Remote Sensing Image Classification Based on Gray System Theory

Gui yufeng, Zhang jixian, Ling zongjian.
(Chinese Academy of surveying & Mapping, Beijing,P.R.China 100039)
E-mail: guiyufeng@263.net

Abstract

The intelligence and automation of image processing and analysis is a bottle problem for photogrammetry and remote sensing. Artificial neural networks is a new solver which imitates brain and gray system theory is a new tool which handles undetermined problem. This paper describe how to combine artificial neural networks with gray system theory to realize classification of remote sensing image . This method can improve the accuracy and effect of traditional classification. The experiments demonstrate the method and give satisfactory results.

Keyword Remote Sensing, Classification, Artificial neural networks, Gray system theory.

1. Introduction

Vast amount of remotely-sensed data are collected to gather information about nature resources by scanning the surface of the earth. Identification, classification, and interpretation of information contained in these data are performed by experienced and skilled personal for tasks such as environmental monitoring and disaster relief. The derivation of land cover maps from remotely-sensed images depends primarily on the classification process in which a mapping is performed from the multispectral image space to the land cover class space. Much effort has been expended in the last decade to develop improved classification algorithms for remotely sensed data using for example statistical, knowledge-based, neural, and hybrid or combined methods (e.g.,Wilkinson et al,1995).

Artificial neural networks have a wide range of application in remote sensing. Most attention has, however, focused on their application in pattern recognition. In particular, the ability to learn by example and generalize make artificial neural networks attractive for the supervised classification of remotely sensed data(Schalkoff 1992,Richards 1993). Furthermore, artificial neural networks have other advantages relative to conventional classifiers. These include an independence of distribution assumptions, ability to handle data acquired at different levels of measurement precision and noise, and, once trained, rapid data processing.

Artificial neural networks first began to used for the classification of remotely sensed imagery around 1998,with the first journal papers appearing one to two years later. The most commonly-used neural network model for image classification in remote sensing is the multi-layer perceptron trained by the back-propagation algorithm. The number of hidden layers and the number of nodes in a hidden layer required for a particular classification problem are not easy to deduce. The neural networks architecture which gives the best results for a particular classification problem can only be determined experimentally, and this can be lengthy process especially for large classification tasks. This is often seen as an objection to neural networks methods. However, some geometrical arguments can be used to derive heuristics to set approximate networks sizes(Lippmann 1987). Ideally, the first hidden layer of a networks with two hidden layers should contain two to three times the number of inputs such that a sufficient number of hyper-planes can be "formed" to define hyper-regions. However, we would caution that it is not possible to rely on such heuristics and that each classification problem needs to be carefully examined in it own right.

2 Neural network classification method based on gray system theory

Gray system theory is invented by Professor Dang Julong,China. It is a new tool which can handles undetermined problem. Gray correlation analysis method measure correlation according to similarity among every factor in a system.

If $\{x_0(k)\}$ is basic sequence and $\{x_i(k)\}$ is i-th correlation sequence, *k* is time, than the correlation coefficient between basic sequence and correlation sequence is

$$\xi_i(k) = \frac{\min\limits_i \min\limits_k \left|x_0(k) - x_i(k)\right| + 0.5 \max\limits_i \max\limits_k \left|x_0(k) - x_i(k)\right|}{\left|x_0(k) - x_i(k)\right| + 0.5 \max\limits_i \max\limits_k \left|x_0(k) - x_i(k)\right|}$$

The correlation degree is :

$$r_i = \frac{1}{n} \sum_{k=1}^{n} \xi_i(k)$$

The correlation degree is given according to "normality", "symmetry", "entirety".

One of the most commonly used neural networks in remote sensing is the multi-layer perceptron(MLP).The back-propagation MLP algorithm described by Rumelhart et al(1986). A neural network consists of a number of interconnected nodes (equivalent to biological neurons).Each node is a simple processing element that responds to the weighted inputs it receives from other nodes. The arrangement of the nodes is referred to as the network architecture(figure 1).

