CLASSIFICATION OF MULTI SPECTRAL IMAGES BY FUZZY AND NEURAL NETWORK APPROACHES

A.Milan lak[a], M.Hamrah[a] , M.GH. Majdabadi [b]

[a] K.N.T University, Tehran, Iran - Milan@ncc.neda.net.ir, Hamrah@ce.kntu.ac.ir
[b] National Cartographic Center, Tehran, Iran - Majdabad@ncc.neda.net.ir

Kew words: fuzzy classification, minimum distance, maximum likelihood

Abstract:

Extraction of information from satellite images, due to better cost & time consumption, all times accessibility and vast area of coverage, is seen as one solution for countries have no overall maps [Valadanzoj M,.Milan A, 1379].
One way of accessing this information is using the classification of satellite images. Nowadays, many experts are working on it and different approach are examined.
One of these approaches is the use of fuzzy technique and another is neural network approach to classify satellite images. In this paper we use these techniques. High quickness comparing existing approaches such as classification based on maximum likelihood and high accuracy are from its advantages.

1. INTRODUCTION

Using up to date and completed information of Earth surface either explanatory or topographic type is very necessary. The more accurate and up to date information from earth surface more reliable decisions and better results. Costs and speed of information generation are of high importance and the lower costs and faster operation leads to more usefulness information.
In up to dating maps, use of more speedy techniques which force less costs to process is very useful. Among other information, information of satellite images due to wide spectrum and wide bands of imagery has absorbed full attention.
Information extraction from satellite images is one of quick and simple ways for end users[Serpico,S.B.and Roli,F., 1995].
Extracting information from multi bands satellites has more importance because of high spectral resolution of them. This leads to accessing more information with less errors and better forms.
All of information computations in remote sensing is divided in two groups as below[Bezdek,J.C.,Enrlich,R.and Full,W.,FCM, 1984]:

1- Point operation such as Histogram modification
2- Neighbor hood operation such as
   a. Convolution
   b. Classification
   c. Fourier transform

In the first case pixels themselves will be discissed one by one but in the second case neighboring pixels are taking part in computationds.

According to these computations, feature classification is done. Classification is distinguishing and grouping features automatically and it is one of the best ways of getting information from satellite images specially for those of high resolution. To do so different algorithm such as Maximum likelihood and Minimum Distance is applied. Each algorithm has its own advantages and disadvantages. Among these algorithms, due to high flexibility of Maximum likelihood algorithm it shows better results but it takes more time and also it assign all pixels to one of existing class whereas it may be of none of them. Minimum Distance algorithm is one of the existing ones that is based on distance between each pixel and classes. This algorithm also assign each pixel to one class.

The advent of fuzzy set theory has comprised a lot of fields such as fuzzy control systems, fuzzy image processing and fuzzy classification of remotely sensed data.
Fuzzy classification, estimates the contribution of each class in the pixel and in computations, it assumes that a pixel is not solid but indecomposable unit in the image analysis and consequently works on a new principle: "one pixel – several class" to provide more information about the pixel [Cannon,R.L.,Dave,J.V.,Bezdek,J.C. and Trivedi,M.M 1986].
A fuzzy c-means clustering algorithms been developed by Bezdek for classification of satellite images [Wang,F., ,1990] and implemented by Cannon in unsupervised classification of remote sensing data [Wang, Lie-Xin].
Wang method modified the traditional maximum likelihood method by implementing a pre calculated fuzzy means and fuzzy covariance matrix [Melgani,F.,Bakir,A.R. and Saleem,M.R. 2000].
Using neural network in classification of RS images latterly is researched and a lot of works is being don on it. Learning possibility is one the important advantages of this approach which can be used in new feature detection.
In this paper we use an explicit fuzzy and neural classification methods for multi spectral remote sensing images and compare these two method with statistical methods.
After explaining the theory, the proposed methods will be applied on 3 bands of Landsat data, bands R,G and B and results will be compared with the minimum distance method.

2. MATHEMATICAL STRUCTURE OF APPLIED REASONING RULES

2.1. MIN reasoning rule

In MIN reasoning rule, we use reasoning based on separate principles with union, MIN Momdani necessity measure, MIN operator for all t-Norms and MAX operator for all s-Norms and we have in brief:

$$\mu_{B^{'}}(y)=\max_{l=1}^{M}\left[\sup_{x\in U}\min(\mu_{A^{'}}(x),\mu_{A_1^l}(x_1),....\mu_{A_n^l}(x_n),\mu_{B^l}(y)\right]$$

This means that operating MIN reasoning rule on fuzzy set A^1 will be lead to fuzzy set B^1.

