Classification of Feature Clustersbased on Fuzzy Rules from Satellite Image

Shrinivas Khandare, Urmila Shrawankar

Abstract—The features on the earth are changing rapidly. Therefore the need arises to detect the changes occurring on the earth and use in different applications simultaneously. This paper proposes fuzzy rule based technique of feature classification from images. In this technique, first the maximum likelihood algorithm is applied by replacing the central pixel with the surrounding pixels reflectance values forming clusters and secondly the standard deviation and mean of the feature cluster obtained are used as an input in framing the fuzzy rules. Then image is given as input to the fuzzy rule based classification technique. The result obtained is compared with the previously existing algorithms. These techniques were performed on hundred images. The obtained accuracy is 80 % correct classification and 20% for misclassification. This can be applied to various applications of image processing and data retrieval systems.

Index Terms—Kappa statistics, Range values, High resolution Pixels, Gaussian distribution function, Satellite image, LISS3, Awifs

1 INTRODUCTION

Image is the basic entity of the any visual interpretation. Visual image is more worth than thousand words. The images contain lot of information of the earth such as land, road, waterbody etc. These are captured by satellite revolving around the earth by the camera fitted on it. The camera sensors are of different types such as LISS3, AWIFS, IKONOS etc. These sensors have different resolution, and vary with different bits and size. These are captured in multiple bands. The wavelength of the visible spectrum ranges from 0.4 um to 0.8 um. The features are needed to be extracted and classified according to the requirement and satellite images availability.

Many of the researchers have developed various techniques developed so far which are elaborated in the literature section. In this work, a method is proposed which extract the multiple road features depending upon the range values, minimum, maximum values and combination of all the results of different steps into final output image.

The motivation of the work is to identify a technique which will classify the given satellite image. And find the applicability of the technique theoretically and experimentally. These existing techniques are Minimum distance to mean, Maximum likelihood, Mahalanobis distance based classifier. For the experiments, different classifiers builton these approaches are used

The most commonly used techniques are supervised and unsupervised classification, while some researchers have used semi-automatic, iterative procedures for extracting the information.

- *Shrinivas Khandare is currently pursuing masters degree program in computer science and engineering in GHRaisoni College of Engineering, MS,India, E-mail: shrinivas.khandare@gmail.com*
- *Dr. Urmila Shrawankar is currently Associate Professor in computer science and engineering in GH Raisoni College of Engineering, MS, India, E-mail: urmila@ieee.org*

It involves perspective human interaction. Supervised classification technique are used to identify the features based on the training samples and in unsupervised classification prior number of clusters are pre-confirmed before applying the classification technique.

The utilization of minimum distance classification methods in solving image classification problems such as landuse classification is considered. In minimum distance classification, a sample training that is group of vectors is classified into the class, whose known or estimated distribution of the sample to be classified. The measure of similarity is distance measure in the space of distribution functions.

For generation of map, the conventional method of collection of data is exhaustive. Therefore, the process of map making is lengthy and it is time consuming. Due to this, the map becomes outdated when it comes to publishing of the map.The development process has increased fastly which leaded to large occupation of features which are not ignorable. Newer feature of roads are developed and old roads have diminished of the roads is required to be noted for the map to inculcate perfect information of the earth surface. These changes have delivered the cartographers new tools for generating and modifying maps and created a path to mapping in sophisticated way. This brought new magnitude of natural paradigm. Hence, there exists a requirement of digital imagery for creating and modifying the street network of the city and for updating the street guide of city using image processing and computer science techniques cannot be over emphasized.

Street guides who are available are out-of- date and misleading. Proposed work is designed to map the streets from satellite image is input taken for the updation of the street networks, newer version of the map is created which motivates young researchers and deep cartographic tasks are possible with limited software involvement which results into accurate results and analysis of the features. By the use of satellite image the results and analysis reveals the importance of the work.

Just as Olaore (2004) the factors affecting the whole process of map making are quality, accuracy. This is observed when GPS data were used for creating and updating the street map of

Kaduna USA. It is observed that satellite data gives the coverage of the area focused, with repeated coverage and synoptic view. It also gives the real time information that can be utilized to update maps. Thus, the use of high-resolution multi spectral images can be used for mapping and updating of urban street networks.

2 METHODOLOGY

2.1 METHODOLOGY AVAILABLE

K-means: In this method, the classification is very sensitive to initial starting values. If the initial values of the two classifications are different and the result of the classifications are also different, than, the classification with the smallest initial values is required to be chosen. Since, this is the objective function which is to be minimized. But, it is found that MSE are generally very small, for two classifications very much. But, by observation, it cannot be judged that the smaller difference value of the MSE is really better than the other.There are various techniques developed so far are listed in the below table. The various existing techniques of Classification are ISODATA, K-Nearest Neighbour, Minimum distance classification, Parallelepiped classifier, Fuzzy c-Mean, Maximum likelihood classification, Spectral angle mapper classification, Spectral Correlation mapper classification, artificial neural network (ANN) classifier, Mahalanobis distance classification, Channeling Techniques.

Supervisedclassification: In this method the strategy is simple: the specialist meaningful (but somewhat Specialist by self experience, with the personal experience of the specialist from a particular region, or by visiting sites, knowledge about the surrounding features chooses and setups distinct classes which are selection based on supervision of the output and allocate them in real category. It is heavely dependent upon the sense and skills of the visual interpreter specialist. The specialist added skill is to find discrete classes on the image. Training sites are group of homogenous pixels of land cover/use class. It is found by similarity of the tones, within the shape boundary covering the feature. Specialist digitize the exact boundary of the features depends upon the tone, shape

Mahalanobis distance measure classification: It is well-known statistical distance function which classifies features based on Mahalanobis distance. The distance between any two data points in the feature space is known as Mahalanobis distance. These two data points are correlated with each other. The variability can be measured by incorporating it into the distance metric directly. Here the corrections of the various data points are taken into consideration. If there, occurs two data points in 2D space, which are not correlated among each other and treated as non-correlated variables than the Mahalanobis distance is the same as the Euclidean distance. In terms of Mathematics, the Mahalanobis distance is equal to the Euclidean distance when the covariance matrix is the unit matrix.

