

Fuzzy Classification of Semi-urban Features from IRS Satellite Imagery

A L Choodarathnakara, Sujith J, Alfiya Nishath H G, Siddiq Shariff, Pradeepa V T, Madhukar G

*^{1,3,4,5,6} Department of Electronics & Communication Engineering, GEC, Kushalnagar, Kodagu, Karnataka, India

²Quadgen Wireless Solutions, Bangalore, Karnataka, India

ABSTRACT

In remote sensing images, a pixel might represent a mixture of class covers, within class variability or other complex surface cover patterns that cannot be properly described by one class. So in order to map a scene's natural fuzziness or imprecision and to provide more complete information through image analysis, a fuzzy logic based classification procedure is necessary. This fuzzy logic is a knowledge based method which makes no assumption about statistical distribution of the data and therefore reduces classification inaccuracies. Also fuzzy logic is interpretable and can combine expert knowledge and training data. Major advantage of fuzzy is that it allows natural description in linguistic terms of problems that should be solved rather than in terms of relationship between precise numerical values. Hence this paper aiming to study fuzzy classifier as an alternative approach to traditional classification techniques for RS data to classify urban features from satellite image. The ERDAS IMAGINE V9.2 remote sensing software is used in this study. The accuracy assessment was conducted based on Overall Classification Accuracy (OCA) and Kappa Statistics. This experiment was conducted using Erdas Imagine V9.2 RS software. Finally, the suitability of Fuzzy classification is verified on Arasikere Semi-urban area and Overall Classification Accuracy 65.83% and 71.85% was obtained for 360 and 720 training sites with 120 validation sites respectively. By increasing validation sites to 240 for 720 training sites, OCA of 73.33% was achieved.

Keywords: Remote Sensing, Semi-urban Area, Fuzzy Classification, Erdas Imagine

I. INTRODUCTION

Remote Sensing refers to the science of identification of earth surface features and estimation of their geophysical properties using electromagnetic radiation as a medium of interaction. Spectral, Spatial, Temporal and Polarization signatures are major characteristics of the sensor or target, which facilitates target discrimination. Earth surface data are seen by the sensors in different wavelengths (Reflected, Scattered and/or Emitted) is radiometrically and geometrically corrected before extraction of spectral information. Image classification is the process of categorizing all the pixels automatically in an image into a finite number of land use/land cover classes. The major operational application themes, in which India has extensively used remote sensing data are agriculture, forestry, water resources, LU/LC, urban sprawl, geology environment, coastal zone, marine resources, snow and glacier, disaster and mitigation, infrastructure development etc.

Image classification is a complex process that may be affected by many factors. Effective use of multiple features of remotely sensed data and selection of suitable classification method are significant for improving classification accuracy. Non parametric classifiers such as fuzzy logic, neural network, decision tree classifier and knowledge based classifiers have increasingly become important approaches for multisource data classification. In general image classification can be grouped into supervised and unsupervised, or parametric and non-parametric, or hard and soft (fuzzy) classification, or pixel, subpixel and perfield.

Designing a suitable image processing procedure is a prerequisite for a successful classification of RS data into a thematic map. Effective use of multiple features of RS data and the selection of a suitable classification method are especially significant for improving classification accuracy. Hence, the present work

examines the complexity of the classification problems where there is a need for highly reliable and accurate classification systems and algorithms. The emphasis is placed on developing advanced algorithms and identifying & reducing uncertainties in the image-processing chain to improve classification performance in terms of accuracy for better understanding and classification of remote sensing images. One important aspect in remote sensing is the categorization and classification of spectral measurements taken from remote sensors into various features on land surface. Improving classification accuracy of digital data, as always been an important concern to extract the real world situation in the form of thematic maps. One of the main problems when generating land cover maps from digital images is the confusion of spectral responses. The possibility, that two or more different features having the same spectral behavior are eventually classified as the same class, not only creates the difficulties in extracting the valid information but also introduces errors in the classification. Therefore, the type of image and consequently the spatial and spectral resolutions influences the classification accuracy, the spatial variability of the land cover types and the attribute to be determined among other factors. Finally any image classification result is influenced by the data itself, data preprocessing, enhancement techniques and classification schemes being used.

