Landsat 8 Data Classification using Fuzzy Inference System

Adithya Kumar[1], Shivakumar B.R.[2]

[1]*M.Tech student, Dept. of ECE, NMAMIT, Nitte, Karnataka, India,*
[2]*Assistant Professor, Dept. of ECE, NMAMIT, Nitte, Karnataka, India.*

Abstract - *In this paper, land use/ land cover classification of Landsat 8 data using Fuzzy Inference System is performed. Supervised classification of remotely sensed data is a challenging task due to overlapping land classes. The Fuzzy technique takes into account the belonging of pixel to several classes. The designing of Fuzzy Inference System using membership functions, writing rules and output of Fuzzy Inference System are discussed. Mamdani Fuzzy Inference System is built to classify the Landsat 8 and accuracy of the output is calculated*.

Key Words: **Remote Sensing, Fuzzy, Confusion Matrix, FIS, Classification**

1. INTRODUCTION

The radiation of the electromagnetic energy emitted and/or radiated from the Earth surface is captured by the satellite sensors. The extracted information are used to construct the image. Image processing techniques are required to enhance and classify required land features. The satellite imagery contains large amount of data for analysis and processing. Since human eyes are not sensitive enough to detect small changes in the features such as intensity, colour, or texture, manual processing of the datasets is a reliable method. To extract the features from images, we require an effective classification methodology. The basic steps of an classification procedure are; selection of relevant classification technique, pre-processing of the imagery, training the algorithm, feature extraction, post-classification processing, and accuracy assessment [2].

Classification is broadly categorized as unsupervised and supervised techniques. In supervised classification the analyst provides training data to the classifier algorithm allowing the classifier to learn [4]. In unsupervised classification techniques, the algorithm groups the pixels into clusters according to their characteristics. This method does not require any training. Clustering is an unsupervised classification technique that groups pixels into classes hence pixels in the same classes are belonging to the same group[1], [3].

Classification can also be categorised as hard and soft classification on the basis of whether a pixel is seen as a decomposable unit or not [1], [3]. In hard classification each pixel is constrained to show membership to a single class. Traditional hard classification techniques are parametric in nature. A pixel is assumed to as belonging to only one class. Hard classifiers are most suitable for homogeneous data sets where, land cover classes are well separated in spectral space. However, studies have shown that they do not perform well on high resolution data in the complex environment of the urban area [5].In soft classification, each pixel may display multiple and partial class membership. Soft classification has been proposed in the literature as an alternative to hard classification because of its ability to resolve mixed pixels issues [6]. A pixel is no longer considered as an indecomposable unit in image analysis. Information about pixel component cover classes becomes available, and a pixel's partial membership enables more accurate statistical parameters [6], [8].

The soft classifiers are useful in classifying mixed pixels in remote sensed data. Since the pixels can belong to different cover classes they are given membership grades to indicate their belonging to certain classes. The properties extracted from RS images are used in Fuzzy set theory to perform classification [10]. The supervised and unsupervised classification both can be performed using Fuzzy sets. Unsupervised classification are done using Fuzzy C-Means clustering algorithm which utilises only the RS data for classification. Supervised classification using Fuzzy sets are done by generating statistical values from the ground data that are used to implement membership function models that perform classification of RS data [11], [18]. For example, to obtain models for vegetation spatial information can be extracted from near infrared and red band images [11], [12]. Several modifications for Fuzzy clustering such as single point iterative weighted Fuzzy C-Means and multi thresholds methods are implemented depending upon the input data to perform classification [13], [14].

2. MATERIALS AND METHODOLOGY

The Landsat 8 data used in this study was acquired from www.earthexplorer.usgs.govwebsite [16]. The size of the study area considered for the research is 500 x 500 pixels. The study area is located in Kumta Taluk, North Canara District, India within the rectangular geographical points- 14.4743255 N and 74.4402695 E to 14.440083 N and 74.4651604 E. By visual inspection six land cover classes were identified on the study area; Scrub Land, Water Body, Evergreen Forest, Kharif, Vegetation and Crop land. The technique is implemented in MATLAB® 2015a.

