A MODIFIED FUZZY CLUSTERING APPROACH IN UNSUPERVISED CLASSIFICATION FOR DETECTING THE MIXED PIXELS OF SATELLITE IMAGES

A.R. SHERWANI, Q.M. ALI, IRFAN ALI*

Department of Statistics and Operations Research, Aligarh Muslim University, Aligarh-202 002, India

Abstract: The major problem of remote sensing images is mixed pixels, available in the data which degrades the quality, accuracy of the image classification and object recognition. To overcome the problem of mixed pixel in a real satellite data a modified K-means clustering algorithm and a modified fuzzy C-means clustering algorithm, are discussed. The algorithms are developed by modifying the membership function of the standard K-means clustering algorithm (FKM) and the standard fuzzy C-means algorithm (FCM). The performance of the proposed algorithms is discussed and compared with the traditional fuzzy K-means algorithm and the traditional FCM algorithm. Results on classification and segmentation of satellite images reveal that the suggestive algorithms are robust and effective.

Keywords: mixed pixel; cluster analysis; fuzzy clustering; fuzzy c-means (FCM) clustering algorithm; modified FCM algorithm; fuzzy k-means clustering algorithm (FKM); modified FKM algorithm.

2010 AMS Subject Classification: 62H30, 62H35, 62H86.

_____

*Corresponding author

E-mail address: irfii.ali@gmail.com

Received November 24, 2021

1. INTRODUCTION

Image segmentation is a significant and difficult issue; it is the first step in image analysis as well as high-level image interpretation and understanding such as robot vision, object recognition, and satellite imaging. Image segmentation is the process of partitioning the original image into homogeneous groups and it plays a vital role in image processing and classification. Many segmentation tools were introduced, and detailed surveys are available from Fu and Mui [1]. Image segmentation approaches, according to Pal and Pal [2], are classified into four types namely: edge detection, clustering, thresh-holding and region extraction.

Clustering is a procedure for categorizing objects or elements so that objects in the same group are substantially more similar with one another than objects in other groups. Many clustering procedures, such as the fuzzy clustering and hard clustering have been used; both have its own unique set of properties. The traditional hard clustering approach assigns each point in the set of data to only one cluster. As a result, the classification results from this approach are frequently quit crisp. That is every pixel of the image belongs to absolutely single class. Nevertheless, in several real-world circumstances, such as those for satellite images, issues such as the presence of mixed pixels in the data, limited spatial resolution, weak contrast, and intensity fluctuation in homogeneities make this hard (crisp) segmentation a challenging task.

Zadeh's [3] introduced the concept of uncertainty (ambiguity), which was only expressed by a membership function. The fuzzy clustering is a soft classification technique that has been extensively studied and effectively used in image classification and image segmentation. In most of the fuzzy clustering approaches, FCM algorithm Bezdek [4] was the most often employed in classification of image as it has robust vagueness features and can preserve more information than hard segmentation methods Bezdek [5]. Now fuzzy clustering has been extensively applied and studied in a variety of substantive areas Bezdek [4], Pal and Majumdar [6]. These approaches become the significant tools in cluster analysis.

The major problem of remote sensing images is mixed pixels, available in the image data which degrades the quality, accuracy of the image classification and object recognition. The

fuzzy approach is being used to process and analyze images in different ways; such as Bibiloni *et al* [7] has used fuzzy mathematical morphology to process the digital images. Dwivedi *et al.* [8] used a fuzzy approach for learning unknown patterns to be used with a neural network for analyzing satellite images to estimate crop area. Thapa and Murayama [9], compared four approaches for mapping that include fuzzy supervised and GIS post-processing. In this paper, to overcome the problem of mixed pixel in a real satellite data a modified fuzzy *K*-means (FKM) clustering algorithm and a modified fuzzy *C*-means (FCM) clustering algorithm are proposed. The algorithms are developed by modifying the membership function of the standard FKM and FCM algorithm. The performance of the proposed algorithms is discussed and compared with the traditional FKM and the traditional FCM algorithm.