Thomas N Lillesand [1] has explained basic concepts and elements necessary to conceptualize an ideal remote sensing and applications in the book Remote Sensing and Image Interpretation. D LU and Q WENG [2] explain image classification process and different techniques of image classification.

Zadeh's development of fuzzy sets motivated many research scholars to use the concepts in image processing. Fuzzy logic is applied to pattern classification and cluster analysis. In 1996 the first paper on fuzzy set theory to pattern classification appeared. Ruspin [3] conducted work on cluster analysis using the concept of fuzzy partition. At about the same time the Rosenfeld [4] used fuzzy sets in image analysis.

András Bárdossy and Luis Samaniego [5] explained fuzzy logic for remote sensing. Timothy J Ross [6] describes the basic concepts of Fuzzy Logic, Fuzzy Sets, Membership Functions, and Fuzzy Classification. In

1973 J C Dunn [7] has introduced fuzzy clustering algorithms, the fuzzy ISODATA and Fuzzy C-means algorithm. I Nedeljkovic [8] has obtained the results of image classification using fuzzy logic. It shows that fuzzy rules for image classification are simple and less time consuming.

Yan Wang and Mo Jamshidi [9] have applied fuzzy logic in classification of remote sensed data. They concluded that fuzzy can incorporate collateral data easily so that some similar land covers can be classified well and it provides membership values in the classification results and explained about fuzzy logic, expert system, and neural networks.

Farid Melgani, Bakir A R Al Hashemy, and Saleem M R Taha [10] describe an explicit fuzzy supervised classification method which consists of three steps. The explicit fuzzification is the first step where the pixel is transformed into a matrix of membership degrees representing the fuzzy inputs of the process. Then, in the second step, a MIN fuzzy reasoning rule followed by a rescaling operation is applied to deduce the fuzzy outputs, or in other words, the fuzzy classification of the pixel. Finally, a defuzzification step is carried out to produce a hard classification.

In this study, developing a suitable Fuzzy supervised classification technique to classify LU/LC features of semi-urban area. The accuracy assessment of the classification process will be done with respect to increase in training sets and validation sets, to check whether there is any improvement in the OCA. There is an increase in OCA of 6.02% and 2.52% for validation sets of 120 and 240 respectively.

II. FUZZY CLASSIFICATION

Fuzzy logic is a tool for embedding structured human knowledge into workable algorithms. In narrow sense fuzzy logic is considered a logical system aimed at providing model for modes of human reasoning that are approximate rather than exact. In wider sense, fuzzy logic is treated as a fuzzy set theory of classes with unsharp or fuzzy boundaries. Fuzzy logic methods can be used to design intelligent system on the basis of knowledge expressed in a common language. Fuzzy logic method permits the processing of both symbolic and numerical information [6].

The fuzzy approach plays an important role in the whole classification procedure. In our project we use fuzzy supervised classification to classify the satellite images. This is a useful approach as it helps in representing the geographical information, in quantifying the change, in measuring uncertainty between the boundaries of two or more classes or uncertainty in a pixel because of mixed classes [11].

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 we cannot measure the uncertainty in an image whereas in fuzzy classification technique; we can get more information from the data.

A fuzzy system used for classification generally comprised of three main steps, as shown in Figure 1.

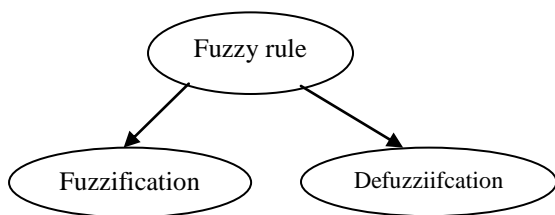


Figure 1: Basic architecture of fuzzy system includes fuzzification, fuzzy rule-base and defuzziifcation

A. Fuzzification

It is the process of making a crisp quantity into fuzzy. We see that many of the quantities are crisp and deterministic but they are actually not deterministic. They have little uncertainty. This uncertainty can be due to imprecision, ambiguity, or, vagueness. Then the variable is probably fuzzy and can be represented by a membership function.