3. FUZZY INFERENCE SYSTEM

Fuzzy Inference System (FIS) are also known as Mamdani system. Fuzzy inference is the process of mapping from a given input to an output using Fuzzy logic. Fuzzy logic is opposite to Boolean logic which not only contains the value 0 and 1, but also contains other values in between 0 and 1. The values between 0 and 1 indicate the extent to which certain input belongs to a particular set. In Fuzzy sets the extent to which an input belonging to all the sets is calculated. Using these values or membership grades the decision are made. Decision making is an important part in this system. The decision is made by the rules and that is why it is also known as Fuzzy rule-based systems. There are two types of FIS namely Mamdani and Sugeno. Mamdani's FIS is the most commonly seen Fuzzy methodology and it expects its output membership functions to be Fuzzy sets. FIS is mainly based on the concept of Fuzzy set theory, Fuzzy IF-THEN rules. FIS uses "IF-THEN" statements, and the connectors present in the rule statement are "OR" or "AND" to make the necessary decision rules.

FIS consists of a Fuzzification interface, a rule base, a database, a decision making unit, and finally a defuzzification interface. The function of each block is as follows:

- A rule base containing a number of Fuzzy IF–THEN rules.
- A database which defines the membership functions of the Fuzzy sets used in the Fuzzy rules.
- A decision-making unit which performs the inference operations on the rules.
- A Fuzzification interface which transforms the crisp inputs into degrees of match with linguistic values.
- A defuzzification interface that transforms the Fuzzy results of the inference into a crisp output.

Therefore FIS first has to perform the Fuzzification of inputs which is done by the membership functions. After calculating the membership grades the rules are applied. The rules consists of Fuzzy logic operators OR and AND. In the rule base the aggregation is performed and Fuzzy set is produced as output. These Fuzzy set are defuzzified using centre of gravity method.

3.1 MEMBERSHIP FUNCTIONS

Membership function is a mathematical function that Fuzzifies an element to obtain membership grades of all Fuzzy sets. The Fuzzy logic toolbox has eleven built-in membership functions. Some of the functions are

- Gaussian membership functions
- Triangular membership functions
- Trapezoidal membership functions
- Piecewise linear functions
- Sigmoid curve

Two types of membership functions that are commonly used in remote sensing for Fuzzification of input data are Gaussian function and triangular function. The graphical and mathematical equations of the membership functions are as follows

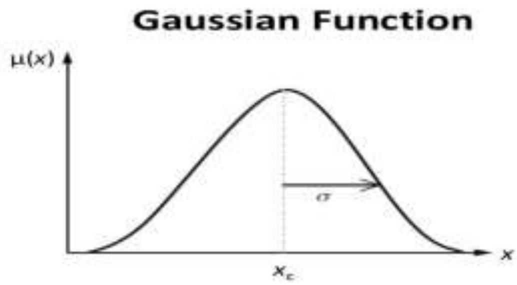

Fig -1: Gaussian membership function

The Gaussian function is defined by a central value m and a standard deviation k>0. The smaller the k is, the narrower the bell is.

$$\mu_A(x) = e^{-\frac{(x-m)^2}{2k^2}}$$

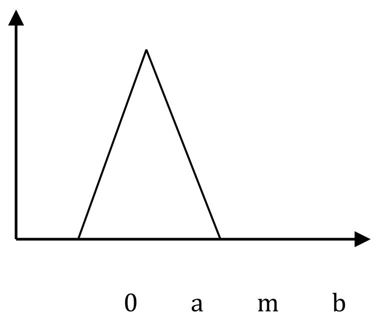

Fig -2: Triangular membership function

The triangular function is defined by a lower limit a, an upper limit b, and a value m, where a< m < b.

