Nonparametric Neural Network Model Based on Rough-Fuzzy Membership Function for Classification of Remotely Sensed Images

Niraj Kumar and Anupam Agrawal

Indian Institute of Information Technology, Allahabad – 211 011

Abstract. A nonparametric neural network model based on Rough-Fuzzy Membership function, multilayer perceptron, and back-propagation algorithm is described. The described model is capable to deal with rough uncertainty as well as fuzzy uncertainty associated with classification of remotely sensed multi-spectral images. The input vector consists of membership values to linguistic properties while the output vector is defined in terms of rough fuzzy class membership values. This allows efficient modeling of indiscernibility and fuzziness between patterns by appropriate weights being assigned to the back-propagated errors depending upon the Rough-Fuzzy Membership values at the corresponding outputs. The effectiveness of the model is demonstrated on classification problem of IRS-P6 LISS IV images of Allahabad area. The results are compared with statistical (Minimum Distance), conventional MLP, and FMLP models.

1 Introduction

Geospatial information, we gather through different sensors and from the concepts of the geographical objects, is generally vague, imprecise and uncertain. Also, the imprecision becomes obvious due to the multi-granular structure of the multi-sensor satellite images and that leads to error accumulation at every stage in geo-processing. It has been observed that the ground truth data, an essential ingredient for a supervised learning, may itself contain redundant, inconsistent, conflicting information.

The geospatial information is received in different windows of the electromagnetic spectrum and at different resolutions. This presents selective look of the geospatial objects under view of the satellite sensor. Therefore, the totality of capturing the truth or facets of the objects seems to be very difficult. This implies that at a given set of parameters of observation, we have limited capability to discern two objects. It is equivalent to say that the knowledge generated from the satellite image at a given resolution and spectrum band, is granular. It is, therefore, imperative to have more observational parameters to decompose this granule, i.e. to obtain finer view of the objects. The effect is that based on the observational parameters, any two objects, may appear same, whereas, the ground truths about the objects force us to have different opinion on them. This phenomenon introduces the rough uncertainty into the

P. Kalra and S. Peleg (Eds.): ICVGIP 2006, LNCS 4338, pp. 106 – 117, 2006.

information system due to imprecision inducted by the observation system. Since the boundaries of various land covers in satellite image are not precise, so fuzzy uncertainty is also associated here.

After the Fuzzy Set theory [4], the Rough Set theory proposed by Z. Pawlak [1], has emerged as another major mathematical approach for managing uncertainty that arises from inexact, noisy, or incomplete information. The focus of rough set theory is on the ambiguity caused by limited discenibility of objects in the domain of discourse.

In this paper, we have attempted to integrate rough sets, fuzzy sets, and artificial neural network (ANN) for designing a nonparametric rough fuzzy neural network model to deal with indiscernibility and fuzziness between patterns. Here we have used the generalized concept of rough membership function in pattern classification tasks to Rough-Fuzzy Membership functions to deal with rough uncertainty [9] in geospatial information gathered by satellites and in ground truth data. Unlike the rough membership value of a pattern, which is sensitive only towards the rough uncertainty associated with the pattern, the rough-fuzzy membership value of the pattern signifies the rough uncertainty as well as the fuzzy uncertainty associated with the pattern.

2 Background

2.1 Rough Sets

Let R be an equivalence relation on a universal set X. Moreover, let X/R denote the family of all equivalence classes introduced on X by R. One such equivalence class in X/R, that contains $x \in X$, is designed by $[x]_R$. For any output class $A \subseteq X$, we can define the lower $\underline{R}(A)$ and upper $\overline{R}(A)$ approximation which approaches A as closely as possibly from inside and outside respectively [9]. Here

$$\underline{R}(A) = \bigcup\left\{[x]_R \mid [x]_R \subseteq A, x \in X\right\} \quad \text{(1-a)}$$

is the union of all equivalence classes in X/R that are contained in A, and

$$\overline{R}(A) = \bigcup\left\{[x]_R \mid [x]_R \cap A \neq \phi, x \in X\right\} \quad \text{(1-b)}$$