B. Fuzzy Rule Base

In the field of artificial intelligence there are different ways to represent knowledge. One of the way is to represent in the form of natural language of the type:

IF Premise , THEN Conclusion (1)

This form of expression (1) is called as IF-THEN rule based form. It says that if we know a fact (premise, antecedent, hypothesis), then we can infer, or derive another fact called a conclusion (consequent).

Fuzzy rules can be combined using fuzzy operators. The basic operators are “and” and “or”. “And” represents the minimum, meaning that the minimum value of all rules defines the return value. “Or” represents the maximum value, meaning that the maximum value of all rules defines the return value. A fuzzy rule base delivers a fuzzy classification, which consists of discrete return values for each of the classes. These values represent the degree of the class assignment. The higher the return values for the most possible class, the more reliable the assignment. The minimum membership value an object needs to have in order to be assigned to a class can be defined [5].

C. Defuzzification

Defuzzification is the conversion of a fuzzy quantity to a precise quantity, just as fuzzification is the conversion of precise quantity to a fuzzy quantity. The output of a fuzzy process can be the logical union of two or more fuzzy membership functions defined on the universe of discourse of output variable. To produce land cover classification results like maps, the fuzzy results have to be translated back to a crisp value, which means that an object is either assigned to a class or not.

III. STUDY AREA AND METHODOLOGY

A. Study Area

The study area considered for our work is semi urban area of Arsikere. It is situated in Hassan, Karnataka, India, its geographical coordinates are 13° 18' 50" North, 76° 15' 22" East and its original name (with diacritics) is Arsikere. It has an average elevation of 807 meters (2647 feet). Arasikere is the taluk center in Hassan district and is known for its coconut production and Malekallu Tirupathi hill. There is also an old temple built by Hoysala rulers by name "Shivalaya", built by Jakanachari and the town is set at the foot of Tirupathi hills. The image dimension of the study area is 607×645 pixels in MS data.

TABLE I
DETAILS OF THE SATELLITE DATA PRODUCTS USED

Satellite and Data type	Spectral Resolution	Spatial Resolution
IRS-P6 (Resourcesat1) Multi-spectral	Green (0.52-0.59µm); Red (0.62-0.68 µm); Infrared (0.77-0.86 µm);	5 m