$$\mu_A(x) = \begin{cases} 0, & x \le a \\ \frac{x-a}{m-a}, & a < x \le m \\ \frac{b-x}{b-m}, & m < x < b \\ 0, & x \ge b \end{cases}$$

3.2 FUZZY LOGIC OPERATORS

The Fuzzy logic operators are AND, OR, NOT. For example, A AND M operator is minimum of (A, M). A OR M operator is maximum of (A, M) and NOT M is 1-M. The main thing about Fuzzy logical reasoning is that it is a superset of standard Boolean logic. Therefore if we keep the Fuzzy values at extremes of 1 (completely true), and 0 (completely false) only standard logical operations will be performed.

3.3 IF-THEN RULES

The Fuzzy logic system utilises the rules to map the input to output. These rules utilise the membership grades and the operators to obtain the required output. The IF part of the rule is called antecedents and THEN part is called consequents. The rule based form uses linguistic variables as its antecedents and consequents. The rule is expressed as follows: if x is A and y is B THEN z is C.

3.4 CLASSIFICATION PROCEDURE

The land cover classes can be represented in the form of membership functions. The land cover classes in the bands have different pixel values in the range of 0 to 65535. The land covers such as Thin Vegetation, Evergreen forests and so on are represented in the form of membership function in each of the bands. The advantage of using Fuzzy is that same pixel value may fall in to two or more classes at the same time. The FIS takes that into consideration when assigning a pixel to a certain class. However, this method requires training values to be given and the classified output depends on training areas. The membership functions are built using Gaussian membership function. From the selected training pixels the mean and standard deviation of all the required land cover classes are obtained to build Gaussian membership functions.

The mean and standard deviations obtained from the training areas is shown in the Table 1.

Table – 1: UNIVARIATE VALUES OF TRAINING DATA FOR FIS

Channel	Mean	Standard deviation
Evergreen Forest		
Band 6*(mf1)*	38170	1198.8
Band 5*(mf1)*	62287	1940.1
Band 7(mf1)	21012	594.8
Scrub Land		
Band 6*(mf2)*	59686	675.6
Band 5*(mf2)*	38692	934.11
Band 7*(mf2)*	58496	840.86
Thin Vegetation		
Band 6*(mf3)*	49272	1611
Band 5*(mf3)*	54227	3801.8
Band 7*(mf3)*	48709	1488.5
Water Body		
Band 6*(mf4)*	38471	1155.6
Band 5*(mf4)*	37581	1229.9
Band 7*(mf4)*	47237	653.2
Crop Land		
Band 6*(mf5)*	63859	651.5
Band 5*(mf5)*	52932	2178.8
Band 7*(mf5)*	63253	735.63

The Fuzzy sets for the land cover classes are built using these values as shown in Fig.3, Fig. 4 and Fig. 5. The values of the land cover classes in each band change because each band is in particular wavelength and the reflection of Earth's surface features is different in different bands. The Fuzzy rules that link these membership functions is given in Table 2

Table – 2: RULES FOR IMAGE CLASSIFICATION PROCEDURE

IF (**band6** is mf1) AND (**band5** is mf1) AND (**band7** is mf1) THEN (class is **Evergreen Forest**)
IF (**band6** is mf2) AND (**band5** is mf2) AND (**band7** is mf2) THEN (class is **Scrub land**)
IF (**band6** is mf3) AND (**band5** is mf3) AND (**band7** is mf3) THEN (class is **Thin Vegetation**)
IF (**band6** is mf4) AND (**band5** is mf4) AND (**band7** is mf4) THEN (class is **Water Body**)
IF (**band6** is mf5) AND (**band5** is mf5) AND (**band7** is mf5) THEN (class is **Crop Land**)

Three bands namely band 6, band 5 and band 7are chosen as the inputs to the FIS. The chosen bands should contain good separability between other classes. The first input is checked with all the membership functions. For example, here mf1 represents membership function of Evergreen Forest. Similarly mf2, mf3, mf4, mf5 represents Scrub Land, Thin Vegetation, Water Body, Crop Land respectively. Each of these membership functions takes different values in other bands. From the training areas these values are extracted.

