HYPERSPECTRAL IMAGES FOR UNCERTAINTY INFORMATION INTERPRETATION BASED ON FUZZY CLUSTERING AND NEURAL NETWORK

Huan Li[a,c], Hongxia Luo[a, c]*, Ziyi Zhu[b], Guangpeng Liu[a]

[a]School of Geographical Sciences, Southwest University, 400715 Chongqing, China, P.R. - (lihuan84, tam_7236)@swu.edu.cn
[b]Computer and Information Science Faculty, Southwest University, 400715 Chongqing, China, P.R. - (zhu_ziyi)@163.com
[c]Key Laboratory of Eco-environments in Three Gorges Reservoir Region, Ministry of Education, Southwest University, Chongqing 400715,China - (tam_7236)@swu.edu.cn

Youth Forum

KEY WORDS: Uncertainty Information, Fuzzy Clustering, Feature Recognition, Hyperspectral Understanding, Neural Network

ABSTRACT:

Effective and understanding exploration of hyperspectral remote sensing data necessitates the development of sophisticated schemes that represent images. Such schemes ideally preserve and recognize significant features. However, uncertainty arises in classification problems when the input pattern is not perfect or measurement error is unavoidable. It would be need to obtain the estimation of uncertainty classification associated with a new observation and its membership within a particular class. Typically, existing methods supplying uncertainty information has monotonic neural network model, back propagation NN (BPNN) and the fuzzy membership model (FMM). The paper describes that an efficient combinational algorithm for uncertainty estimation on spectral dimension in hyperspectral remote sensing images is proposed for classification accuracy improvement and computing efficiency. The combination of fuzzy clustering and neural network adopted, which input patterns are divided into several small neural networks based on fuzzy clustering, is provides the classification boundaries based on the degree-of-dissimilarity measurement of the input pattern associated with each classification class. And we proposed the neural network with dynamical neuron created during learning NN algorithm. In the experiment, We tested three methods for misclassification using the same data and compared the performances with our method.

1. INTRODUCTION

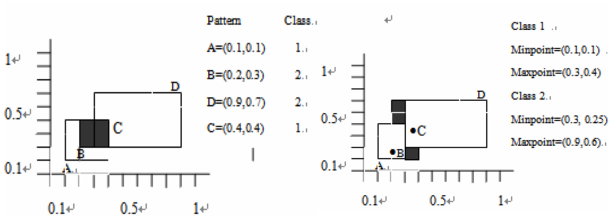
Uncertainty estimation plays an important role in data classification. Over the past 30 years, many different data classification techniques have been proposed, motivated by the increased applications of human brain intelligent requiring the recognition of uncertainty information (Bezdek, J.C., 1993; Kynan, E., 2007; Tan, X., 2005; Valentin, D., 2005; Wu, X., 2000). One technique that has been used for many classification problems is the artificial neural network(ANN) (Jain, A.K, Duin, R.P.W. and Mao, J., 2000). ANNs, as new means of implementing various classifiers consisted of a massively parallel distribution of neurons with many interconnections, have been proven to be suitable for classification because of the ability to learn from representative pattern data, good generalization, and highly nonlinear decision boundaries(Haertel V.F., 2005). These boundaries may be determined according to some predefined measures, such as minimizing the misclassification rate(Archer, N.P., 1993; Wang, S., 1993). In traditional classification, using ANN, the number of output nodes corresponds to the number of pattern classes and, during training, the output node corresponding to the class of the training pattern vector is clamped at '1', which all other output nodes are clamped at '0'(Pal, S. K. ; Mitra, S., 1992).

However, in many hazardous situations, classes are often fuzzy, ambiguity, or ill defined, therefore, traditional classifier may not provide an adequate representation of the relationship between a pattern vector and its "belongingness" to a particular class. There are several methods that have been used to deal with uncertainty, and represent alternatives with respect to ambiguous or diffuse evidence. These methods include expert-system, neural networks, fuzzy set theory, belief function theory, and Bayesian networks.