The remaining Sections of this study are organised as; Section 2 explain the problem of interest. Image classification techniques and their types are briefly describes in Section 3. Description of framework for the study is given in Section 4. Finally, results of the study and their conclusions are explained in Section 5 and Section 6.

2. PROBLEM TO BE ENTERTAINED

The presence of large pixels size in land cover satellite images raises the possibility of the presence of mixed pixels. In digital image analysis, mixed pixel problem is created by those pixels which are not totally occupied by a single, homogeneous class. When a pixel area is occupied by two or more classes that are differ in terms of brightness. For instance, a pixel in land sat imagery has a size of 10m $\times$ 10m whereas the pixel size of Quick bird imagery is 0.60m $\times$ 0.60m. This difference in resolution increases the chances of encountering mixed pixels in Land sat imagery as compared to Quick bird image. Thus, mixed pixels can be one of the sources of error in per pixel classification and should be treated accordingly.
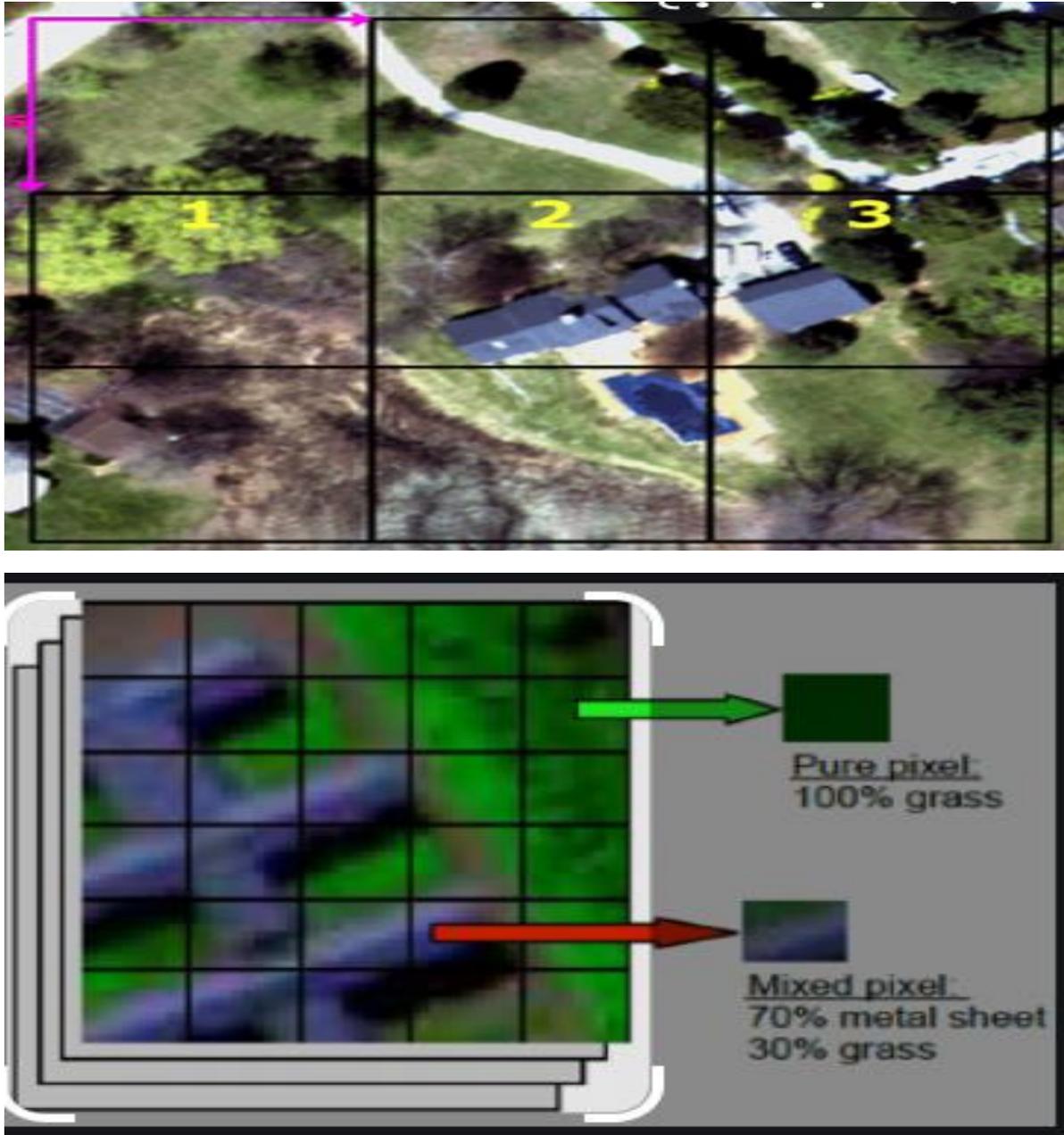
Fig1: Mixed pixels due to the presence of small, sub-pixel targets with in the area it represents.