2.2 METHODOLOGY PROPOSED

2.2.1 System input satellite data imagery:

For this work high –IKONOS and mid resolution- LISS3 satellite data is input to the system.

2.2.2 **Pre-processing of high resolution satellite data:**Input image is processed for errors which occur due to atmosphere and radiation of the objects on the earth. The corrections are applied as mentioned below:

2.2.2.**1 Geometric corrections of an image**

The image is captured from the satellite. But due to instability in the movement of the satellite, the angle is change by some degrees. The image is corrected for the camera angle distortions.

2.2.2.2**Radiometric corrections of an image**

The radiometric corrections are due to radiometric errors such as noise in the atmosphere, camera defects, loss of data due to stripping etc. Input image is corrected to show the emitted reflectance as it is on the ground by IHS correction technique.

2.2.2.3**Contrast Enhancement of an image**

From the available range of 8-bit or 256 levels in an input image. The histogram equalization technique identifies the input image to represent the original data in a useful way. Since only a small percentage of data is reflected from the available range and to increase the divergence between background data and the target data.

2.2.3**Identification of target class:**

Trainer should have good knowledge of the objects on the earth in terms of types and location of features. The output classes are Major road, Waterbody, Settlement, All road, Canal, Open-land etc. The output image is required to be converted to *.tiff format.

2.2.4 Implementation of Semi-Automatic traditional algorithms

2.2.4.1 **K-means classifier**

1.1 x1,…, xN are data points or vectors of pixels/observations [11][29]
1.2 Each pixels/observation (vector xi) will be assigned to one and only one cluster [11][29]
1.3 C(i) denotes cluster number for the ith pixel/observation [11][29]
1.4 Dissimilarity measure: Euclidean distance metric
1.5 K-means minimizes within-cluster point scatter:

$$W(C)=\frac{1}{2}\sum_{k=1}^{K}\sum_{C(i)=k}\sum_{C(j)=k}\left\|x_i-x_j\right\|^2=\sum_{k=1}^{K}N_k\sum_{C(i)=k}\left\|x_i-m_k\right\|^2$$

[11][29]

Where

i. mk is the mean vector of the kth cluster
ii. Nk is the number of pixel/observations in kth cluster

1.6 For a given cluster assignment C of the data points, compute the cluster means mk:

$$m_k=\frac{\sum_{i:C(i)=k}x_i}{N_k},\ k=1,\ldots,K. \quad [11][29]$$

For a current set of cluster means, assign each pixel/observation as:

$$C(i) = \arg\min_{1\le k\le K} \left\| x_i - m_k \right\|^2,\ i = 1,\ldots,N \qquad [11][29]$$

1.7 Iterate above two steps until convergence

2.2.4.2 Isodata classifier

i. An enhancement of the approach taken by the k-means algorithm. [45]
   a. "k" is allowed to range over an interval
   b. Discard clusters with too few elements
   c. Merge clusters [45]
      i. Too large or too close
   d. Split clusters
      i. Too few or containing very dissimilar samples[45]

2.2.4.3 MDM (Minimum Distance to Mean) classifier

Step 1: The user supplies the spectral class means in n-dimensional space and calculate the distance between the candidate pixel and each of the class.
Step 2: The candidate pixel is assigned the class with the smallest Spectral Euclidean distance (minimum distance) to the candidate pixel.
Step 3: The distance is calculated using either an n-dimensional Pythagorean Theorem.

$$\mathrm{Distance}(X,\mu_k) = \sum_{i=1}^{n}\left((X-\mu_K)^{2}\right)^{1/2}$$

[43]

Step 4: For (Class land cover=0) to length (Class)

Calculate the mean vector in (Red, Green, Blue) band.

Step 5: if every pixel = nearest class/cluster than Class=pixel;
Step 6: Define a limit beyond which a pixel remains unclassified.
For each pixel, finding the closest cluster mean is simple and fast, but what about the data points which are far from the mean of the cluster. Minimum distance to mean does not take variance of the data points into consideration which is insensitive to the variance of clusters. Therefore, variance and covariance can be considered for better allocation of pixels to their appropriate class [43].

2.2.4.4 MD (Mahalanobis distance) classifier

Step 1: The user supplies the spectral class means in n-dimensional space and the algorithm calculates the distance between the candidate pixel and each of the class.
Step 2: The candidate pixel is assigned the class with the Spectral Euclidean distance (Mahalanobis distance) to the candidate pixel.
Step 3: The distance is calculated using either an n-dimensional Pythagorean Theorem.

Step 4: For (Classland cover=0) to length (Class)

Calculate the mean vector in (Red, Green, Blue) band.

$$d_{mahalanobis} = \sqrt{(x-y)^T \textstyle\sum^{-1}(X-Y)} \quad [40]$$

Step 5: if every pixel = nearest class/cluster than Class=pixel;
Step 6: Define a limit beyond which a pixel remains unclassified [40]