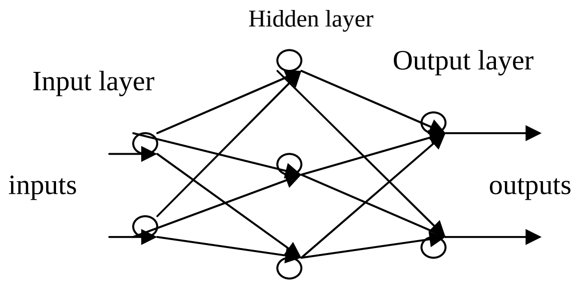

Figure 1. The MLP algorithm

In the MLP algorithm an input pattern is presented to the network via the input layer and the input signals are passed to the nodes in the next layer in a feed-forward manner. As the signal passes from node to node, it modified by the weights associated with the connection .The receiving node sums the weighted signals from all nodes to which it is connected in the preceding layer. Formally, the input that a single node receives is weighted according to :

$$net_j = \Sigma \omega_{ji} o_i$$

where ω_{ji} represents the weights between node *i* and node *j* ,and o_i is the output from node *i*. The output from a given node *j* is then computed from:

$$o_j = f(net_j)$$

The function f is usually a non-linear sigmoid function that is applied to the weighted sum of inputs before the signal passes to the next layer.

The aim of network training is to build a model of the data generating process so that the network can generalize and predict outputs from inputs that it has not seen before. For the MLP a training pattern is presented to the network and the signals are fed- forwards as described above. The weights of the connections are altered according to what is known as the generalized delta rule:

$$\Delta\omega_{ji}(n+1) = \eta(\delta_j o_i) + \alpha\Delta\omega_{ji}(n)$$

where η is the learning rate parameter, δ_j is an index of the rate of change of the error, and α is the momentum parameter. This process of feeding forward signals and back-propagating the error is repeated iteratively until the error of the network as a whole is minimized or reaches an acceptable magnitude. It is through the successive modification of the (adaptive)weights that the neural network is able to learn.

IF the hidden layers contain m nodes, train the network by training sample. After p times training, the output sequence of hidden layers is

$$o_i = (o_{i1}, o_{i2}, \cdots, o_{ip}) \qquad (i = 1,2,\cdots,m)$$

We regard the output o_i of i-th node in the hidden layer as the basic sequence and the others output o_j $(j = 1,2,\cdots,m, j \neq i)$ of j-th node as correlation sequence. According to above method, find the correlation degree r_{ij}. If the correlation degree $r_{ij} \geq c$ (c is a threshold), than we regard that output of i-th node and output of j-th node is near, and i-th node and j-th node combine to make a node so the number of nodes of hidden layer can decrease.

If s-th node and r-th rode of hidden layer combine to make q-th node, than weight corresponding q-th node is

$$w_{qj} = (w_{rj} + w_{sj})/2 \quad (j = 1,2,\cdots,n)$$ n is the number of nodes of input layer.

$$v_q = (v_r + v_s)/2$$ v_r is output of the r-th node , v_s is output of s-th node and v_q is output of q-th node.

By using above method, the architecture of the network can be simplified. We can gain a reasonable network by training and simplifying continually.

3 experiment and result

The prosed classification method is applied to Landsat Thematic Mapper(TM) data. An image sampled from the satellite Landsat-5 consists of TM band which includes 7 bands. They are from band 1 to band 7,but only bands 3,4,and5 are used. The others bands are omitted. The size of the study area is 256 by256 pixels. We define five categories: water; forest; farm, density building district; bare land. The number of nodes of the input layer is 3. The original number of the nodes of the hidden layers is 15 .The number of the nodes of the output layer is 5. We select 500 training samples. The network was trained for 26 000 iteration and the final number of nodes of the hidden layer is 8.The average classification accuracy was 82 percent.

4 Conclusions

We have presented a new method to find the number of the nodes of the hidden layer. The experimental results show that the method is effective.

Reference

[1] Atkinson,P.M.and Tatnall,A.R.L.,1997,Neural networks in remote sensing. International Journal of remote sensing,18,699-709.

[2] Givco,D.l.,1993,Artificial neural networks for land-cover classification and mapping. International Journal of geographical information system, 7,173-186.

[3] K.Funahashi,1998,Multi-layer neural networks and bayes decision theory, Neural network,11,209-213

[3] Paola,J.D. and Schowengerdt,R.A.,1995,A review and analysis of back-propagation neural networks for classification of remotely-sensed multi-spectral imagery. International Journal of remote sensing,16,3033-3058.

[4] Serpico,S.B.and Roli,F.,1995,Classification of multisensor remote-sensing images by structured neural network. I.E.E.E. Transaction on geoscience and remote sensing,33,562-578

[5] Zhuang, X., Engel,B.A.,Xiong,X. And Johannsen,C.J,1995,Analysis of classification results of remotely sensed data and evaluation of classification algorithms.Photogrammetric engineering and remote sensing,61,427-433.