is the union of all equivalence classes in X/R that overlap with A. A rough set $R(A) = \left\langle \overline{R}(A), \underline{R}(A) \right\rangle$ is a representation of the given set A by $\overline{R}(A)$ and $\underline{R}(A)$. The set $BN(A) = \overline{R}(A) - \underline{R}(A)$ is a rough description of the boundary of A by the equivalence classes of X/R. The approximation is rough uncertainty free if $\overline{R}(A) = \underline{R}(A)$. Thus, when all the patterns from an equivalence class do not carry the same output class labels, rough ambiguity is generated as a manifestation of the one-to-many relationship between that equivalence class and the output class labels.

The rough membership function $r_A(x): A \to [0,1]$ of a pattern $x \in X$ in the output class A is defined by

$$r_A(x) = \frac{\|[x]_R \cap A\|}{\|[x]_R\|} \tag{2}$$

where $\|A\|$ denotes the cardinality of the set A.

2.2 Fuzzy Sets

In traditional two-state classifiers, where a class A is defined as a subset of a universal set X, any input pattern $x \in X$ can either be a member or not be a member of the given class A. This property of whether or not a pattern x of the universal set belongs to the class A can be defined by a characteristic function $\mu_A(x): X \to \{0,1\}$ as follows

$$\mu_A(x) = \begin{cases} 1 & \text{if and only if } x \in A \\ 0 & \text{Otherwise} \end{cases} \tag{3}$$

In real life situations, however, boundaries between the classes may be overlapping. Hence, it is uncertain whether an input pattern belongs totally to the class A. To take care of such situations, in fuzzy sets the concept of characteristic function has been modified to membership function $\mu_A(x): X \to [0,1]$. This function is called membership function, because larger value of the function denotes more membership of the element to the set under consideration.

2.3 Rough Fuzzy Sets

Let X is a set, R is an equivalence relation defined on X and the output class $A \subseteq X$ is a fuzzy set. A rough-fuzzy set is a tuple $\langle \overline{R}(A), \underline{R}(A) \rangle$, where the lower approximation $\underline{R}(A)$ and the upper approximation $\overline{R}(A)$ of A are fuzzy sets of X/R, with membership functions defined by

$$\mu_{\underline{R}(A)}([x]_R) = \inf\{\mu_A(x) \mid x \in [x]_R\} \tag{4-a}$$

$$\mu_{\overline{R}(A)}([x]_R) = \sup\{\mu_A(x) \mid x \in [x]_R\} \tag{4-b}$$

Here, $\mu_{\underline{R}(A)}([x]_R)$ and $\mu_{\overline{R}(A)}([x]_R)$ are the membership values of $[x]_R$ in $\underline{R}(A)$ and $\overline{R}(A)$, respectively.

2.4 Rough-Fuzzy Membership Function

The rough-fuzzy membership function of a pattern $x \in X$ for the fuzzy output class $C_i = A \subseteq X$ is defined by [9]

$$^{l}C_i = \frac{\|F \cap C_i\|}{\|F\|} \tag{5}$$

where $F = [x]_R$ and $\|C_i\|$ means the cardinality of the fuzzy set C_i. One possible way to determine the cardinality is to use: $\|C_i\| \stackrel{def}{=} \sum_{x \in X} \left(\mu_{C_c}(x)\right)$ For the '$\cap$' (intersection) operation, we can use

$$\mu_{A \cap B}(x) \stackrel{def}{=} \min\{\mu_A(x), \mu_B(x)\} \; \forall x \in X \tag{6}$$

It must be noted that, the concept of rough-fuzzy set is necessary when dealing with ambiguous concepts, whereas rough-fuzzy membership function is needed when uncertain data are considered.

3 Related Work

The main approaches to classification of remote sensing images are statistical methods [10], Artificial Neural Network methods [11], Fuzzy methods [12], [16], Fuzzy neural networks [13], Multi-source classification methods [14] and Hybrid approaches [15]. Statistical methods like Parallelepiped method, Minimum distance classifier, and Maximum likelihood classifier are very much dependent on the distribution of classes.