A further method FMNN proposed provided uncertainty information in the form of fuzzy measures. This is achieved by removing the error process considering an emerging consensus that the contraction procedure is best avoided in min-max classification (Bargiela, A., 2003; Joshi, A,1997; Simpson, P. K., 1993). An example of a two-class classification situation illustrated in Figure 1, in which the N-dimensional pattern vectors share an overlapping removing process in NN. Figure 1(a) is used for training FMNN classifier, two hyper boxes are created with an overlap.
To remove the overlap, hyper boxes are contracted and the results are depicted in Figure 1(b) . Note that after contracted training sample B and C are contained in the hyper boxes of class 1 and 2, respectively.

*Corresponding author : Hongxia Luo; tam_7236@swu.edu.cn; phone +86-23-68382061;

Natural geography doctorate opening funds(SUNUB2005036)



(a)Overlap case (b) overlap removing process in

Figure 1. Point B is classified as class 1 and similarly point C gets full membership in class 2, which causes a classification error.

To tackle the uncertainty, we introduce a concept of active neuron and complementation neuron, which are created dynamically during learning process. These dynamical neuron control membership in the overlapped region maintain high dimension and increase robustness in data classification. The dynamical and complementation neuron active only if the test sample falls in the overlapped region of the two patterns representing different classes, when there is an ambiguity about deciding the class or the test data. The proposed algorithm minimizes errors in the learning process by removing the contraction process form learning algorithm.

The paper is organized as follows. Section 2 analyses hyperspectral data handling errors in classification and elaborates the architecture of fuzzy clustering and neural network with neuron. Section 3 provides a new learning algorithm for uncertainty estimation, shown experiment on the classification of hyperspectral data and a comparison of BPNN and FMM. Section 4 concludes the work with summary.

2. ARCHITECTURE OF FUZZY CLUSTERING AND NEURAL NETWORK WITH DYNAMICAL NEURON

2.1 Analysis Data Error

It's analyzed that process involved in the learning algorithm modifies these min-max points to remove ambiguity in the overlapped classes. The classification error for the training data due to contraction process is discussion in details. Figure 3 and 4 illustrate the problem of hyper box full and partial containments.

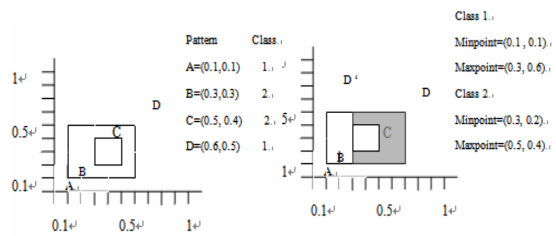


Figure2. Hyper box full containments problem.

For full and partial containment, FMNN solves the problem using contraction as figure 2 and 3. After contraction, the max point of class 1 hyper box is changed from D to D'. In other words, due to contraction, hyper box min-max points represent the acquired knowledge, which is tampered and may lead to gradation or classification errors.

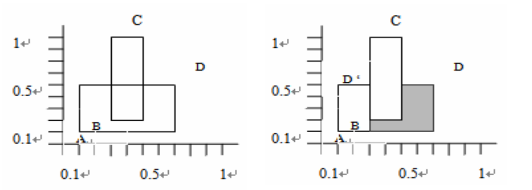


Figure 3. Hyper box partial containments problem.

2.2 Fuzzy Clustering

Fuzzy clustering is a data clustering algorithm in which each data point is associated with a cluster through a membership degree(Lu Jiaming; Yuan Xue; Yahagi Takashi., 2006). It is assumed that fuzzy clustering in which each similar data points is defined by having high membership degrees in the same cluster and the all chosen degrees-membership sum is one. Therefore, fuzzy clustering is a partition of the set of data and a collection N data points can be divided into γ fuzzy clusters, which are described by an matrix $U(\mu_{ik})$, where μ_{ik} is the membership between zero and one, and the sum of $U\mu_{ik}\times\gamma_j$ in each column j (form 1 to N) is one. Fuzzy membership is to supply uncertainty information for multiclass classification. In this paper, membership degree is calculated using Manhattan, Euclidean, Chebychew *L(∞)*, Cosine distance metrics and the Pearson correlation coefficient, in which the function are given in section 3. A fuzzy of input feature vector $X=\{x_1, x_2, ..., x_N\}$ is represented by a matrix $U(\mu_{ik})$. The μ_{ik} satisfy the constraints:

$$\mu_{ik}\in[0,1] \qquad 1\le\mu_{ik}\le\gamma \tag{1}$$

$$\sum_{i=1}^{N}\mu_{ik}=1 \qquad 1\le k\le N \tag{2}$$

$$0\le\sum_{i=1}^{N}\mu_{ik}\le N \qquad 1\le i\le\gamma \tag{3}$$

where N is a collection of N data points, γ is fuzzy groups. Then, the dissimilarity measure which is a cluster centre is minimized.

2.3 Fuzzy and Neural Network with Dynamical Neuron Architecture

The neural network is composed of three-layer simulated input neuron, hidden neuron and dynamical neuron. The number of hidden units 2, 3 and 4 were selected by sixfold cross validation from 6 to 300 units based on the correct classification rate and distance metric(Setinono, R., 2001). The architecture of FMNN is shown in figure 4. The middle layer neurons and output layer nodes are partitioned into two parts: classifying neuron and dynamical neuron. The classifying and dynamical neuron is decide membership in the uncertainty information of classification. All middle layer neurons are created during the training process.

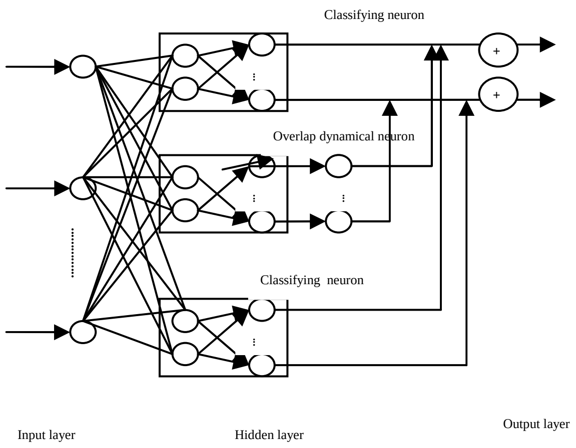


Figure 4. Structure of proposed Architecture

2.4 Algorithm of Matrix U/Membership Degree

The membership degree is calculated using formula (4)~(8), so the algorithm consists of a series of iterations. The algorithm converges to minimized the dissimilarity measure which is defined as follows.

Step 1) Initially. According to the constraints of matrix *U*, the *U* is constructed randomly between 0 and 1 that satisfied (1)~(3).

Step 2) Computed membership function. For each cluster i, those formulas are computed respectively. Stop if its improvement over the previous iteration below a threshold.

Than a new *U* is computed and repeated Step 2)

Step 3) a. Evaluate the classification boundaries based on defuzzification of matrix U .
b. Find misclassification sets for each class $C_{\lambda 1 mis}, C_{\lambda 2 mis} \ldots C_{\lambda k mis}$
c. Calculated the membership value for the misclassified individuals in each class according to the fuzzy membership function. The number of membership function is decreased based on dufuzzification. The network should provide more information about uncertainty in classification problems.

3. ALGORITHM AND EXPERIMENTAL RESULTS

Experiments have been carried out using MODIS, which choose 100 individuals in Figure 5, Wuhan university, China (50 individuals were known and 50 individuals were unknown). Each known individual provided 12 frontal samples which show different features, which 6 samples were selected as the training set, and 6 samples were used the test pattern for recognition. The image was cropped and rescaled covering 0.459-2.135 um wavelength with 250~1000m spatial resolution in 36 spectral bands. Five classes, evergreen, serpentine, green-strone, chaparral and soil were selected. The implementation relied on Matlab including the optimization toolbox and ANN.