Fuzzy set theory provides useful concepts and tools to deal with mixed pixel problem in land cover classification. The concept of partial membership allows that the information about more complex situations, such as cover mixture or intermediate conditions, can be better represented and utilized. Classifiers which assume fuzzy membership of multiple classes in a pixel are called as fuzzy classifiers.

3. IMAGE CLASSIFICATION TECHNIQUES

Image classification is the basic task for processing of remotely sensed images. Classification techniques provide frameworks for organizing and categorizing information that can be extracted from image data. Spatial unit to be classified may be a pixel or an object in the imaged scene. Each pixel of an image is represented by a vector, consisting of a set of measurements (e.g. spectral bands, textural features etc.). The general image classification methods of remote sensing such as, pixel-wise classification methods consider each pixel is pure and categorized as a single land use land cover type Fisher [10]. Generally, pixel-wise classification algorithms are divided into two groups: unsupervised classification and supervised classification.
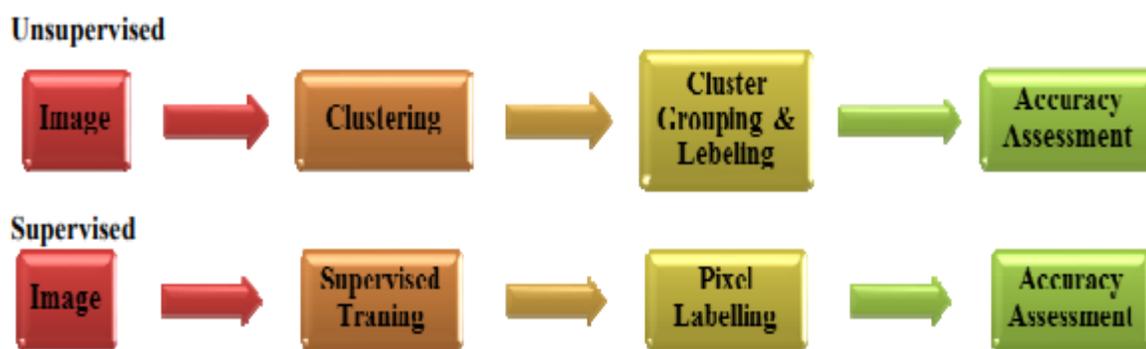

**Fig 2.** Flow chart for Unsupervised and Supervised Classification.

3.1. UNSUPERVISED CLASSIFICATION

Unsupervised image classification is a method by which the large numbers of unknown pixels in an image are separated by an image interpreting software based on their reflectance values into classes or clusters with no training from the analyst Tou and Gonzalez [11]. The most common unsupervised clustering methods are: Fuzzy K-means clustering and fuzzy C-means clustering methods. These methods are purely based on spectrally pixel-based statistics and include no prior information of the features of the themes being studied. On the other hand, supervised classification is a method in which the user defines small areas called training sites on the image, which contain the predictor variables measured in each sampling unit, and assigns prior classes to the sampling units.