There has been a spurt of activity to integrate different computing paradigms such as fuzzy set theory, neural networks, genetic algorithms, and rough set theory, for generating more efficient hybrid systems that can be classified as *soft computing* methodologies. The purpose is to provide flexible information processing systems that can exploit the tolerance for imprecision, uncertainty, approximate reasoning, and partial truth in order to achieve tractability, robustness, and low cost in real-life ambiguous situations [2]. Neuro-fuzzy computing [3] captures the merits of fuzzy set theory [4] and artificial neural networks [5]. This integration promises to provide, to a great extent, more intelligent systems (in terms of parallelism, fault tolerance, adaptivity, and uncertainty management) to handle real-life recognition/decision making problems. But all these models only deal with fuzzy uncertainty.

Artificial Neural networks are generally described as nonparametric. The performance of the neural network depends to a significant extent on how well it has been trained and not on the adequacy of assumptions concerning the statistical distribution of the data. The most popular neural network classifier in remote sensing is the multilayer perceptron. Classification can also be carried out by other main type of neural networks such as SOM and fuzzy ARTMAP [15].

A fuzzy classification is a soft classification, which is used to find out uncertainty in the boundary between classes and to extract the mixed pixel information. This is achieved by applying a function called "membership function" on remotely sensed images. Using "hard" classification methods, we cannot measure the uncertainty in an image whereas in a fuzzy classification technique, we can get more information from the data [12]. Fuzzy multilayer perceptron (FMLP) introduced by Pal and Mitra [13], is a fuzzy version of MLP having advantages of both neural network and fuzzy logic. It employs the supervised back propagation learning algorithm and incorporates fuzzy set-theoretic concepts in both input and output stages.

Many researchers have looked into the implementation of decision rules extracted from operation data using rough set formalism, especially in problems of machine learning from examples and control theory [6]. In the context of neural networks, an attempt of such implementation has been made by Yasdi [7]. The intention was to use rough sets as a tool for structuring the neural networks. The methodology consisted of generating rules from training examples by rough-set learning, and mapping the dependency factors of the rules into the connection weights of a four-layered neural network. Application of rough sets in neurocomputing has also been made in [8]. However, in this method, rough sets were used for knowledge discovery at the level of data acquisition, (*viz.*, in preprocessing of the feature vectors), and not for structuring the network.

4 Proposed Method

As explained in section 1 due the multi-granular structure of the multi-sensor satellite images, we have limited capability to discern two objects. The effect is that based on the observational parameters, any two objects, may appear same, whereas, the ground truths about the objects forces us to have different opinion on them, i.e. we must have to deal with rough uncertainty in association with fuzzy uncertainty to obtain better classification accuracy. The fuzzy MLP model explained in [13] and fuzzy classification model proposed by Farid [16] only deals with fuzzy uncertainty. Apart from that, extraction of class-conditional spectral parameters using mean and standard deviation from supervised training sites of pure pixels used in the FMLP is dependent on the distribution of the reflectance values. In [17] it is observed that Neural Network classifiers as compared to statistical classifiers are nonparametric (distribution free). Statistical classifiers give incorrect results when reflectance values of classes are very close.

The proposed method effectively copes up with these two problems and provides better classification accuracy. The steps of the proposed method are described below.

Step 1: Generating equivalence classes for M_1 pure labeled training vectors

Let U be the set of M_1 pure labeled pixels i.e. $U = \left\{\left(x_1, y_1\right), ..., \left(x_{M_1}, y_{M_1}\right)\right\}$ where $\left(x_i, y_i\right)$ represents a pure labeled pixel and $A = \left(U, A \cup \left\{d\right\}\right)$ be a decision system, where $A = \left\{a_1, a_2, ..., a_D\right\}$ is the set of conditional attributes such that $a_j\left(X_i\right) = f_{i,j}$, D is the dimensionality of the input feature, d is the decision attribute such that $d\left(X_i\right) = Y_i$ where $Y_i = \left(y_{i,1}, ...y_{i,L}\right)$, L is the total number of land cover classes and $\left|U\right| = M_1$.