3.1 Pre-processing Image

MODIS image was re-projected from the native SIN projection to a UTM-WGS84 reference system and resized on the study area by means of the "MODIS Reprojection Tool" software(Andrew, J. Elmore, et al., 2000). MODIS was assumed to be well co-registered, so that no further geometric correction was carried out. Cloud-contaminated pixels were selected by fixing an arbitrary threshold on Quality Assurance values(Cloud State > 0) for daily images, usefulness index > 3 for composite images(L. Busetto, M. Meroni, R. Colombo., 2008). Due to the high dimensionality of the data, principal component analysis(PCA) approach was used to reduce dimensionality. For PCA method, the image space were projected to a 32-dimensional feature subspace for classification.

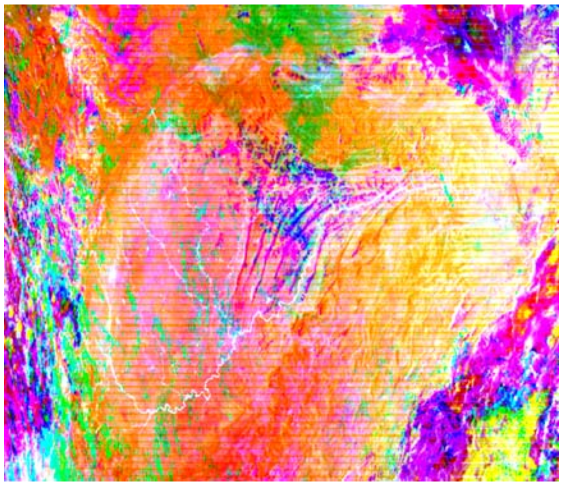


Figure 5. MODIS image

3.2 Evaluated Each Class Similarity

It is proposed the relationship function between each input *X* and each of χ_c classes. In other words, relationship function *F* provide a measure of similarity between the input and each of classification class. So, the functions Euclidean distance was used, which given in formulation (4)~(8).

$$f_{Manhat\tan}(x_i) = 1 - \delta\left(\frac{1}{N}\sum_{n=1}^{N}\left|x_{in} - \overline{x}_{in}\right|\right) \tag{4}$$

$$f_{Euclidean}(x_i) = 1 - \delta\left(\frac{1}{N}\sum_{n=1}^{N}\sqrt{(x_{in} - \overline{x}_{in})^2}\right) \tag{5}$$

$$f_{Chebychew}(x_i) = 1 - \left(\max_{n=1}^{N}\left|x_{in} - \overline{x}_{in}\right|\right) \tag{6}$$

$$f_{Co\sin e}(x_i) = 1 + \delta\left(\frac{\sum_{n=1}^{N}(x_{in} \cdot \overline{x}_{in})}{\sqrt{\sum_{n=1}^{N} x_{in}^{2}} \bullet \sqrt{\sum_{n=1}^{N} \overline{x}_{in}^{2}}} - 1\right) \tag{7}$$

$$f_{PearsonCorrelation}(x_i) = \left(\frac{\sum_{n=1}^{N} (x_{in} - \frac{1}{N}\sum_{n=1}^{N} x_{in})(\overline{x}_{in} - \frac{1}{N}\sum_{n=1}^{N} \overline{x}_{in})}{\sqrt{(\sum_{n=1}^{N} x_{in} - \frac{1}{N}\sum_{n=1}^{N} x_{in})^2} \bullet \sqrt{(\sum_{n=1}^{N} x_{in} - \frac{1}{N}\sum_{n=1}^{N} \overline{x}_{in})^2}} \right) \quad (8)$$

where δ is a normalizing parameter ($0 < f(x_i) < 1$), and is evaluated 3.90, 9.90, 5.96 for Manhattan, Euclidean, Cosine measures, respectively. x_{in} is coefficient of input X and $\overline{x}_{in}$ is mean coefficient, n = 1,2 … N. Lowering the threshold value raises the correct classification rate but lowers easily causing unknown to be judged as known class. In contrast, raising the threshold cause unknown class to be judged known. The similarity threshold value is set to 0.96 in our experiment, which can be achieved the best performance.