When the first input i.e Band 6 pixel value is given, the membership of that pixel with all the five membership function is checked. Therefore this step is called Fuzzification because the probability of the pixel in all the classes is checked. These probability values aare called membership grades. The same steps are repeated for inputs Band 5 and Band 7. The Fuzzified pixels are employed in rules to get a final Fuzzy output.

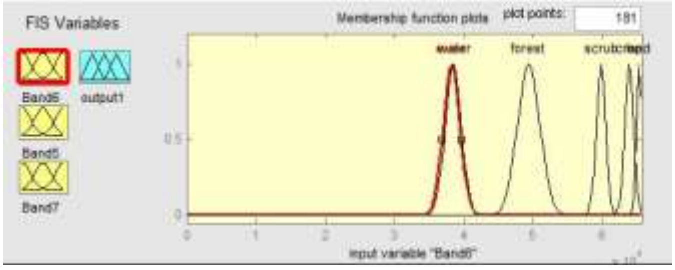

Fig -3: Membership functions for band 6

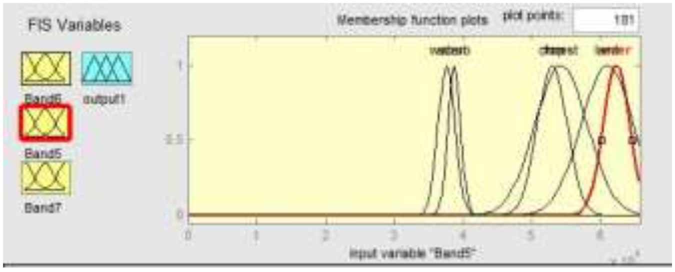

Fig -4: Membership functions for band 5

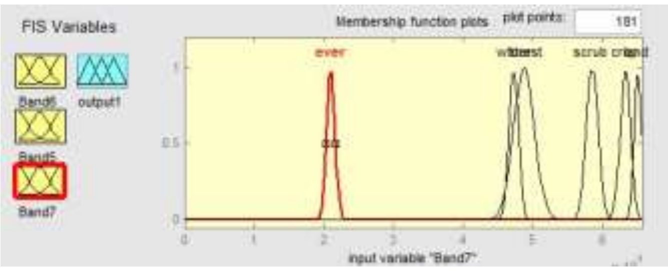

Fig -5: Membership functions for band 7

4. RESULTS AND DISCUSSION

The land cover classification using FIS technique was done for classifying five classes namely Evergreen Forest, Scrub Land, Thin Vegetation, Water Body and Crop Land. The study area of a region Kumta Taluk is shown in Fig. 6. The selected training areas are used to build membership functions and the rule base connects the membership functions. Based on the rules written the input pixels are classified.

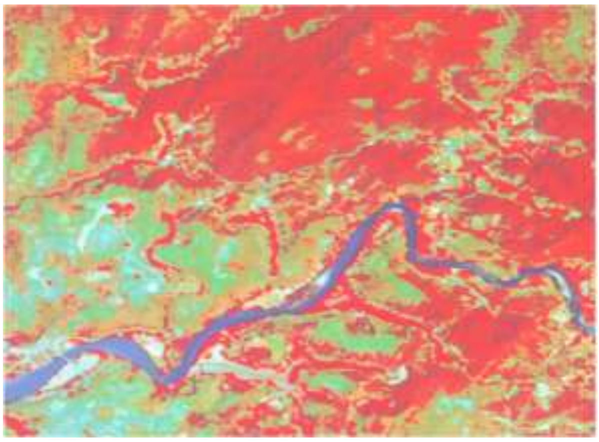

Fig -6: Study area located in Kumta Taluk

Five land cover classes are classified using FIS as shown in Fig. 7. The accuracy of the classifier depends on the training values provided to build the membership function. The accuracy assessment for the classifier is done by building a confusion matrix shown in table. The rule viewer shown in Fig. 8 shows the connection of inputs and outputs. The first column in the Fig. 8 contains the membership functions of band 6 and it checks the belonging of the input to all the membership functions. Similarly the second column shows the membership functions for band 5 and third column shows the membership functions of band 7. In this way the probability of pixels belonging to all the classes are checked. TABLE III shows the confusion matrix which is the statistical assessment of the classifier.