The equivalence relation $IND(A)=\{(x_i,x_j)\in U\times U:\forall_{a\in A}(a(x_i)=a(x_j))\}$ divides the set of objects U into equivalence classes and two objects belong to the same class if and only if they have the same values on attributes from A. It is to be noted that if each equivalence class generated by $IND(A)$ contains the objects belonging to same output class then there is no rough uncertainty. Here we can also use rough set to find the reduct $B\subseteq A$ (B is the reduced set of attributes while maintaining the decision capability of the decision table) and then generate the equivalence classes corresponding to $IND(B)$. Let $[x]_{IND(A)}$ represents the equivalence class to which x belongs. The equivalence class $[x]_{IND(A)}$ can be understood as a degenerated fuzzy set with those elements belonging to the class possessing a membership of one, and zero otherwise.

Step 2: Assigning fuzzy membership grade to each fuzzy output class for M_1 pure labeled pixels

The M_1 pixels under consideration are pure labeled pixels i.e. for them we are confirmed that to which class they belongs, so here we are taking the fuzzy membership value to the appropriate class as 0.9 and to others as 0.1/(L-1), where L is the number of land cover classes. Thus for this stage membership function can be defined as

$$\mu_{C_i}(x)=\begin{cases}0.9 & \text{if Ci is the appropriate class of x}\\ 0.1/(\text{L-1}) & \text{Otherwise}\end{cases} \tag{7}$$

where C_i, i=1,2,….,L is the fuzzy output class.

Step 3: Calculating Rough-Fuzzy membership grade to each fuzzy output class for M_1 pure labeled pixels

The Rough-Fuzzy membership value ${}^{l}C_i(x)$, to class C_i for input vector x $=(f_1,...f_D)$ is calculated using the equation (5) described in section 2.4.

Step 4: Designing and training neural network for M_1 pure labeled pixels

A (H+1) layered MLP with D neurons in input and L neurons in output layer, consisting of H-1 hidden layers, is trained by clamping the input vector $X_i=(f_{i,1},....f_{i,D})$ at input layer and the desired L-dimensional output vector with components ${}^{l}C_i(x)$ at the output layer.

Step 5: Computation of Rough-Fuzzy membership grades for M_1 pure and M_2 unlabeled mixed training vectors

In this step M_2 unlabeled mixed training samples are included in the training set. For (M_1+M_2) training vectors we calculate the rough-fuzzy membership grade to each class by clamping the input vector $X_i=(f_{i,1},....f_{i,D})$ i=1, 2,...., M_1+M_2 at the input layer of neural network trained in the previous step. The output values of the training vectors are normalized such that all the membership values of the classes sum up to 1.

Normalization is done by first obtaining the total output sum and then all the output values are divided by the total sum. Obtaining membership values using Neural network makes the Rough Neuro-Fuzzy classifier independent of class reflectance distribution and inclusion of unlabeled pixels helps in increasing the classification accuracy.

Step 6: Input data fuzzification

The (M_1+M_2) labeled training vectors are fuzzified before being employed in the FMLP training session. This means that every non-normal component of input pattern $X_i = \left(f_{i,1}, ...f_{i,D} \right)$, i=1,........, M_1+M_2 is converted into normal degrees of membership to fuzzy linguistic sets low, medium, and high as explained in [13].

Step 7: Supervised training of final Rough Neuro-Fuzzy network via backpropagation algorithm

The complete training set, consisting of (M_1+M_2) training vectors, is employed by the traditional error backpropagation algorithm to train Rough Neuro Fuzzy Network. The proposed Rough Neuro Fuzzy Network is an (H+1)-layered MLP with $3 \times D$ neurons in the input layer and L neurons in output layer, such that there are H-1 hidden layers. The input vector, with components fuzzified as in [13], is clamped at the input layer while the desired L-dimension output vector obtained in step 5 is clamped during training at the output layer.