2.2. Product Reasoning Rule

In product reasoning rule, we use reasoning based on separated principle in combination with union, product Momdani necessity measure, algebraic product for T-Norms and MAX operator for S-Norms and in brief we write the product inference engine as:

$$\mu_{B^1}(y)=\max_{l=1}^{M}\left[\sup(\mu_{A^1}(x)\prod_{i=1}^{n}\mu_{A_i^l}(x_i)\mu_{B^l}(y))\right]$$

This means that MIN inference engine produce B^1 fuzzy set from A^1 fuzzy set.

PRODUCT and MIN inference engine are the most important inference engines used in fuzzy system and control and easiness of computation of fuzzy is the most important advantage of them especially in PRODUCT inference engine.

3. EXPLICIT FUZZY CLASSIFICATION METHOD

As illustrated in fig1, fuzzy classification in first step use the Gaussian distribution fuzzyfication process to input data.

In next step MIN inference engine is used to classify . afterward again measurement is done. Finally a hard classification is provided by performing a MAX operation to defuzzyfy the fuzzy output into a hard output .

In fig1,block diagram of the explicit fuzzy classification illustrated.

As illustrated in fig1, fuzzy classification in first step use the Gaussian distribution fuzzyfication process to input data.

As pointed above, we have oriented the choice of membership function to the Gaussian distribution because it represent a powerful general distribution model.

The liner models are not able to express correctly the natural non linear distribution of classes and the class distributions are totally independent from each other. Consequently we can write the membership function of classes:

$$f_{b,c}(x_b)=\exp(-\frac{(x_b-\mu_{b,c})^2}{2*\delta^{*2}_{b,c}})$$

Where $\mu_{b,c}$ is the mean of class c in band b and x_b is the brightness value of the spectral pixel X in the band b. The pixel vector X in the B dimensional space is :

$$x_b=[x_1, x_2, ... x_b, ..., x_B]^T$$

And δ^* is the modulated standard deviation of class c in band b and is: $\delta^*=\alpha_{b,c}\delta_{b,c}$

Where $\delta_{b,c}$ is the standard deviation of class c in band b and $\alpha_{b,c}$ Is the modulation factor and it can be approximated by a linear function:

$$\alpha_{b,c}=0.6*P_{b,c}+0.22$$

Where $P_{b,c}$ is the expected extent of class c in band b

$$P_{b,c}=\frac{T_{b,c}}{\sum_{i=1}^{N}T_{b,i}}$$

Where N is the number of classes and T represents the expected number of pixels in the signature histogram, corresponding to mean …… of the class c then with Gaussian membership function and its characteristic for a multi spectral pixel X , the matrix of fuzzy inputs Fip can be written as:

$$F_{ip}=\begin{bmatrix} f_{1,1}(x_1) & f_{1,2}(x_1) & f_{1,N}(x_1) \\ f_{2,1}(x_2) & f_{2,2}(x_2) & f_{2,N}(x_2) \\ f_{B,1}(x_B) & f_{B,2}(x_B) & f_{B,N}(x_B) \end{bmatrix}$$

The dimension of this matrix is B * N where N is number of classes and B is number of bands. This matrix is input of inference engine and in the data analysis step is used.

4. INFERENCE ENGINE

In dealing with everyday problems, humans often employ a simple (but efficient) principle which can be summarized in these few words:

"work on the worst assumptions to find the best results".

The fuzzy reasoning rule that satisfies this principle is the MIN rule. On the other hand, the absence of covariance information makes the fuzzy partitions not necessarily optimal because classes with a high degree covariance will not be limited to the high fuzzy membership areas. Therefore high fuzzy membership values more closely represent a maximum possible value rather than a high prior probability of class membership. Applying inference fuzzy engine on matrix which is produced in fuzzyfication process , we obtain a primitive fuzzy output vector:

$$F^{'}op=[F_1^{'}(\underline{x}), F_2^{'}(\underline{x}),...., F_N^{'}(\underline{x})]^T$$