3.1.1. CLUSTER ANALYSIS

Cluster analysis is one of the major techniques used in pattern recognition and image classification. The unsupervised technique that is most readily used by the analysts is cluster analysis in different areas such as remote sensing, taxonomy, medical science, engineering systems, robotics and image processing etc,. Clustering is a technique to identify the number of sub-classes of c clusters in a data set X consisting of n data samples, and partitioning the data set X into c clusters $(2 \leq c \leq n)$. It is to be noted that $c = 1$ denotes the rejection of the presence of clusters in the data, where $c = n$ represents the trivial case where each sample is in a cluster by itself. There are two types of c-partitions of the data set: hard or crisp and soft or fuzzy. In numerical data interpretation one supposes that the elements of each cluster bear more mathematical similarity to each other than to elements of other clusters. Two significant issues to consider in this regard are how to compute the similarity between the pairs of observations and how to calculate the partitions once they are formed. If one can determine a suitable distance measure and compute the distance between all pairs of observations, then one may expect that the distance between points in the same cluster will be considerably less than the distance between points in different clusters.

The traditional (hard) clustering techniques restrict that every pixel of the data belongs to exactly one class or one cluster. Zadeh [3] proposed fuzzy set theory gave an idea of belongingness described by a membership function and provides imprecise class membership information. Bellman, *et al*., [12] and Ruspini [13] were early proposed the applications of fuzzy set theory in cluster analysis. The two unsupervised classifiers that are most commonly used are classifiers based on fuzzy K-means clustering and fuzzy C-means clustering.
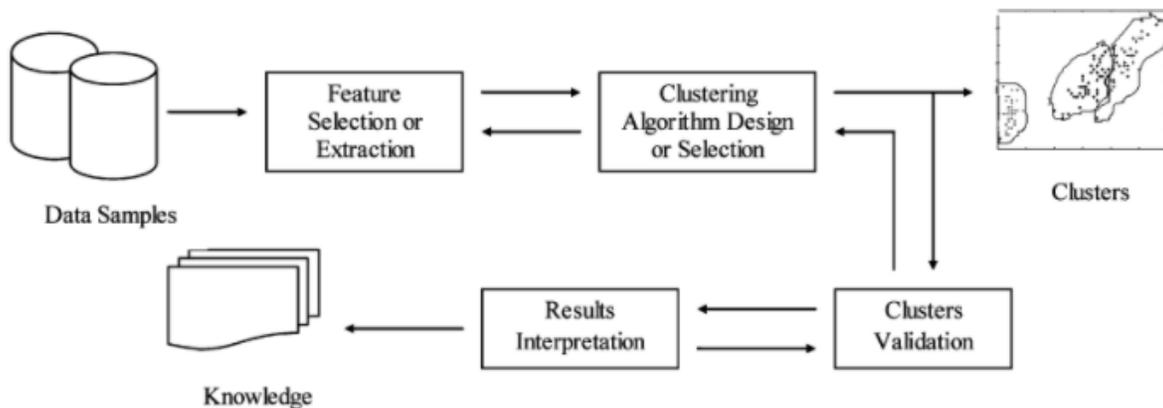
Fig 3. Flow chart for Clustering.

3.2.   FUZZY *K*-MEANS CLUSTERING ALGORITHM

Fuzzy *K*-Means (FKM) is exactly the same algorithm as *K*-means, which is most common clustering technique. The only difference is that instead of assigning a point exclusively to only one cluster, it can have some fuzziness or overlap between two or more clusters.