3.3 Computation Defuzzification

According to the fuzzy clustering algorithm in section 2, X individuals divided into $\chi_c \times X$ matrix U, where the i , kth entry μ_{ik} is the membership between 0 and 1, and the sum of the entries in each column is one, and the last number of clusters is ten. The lines of matrix U determine how many individuals one cluster may contain. If the μ_{ik} that is not zero than it is saved, and it's to guarantee that an μ_{ik} element belongs to at least one class/ cluster. The computation defuzzifiication of U using data fuzzy clustering is shown in Table 1. We preformed a experiment to determine the value of max cluster number. When the maximum number of cluster member is set to 8 ~10, the algorithm achieved higher correction classification rate in Figure 6. The fuzzy membership function for class evergreen was presented in Figure 7.

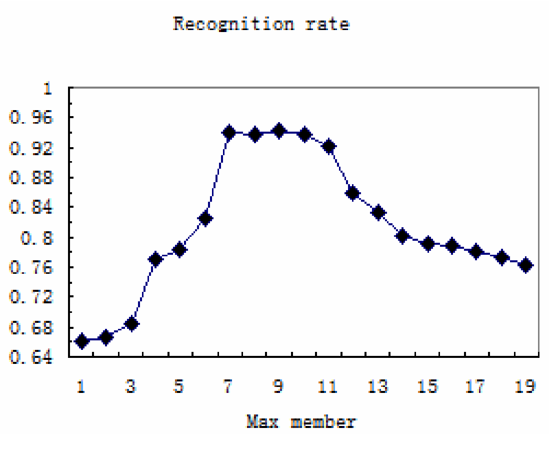


Figure 6. Numbers of Max cluster as correction classification rate

Class/Cluster	Cluster Member
Evergreen	1, 7, 8, 12, 13, 16, 21
Grass	3, 4, 6, 10, 11, 30, 32, 33, 38, 40, 48, 49
Green-strone	5, 9, 14, 31, 37, 45, 46, 47, 50
Chaparral	2, 15, 17, 20, 22, 23, 27, 28, 29,36,41, 44
Soil	18, 19, 24, 25, 26, 34, 35, 39

Table 1. Determine matrix U using computation defuzzification

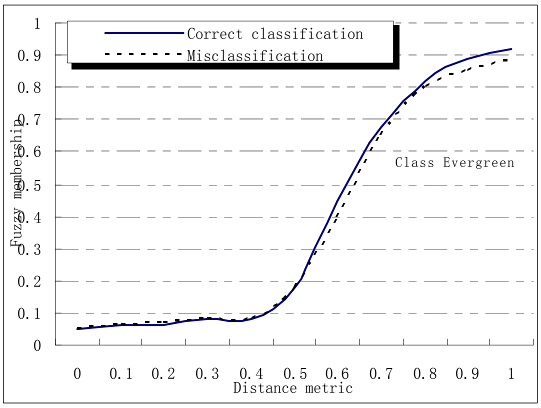


Figure 7. Fuzzy membership function for class evergreen

3.4 Classification by Five Distance Metric

The possibility networks which were trained under the monotonic constraint, which was evaluated during training according to the formula (9) (Wang S., 1994)

$$\sum_{h=1}^{H} w_{hi} v_h e^{-whf(xi)} [1 + e^{-whf(xi)}] \geq 0 \qquad i = 1, 2, \ldots N. \qquad (9)$$

where W_h =[W_{h1} , W_{h2} ,…, W_{hL}] were the weight matrices for the connections between the input and hidden nodes, $f(x)$ =[$f(x1)$, $f(x2)$,.. $f(xN)$] were input vector.