2.2.4.5ML (Maximum Likelihood) classifier

Step 1: One of the supervised classification technique is Maximum Likelihood (ML) which is derived from the Bayes theorem. In this, a posteriori distribution $P(\frac{k}{x})$, i.e. pixels having denoted by vector x belong to class is given by equation (1).
Step 2: The pixel with the maximum likelihood is assigned to the class is determined by discriminant function.
Step 3: First the prior probabilities of all the classes are calculated from the feature space of the population
Step 4: Find what percentage of pixels of the class have occurrence in the population.
Step 5: The inputs of the discriminant function are 1) Mean of the class 2) Covariance of the class. These values are estimated out of the training pixels of a class.
Step 6: After, the prior probabilities of the classes are obtained, than the pixels which are required to be classified in their particular class are identified. The probability of the pixel is calculated by the probabilities that the pixel belong to that particular class divided by the summation of all the probabilities of this pixel in all other classes. This probability is given by P(k). It is also the initial probability of class k which is calculated as a decision rule with the percentage of occurrence of these types of homogenous pixels in the overall population. $P\left(\frac{X}{k}\right)$is calculated as the observed value of X from the class k [43]. It is also known as probability density function. $P(i) * P\left(\frac{X}{i}\right)$is the pixel probability to belong to class i [43]. The sum of all such probabilities are calculated and given as

$$P(k) * P\left(\frac{X}{k}\right) \qquad [43][34] \qquad (1)$$

$$\sum P(i) * P\left(\frac{X}{i}\right) \qquad [43][34] \qquad (2)$$

Thus, the likelihood is calculated as

$$L_k = P\left(\frac{k}{X}\right)\frac{P(k)*P(\frac{X}{k})}{\sum P(i)*P(\frac{X}{i})} \qquad (3)[34][43]$$

Step 7: The population data is assumed with Gaussian function to find the log likelihood or discriminant function [17]. This multivariate distribution function works as probabilities density function.

IJSER

$X \varepsilon i$ if $P(i/x) > P(j/x)\ \forall j \neq i$; (4) [17] [43][34]

The equation for the likelihood is given as for normal distributions. In this equation 4, is the number of n number of bands.

$$L_k(X) = \frac{1}{(2\pi)^{\frac{n}{2}} |\sum_k|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(X-\mu_k)\sum\nolimits_k^{-1}(X-\mu_k)^t\right)$$

[43][34][7] (4)

Where: "m" → number of bands, "n" → number of pixels.

1. X → Image data which belongs to n number of bands/channels.
2. $L_k(X)$ → Likelihood of X vector fitting to class k.
3. μ_k → Mean vector of class k.
4. "X" → Image data which belongs to "n" number of bands.
5. L_k→ Likelihood of "X" vector fitting in class k.
6. u_k → Class (k) mean vector.
7. $\sum k$ → Sum of (variance and covariance) matrix of class (k).
8. $|\Sigma k|$ →Determinant of Σk.

Eq. (4) can be written as follows, Log in the above eq. (4) represents monotonic of a function:

$X\ \varepsilon\ i$ if $g_i(i/x) > g_j(j/x)$ Y $\forall j \neq i$; [17] [43][34]

Mean = $\frac{1}{n}\sum_{j=1}^{n} X_{ij}$ where i ranges from 1 to m. (5)

Variance and Covariance $(\sum e) = \frac{1}{m}\sum_1^m (X_i - \mu_e)$ [34] [43][17] (6)

Step 8: Every individual pixel is allocated to the desired class with the maximum likelihood of belonging to that class. And the pixel is unclassified if the probability is not equal or below the threshold value given by the user [42].

The steps are as follows:

1. The total number of feature is determined for the satellite image.

2. The sample sites are found for each output class. For this the Jeffries-Matusita (JM) distance is utilised which thresholds the class separability of the sample sites taken [44].

3. The sample sites as taken to estimates the mean vector and covariance of the class.

4. Pixels are classified as desired output of land feature class or treated as unclassified.

Step 9: From the above equation (5), (6), estimation is

calculated before using the maximum likelihood algorithm whether the distribution of the sample data will fit into the normal distribution of the population or not [34].

2.2.4.5 Implementation of Proposed Fuzzy based classification

In the above classification Rules are framedfor Vegetation, Field plot, Settlement, Major road, Waterbody, all roads. The corresponding rules areas follows:

Rule 1: If Green→ g_1 and Red → g_1 and NIR→g_1 then Class Vegetation
Rule 2: If Green→g_2 and Red →$g_2$and NIR→$g_2$then Class Field plot
Rule 3: If Green→$g_3$and Red →$g_3$and NIR→$g_3$then Class Settlement
Rule 4: If Green→$g_4$and Red →$g_4$and NIR→$g_4$then Class Major Road
Rule 5: If Green→$g_5$and Red →$g_5$and NIR→$g_5$then Class Waterbody
Rule 6: If Green→$g_6$and Red →$g_6$and NIR→$g_6$then Class All roads)

2.2.4.6 Algorithm for Fuzzy rule based classification

Step 1: Create membership functions (gi) for each class.
Step2: Membership function is Gaussian Normal distribution, as it uses mean and standard deviation as the input.
Step3: The names of the model and names of the membership functions are set as input.
Step 4: For each class the Red, Green, NIR membership functions are newly assigned.
Step 5: The bands (Red, Green, NIR) are assigned rules as Red= g1, Green= g1, NIR= g1, Output= vegetation.
Step 6: Similarly, Bands are assigned Rules for other classes.
Step7: The output variable (Vegetation, Field plot, Settlement, Major road, Waterbody, all roads).
Step 8: The surface view is generated.

2.2.4.7 Comparison of Result:

The Maximum likelihood, Mahalanobis distance and Minimum distance to mean classifier are compared theoretically and experimentally to generate the classified image with the feature class as shown in figure 1.

2.2.4.8 Analysis:

The results are obtained from Maximum likelihood, Mahalanobis distance; Minimum distance to mean classifier, k-means, Isodata, fuzzy is analysed to find the number of pixels covered by the classes.

2.3 Output :

The desired output image will be converted into .img or .tif.