5 Results and Discussion

Two study areas from high resolution multi-spectral IRS-P6 LISS-IV satellite image of Allahabad region acquired in April 2004 are selected for classification purpose. The spatial resolution of the images is 5.8 m. Three bands available in IRS-P6 LISS-IV are taken into consideration for analysis. The two LISS-IV satellite images are first geo-referenced using 15 well distributed Ground control points (GCP) for each image collected using Leica GS5 GPS receiver and then the images were converted to Geo-tiff image format. This is just to make the analysis work easier. The first study area has a geographical extent of $81^{\circ}45'36.07"E$ to $81^{\circ}47'23.67"E$ and $25^{\circ}26'23.92"N$ to $25^{\circ}25'12.42"N$ and second study area has a geographical extent of $81^{\circ}51'49.53"E$ to $81^{\circ}53'17.57"E$ and $25^{\circ}26'30.63"N$ to $25^{\circ}27'27.37"N$.

For study area 1 there are totally seven predefined classes for three bands in the image which are used for analysis. For study area 2 there are totally six predefined classes for three bands in the image which are used for analysis.

MATLAB is used for writing program for classification. MATLAB Mapping toolbox is used to read shape files and geo registered images. A well distributed 90 ground truth pixels were collected using the GPS receiver for each class in each study area. Out of these 90 pixels, 50 pixels were used from training of ANN and remaining 40 were used for accuracy assessment. Pixels were collected as 2D point shapefiles. Overall accuracy, User's accuracy, Producer’s accuracy and Kappa Coefficient [10] of the proposed methodology are compared with statistical, neural network, and FMLP models. Fig. 1(a) and Fig. 1(b) show the study area 1 and study area 2 respectively.

Fig. 2(a) and Fig. 2(b) show the classified images of study area 1 and study area 2 respectively, using statistical method (Minimum Distance).

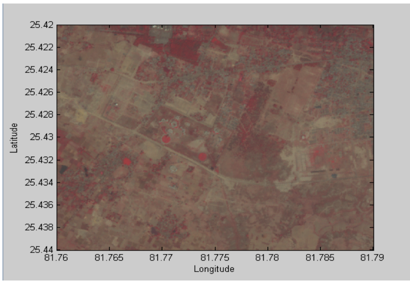

Fig. 1(a). Study area 1

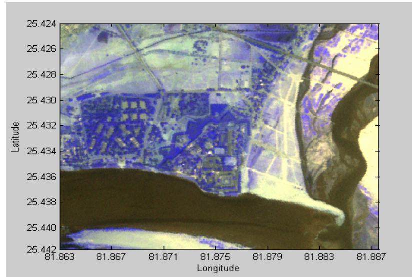

Fig. 1(b). Study area 2

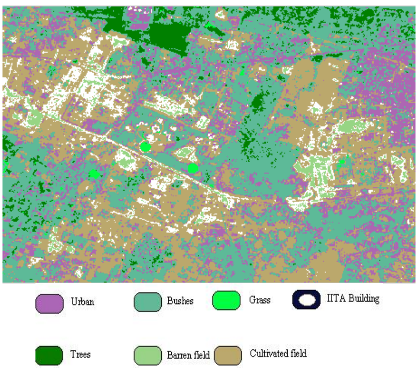

Fig. 2(a). Classified Study area 1 byStatistical method

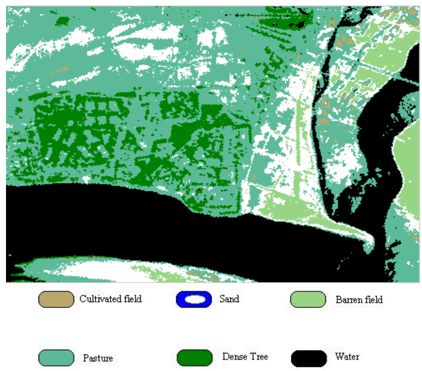

Fig. 2(b). Classified Study area 2 by Statistical method

Fig. 3(a) and Fig. 3(b) show the classified images of study area 1 and study area 2 respectively, using Neural Network method.