The initial network weights had a uniform, random distribution between -0.3 and 0.3, and training was undertaken the constraint formula (9), the training rate was initially η = 0.2 for each sample, and was each interation in which a sample did not conform to the monotonic constraint(Archer, N.P. and Wang, S., 1993). The possibility networks NN were trained according to 50 training pattern vectors, until the mean absolute error for the possibility networks reduced to a value of 0.005. The number of neurons in the hidden layer H was vary set to 2, 3 and 4. The classification percentage performance was shown Table 2.

Distance Metric	Hidden Neuron (H)		
	2	3	4
Manhattan	93.3	93.2	93.3
Euclidean	95.6	96.85	96.76
Chebychev	80.00	83.20	83.36
Cosine	91.20	91.20	91.10
Correlation	91.10	92.00	91.00

Table 2. Classification percentage of the training data (Deviation is defined as $\eta_{x+1} = (1 - 0.01)\eta_x$)

The weights were adapted according to the minus gradient of the squared Manhattan, Euclidean, Cosine, Chebychev and Pearson Correlation distance between the desired and obtained outputs. The Chebychev distance metric had a significantly lower classification performance, so the choice of its to reduce the computational complexity is unjustified for the performances. It was found that the choice of the Euclidean distance had higher classification performance and lower computational complexity. It was also found that the training methods had only a small influence of the classification application.

3.5 Learning Neuron Network Algorithm

In learning process, we experiment the individuals 1, 7, 8, 12, 13, 16, 20, 21 data in the evergreen class. For 50 known individual in MODIS data, 600 sample are used for recognition certainty. And the 600 sample are also prepare for 50 unknown individual to determine whether the FNN can estimate the uncertainty information. After learning algorithm, one evergreen clustering class output results is presented in Table 3.

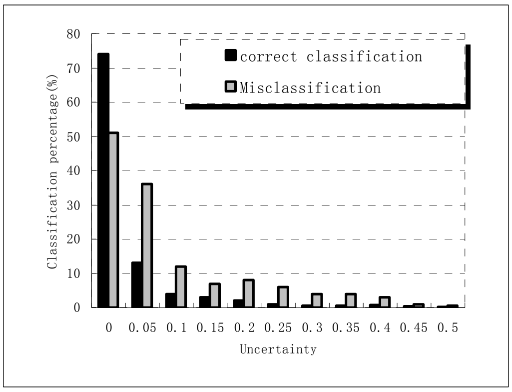


Figure 8. Classification uncertainty measure for test data

Cluster member	sample	output of						
1	traninig1	0.9900	0.0147	0.0100	0.0100	0.0127	0.0160	0.0100
	.	0.9809	0.0137	0.0174	0.0157	0.0010	0.0190	0.0186
	.	0.9756	0.0149	0.0100	0.0205	0.0010	0.0103	0.0110
	.			.	.	.		
	.			.	.	.		
	.	0.9696	0.0200	0.0180	0.0264	0.0159	0.0225	0.0165
	testing12	0.9769	0.0110	0.0100	0.0100	0.0130	0.0127	0.0100
7	traninig1	0.0146	0.9689	0.0199	0.0100	0.0100	0.0100	0.0209
	.	0.0130	0.9686	0.0100	0.0100	0.0100	0.0098	0.0205
	.							
	.			.	.	.		
	.			.	.	.		
	.							
	testing12	0.0199	0.9742	0.0100	0.0100	0.0100	0.0100	0.0100
8	traninig1	0.0100	0.0100	0.9759	0.0100	0.0126	0.0169	0.0126
	.	0.0186	0.0189	0.9768	0.0100	0.0100	0.0100	0.0100
	.							
	.			.	.	.		
	.			.	.	.		
	.							
	testing12	0.0124	0.0159	0.9769	0.0100	0.0224	0.0100	0.0111
12	traninig1	0.0124	0.0100	0.0178	0.9812	0.0100	0.0126	0.0190
	.	0.0226	0.0100	0.0132	0.9713	0.0100	0.0269	0.0236
	.							
	.			.	.	.		
	.			.	.	.		
	.							
	testing12	0.0100	0.0100	0.0100	0.9728	0.0100	0.0100	0.0100
13	traninig1	0.0191	0.0221	0.0222	0.0112	0.9758	0.0100	0.0100
	.	0.1130	0.1230	0.0100	0.0100	0.9776	0.0164	0.0198
	.							
	.			.	.	.		
	.			.	.	.		
	.							
	testing12	0.0100	0.0100	0.0100	0.0100	0.9900	0.0125	0.0100
16	traninig1	0.0191	0.0226	0.0100	0.0140	0.0100	0.9900	0.0100
	.	0.1152	0.0100	0.0220	0.0286	0.0100	0.9800	0.0100
	.							
	.			.	.	.		
	.			.	.	.		
	.							
	testing12	0.0100	0.0140	0.0100	0.0120	0.0125	0.9758	0.0100
20	traninig1	0.0143	0.0203	0.0100	0.0203	0.0185	0.0100	0.9560
	.	0.0103	0.0100	0.0100	0.0100	0.0100	0.0120	0.9700
	.							
	.			.	.	.		
	.			.	.	.		
	.							
	testing12	0.0100	0.0100	0.0110	0.0120	0.0100	0.0100	0.9810