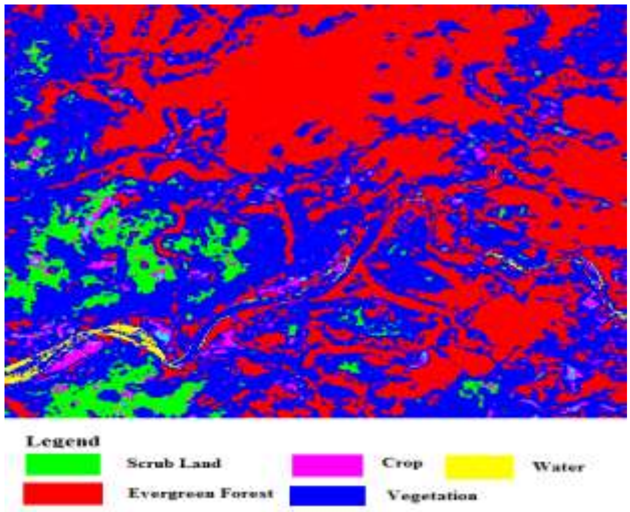

Fig -7: Classification method of FIS method

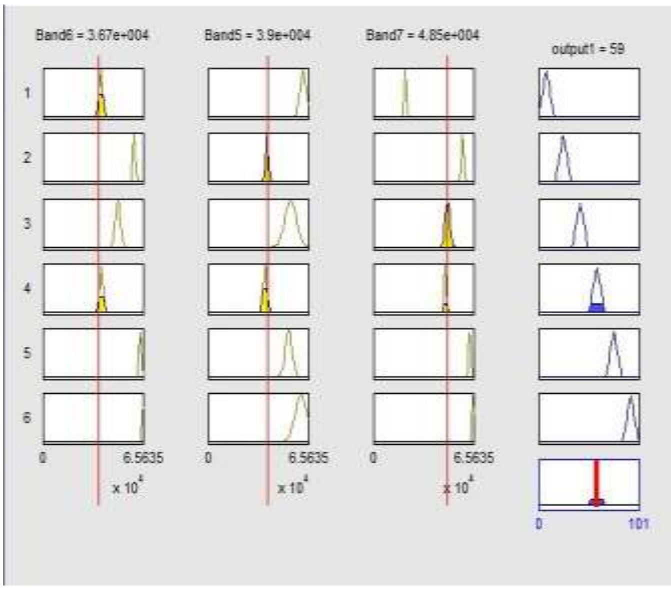

Fig -8: Rule viewer showing connection between input and output

Table -3: Confusion matrix for FIS

A \ P	Evergreen Forest	Scrub Land	Thin Vegetation	Water Body	Crop Land
Evergreen Forest	9136	1	862	0	1
Scrub Land	0	34	27	0	0
Thin Vegetation	0	0	41	0	0
Water Body	0	0	2	23	0
Crop Land	0	0	2	0	39
Overall Classification Accuracy=91.19 %					

A: Actual, P: Predicted

The ground truth data are obtained from cross referencing from websites like Wikimapia and Google maps. The rows of the confusion matrix show the actual classes of the pixels and columns shows the predicted classes by the classifier. For example, 862 pixels of Evergreen Forest have been misclassified to Thin vegetation. The properly classified pixels are shown along the diagonals of the matrix. The overall classification accuracy obtained is 91.19 %.