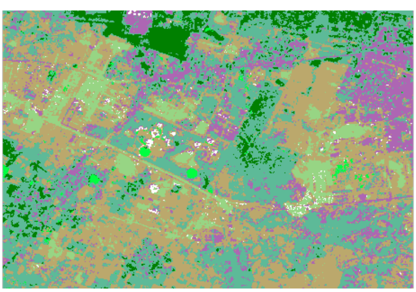

Fig. 3(a). Classified Study area 1 by Neural Network method

Fig. 3(b). Classified Study area 2 by Neural Network method

Fig. 4(a) and Fig. 4(b) show the classified images of study area 1 and study area 2 respectively, using the FMLP method.

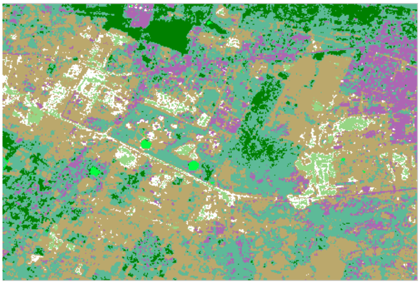

Fig. 4(a). Classified Study area 1 by FMLP method

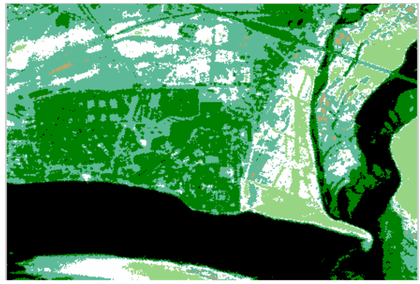

Fig. 4(b). Classified Study area 2 by FMLP method

Fig. 5(a) and Fig. 5(b) show the classified images of study area 1 and study area 2 respectively, using the proposed method.

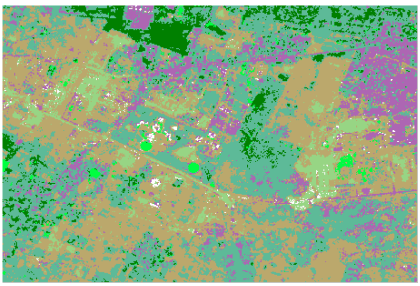

Fig. 5(a). Classified Study area 1 by proposed method

Fig. 5(b). Classified Study area 2 by proposed method

Table 1 briefly shows the overall accuracy and Kappa coefficients for the Statistical (Minimum Distance), Neural Network, FMLP, and the proposed method in cases of study area 1 and study area 2.

Fig. 6(a) and Fig. 6(b) show the plots of overall accuracy for study area 1 and study area 2 respectively, in case of various classifiers.

Table 1. Overall Accuracy and Kappa coefficients for study area 1 and study area 2

	Overall Accuracy	**Kappa Coefficient**	**Overall Accuracy**	**Kappa Coefficient**
	Study area 1		**Study area 2**	
Statistical	94.29%	93.31%	93.33%	92.04%
Neural Network	94.64%	93.68%	94.17%	92.99%
FMLP	92.86%	91.38%	86.67%	84.25%
Proposed	97.14%	96.61%	97.5%	96.95%

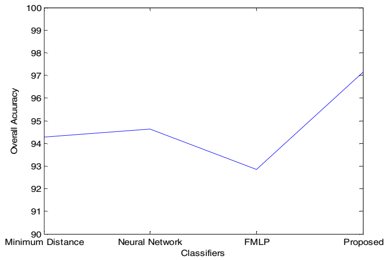

Fig. 6(a). Plot of Overall Accuracy for study area 1

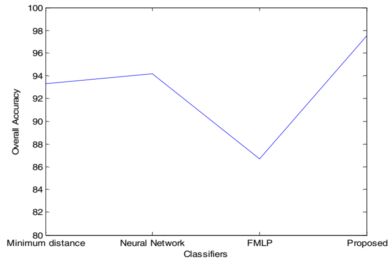

Fig. 6(b). Plot of Overall Accuracy for study area 2

Fig. 7(a) and Fig. 7(b) show the plots of Kappa Coefficients for study area 1 and study area 2 respectively, in case of various classifiers.