B. Methodology

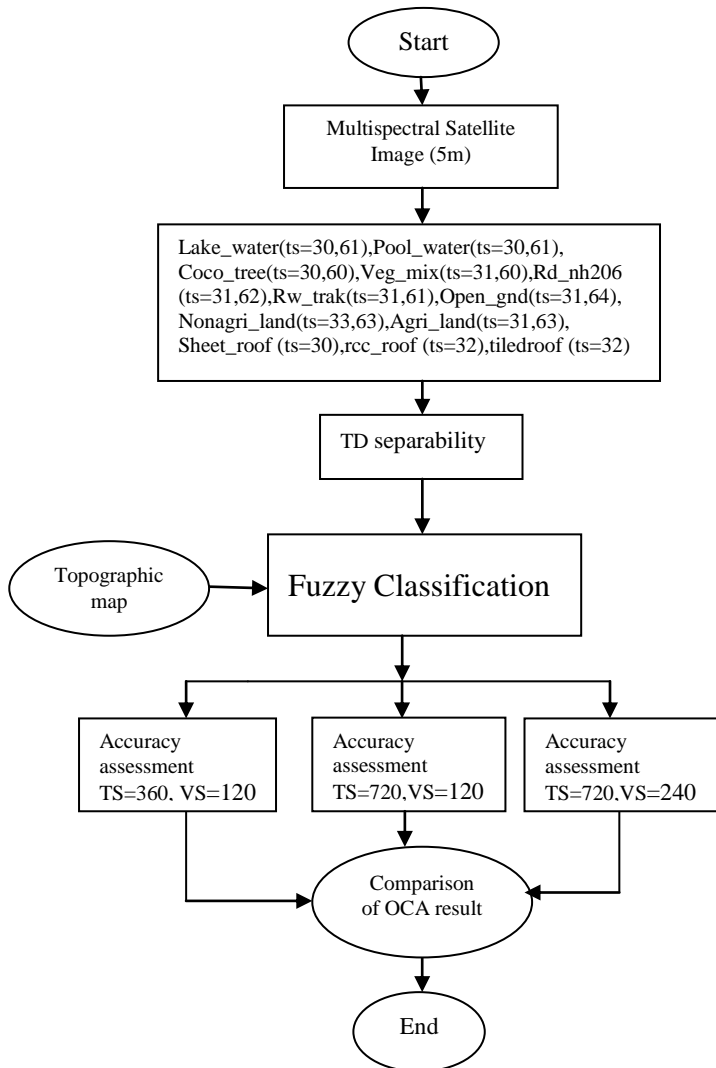


Figure 2: Flow chart of Methodology used in fuzzy classification

IV. RESULTS AND DISCUSSIONS

Supervised classification is performed using 5m LISS-IV data. In supervised classification, the basic steps followed are: (1) select training samples which are representative and typical for that information class; (2) perform classification after specifying the training samples set and classification algorithms; (3) assess the accuracy of the classified image through analysis of a confusion matrix which is generated either through random sampling or using test areas as reference data. Accuracy assessment was carried out in ERDAS IMAGINE V 9.2 RS image processing software. The classifiers was trained with two sets of data viz. 360 and 720 instances and with two sets of validation points viz. 120 and 240 and error matrices were computed for performance analysis.



Figure 3: Snapshot of the Study area: Arsikere, Hassan, Karnataka, India. (5m MS data)

From the field survey of the study area and the visual interpretation of data, the following 12 classes were identified, which included major classes and sub-classes, for the present study. The 12 classes are: Pool water, Coco_Tree, Open Ground, Lake Water, Nonagri Land, Vegmix, Railway Track, Road_NH, Agri Land, RCC Roof, sheet Roof and Tiled as shown in Figure 3. The screen shot of training samples considered is as shown in Figure 4.



Figure 4: Training samples for classification

The screenshots of transformed divergence for 360 training samples are as shown in Figure 5 and 6 and that for 720 samples is as shown in Figures 7 and 8.