2.3.4.4 Accuracy assessment:

Method 1: The image contains total 13, 29,409 numbers of pixels. These pixels are required to be classified. The values are obtained from the table. The output image with class (Vegetation, Field plot, Settlement, Major road, waterbody, canal, all roads, and open land) is obtained with each class having individual number (quantity) of pixels statistics in the form of table etc. As a result, we get number of pixels (out of 13, 29,409 pixels) in each class. To show in percentage, we have calculated the percentage value from the total value (13,29,409 pixels). The results are vegetation with (261705 pixels) which is equal to 19% of total 13, 29,409 numbers of pixels). Similarly for other class we obtain values.

By averaging:
Step 1: For eg; vegetation class
In Table 1: calculate: SUM of (A) + (B)+(C)+(D)+(E) = 143.30 mentioned in Column (G);
Average = $\frac{\text{Sum}}{5}$ = 28.66;
Step 2: Check for Fuzzy (F) what is the increase/ decrease of pixels by subtracting (G) – (F) = -3.03;
Similarly,
Step 3: For eg; Field Plot class
Follow step: 1 for other class
Step 4: Total increase of pixels is 19.42% from the total number of pixels in all the class.
Step 5: Therefore, the Accuracy achieved is given by total (100) -19.42 = 80.58%

Method 2: It considers user, producer accuracy and kappa statistics for accuracy calculation in detail. Accuracy of classification is done to assess the classified output. Before proceeding to accuracy assessment, reference data is identified for feature class type at defined regions. The gis layers is utilised for referencing the data. Stratified sampling is done to place the observed values. Suggested samples are 50 per class [16] as a thumb rule. This determines class types from reference source and classified map as shown in table (3),(5). Error matrix is used to quantify the comparison. [16]

The total accuracy is given by:

$$(\text{Number of correct plots})/(\text{Total number of plots})$$

[16][37] Eq. (1)

Diagonal element values in Table (3) (5), represents class that are correctly classified as per the reference data. Non- diagonal elements are misclassified sites. But, total accuracy is calculated by the average of the diagonal elements, but it does not infer that the error was distributed evenly between the classes or cannot determine, really few classes are good and bad. Therefore, introduced new type of accuracy such as accuracy of user and accuracy of producer. User's accuracy (corresponds to error of omission-(inclusion)) and producer's accuracy (corresponds to error of commission (exclusion). In producers's accuracy, from the perspective of the maker of the classified map, how accurate is the map? For a given class in reference map, how many of the pixels on the map are labelled correctly? [16]. It is given by:

Users accuracy=

$$(\text{in a given map class Numbers identified correctly})/(\text{Number claimed to be in that map class})$$

[16] Eq. (2)

Producers accuracy=

$$(\text{in ref. plots of a given class Number correctly identified})/(\text{Number actually in that reference class.})$$

[16] Eq. (3)

Kappa statistics ($\widehat{K}$) measure the agreement between two or more observers, include a statistic that takes into account the fact that observers will sometimes agree or disagree simply by chance [16]. It is given by:

$$\hat{K} = (\text{Observed accuracy} - \text{chance agreement})/(1 - \text{chance agreement})$$

[16][37]

Where:

a) Diagonal elements represents observed accuracy in error matrix.
b) Off diagonal represents chance agreement- sum(for each class, product of row and column total) [16] [37]

Kappa statistics ($\widehat{K}$) is calculated by:

$$\widehat{K} = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}(x_{i+} \times x_{+i})}{N^2 - \sum_{i=1}^{r}(x_{i+} \times x_{+i})}$$

[16] [37] Eq. (4)

Where:

1. N → Total number of sites in the matrix,
2. r → Number of rows in the matrix,
3. xii → Number in row i and column i,
4. x+i →Total for row i
5. xi+ → Total for column i

Higher the value Kappa obtained better classification results obtained. [16][37]

Table. 1 Result of classification of different classifiers with % of pixels occurring in each class

Class	Mahalanobis Distance (%) of total number of pixels. (A)	Maximum Likelihood (%) of total number of pixels(B) (Benchmark)	Minimum distance to Mean (%) of total number of pixels(C)	K-means, (%) of total number of pixels(D)	Isodata (%) of total number of pixels(E)	Fuzzy classification rule based (%) of total number of pixels (F).	(G) Avg. of class (A)+ (B)+ (C)+ (D)+ (E)	(H) Percentage of increase/ decrease of pixels in each class (%) (G) –(F)
Vegetation	19.68581528	36.799	25.74941	25.50	35.55	18.82	28.66	-3.03%
Field Plot	8.319035	3.278299	10.94464	9.56	7.52	11.615	07.92	+3.67%
Settlement	8.067043325	8.727487	9.364161	7.56	8.52	9.67	08.45	+1.22%
Major Road	29.82708858	11.80179	19.19387	10.23	9.6523	22.56	16.14	+6.42%
Waterbody	3.972500	7.939844	16.53893	4.35	5.688	12.65	07.69	+4.69%
Canal	10.35174277	6.003344	2.695559	5.36	6.574	9.58	06.19	+3.39%
All Roads	18.65678659	7.109099	10.2183	17.24	8.57	13.925	12.36	-1.84%
Open Land	1.119971356	18.34101	5.295135	20.2	17.97	1.18	12.59	-11.41%
Total Number of Pixels (I)	1329409	1329409	1329409	1329409	1329409	1329409	100	+19.42% (total increase)

IJSER

Total increase of pixels in each class is 19.42% from (I). Therefore, the Accuracy is 80.58%