Where 1- using MIN fuzzy reasoning rule we have:

$$F_i^{'}=Min(f_{b,i}(x_b)) \quad b=1,2,.... B.$$

2- using product reasoning rule we have:

$$F_i^{'}=\prod(f_{b,i}(x_b)) \quad b=1,2,...., B$$

after this step , measurement of samples is done to improve the results in study (faze) and finally the vector fuzzy output will be as:

$$\underline{F}\ op=[F_1(\underline{x}), F_2(\underline{x}),....., F_N(\underline{x})]^T$$

where:

$$F_i(\underline{x})=\frac{Fi^{'}(\underline{x})}{\sum_{j=1}^{N}F^{'}j(\underline{x})}$$

5. Defuzzyfication

In the next step of fuzzy classification by performing a MAX operator, we defuzzyfy the fuzzy output to obtain a hard classification, we select among the classes mixed in the pixel the class c with the highest extent such that:

$$\forall\ i \in 1,2,\ldots,N\ ,\ i \neq c\ ,\ F_c(\underline{x}) \geq F_i(\underline{x})$$

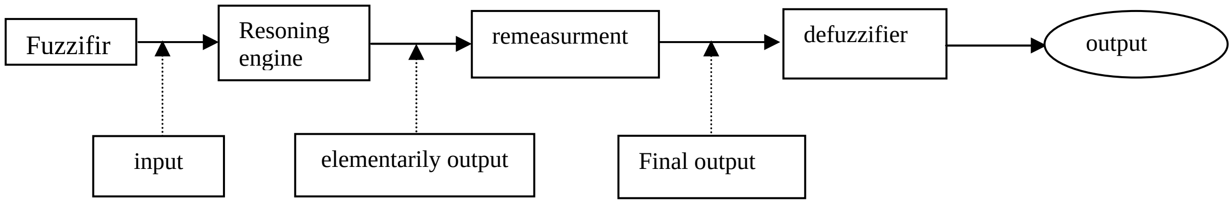

Fig 1- ,block diagram of the explicit fuzzy classification illustrated.

5. USING NEURAL NETWORKS

Artificial Neural Network can be seen as a new generation information processing systems that generally are based on using some architectural principles of human being brain[Lin and Lee 1966]. The main research topic in Neural Network domain is emphasis on modeling brain as a computing tools. Since ANNs are gotten from modeling of real biologic Neurons in brain, the processing elements are called neuron or artificial neuron as well. As mentioned above learning possibility is one the important advantages of this approach which can be used in new feature detection and classification. ANNs are consist of sets of related neurons thus each neuron associates to itself or others through weights of them. That is why sweeping architecture of neurons and geometry of relationship between them are of main parameter of ANNs systems definition. Fig 3 shows the five main geometry of relationship[Lin and Lee 1966].

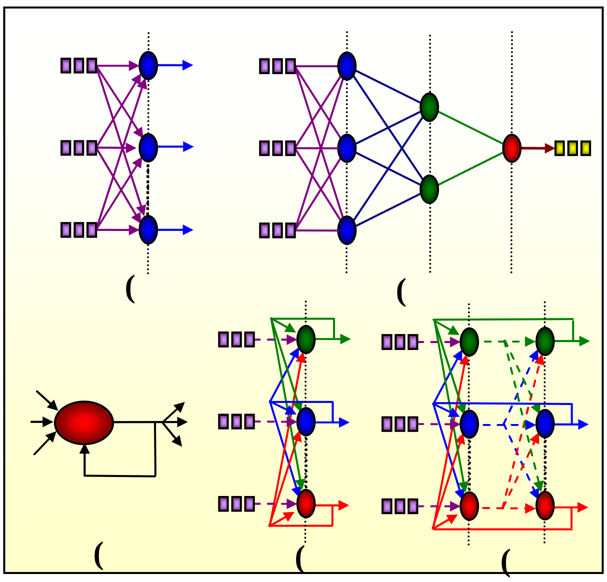

Fig2 . shows the synaptic architecture used in ANNs

We can imagine one layer of nodes consist of several nodes. And we can relate input with other nodes with different weights then forming one network of one layer feed forward (Fig2 a). we can relate several layer which leads to one multilayer feed forward network (fig 2 b). input layer is receiving layer of inputs and no computation except buffering of input signal is done by it. Outputs of network are provided by output layer. Every layer among input and output layer which have no direct relation with external media is called hidden layer. In a neural network we have from zero to several hidden layer.