Table 3. In learning produce evergreen class correct classification results

To illustrate the generalized performance of the NN with dynamical neuron, the uncertainty information for the correctly classified and misclassified test pattern vectors is presented in Figure 8. The number of test pattern is 50 and chooses the Euclidean distance measure according to section 3.4 experiment. It can be see that the test patterns were misclassified had a large uncertainty compare to the correctly classified and there was an increase in the uncertainty information for misclassified test patterns.

3.6 Comparison

3.6.1 Comparison with Traditional Classifiers

For training and testing data, Evergreen class was randomly selected 50 training samples and other remaining was testing data. When the coefficient is set to 0.04, no training and testing errors were observed for the proposed approach shown in Table. 4. Table 4 illustrated that fuzzy cluster and neural network with dynamical neuron classifies all the data correctly.

The number of test pattern is 50 and chooses the Euclidean distance measure according to section 3.4 experiment. It can be see that the test patterns were misclassified had a large uncertainty compare to the correctly classified and there was an increase in the uncertainty information for misclassified test patterns.

Method	Mis-classification	Error rate
k nearest neighbourhood	3	3.91
Bayes classifier	6	8.2
Back propagation NN	2	6.25
Fuzzy membership model	4	4.17
Proposed Method	1	1.67

Table 4. illustrated that fuzzy cluster and neural network with dynamical neuron classification

4. DISCUSSION AND CONCLUSION

In this paper, an efficient approach for uncertainty information was presented. In order to assess this system, we tested four existing approaches for uncertainty estimation using the same data and compared the performances with our method. The four method were *k* nearest neighbourhood, Bayes classifier , Back propagation NN, Fuzzy membership model, which carried out using patterns of 100 individuals (50 individuals were known and 50 individuals were unknown) as the same group in the our research. In the analysis, it was shown that it was possible to use neural network models, trained under the NN constraint, to represent the degree-of-membership each new observation for each class, so that a measure of uncertainty due to misclassification could be obtained. The error analysis of learning algorithms cause errors in training phase based on the principle of minimal disturbance. The dynamical neuron can handle correct classification and misclassification more efficiently and is capable to approximate the complex data more accurately.

REFERENCES

Archer, N.P. and Wang, S., 1993. Application of the back propagation neural network algorithm with monotonicity constraints for two-group classification problems. *Decision Sci.,* vol. 24, no. 1, pp. 60-75.

Archer, N.P. and Wang, S., 1991. Fuzzy set representation of neural network classification boundaries. *IEEE Trans. Syst.*vol. 21, no.4, pp. 735-742.

Andrew, J. Elmore, et al., 2000. Quantifying vegetating change in Semiarid environment : precision and accuracy of spectral mixture analysis and the normalized difference vegetation index. *J. Remote Sensing of Environment,* vol. 73, pp. 87-102.