Table 3. Error Matrix of Maximum Likelihood algorithm

ERROR MATRIX-MAXIMUM LIKELIHOOD ALGORITHM												
	Reference Data											
Classified Data	0	Shadow	Trees	Waterbody	Maj-road	Road	Pathway	Open-land	Gullies	Minor-road	Building	Row Total
0	1	0	0	0	0	0	0	0	0	0	0	1
Shadow	0	15	1	0	0	0	0	0	0	0	0	16
Trees	0	0	6	0	0	0	0	0	0	0	0	6
Waterbody	0	0	0	7	1	0	0	0	0	0	0	8
Major road	0	0	0	1	7	0	0	0	0	0	0	8
Road	0	0	0	0	0	11	0	1	0	0	0	12
Pathway	0	0	0	0	0	1	12	0	1	0	0	14
Open-land	0	0	0	0	0	0	2	11	1	0	0	14
Gullies	0	0	0	0	0	0	0	2	6	0	0	8
Minor-road	0	0	0	0	0	0	0	0	0	3	0	3
Built-up	0	0	0	0	2		1	0	1	5	0	10
Column Total	1	15	7	8	10	12	15	14	9	8	1	100

Dark colored box represents diagonal values for each class which are input in the calculation of overall accuracy in equation (1). Kappa from Equation (4) is derived as

$$\widehat{K}= ([100 \times (1+15+6+7+7+11+12+11+6+3+0)] - [1\times1+15\times16+7\times6+8\times8+10\times8+12\times12+15\times14+14\times14+9\times8+8\times3+1\times10])/(100\times 100 - [1\times1+15\times16+7\times6+8\times8+10\times8+12\times12+15\times14+14\times14+9\times8+8\times3+1\times10] = 0.764$$

Table 4. Accuracy totals- Maximum Likelihood algorithm

ACCURACY TOTALS- MAXIMUM LIKELIHOOD ALGORITHM					
Name of Class	Totals No of Reference sites	Totals No of Classified sites	Number of sites Correct	Accuracy of producers -eq.(3)	Accuracy of Users -eq.(2)
0	1	1	1	--	--
Shadow	15	16	15	100.00%	93.75%
Trees	7	6	6	85.71%	100.00%
Waterbody	8	8	7	87.50%	87.50%
Major road	10	8	7	70.00%	87.50%
Road	12	12	11	91.67%	91.67%
Pathway	15	14	12	80.00%	85.71%
Open-land	14	14	11	78.57%	78.57%
Gullies	9	8	6	66.67%	75.00%
Minor-road	8	3	3	37.50%	100.00%
Built-up	1	10	1	100.00%	10.00%
Column Total	100	100	80		
Overall Classification Accuracy = 80.00% eq.(1)					
End of Accuracy Totals					

Overall Classification Accuracy = 80.00% by using equation (1).

Table 5. Error Matrix of Fuzzy algorithm

ERROR MATRIX-FUZZY ALGORITHM												
	Reference Data											
Classified Data	0	Shadow	Trees	Waterbody	Maj-road	Road	Pathway	Open-land	Gullies	Minor-road	Building	Row Total
0	1	0	0	0	0	0	0	0	0	0	0	1
Shadow	0	15	1	0	0	0	0	0	0	0	0	16
Trees	0	0	6	0	0	0	0	0	0	0	0	6
Waterbody	0	0	0	7	1	0	0	0	0	0	0	8
Maj-road	0	0	0	1	7	0	0	0	0	0	0	8
Road	0	0	0	0	0	11	0	1	0	0	0	12
Pathway	0	0	0	0	0	1	12	0	1	0	0	14
Open-land	0	0	0	0	0	0	1	12	1	0	0	14
Gullies	0	0	0	0	0	0	0	2	6	0	0	8
Minor-road	0	0	0	0	0	0	0	0	0	3	0	3
Building	0	0	0	0	2	0	1	0	1	5	0	10
Column Total	1	15	7	8	10	12	14	15	9	8	1	100

Dark coloured box represents diagonal values for each class which are input in the calculation of overall accuracy in equation (1). Kappa from Equation (4) is derived as

$$\hat{K} = ([100 \times (1+15+6+7+7+11+12+12+6+3+0)] - [1 \times 1 + 15 \times 16 + 7 \times 6 + 8 \times 8 + 10 \times 8 + 12 \times 12 + 15 \times 14 + 14 \times 14 + 9 \times 8 + 8 \times 3 + 1 \times 10])/(100 \times 100 - [1 \times 1 + 15 \times 16 + 7 \times 6 + 8 \times 8 + 10 \times 8 + 12 \times 12 + 15 \times 14 + 14 \times 14 + 9 \times 8 + 8 \times 3 + 1 \times 10]) = 0.775$$

Table 6. Accuracy totals of Fuzzy algorithm

ACCURACY TOTALS-FUZZY algorithm					
Name of Class	Totals No of Reference sites	Totals No of Classified sites	Number of sites Correct	Accuracy of producers -eq.(3)	Accuracy of Users -eq.(2)
0	1	1	1	--	--
Shadow	15	16	15	100.00%	93.75%
Trees	7	6	6	85.71%	100.00%
Waterbody	8	8	7	87.50%	87.50%
Major road	10	8	7	70.00%	87.50%
Road	12	12	11	91.67%	91.67%
Pathway	14	14	12	85.71%	85.71%
Open-land	15	14	12	80.00%	85.71%
Gullies	9	8	6	66.67%	75.00%

ISSN 2229-5518

Minor-road	8	3	3	37.50%	100.00%
Built-up	1	10	1	100.00%	10.00%
Column Total	100	100	81		
Overall Classification Accuracy = 81.00% eq.(1)					
End of Accuracy Totals					

Overall Classification Accuracy = 81.00% by using equation (1).

Table 2. Total accuracy of the classification

Method	Percentage of total Classified pixels	Percentage of Total Misclassified pixels	Accuracy (Percentage of Total pixels of image)
Maximum likelihood classification(ML)	78	22	78
Fuzzy Rule based classification	80.58	19.42	80.58
Mahalanobis Distance based classification(MD)	75	25	75
Minimum distance to Mean based(MDM) classification	76	24	76
K-means	74	26	74
Isodata	74	26	74

Table 7. COHEN'S KAPPA ($\widehat{K}$) STATISTICS OF ALGORITHMS.