5. CONCLUSION

The success of an image classification in remote sensing depends on many factors, the availability of high-quality Remotely Sensed (RS) imagery and ancillary data, the design of a proper classification procedure, and the analyst's skills and experiences. The advantage of the FIS is the ability to implement the classification using simple IF-THEN rules. This system allows adding any number of input bands (multiband). It helps to make use of the all the Landsat 8 bands which contain more information than the conventional three band combination input.

REFERENCES

[1] J. A. Richards, Remote Sensing Digital Image Analysis, 5th ed. Springer, 2013.

[2] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," Int. J. Remote Sens., vol. 28, no. 5, pp. 823–870, 2007.

[3] J. R. Jensen, Introductory digital image processing: a remote sensing perspective, 3rd ed. New Jersey: Prentice-Hall Inc., 2000.

[4] U. Maulik and D. Chakraborty, "Remote Sensing Image Classification: A survey of support-vector-machine-based advanced techniques," IEEE Geosci. Remote Sens. Mag., vol. 5, no. 1, pp. 33–52, 2017.

[5] N. K. Patel, T. T. and Medhavy, C. and Patnaik, and A. and Hussain, "Multi temporal ERS-1 SAR data for identification of rice crop," J. Indian Soc. Remote Sens., vol. 23, no. 2, p. 33, 1995.

[6] M. Ji and J. R. Jensen, "Fuzzy Training in Supervised Digital Image Processing," Geogr. Inf. Sci., vol. 2, no. 1, pp. 1–11, 1996.

[7] G. Droj, "The applicability of fuzzy theory in remote sensing image classification," Stud. Univ. Babes, Bolyai, Inform., vol. LII, no. 1, pp. 89–96, 2007.

[8] F. Wang, "Fuzzy Supervised Classification of Remote Sensing Images," IEEE Trans. Geosci. Remote Sens., vol. 28, no. 2, pp. 194–201, 1990.

[9] L. A. Zadeh, "Fuzzy sets," Inf. Control, vol. 8, no. 3, pp. 338–353, 1965.

[10] J. C. Bezdek, Pattern Recognition with Fuzzy Objective Function. New York: Plenum Press, 1981.

[11] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," IEEE Trans. Geosci. Remote Sens., vol. 42, no. 8, pp. 1778–1790, 2004.

[12] R. Jeansoulin, Y. Fontaine, and W. Frei, “Multitemporal Segmentation by Means of Fuzzy Sets,” in Seventh International Symposium on Machine Processing of Remotely Sensed Data with special emphasis on Range, Forest and Wetlands Assessment, 1981, pp. 336–339.

[13] R. L. Cannon, J. V. Dave, J. C. Bezdek, and M. M. Trivedi, “Segmentation of a Thematic Mapper Image Using the Fuzzy c-Means Clusterng Algorthm,” IEEE Trans. Geosci. Remote Sens., vol. GE-24, no. 3, pp. 400–408, 1986.

[14] J. T. Kent and K. V. Mardia, “Spatial classification using fuzzy membership models,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 10, no. 5, pp. 659–671, 1988.

[15] C. Li, J. Yin, and J. Zhao, “Extraction of urban vegetation from high resolution remote sensing image,” in 2010 International Conference On Computer Design and Applications, 2010, pp. V4-403-V4-406.

[16] USGS, “Earthexplorer.” [Online]. Available: https://earthexplorer.usgs.gov/. [Accessed: 01-Jan-2017].

[17] J. R. Jensen et al., “Remote sensing agricultural crop type for sustainable development in South Africa,” Geocarto Int., vol. 21, no. 2, pp. 5–18, 2006.

[18] B.R. Shivakumar, S.V. Rajashekhararadhya, "Fuzzy Methodology for Enhancing the Classification Accuracy of Spectral Correlation Mapper Classifier: A Case Study over North Canara District, India," ISOR journal of Applied Geology and Geophysics, vol. 5, pp. 26-33, 2017