	930	87	3
Sandy land	0	0	52	2	1374	0
waters	0	0	4	0	0	822
total	4538	4024	291	1100	1470	826
Kappa	0.8132		Total precision	86.5947%		

Pixel statistics validate precision.

In order to verify the correctness of the result, taking field survey, respectively making an investigation of wetland landscape to investigate several typical wetland landscape of field, then revise the mistake of first classifying and to interpret detailed. Eventually getting visual interpreting vector diagram, converting to raster images, regarding visual interpreting as a standard, taking statistics and comparison by pixel statistics and two kinds of classification method.

Tab 3 Pixel contrast statistics of Different classification methods

Land types / Number of Pixels	forest land	waters	Residential area	grassland	Cultivated land	sandy land	total
Pixels of BP neural network	2876	265559	4887	681017	425	332945	1287709
percentage	0.22%	20.62%	0.38%	52.89%	0.03%	25.86%	100.00%
Pixels of maximum likelihood	3144	253223	5002	631942	376	394022	1287709
percentage	0.24%	19.66%	0.39%	49.07%	0.03%	30.60%	100.00%
Pixels of Visual interpreting	2936	268264	4591	693300	398	318220	1287709
percentage	0.23%	20.83%	0.36%	53.84%	0.03%	24.71%	100.00%

From the above table it can be seen that different classification methods for different land types, the classification accuracy is different. also, two kinds of classification methods for water classification accuracy are higher, because the reflectivity of waters is lower, distinguished easily from other land types, so the classification effects of two classification methods for waters almost are the same; In the

classification of woodland and cultivated land, the method of maximum likelihood mistakenly classify cultivated land, However BP neural network for both classification precision obviously improved; In the classification of woodland and grassland ,the method of maximum likelihood for woodland and grassland is not obvious , while classification accuracy of BP neural network is higher.

Through comparison of classification accuracy of the forest land, grassland and farmland, it shows that the method of maximum likelihood is not as effective as the BP neural network in distinguishing tiny differences of features , because the spectral features of those are special similar, BP neural network is training repeatedly by the sample of study, is a kind of infinite approximation to the classification results, and maximum likelihood is based on statistics, is a method of assumed probability distribution, so the method of maximum likelihood is obvious not as effective as the BP neural network in distinguishing tiny differences of features.

Two methods for the classification effect of the sandy land, from the confusion matrix effect, the situation that sandy land mistakenly classified as cultivated land, forest land and grassland is more appearing, it shows the Vegetation coverage of sandy land is higher in Dalinor ecological protection zone that the vegetation cover of sandy land is taller, so spectral features are closer, resulting in the distinction degree of the sandy land and grassland, cultivated land, woodland is not high.

From the above analysis, we can draw the conclusion that the classification accuracy of BP neural network is higher than the method of maximum likelihood, especially in distinguishing tiny differences of features. Two kinds of classification methods for classification effect of water are higher, but BP neural network for distinguishing mountain shadow from water is worse, it result in mountain shadow mistakenly classify water. Forest land, grassland and cultivated land are difficult to distinguish for more closely spectral features, and need to fix classification results by visual interpreting and evidence.

From the results of pixel statistics to see, based on the BP neural network classification results, the pixels of forest land is 16.58%, for the classification results of method of maximum likelihood, the pixels of woodland is 18.15%, while the number of pixels for visual interpreting is 15.57%, apparently the classification result based on the BP neural network is closer to visual interpreting result.

The classification results for waters, BP neural network and method of maximum likelihood are 20.68% and 19.72% respectively, visual interpreting result is 20.89%. It shows that the two methods for the classification results precision of waters is basically the same, it close to the actual value.

Calculating according to pixel statistics, the pixels proportion of grassland farmland and sandy land are closer to the result of visual interpreting by LM algorithm BP neural network. From this aspect, the classification accuracy of LM algorithm on the BP neural network is higher than the method of maximum likelihood.

Conclusion

Using band selection, band synthetic and geometric correction, this paper mainly preprocess data of satellite remote sensing image of study area, constructs the training sample by the method of region of not interest decimation. It is effectively improved the classification accuracy, takes remote sensing satellite images automaticly classified by the traditional method of maximum likelihood and BP neural network based on improved LM algorithm by Neural Network Toolbox provided by Matlab. taking visual interpreting of field investigation measure and precision contrast analysis by confusion matrix and pixel statistics, the Kappa coefficient of BP neural network is 0.8781, the classification accuracy is 90.71%, Kappa coefficient of method of maximum likelihood is 0.8132, the classification accuracy is 86.5947%. The BP neural network based on LM algorithm has higher classification precision than method of maximum likelihood on classification of remote sensing image.

References

[1] Http://baike.baidu.com/view/74463.htm

[2] Neural network theory and MATLAB7 realized , Fly think technology product research and development center,Publishing House of Electronics Industry.

[3] RuihongYue in:Research on Land Cover Classification in Mongolian Plateau Based on MODIS data, Master's thesis: Inner Mongolia Normal University(2009).

[4] Yinhui Zhang,Gengxing Zhao in:Classification Methods Of Land Use/Cover Based On Remote Sensing Technologies, Journal of China Agricultrural Resources and Regional Planning, 21-25(2002),23(3).

[5] Xian-cai Gui in:Realization of BP Networks and Their Applications on MATLAB, Journal of Zhanjiang Teachers College, 25(2004).

[6] YeijunDu in:Application and Contrast Research on Remote Sensing Image Classification Based on Artificial Neural Network, Master's thesis: Inner Mongolia Normal University(2009).