ML	Overall Kappa Statistics ($\widehat{K}$)	0.764
MD	Overall Kappa Statistics($\widehat{K}$)	0.749
K-MEANS	Overall Kappa Statistics($\widehat{K}$)	0.676
FUZZY	Overall Kappa Statistics($\widehat{K}$)	0.775
MDM	Overall Kappa Statistics($\widehat{K}$)	0.758
ISODATA	Overall Kappa Statistics($\widehat{K}$)	0.685

4 RESULT AND ANALYSIS

As a result, number of pixels in each class as shown in Table 1 above with percentage of pixels. Considering each algorithm iteratively. The output of each algorithm in the result obtained, are the group of pixels (from the total number of pixels in the image) in each class. In Table 1, it is shown that there are total 13,29,409 pixels in the last row. Out of these 13, 29,409 total, pixels were separated output class. If we count number of pixels of each class than total number of pixels which summed up are 13,29,409 pixels. In Table 1 the numbers of pixels are shown with the percentage (%) symbol. Hence, in each row of table the values show the percentage of pixels in the class from the total number of pixels in the image that is 13,29,409 pixels.
The results are obtained from the proposed and the existing algorithms as shown in Table 1 and its graphical representation in figure 1. The results are compared with the previous algorithms of the researchers such as k-means, Isodata, Minimum distance to mean, Mahalanobis distance, Maximum likelihood etc. The compared results are shown in Table 1 above in rows and columns.

Different classification techniques are applied on the LISS3 imagery. The desirable classes are vegetation, field plot, settlement, canal, major-road, waterbody, all roads, openland are obtained. The histogram values are also obtained as shown in table 2 for all the algorithms in all the classes. The percentage of histogram value from the total number of 13492 pixels as shown in Table 1 is calculated.
From the Table 1, it is inferred that Maximum likelihood shows 36.799% of maximum % of pixels as compared to Mahalanobis- 19.685% and 25.749% for Minimum distance to mean classifiers. For field plots Minimum distance to mean shows maximum of 10.944% as compared to Mahalanobis of 8.319% and 3.278% of ML. For settlement it is highest with the Minimum distance to mean with 9.364% as compared to 8.727% for Mahalanobis distance classifier. For Major road, Mahalanobis distance classifier shows 29.827% of pixels as compared to Maximum likelihood with 11.801% of pixels which is an increase of 10.7% and increase of 10.633% of pixels in Minimum distance to mean. For waterbody, Minimum distance to mean shows 16.538% of pixels which is an increase of 8.599% and 12.564% of increase from Mahalanobis of 3.972% of pixels. For Major roads 10.35% is shown with maximum number of pixels which is an increase of 4.351% Maximum likelihood and increase of 7.656% from Minimum distance to mean of 2.695% of pixels. For all roads, 18.656% for Mahalanobis distance, 7.109% for Maximum likelihood, 10.218% for Minimum distance to mean. It is the class other than the major road, canal, waterbody features. For open land Maximum likelihood 18.341%, 5.295% for Minimum distance to mean, and

IJSER

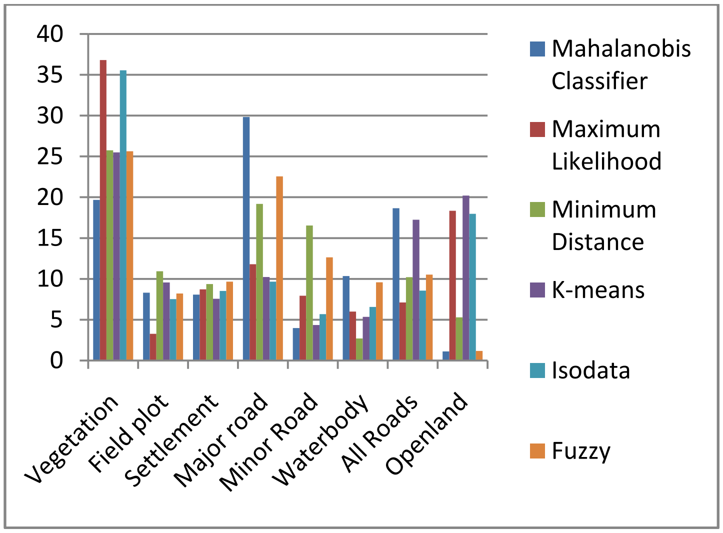

Fig.1 Comparison of Results of different classifiers techniques

Y- axis: Percentage of pixels in desired class X- axis: Names of the desired class

1.119% for Mahalanobis distance classifier which is least number of pixels among all the classes. In the above table 1, proposed method shows higher increase of pixels in major road category with 22.56%. Considerable decrease in vegetation category. Considerable decrease in open land category.

The graphical representation of analysis of comparison of various classifiers is shown in Figure 1. In figure 1 above, vegetation in Maximum likelihood classification is shown above 36.799% due to mixing of the vegetation and the field plot pixels which shows similar signature due to different spectral characteristic and similar reflectance characteristic. While in Fuzz rule based classification, the values are distributed across all the class of all roads, waterbody and minor roads. Class of all roads, waterbody and minor roads are not prominent in other methods of classification as shown in figure 1. Openland was found to be prominent in all the methods of classification except fuzzy rule based classification because it considers the maximum degree of fitness of the pixel to its natural class. The histogram values are taken into consideration for all the class and compared to original pixels in the image for sample feature. This algorithm is implemented on high resolution satellite image. Shows high resolution satellite result of classes in Table 1 &2.