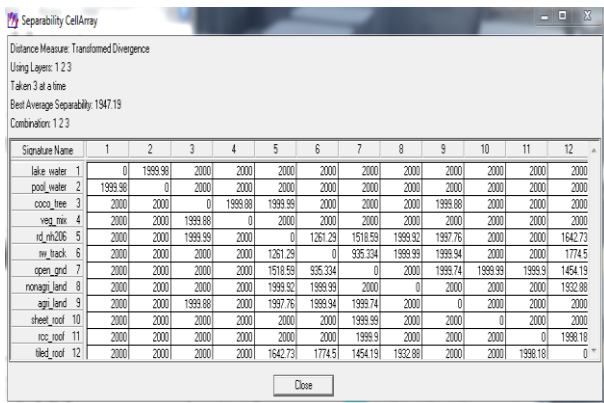


Figure 5: TD for MS bands: Best Average Separability: 1947.19

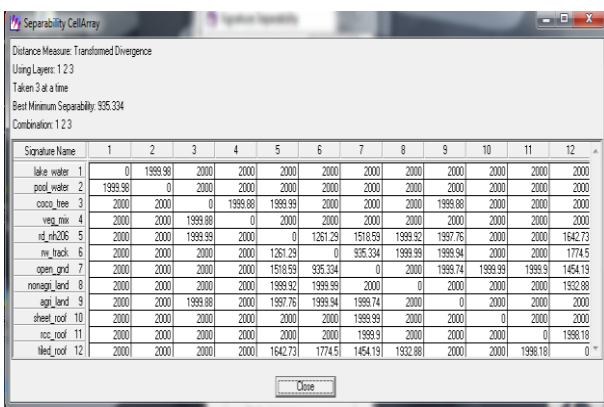


Figure 6: TD for MS bands: Best Minimum Separability: 935.334

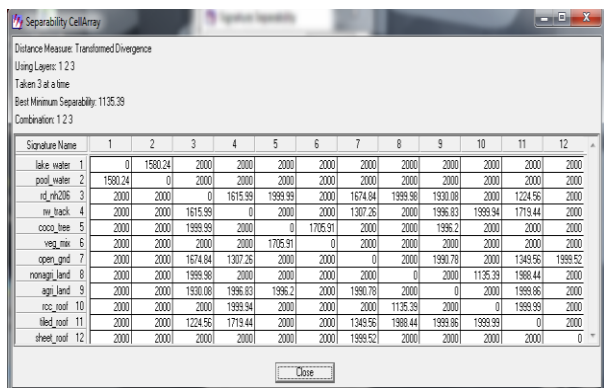


Figure 7: TD for MS bands: Best Minimum Separability: 1135.39

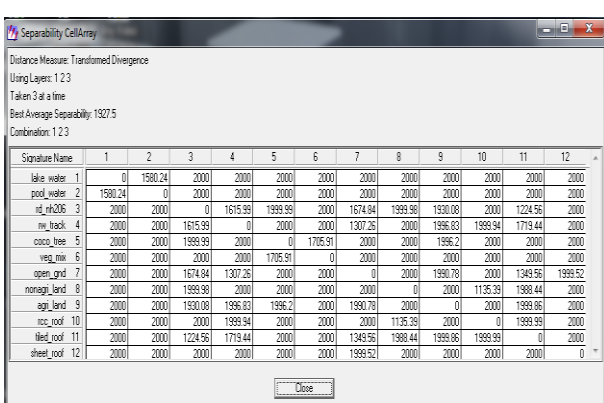


Figure 8: TD for MS bands: Best Average Separability: 1927.5

The classified image for Fuzzy after having trained with 360 training samples is as shown in Figure 5 and 6. As it can be seen from Figure 7 and 8 that NH road is classified as RCC, and the regions around water is also classified as NH road. This is mainly attributed to the fact that the classes exhibit similar spectral reflectance value and poor TD between them (Figure 5). The Railway track show a TD of 935.334 and, Open ground and Tiled roof, exhibit a TD of 1454.19 – both are considered to be very low. The best TD is supposed to be more than 1750.

Further, the classifier was also trained with less number of training samples i.e., 720 samples, and thus obtained classified image is shown in Figure 10. From the figure 7 and 8 it can be noted that there are large numbers of misclassifications. Tiled roof is classified as sheet roof which indicates the importance of the training samples required to train the classifier. Non agri land and rcc roof shows an average separability of 1135.39, which is very low. As a result of which non agri land is not correctly classified. The NH road pixels are misclassified in large amount. Water exhibits the maximum TD of 2000 as it is distinct from other classes.

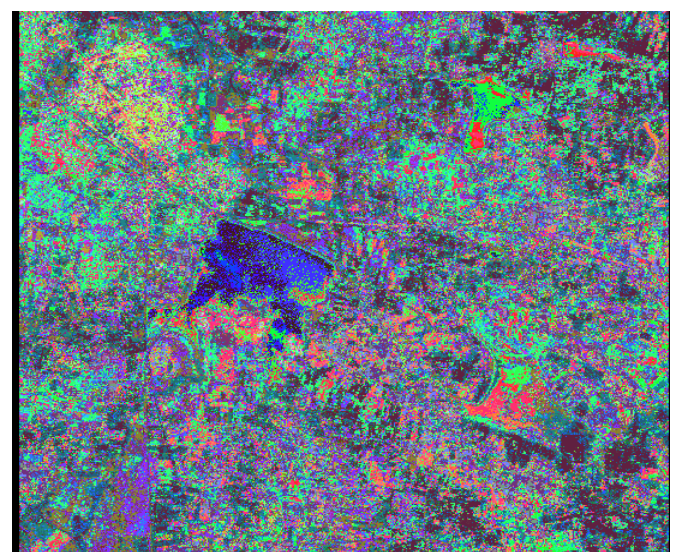


Figure 9: Fuzzy Supervised Classified Image obtained Using 360 training samples

Legends:

- lake_water
- pool_water
- coco_tree
- veg_mix
- rd_nh206
- rw_track
- open_gnd
- nonagri_land
- agri_land
- sheet_roof
- rcc_roof
- tiled_roof

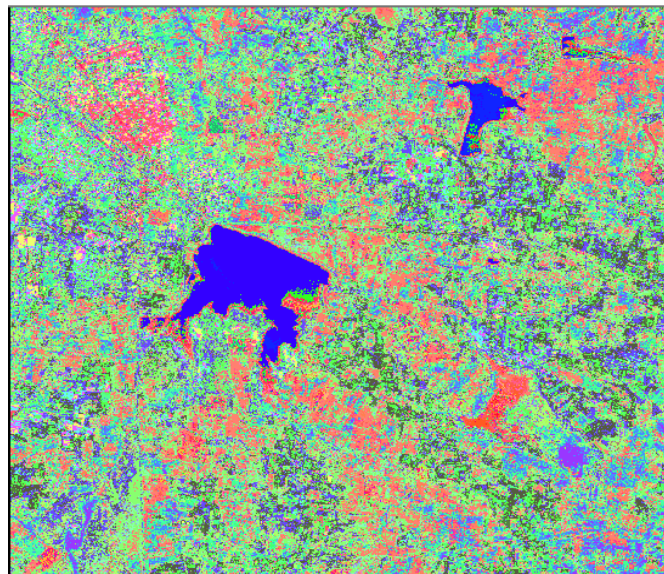


Figure 10: Fuzzy Supervised Classified Image obtained using 720 training samples

Legends:

lake_water		pool_water		coco_tree	
veg_mix		rd_nh206		rw_track	
open_gnd		nonagri_land		agri_land	
sheet_roof		rcc_roof		tiled_roof	

The Figure 9 shows fuzzy classified image for 360 training set. The Figure 10 shows fuzzy classified image for 720 training set. In fuzzy classification, check out the classes that are made up of mixed pixels and the land cover are not homogenous. The colors are different for the mixed classes from the colors assigned to the classification.

TABLE II

ERROR MATRIX FOR 360 TRAINING SAMPLES AND 120 TEST POINTS

Reference Data

C	1	2	3	4	5	6	7	8	9	10	11	12	RT
1	11	0	0	2	0	0	0	0	0	0	0	0	13
2	0	2	0	0	0	6	0	0	0	0	0	0	4
3	0	0	13	0	0	0	0	1	0	0	0	0	14
4	1	0	0	6	0	0	0	0	0	0	0	0	7
5	0	0	0	0	9	0	0	5	0	0	0	0	14
6	0	1	0	0	0	5	0	0	0	0	0	0	6
7	0	0	0	0	0	0	6	0	2	0	0	0	8
8	0	0	3	0	0	0	1	4	0	0	0	0	8
9	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	16	6	6	28
11	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	1	0	2	0	4	2	2	0	0	6	17
CT	12	3	17	8	11	7	11	12	4	16	6	12	119

Legends: C=Classes, RT=Row Total, CT=Column Total, 1=Pool water, 2=Coco_Tree, 3=Open Ground, 4=Lake Water, 5=Nonagri Land, 6=Vegmix, 7=Railway Track, 8=Road_NH, 9=Agri Land, 10=RCC Roof, 11=sheet Roof, 12=Tiled Roof

TABLE III

ACCURACY AND KAPPA TABLE FOR DATA GIVEN IN TABLE 2

Class Name	PA (%)	UA (%)	Kappa
1	91.67	84.62	0.8291
2	66.67	50.00	0.4872
3	76.47	92.86	0.9168
4	75.00	85.71	0.8469
5	81.82	64.29	0.6068
6	71.43	83.33	0.8230
7	54.55	75.00	0.7248
8	33.33	50.00	0.4444
9	---	---	---
10	100	57.14	0.5055
11	---	---	---
12	50.00	35.29	0.2810