5-1 Learning Principle in Neural Networks

one of the important characteristic of one Neural Network is learning procedure. Generally two type of learning is discussed in NNs. First is parameter learning which deals with updating joining weights in NNs and second is structure learning which deals with structure of network changes including number of nodes types of joints. These two type can be done simultaneously or independently by three different ways (fig 3) of supervised learning or reinforcement learning or unsupervised learning.

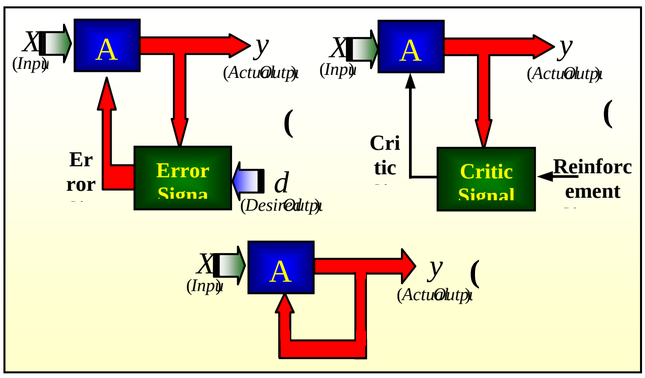

Fig 3. shows the different procedures of in ANNs[Lin and Lee 1996]

In supervised learning, in each epoch when inputting to ANN, system respond a desired corresponding response. In other word in supervised learning, one system is inputted by samples of couples of input-output like

$$(x^{(1)},d^{(1)}),(x^{(2)},d^{(2)}),\ldots,(x^{(k)},d^{(k)}),\ldots$$

When input is $x^{(k)}$ then the $d^{(k)}$ will be the output. As illustrated in figure 3 (a), the difference between actual output of $y^{(k)}$ and desired output of $d^{(k)}$ is measured in error producer signals and by defining error signals and correcting their weights this difference inclined to zero.

In supervised learning, assumption is that correct values of target for each input is defined but in some occasion there is very little information in access for example could be said output is very high or 50% is correct. Learning based on this type of information is called reinforcement learning and reinforcement signal. As illustrated in figure 3(b) reinforcement learning is a supervised learning why network get some feedback from media but reinforcement learning is just evaluative not advisor. It means that signal defines how much it is good or bad and there is no pointing to what should be the correct answer.

In unsupervised learning there is no trainer for feedback information. see figure 3(c) and network itself must find its archetype, characteristics, discipline, correlations and other needed issues and apply in output.

5-2 -Learning Principal in Feed Forwarded Neural Network Based on Back Propagation Approach

Among existing Neural Network models, Back propagating networks are the most important models which leads to spread the application of neural network in different disciplines. Back propagation network is a multi layer feed forwarded with a supervised learning algorithm from gradient descent which is named back propagation learning discipline[Lin and Lee 1996]. In this type of learning, existing error in output layer forwarded to the back in order to adjust the weights of previous layers and minimize the errors. In this approach adjustment of weights for minimizing value of one function is done by adjusting their values according to negative values of relative gradient which is called gradient descent. See Figure 5. if $w_{\otimes,\otimes,\otimes(l,n,i)}$ = weight of *i-th* of *n-th* neurons in layer *l* , $o_{\otimes,\otimes(l,n)}$ = *n-th* outputs of neurons in layer *l* and $Net_{\otimes,\otimes(l,n)}$ = weighted summation of all *n-th* input neurons in layer *l* , then we have

$$o_{l,n} = f\left(\sum_{i'=1}^{N_{l-1}} w_{l,n,i'} \cdot o_{l-1,i'} - \theta_{l,n}\right) = f\left(Net_{l,n}\right)$$

where $w_{l,n.i}$ express the weight of synapse connecting origin neuron of i to target neuron of *n* in layer *l* thus we have:

$$\Delta w_{l,n,i} = -\eta \frac{\partial E}{\partial w_{l,n,i}}$$

where η is a positive constant value which controls adjustment amounts. Adjusted weights of network in each iteration is computed by following formula:

$$w_{l,n,i}^{k+1} = w_{l,n,i}^{k} + \Delta w_{l,n,i}^{k}$$

$$\text{For Ourput Layer}: \left\{\Delta w_{l,n,i}^{k} = \eta \cdot o_{l,n} \cdot (1-o_{l,n}) \cdot o_{l-1,i} \cdot (t_{l,n} - o_{l,n})\right.$$

$$\text{For Hidden Layer}: \begin{cases} \Delta w_{l,n,i}^{k} = \eta \cdot o_{l,n} \cdot (1-o_{l,n}) \cdot o_{l-1,i} \cdot \sum_{n'=1}^{N_{l+1}} \delta_{l+1,n'} \cdot w_{l+1,n',n} \\ \therefore \delta_{l,n} = -\dfrac{\partial E}{\partial Net_{l,n}} \end{cases}$$