5 Conclusion

The analysis of different techniques of image classification is done using high resolution satellite imagery. Maximum likelihood works better for vegetation; Mahalanobis distance classifier works best on roads, Minimum distance to mean worked best for open land features. The fuzzy rule based classification works better on all the class. The application of these techniques on high resolution satellite imagery depends upon how correctly the training samples sites have been taken. As good as the distinct signatures are that much accurate is the classification. Therefore, accuracy of signature is required to be considered. The algorithm varies from class to class. If the data is homogenous, the inverse of variance-covariance matrix is unstable. If the distribution of the population is not normal distributed the Maximum likelihood cannot be applied. The complete algorithm works on the range values of the class used in the classification. They are required to be very distinct from each other. The accuracy depends upon how correctly the range values are obtained. The maximum likelihood and fuzzy rule based classification algorithm show satisfactory results for classification of features. The maximum likelihood supervised classification method could be used for different feature classification. The quality of the output depends upon the image enhancement and stretching parameters which are to be taken care before processing. This technique is required to be tried on various other satellite data sets and other digital imagery.

References

[1] N. Elavarasan, K.Mani,"A Survey on Feature Extraction techniques," International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE), Vol. 3, no. 1, pp. 52-55, Jan 2015.

[2] S.A. Medjhaid, "A Comparative Study of Feature Extraction Methods in Images Classification," I.J. Image, Graphics and Signal Processing, DOI: 10.5815/ijigsp.2015.03.03,2015, Vol. 3, pp. 16-23, 2015.

[3] Chen, Z., Wu, Q. and Yang, W,"Target Image Classification through Encryption Algorithm Based on the Biological Features," International Journal of Intelligence Science, doi: 10.4236/ijis.2015.51002, Vol. 5, pp. 6-12, 2015.

[4] S.R. Bhatt, P.K. Mishra,"Analysis of Global Kernels using fuzzy c means algorithm," International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 4, no. 6, pp. 79-82, June 2014.

[5] M. Li, S. Zang, B. Zhang, S. Li, and C. Wu,"A Review of Remote Sensing Image Classification Techniques: the Role of Spatio-contextual Information," European Journal of Remote Sensing, Vol. 47, pp. 389-411, 2014.

[6] R. Dwiwedi, A. Kumar, and S.R. Ghosh,"Study of fuzzy based classifier parameters across Spatial Resolution," International journal of Computer application, Vol. 50, no. 11, pp. 17-24, July 2012.

IJSER

[7] U. Maulik,"Automatic Fuzzy Clustering Using Modified Differential Evolution for Image classification," IEEE Transactions on Geoscience and Remote Sensing, Vol. 48, no. 9, pp.3503-3510, Sept. 2010.
[8] W. Lu, W. Björn, and J. A.Benediktsson,"Ensemble methods or spectral-spatial classification of urban hyperspectral data," IEEE International Geoscience and Remote Sensing Symposium, IGARSS2009, Vol.4, pp. 944 – 947, 2009.
[9] F. Masulli and A. Schenone,"A fuzzy clustering based Segmentation system as support to diagnosis in medical imaging, "Artificial Intelligence in Medicine, Vol.16, no. 2, pp. 129-147, June 1999.
[10] I. Couloigner, and T. Ranchin,"Mapping of urban areas: A multiresolution modeling approach for semiautomatic extraction of streets," Photogrammetric Engineering & Remote Sensing, Vol. 66, no. 7, pp. 867–874, 2000.
[11] Siddheswar Ray and Rose H Turi,"Determination of number of clusters in k-means clustering and application in colour image segmentation,"The Fourth International Conference on Advances in Pattern Recognition and Digital Techniques (ICAPRDT'99), pp. 137-143, 1999
[12] A. David,"Efficient fuzzy clustering of multi-spectral images," IEEE International Conference on Geoscience and Remote Sensing Symposium, IGARSS '99 Proceedings, Vol. 3, pp. 1594 – 1596, 1999.
[13] Q. Cheng, L. Jiancheng, S. Zhanfeng, Z. Zhiwen and W.Wei,"Spatial knowledge based complicated urban area classification from high-resolution remote sensing image," IEEE International Conference on Spatial Data Mining and Geographical Knowledge Services (ICSDM) DOI: 10.1109/ICSDM.2011.5969075, pp. 405 – 408, 2011
[14] J.Chanussot, G. Mauris, and P. Lambert,"Fuzzy fusion techniques for linear features detection in multitemporal SAR images," IEEE Transactions on Geoscience and Remote Sensing, Vol. 37, no. 3, pp.1292–1305, 1999.
[15] A.Baraldi and F. Parmiggiani, "A NagaoMatsuyama approach to high resolution satellite image classification," IEEE Transactions on Geoscience and Remote Sensing, Vol. 32, no. 4, pp. 749–758, 1994.
[16] B.Sowmya,"Colour Image Segmentation Using Soft Computing Techniques," International Journal of Soft Computing Applications ISSN: 1453-2277, Issue. 4, pp.69-80, 2009.
[17] T.M.Lillesand and R.W. Kieffer, Remote Sensing and Image Interpretation, 3rd ed., John Wiley and Sons, New York, pp. 356-378, 550-586, 2000.
[18] K.N. Ndukwe, "Technology in surveying and mapping," RhyceKerex Publishers, 2001
[19] A.Y. Ogunleyeand A.A. Obiniyi,"Mapping and cartography," Proceedings of the 29th Annual conference of NCA, Kaduna, 2007.
[20] K. R. Salla and S.S. Kadam,"A fuzzy inference system based approach for detection, classification and grading of cataract. International journal of computer science and application. ISSN: 0974-0767, pp. 18-23, 2010.
[21] I. Nedeljkovic,"Image classification based on fuzzy logic," The International archive of the photogrammetry remote sensing and spatial information sciences, Vol. 24, Part xxx, Jan 2004.
[22] J. Kaur, P. Sethi,"Evaluation of fuzzy inference system in image processing," International journal of computer applications (ISSN: 0975-8887), Vol.68, pp.1-4, 2013.
[23] Borkar, M. Atulkar,"Fuzzy inference system for image processing," International journal of advanced research in computer engineering and technology (IJARCET), Vol.2, no. 3, pp.1007-1010, 2013.
[24] M.Sharma, R. Gupta et.al, "Efficacious approach for satellite image classification," Journal of electrical and electronics engineering Research, Vol. 3, no. 8, pp. 143-150, 2011.
[25] Songcan Chen, Daoqiang Zhang, "Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure," IEEE Transactions on Systems, Man and Cybernetics, Vol.34, no.4, pp: 1907-1916, Aug 2004.
[26] Wang, Z.; Bovik, A. C.; Sheikh, H. R. & Simoncelli E. P, "Image quality assessment: From error visibility to structural similarity," IEEE Trans. Image Processing, Vol. 13, no. 4, pp. 600-612. (2004),
[27] Mona Gamal, Elsayed Radwan, and Adel M.A. Assiri,"Credible Fuzzy Classification based Technique on Self Organised Features Maps and FRANT IC-RL," (IJACSA) International Journal of Advanced Computer Science and Applications www.ijacsa.thesai.org, Vol. 5, no. 10, pp. 71-81, 2014.
[28] Malik V et al,"Satellite image classification using fuzzy logic," International journal of recent technology and engineering, Vol.2 , no. 2, pp. 204-207, 2013.
[29] Ghosh S and Dubey S.N,"Comparative analysis of k-means and fuzzy c-means algorithms," International journal of Advance computer science and applications, Vol. 4, no. 4, pp.35-39, 2013.
[30] Wen-hui Zhao; Hui-li Gong; Wen-ji Zhao; Tao Tang,"A fuzzy object-based data processing for high resolution remote sensing image," 7th IEEE Conference proc. Of International conference on fuzzy system and knowledge discovery (FSKD) DOI: 10.1109/FSKD.2010.5569304, Vol. 5, pp.2299-2302, 2010.
[31] Mokhtarzade. M, Zoej M.J.V, Ebadi H,"Automatic road extraction from high resolution satellite images using Neural networks, texture analysis, Fuzzy clustering and genetic algorithms," The International archives of the Photogrammetry Remote sensing and spatial information. Vol. xxxvi, part 33b, pp. 549-556, 2008.
[32] Chaudhari S.N, Ghosh, A, "Feature extraction using fuzzy rule based system," International journal of