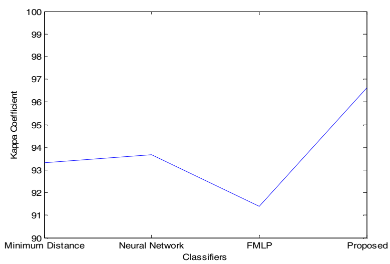

Fig. 7(a). Plot of Kappa Coefficient for study area 1

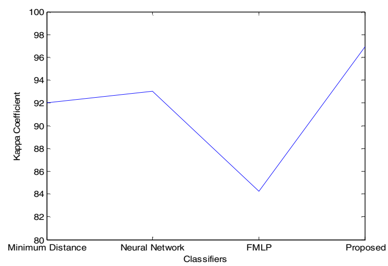

Fig. 7(b). Plot of Kappa Coefficient for study area 2

Fig. 8(a) and Fig. 8(b) show the bar charts of User's accuracy for study area 1 and study area 2 respectively, in case of various classifiers.

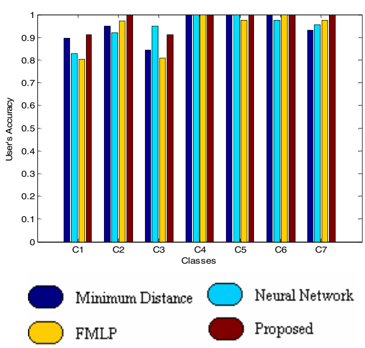

Fig. 8(a). Bar chart of User's Accuracy for study area 1

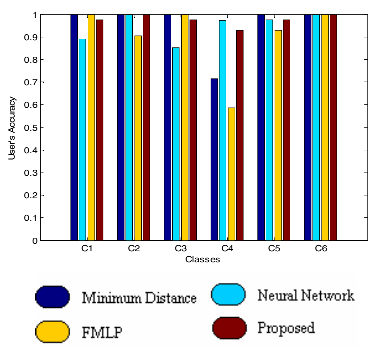

Fig. 8(b). Bar chart of User's Accuracy for study area 2

Fig. 9(a) and Fig. 9(b) show the bar charts of Producer's accuracy for study area 1 and study area 2 respectively, in case of various classifiers.

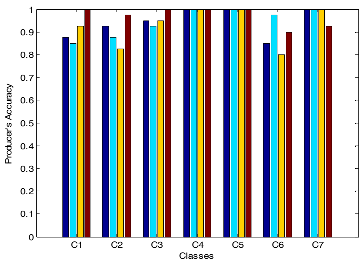

Fig. 9(a). Bar chart of Producer's Accuracy for study area 1

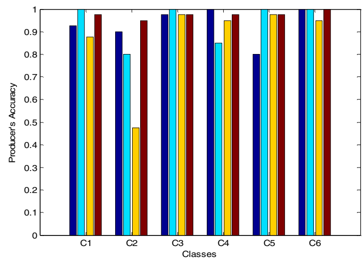

Fig. 9(b). Bar chart of Producer's Accuracy for study area 2

From the comparison of overall accuracy and Kappa Coefficients it becomes clear that the proposed method is giving better results in comparison to other ones. The proposed method also gives better User's and Producer's accuracy in case of most of the classes. The proposed method is distribution free, and is capable enough to handle fuzzy uncertainty as well as rough uncertainty associated with the satellite image classification process.