6. RESULTS

Fig 4 shows the study area through the three bands(bands 3,4 and 5) LandsatTM with spatial resolution of 30 * 30 m2. this image is a piece of one1141 * 961 pixels of Ahvaz region include three major covering classes which are river, barren lands, agricultural lands. To distinguish the classes in the picture a Mathlab program using fuzzy and neural network method is written.

In the first step for each class, we have selected a set of pixels according to our knowledge of region. Fig 5 shows the selected regions.

These parts were used as training site. The results of these methods were compared with the results of minimum distance method.

The program uses three methods for classification. The results of these three methods are shown in fig 6, 7, 8 and 9. according to these figures this can be deduced that fuzzy method has better accuracy than neural network and minimum distance methods in extracting river class. Also in extraction of barren land cover and agriculture land cover, fuzzy method show better accuracy in accordance with neural network and minimum distance methods.

Another point is the fuzzy method produce more smooth results but in minimum distance method , borders extracted more accurate. Comparing Minimum distance and Neural Network shows better accuracy of classification of Neural Network Method after Fuzzy method. Also in Fuzzy classification using product or minimum inference engine makes no difference. Also in a time point view, the methods are very near.

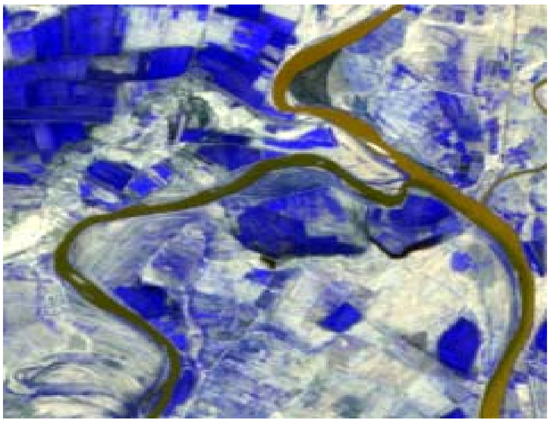

fig 4 –study area

Fig 5 shows different classes

Fig 6 shows the results of Fuzzy classification by MIN reasoning engine and fig 7 belongs to fuzzy classification by product reasoning engine and fig 8belongs to classification by Minimum distance method.