Bezdek, J.C., 1993. A review of probabilistic, fuzzy, and neural models for pattern recognition. *J. Intell. Fuzzy Syst.*, vol. 1, no. 1, pp.1-25.

Bargiela, A., Pedrycz, W. and Tanaka, M., 2003. Exclusion/inclusion fuzzy classification network. *Proc. Knowledge-Based Intell. Eng. Syst. Conf.* Oxford, U.K., pp. 1236-1241.

Haertel V.F., Shimabukuro, Y.E., 2005. Spectral linear mixing model in low spatial resolution image data. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, pp. 2555-2562.

Jain, A.K, Duin, R.P.W. and Mao, J., 2000. Statistical pattern recognition: a review. *IEEE Trans. Pattern Anal. Machi. Intell.*, vol.22, no.1, pp.4-37.

Joshi, A., Ramakrishman, N., Houstis, E.N. and Rice, J.R., 1997. On neuro-biological, neuro-fuzzy, machine learning, and statistical pattern recognition techniques. *IEEE Trans. Neural Netw.*, vol. 8, no. 1, pp.18-31.

Kynan, E., 2007. Uncertainty estimation using fuzzy measures for multiclass classification. *IEEE Trans. Neural Netw.*, vol. 18, no. 1, pp.128-139.

Lu Jiaming, Yuan Xue, and Yahagi Takashi., 2006. A method of face recognition based on fuzzy c-means clustering and associated sub-NNs. *IEEE Trans. Neural Netw.*, vol. 18, NO. 1, pp. 150-159.

Lorenzo Busetto, Michele Meroni, Roberto Colombo., 2008. Combining medium and coarse spatial resolution satellite data to improve the estimation of sub-pixel NDVI time series. *Remote Sensing of Environment*, 112, pp. 118-131.

Pal, S. K. and Mitra, S., 1992. Multilayer perceptron, fuzzy sets, and classification. *IEEE Trans. Neural Netw.*, vol. 3, no. 5, pp. 683-697.

Simpson, P. K., 1993. Fuzzy min-max neural network-Part Ⅱ: Clustering. *IEEE Trans. Neural Netw.*, vol. 1, no. 1, pp.32-45.

Setinono, R., 2001. Feedforward neural network construction using cross validation. *Neural comput*., vol. 13, pp. 2865-2877.

Tan, X., Chen, S., Zhou Z. H. and Zhang, F., 2005. Recognizing partially occluded, expression variant faces from single training image per person with SOM and soft k-NN ensemble. *IEEE Trans. Neural Netw.*, vol.16, no. 4, pp. 875-886.

Valentin, D., Abdi, H., O'Toole, A. J.O. and Cottrell G. W., 1994. Connectionist models of face processing: A survey. *Pattern Recognit.*, vol. 27, no. 9, pp. 1209-1230.

Wu, X. and Er, M. J., 2000. Dynamic fuzzy neural networks: A novel approach to function approximation. *IEEE Trans. Syst., Man, Cybern., B.Cybern.*, vol.30, no.2, pp. 358-364.

Wang S., 1994. Generating fuzzy membership function a monotonic neural network model. *Fuzzy Sets Syst.*, Vol. 61, pp. 71-81.

Xie, X., Sudhakar, R. and Zhuang, H., 1993. Corner dectection by a cost minimization approach. *Pattern Recognit.*, vol. 26, no. 12, pp. 1235-1243.

Yuan, J. L. and Fine, T. L., 1998. Neural-network design for small training sets of high dimension. *IEEE Trans. Neural Netw.*, vol. 9, no. 1, pp. 266-280.

ACKNOWLEDGMENT

I would like to thanks my advisor professor Hongxia Luo for providing valuable insightful suggestions and encouragement. The research discussed in the paper was supported by Natural Geography Doctorate Opening Funds (SWNUB2005036) and Doctoral Funds Administered by Southwest University.