6 Conclusions and Future Scope

From this experimentation we found that the concept of rough set plays an important role for getting better accuracy in case of multispectral image classification. We see that the knowledge generated from the satellite image at a given resolution and spectrum band is granular, which generates rough uncertainty in data/information. This uncertainty can not be dealt with by simply considering the overlapness in terms of fuzzy logic. To deal with this situation, we have to consider the vagueness which is generated due to insufficiency of knowledge about the event, data or world. FMLP only considers the fuzzy uncertainty associated with data. Moreover, the initial fuzzy membership value calculation is also dependent on distribution of data. In the proposed Rough Neuro-Fuzzy method we tried to deal with both of these flaws. By the use of Rough-Fuzzy membership value in place of simple fuzzy membership value, we can effectively model rough uncertainty as well as fuzzy uncertainty. The use of neural network to generate Rough-Fuzzy Membership value for final training vectors makes the whole model distribution free. The experimental results shown in the previous section are supporting for the same.

As a further improvement in the proposed approach, instead of using classical equivalence classes we can use fuzzy equivalence classes based on weak fuzzy similarity relation. By using this we can model the fuzziness associated in multispectral image classification more effectively.

References

1. Pawlak, Z.: Rough Sets - Theoretical Aspects of Reasoning about Data. Kluwer Academic Publishers, Dordrecht (1991)
2. Zadeh L.A.: Fuzzy Logic, Neural Networks, and Soft Computing. Comm. ACM, 37 (1994), 77-84
3. Bezdek, J.C., Pal, S.K.: Fuzzy Models For Pattern Recognition: Methods that Search for Structures in Data. IEEE Press, New York (1992)
4. Zadeh, L.A.: Fuzzy sets. Inform. Contr., 8(1965) 338–353
5. Lippmann, R.P.: An Introduction to Computing with Neural Nets. IEEE Acoust., Speech, Signal Processing Mag., 4(1987) 4–22
6. Slowinski, R. (Eds.): Intelligent Decision Support - Handbook of Applications and Advances of the Rough Sets Theory. Kluwer Academic Publishers (1992)
7. Yasdi, R.: Combining Rough Sets Learning and Neural Learning Method to Deal with Uncertain and Imprecise Information. Neuro-Computing, 7(1995) 61-84
8. Czyzewski, A., Kaczmarek A.: Speech Recognition Systems Based on Rough Sets and Neural Networks. Proc. 3rd Wkshp. Rough Sets and Soft Computing (RSSC'94), San Jose, CA, (1994) 97–100
9. Sarkar, M., Yegnanarayana B.: Rough-Fuzzy Membership Functions. Proc. IEEE World Congress on Computational Intelligence, Alaska, USA, 1(1998) 796 – 801
10. Jensen, John R.: Introductory Digital Image Processing, a Remote Sensing Perspective. 2nd Edn. Prentice Hall series in Geographic Information Science (1995)
11. Gopal, S., Fischer, M.: A Comparison of Three Neural Network Classifiers for Remote Sensing Classification. IEEE Geoscience and Remote Sensing Symposium, 1(1996) 787-789
12. Wang, F.: Fuzzy Supervised Classification of Remote Sensing Images. IEEE Transactions on Geoscience and Remote Sensing, 28(1990) 194-201
13. Pal, S., Mitra, S.: Multi-layer Perceptron, Fuzzy Sets, and Classification. IEEE Transaction on Neural Networks, 3 (1992) 683–697
14. Benediktsson, J.A., Sveinsson, J.R.: Multisource Remote Sensing Data Classification Based on Consensus and Pruning. IEEE Transactions on Geoscience and Remote Sensing, 41 (2003), 932-936
15. Tso, B., Mather, P.M.: Classification Methods for Remotely Sensed Data. CRC Press (2001)
16. Melgani, F., Al Hashemy B.A.R., Taha, S.M.R.: An Explicit Fuzzy Supervised Classification Method for Multispectral Remote Sensing Images. IEEE Transactions on Geoscience and Remote Sensing, 38(2000) 287-295
17. Paola, J.D., Schowengerdt, R.A.: Detailed Comparison of Backpropagation Neural Network and Maximum-Likelihood Classifiers for Urban Land Use Classification. IEEE Transactions on Geoscience and Remote Sensing, 33(1995) 981-996