Fig 6- show the result of Fuzzy classification by Min reasoning engine

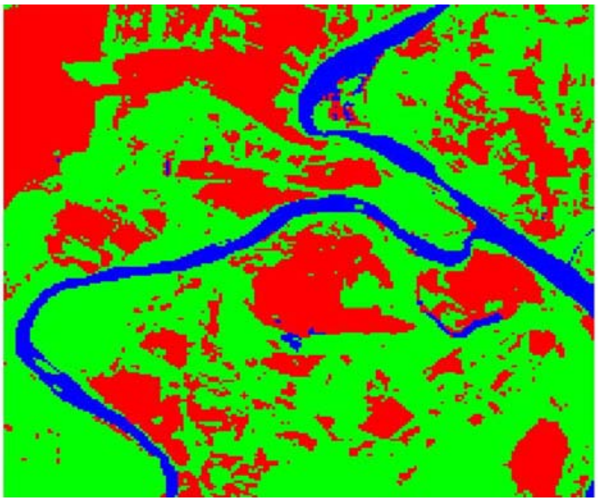

Fig 7- show the result of Fuzzy classification by proiduct reasoning engine

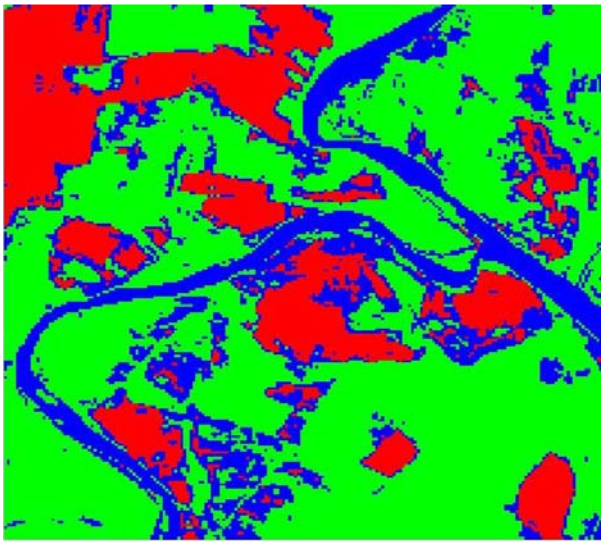

Fig 8- shows the result of to classification by Minimum distance method.

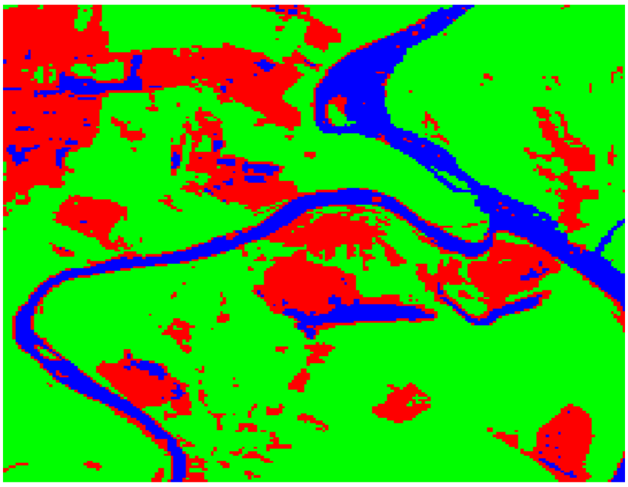

Fig 9- shows the result of to classification by Neural Network method.

7. CONCLUSION

We have compared in this paper methods of fuzzy classification and neural network classification with each other and both of them with Minimum distance method which reveals promising performance of fuzzy method particularly in time of classification.
Ability of these methods comparing statistical ones is explicitly visible. Continuity and power of Fuzzy method is of simplicity and optimum extraction of data. Another advantage of Fuzzy method is in deletion of one band and adding new band without any need to structural changes.
With advent of new generations of sensors providing broad collections of data (for example, HIRIS with 192 spectral bands) deliberation of using these sensors by these methods individually or in combination is proposed.

ACKNOWLEDGMENT

The authors would like to thank the National Cartographic Center of Iran(NCC) for providing the raw images.

REFERENCES

Bezdek,J.C.,Enrlich,R.and Full,W.,FCM, 1984, The fuzzy c-means clustering algorithm, Comp., vol.10,pp.191-203,1984

Cannon,R.L.,Dave,J.V.,Bezdek,J.C .and Trivedi,M.M 1986, Segmentation of a thematic mapper image using the fuzzy c-means clustering algorithm, Trans.Geosci.Remote Sensing,vol.GE-24,pp.400-408,May 1986

Lin C. T., Lee, C. G., 1996. *Neural Fuzzy Systems*. Prentice Hall Publication.Lowe, D.G., 1991. *Fitting parameterized three-dimensional models to images*, IEEE Transactions on Pattern Analysis and Machine Intelligence 13(5), 441-450

Melgani,F.,Bakir,A.R. and Saleem,M.R. 2000, ,An explicit fuzzy supervised classification method for multispectral remote sensing images,IEE Trans.Geosci.Remote Sensing, vol.38.NO.1. January 2000

Serpico,S.B.and Roli,F., 1995, Classification of multisensor remote-sensing images by structured neural networks,IEEE Trans.Geosci.Remote Sensing,vol.33,pp.562-578,May 1995.

Valadanzoj M,.Milan A, 1379, Evaluation of DLT model and orbital parameter model for geometer correction satellite imageries ,five civil engineer conference,pp.27-33,1379

Wang, Lie-Xin, "A course in fuzzy systems and control" , translated in Farsi

Wang,F., ,1990, ,Fuzzy supervised classification of remote sensing images,IEEE Trans. Geosci. Remote sensing,vol.28,pp.194-201,1990