Overall Classification Accuracy = 65.83%
Overall Kappa Statistics= 0.6188

TABLE IV

ERROR MATRIX FOR 720 SAMPLES AND 120 TEST POINTS

Reference Data

C	1	2	3	4	5	6	7	8	9	10	11	12	RT
1	8	0	1	0	0	0	0	0	0	0	0	0	9
2	0	7	0	0	0	0	4	0	0	0	0	0	11
3	1	0	10	0	0	0	0	0	0	0	0	0	11
4	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	11	0	0	0	0	0	0	1	12
6	0	0	0	0	0	6	1	0	0	0	0	0	7
7	0	0	0	0	0	2	3	0	0	0	0	0	5
8	0	0	0	0	0	1	0	9	0	0	0	0	10
9	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	1	0	0	0	0	0	0	0	5	0	0	6
11	0	0	0	0	0	0	0	0	0	0	1	1	2
12	0	0	0	0	6	1	3	0	0	0	9	27	46
CT	9	8	11	0	17	10	11	9	0	5	10	29	119

Legends: C=Classes, RT=Row Total, CT=Column Total, 1=Pool water, 2=Coco_Tree, 3=Open Ground, 4=Lake Water, 5=Nonagri Land, 6=Vegmix, 7=Railway Track, 8=Road_NH, 9=Agri Land, 10=RCC Roof, 11=sheet Roof, 12=Tiled Roof

Classified Data

TABLE V
ACCURACY AND KAPPA TABLE FOR DATA GIVEN IN TABLE 4

Class Name	PA (%)	UA (%)	Kappa
1	88.89	88.89	0.8799
2	87.50	63.64	0.6104
3	90.91	90.91	0.8999
4	...	---	---
5	64.71	91.67	0.9029
6	60.00	85.71	0.8442
7	27.27	60.00	0.5596
8	100	90.00	0.8919
9	...	---	---
10	100	83.33	0.8261
11	10.00	50.00	0.4545
12	93.10	58.70	0.4553

Overall Classification Accuracy = 71.85%
Overall Kappa Statistics= 0.6836

TABLE VII
ACCURACY AND KAPPA TABLE FOR DATA GIVEN IN TABLE 6

Class Name	PA (%)	UA (%)	Kappa
1	85.00	88.89	0.8787
2	72.22	92.86	0.9227
3	90.48	95.00	0.9452
4	11.11	66.67	0.6394
5	76.67	95.83	0.9523
6	80.00	88.89	0.8840
7	65.00	92.86	0.9220
8	94.74	94.74	0.9428
9	75.00	75.00	0.7457
10	42.11	88.89	0.8792
11	54.55	92.31	0.9152
12	97.22	38.89	0.2800

Overall Classification Accuracy = 73.33%
Overall Kappa Statistics= 0.6880

TABLE VI
ERROR MATRIX FOR 720 SAMPLES AND 240 TEST POINTS

Reference Data

C	1	2	3	4	5	6	7	8	9	10	11	12	RT
1	16	0	2	0	0	0	0	0	0	0	0	0	18
2	1	13	0	0	0	0	0	0	0	0	0	0	14
3	1	0	19	0	0	0	0	0	0	0	0	0	20
4	0	0	0	2	0	0	0	1	0	0	0	0	3
5	0	0	0	0	23	0	0	0	0	0	0	1	24
6	0	0	0	1	0	8	0	0	0	0	0	0	9
7	0	0	0	0	0	1	13	0	0	0	0	0	14
8	0	0	0	0	0	1	0	18	0	0	0	0	19
9	0	0	0	0	0	0	1	0	3	0	0	0	4
10	0	1	0	0	0	0	0	0	0	8	0	0	9
11	0	0	0	0	0	0	0	0	0	1	12	0	13
12	2	4	0	1	7	0	6	0	1	10	10	35	90
CT	20	18	21	18	30	10	20	19	4	19	22	36	237

Classified Data

Legends: C=Classes, RT=Row Total, CT=Column Total, 1=Pool water, 2=Coco_Tree, 3=Open Ground, 4=Lake Water, 5=Nonagri Land, 6=Vegmix, 7=Railway Track, 8=Road_NH, 9=Agri Land, 10=RCC Roof, 11=sheet Roof, 12=Tiled Roof

From the above results, fuzzy supervised classification gives better accuracy with the increase in number of training samples. Fuzzy shows a great improvement where there were mixed pixels exist.