computer science and applications, Vol.5, no. 3, pp. 1-8, July, 2008.

[33] Miosso J.C, Bauchspiess A, "Fuzzy inference system applied to edge detection in Images," Proceedings of the V. Brazilian conference on Neural Networks, pp.481-483, April 2001.

[34] Shackelford K.A, Davis H.C, "A Hierarchical fuzzy classification approach for High resolution multispectral data over urban areas," IEEE Transaction on Geoscience and Remote Sensing, Vol. 41, no. 9, pp.1920-1932, Sep 2003.

[35] Rui Xu, Wunsch D,"Survey of clustering algorithms," IEEE Transactions on Neural networks, Vol.16, no. 3,pp.645-678, May 2005.

[36] Wacker A.G and Landgrebe, D.A, "Minimum distance classification in remote sensing," 1972. LARS Technical reports Print # 030772. Paper 25, http://docs.lib.purdue.edu/larstech/25, 1-1-1972.

[37] M.H FazelZarandi, M. Zarinbal, I.B. Turksen,"Type-II Fuzzy Possibilistic C-Mean Clustering," Proc. Of IFSA-EUSFLAT-2009, ISBN: 978-989-95079-6-8, pp. 30-35, 2009.

[38] W. Lu, W. Bjorn. J., A. Benediktsson,"Ensemble methods or spectral spatial classification of urban Hyperspectral data,"2009 IEEE Geoscience and Remote Sensing Symposium, IGARSS, Capetown,Vol. 4, pp. 944-947, 12-17 July 2009.

[39] Wafabarkhoda, Fardin Akhlaqian Tab, Om-Kolsoom Shahryari, "Fuzzy Edge Detection Based on Pixel's Gradient and Standard Deviation Values," Proceedings of the International Multiconference on Computer Science and Information Technology, ISBN 978-83-60810-22-4 ISSN 1896-7094, Vol. 4, pp. 7 – 10, 2009.

[40] K.Perumal, R.Bhaskaran, "Supervised classification performance of Multispectral images," Journal of computing ISSN 2151-9617, Vol.2, no. 2,pp-124-129, February 2010,

[41] M.Renuka Devi, Dr. S. Santhosh Baboo, "Land use and Landcover classification using RGB&L Based supervised classification algorithm," International Journal of Computer Science Engineering, Technology, ISSN: 2229-3345., Vol.2, no. 10, pp. 167-180, October 2011.

[42] J.A. Richards, "Remote sensing remote image analysis: An introduction," Springer-Verlag, Berlin, Germany, 1999.

[43] Kanika Kalra, Anil Kumar Goswami, Rhythm Gupta,"A comparative study of supervised image classification algorithms for satellite images," International Journal of Electrical, Electronics and Data Communication, ISSN: 2320-2084,Vol.1, no. 10, pp.10-16, Dec, 2013.

[44] S.Padma, S.Sanjeevi, "Jeffries Matusita-spectral angle mapper (jm-sam) spectral matching for species level mapping at bhitarkanika, muthupet and pichavaram mangroves," The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XL-8, 2014 ISPRS Technical Commission VIII Symposium, 09 – 12 December 2014, Hyderabad, India

[45] Ball, G.H.; Hall, D.J,"Isodata, a Novel Method of Data Analysis and Pattern Classification," Stanford Research Institute: Menlo Park, CA, USA, 1965

[46] Ndukwe, K.N., "Principles of environmental remote sensing and photo Interpretation," New concept publishers, Enugu, Nigeria, 1997.

IJSER