TABLE VIII
ACCURACY WITH INCREASE IN TRAINING AND VALIDATION SITES IN FUZZY SUPERVISED CLASSIFICATION

No. of Training Sites	No. of Validation Sites	Fuzzy Supervised Classification
360	120	65.83%
720	120	71.85%
720	240	73.33%

The Table VIII shows that in all the three cases considered, fuzzy is more accurate even with increase in validation points. This OCA can still improve the accuracy by applying fuzzy classification to fused images. Using GIS manipulations, classification results of pixel based image analysis can be improved to obtain higher accuracy.

V. CONCLUSION

In this work the study area considered is Arsikere taluk in Hassan district. It is a semiurban area with moderate rainfall. This place is connected to various important cities in the state via bus and rail transport. The township is undergoing lot of changes. The objective was to study this area for classification purpose using fuzzy logic. The accuracy obtained for 360 and 720 samples with 120 test points 65.83% and 71.85% respectively and the accuracy obtained for 720 samples with 240 test points is 73.33% using fuzzy logic based classifier. As training samples were increased accuracy

also increased by 6.02% and 2.52% for increase in validation respectively. It is found that higher the training samples higher the OCA in both the classifiers. It is to be noted that the number of training samples depends upon the complexity of the study area considered. If the study area is simple then it consists of well defined crisp classes then less number of pixels can also give better accuracy. This study concludes that application of fuzzy gives better accuracy than the conventional methods.

VI. REFERENCES

- [1] Thomas M Lillesand, Ralf W Kiefer and Jonathan W Chipman, Remote Sensing and Image Interpretation, 5th Edition, John Wiley & Sons, 2004.
- [2] D.LU and Q.Weng, A survey of image classification methods and techniques for improving classification performance, *International Journal of Remote Sensing*, Vol.28, No.5, 10 March 2007, 823-870.
- [3] Ruspín A New Approach to Fuzzy Clustering Information and Control, Vol.15 pp 22- 32, 1969.
- [4] A. Rosenfeld, Fuzzy Graphs in Fuzzy Sets. Their Application to Cognitive & Decision Process Academic Press, 1975.
- [5] Andras Bardosy and Luis Samaniego, Fuzzy Rule-Based Classification of Remotely Sensed Imagery, *IEEE Transactions on Geoscience and Remote Sensing*, Vol.40, No. 2, February 2002.
- [6] Timothy J Ross, Fuzzy Logic with Engineering Applications, John Wiley, 2nd Edition, 2005.
- [7] J.C.Dunn, Fuzzy Relative of the ISODATA Process and its using Detecting Compact well Separated Clusters, *Journal of Cybernetics*, VOL.3, pp 32-57, 1973.
- [8] I.Nedeljkovic, Image Classification Based on Fuzzy Logic, Mapsoft Ltd.
- [9] Yan Wang and Mo Jamshidi, Fuzzy Logic Applied in Remote Sensing Image Classification, *IEEE International Conference on Systems, Man and Cybernetics*, 2004.
- [10] Farid Melani, Bakir. A.R. Al. Hashemy, and Saleem M.R. Taha, An Explicit Fuzzy Supervised Classification Methods for Multispectral Remotely Sensed Images, *Transactions on Geoscience and Remote Sensing*, VOL.38, January 2000, pp 287-295.
- [11] Chi-Farn Chen, Fuzzy Training Data for Fuzzy Supervised Classification of Remotely Sensed Images, *Center for Space Remote Sensing Research*, National Central University.