Suppose $r \geq 1$ patterns $X_i = X_1, X_2, ..., X_r$ are given. The FKM clustering algorithm partition data points $X_i = X_1, X_2, ..., X_r$ into *k* clusters C_j (*j* = 1, 2, …, *k*) and clusters C_j are associated with centroid of the cluster v_{jp}.

The FKM algorithm is as follows:

Step1: For $j = 1, 2, ..., k$, do

For $p = 1, 2, ..., M$, estimate a value of the centroids of the *k* clusters will be iteratively modified to converge to the centroids obtained from the procedure.

Step2: For $j = 1, 2, ..., k$, do

For $p = 1, 2, ..., M$, do $v'_{jp} \leftarrow v_{jp}$

v'_{jp} saves the value of v_{jp} before the iteration. During the iteration, the value of v_{jp} will change. At the end of the iteration, v_{jp} will be compared with v'_{jp}.

Step3: The relationship between centre of the cluster and a data point is fuzzy, i.e., a membership

$$\mu_{C_j}(X_i) = \left[\sum_{g=1}^{k} \left(\frac{\sum_{p=1}^{M}(x_{ip} - v_{jp})^2}{\sum_{p=1}^{M}(x_{ip} - v_{gp})^2} \right)^{\frac{1}{m-1}} \right]^{-1}$$ ∈ [0, 1] is used to represent the degree of belongingness of

data point X_i and cluster centre v_{jp}. The value of variable *m* is chosen by the user for the first iteration and remains the same for the subsequent iterations. There is no known theoretical basis for choosing a value of *m*. It is however, mathematically established that procedure would converge for any value of *m* lying in the open interval (1, ∞).

Moreover, it is mathematically established that $\sum_{j=1}^{k} \mu_{Cj}(X_i)=1$ for $1 \leq i \leq r$, that is, the

sum of the membership values of pattern in the k clusters is equal to 1.

Step 4: For $j=1,2,...,k$, do

For $p=1,2,...,M$, do

$$v_{jp} = \frac{\sum_{i=1}^{r} x_{ip}\left[\mu_{Cj}(X_i)\right]^m}{\sum_{i=1}^{r}\left[\mu_{Cj}(X_i)\right]^m}$$

New coordinates of the centroids V_1 to V_k of the clusters are calculated. The value of m is as

chosen in the previous step.

Step 5: For $j=1,2,...,k$, calculate the Euclidean distances between the current centroid V_j and

the centroid V_j' at the beginning of the iteration, as recorded in Step 2. Let the largest of these

distances be *D*.

Step 6: If $D > \varepsilon$, go to Step 2. Otherwise, return from the procedure with

$$\mu_{Cj}(X_1),\ \mu_{Cj}(X_2),...,\ \mu_{Cj}(X_r)$$

For $1 \leq j \leq k$.

Thus, the membership value of every pattern in each cluster is obtained. The variable ε is

the convergence criterion of the procedure; it is a small value, say 0.01, chosen by the user.

3.3. FUZZY *C*-MEANS CLUSTERING ALGORITHM

Dunn [14] proposed the Fuzzy *C*-Means (FCM) clustering algorithm and later on it was extended

by Bezdek [15]. The algorithm is an iterative clustering technique that gives an optimal *c*

partition by minimizing the weight within group sum of the squared error objective function

J_{FCM} :

$$J_{FCM} = \sum_{k=1}^{n}\sum_{i=1}^{c}(u_{ik})^q d^2(x_k, v_i) \tag{1}$$

where, $X = \{x_1, x_2, ..., x_n\} \subseteq R^p$ is the p-dimensional vector space of the data set, n is the number of data items, number of clusters are c with $2 \leq c < n$, u_{ik} is the degree of membership of x_k in the ith cluster, q is the weighting exponent on each fuzzy membership, v_i is the prototype of the centre of the cluster i, $d^2(x_k, v_i)$ is a distance measure between cluster centre v_i and object x_k. The objective function J_{FCM} gives the solution by iterative process, which is explained as follows:

1. Give the values for $c, \varepsilon,$ and q

2. Obtain partition matrix (fuzzy), $U = [u_{ik}]$.

3. Set counter loop as, $b = 0$.

4. Calculate c, centres of cluster $\left\{v_i^{(b)}\right\}$ with $U^{(b)}$

$$v_i^{(b)} = \frac{\sum\limits_{k=1}^{n}(u_{ik}^{(b)})^q x_k}{\sum\limits_{k=1}^{n}(u_{ik}^{(b)})^q} \tag{2}$$

5. Calculate the membership $U^{(b+1)}$. For $k = 1$ to n, calculate the following: $I_k = \{i | 1 \leq i \leq c,$ and $d_{ik} = \|x_k - v_i\| = 0\}$, $/ I$; for the kth column of the matrix, compute new membership values:

(a) If $I_k = \phi$, then

(b) Else $u_{ik}^{(b+1)} = 0 \ \forall \ i \notin I$ and $\sum_{i \in I_k} u_{ik}^{(b+1)} = 1$; then for next k.

6. If the value of $\left\|U^{(b)} - U^{(b+1)}\right\| < \varepsilon$, stop; or else, set $b = b + 1$ and move to step 4.

3.4. MODIFIED *K*-MEANS CLUSTERING ALGORITHM

In this proposed modified *K*-means clustering algorithm we use trapezoidal membership function instead of the membership function used in the conventional *K*-means algorithm. Since the trapezoidal membership function is based on the threshold values or the values lies within the

range or interval. Due the flexibility in terms of membership the trapezoidal membership function more accurately assign the membership value of belongingness of mixed pixel present in that particular class. Let us suppose that $r \geq 1$ patterns $X_i = X_1, X_2,..., X_r$ are given. The modified FKM clustering algorithm partitions the data points $X_i = X_1, X_2,..., X_r$ into k clusters C_j ($j = 1, 2,..., k$) and clusters C_j are associated with centroid of the cluster v_{jp}.

The proposed modified FKM is as follows:

Step1: For $X_i = X_1, X_2,..., X_r$ data points we assign the membership to each data points X_i(s). The relationship between cluster centre and a data point is fuzzy. That is, a membership $\mu_{Cj}(X_i) \in [0, 1]$ is used to represent the degree of belongingness of cluster centre v_{jp} and the data point X_i. The membership value can be found as. Let *a, b, c* and *d* represents the X coordinates of the membership function $\mu_{Cj}(X_i)$. Then, trapezoid (X ; *a, b, c, d*) = 0 if $X \leq a$;

$$\mu_{Cj}(X_i) = \begin{cases} 0 & if \quad X < a \\ \dfrac{(X-a)}{(b-a)} & if \quad a < X \leq b \\ \dfrac{(d-X)}{(d-c)} & if \quad c \leq X \leq d \\ 0 & if \quad X \geq d \end{cases}$$

$$\mu_{Cj}(X_i) = \max\left(\min\left(\frac{X-a}{b-a}, 1, \frac{d-X}{d-c}\right), 0\right)$$

Moreover, it is mathematically established that $\sum_{j=1}^{k} \mu_{Cj}(X_i) = 1$ for $1 \leq i \leq r$, that is, the sum of the membership values of pattern in the k clusters is equal to 1.

Step2: For $j = 1, 2,..., k$, do

For $p = 1, 2,..., M$, do $\qquad v'_{jp} \leftarrow v_{jp}$

v'_{jp} saves the value of v_{jp} before the iteration. During the iteration, the value of v_{jp} will change. At the end of the iteration, v_{jp} will be compared with v'_{jp}.

Step 3: For $j = 1,2,..., k$, do

For $p = 1,2,..., M$, do

$$v_{jp} = \frac{\sum_{i=1}^{r} x_{ip} \left[\mu_{Cj}(X_i)\right]^m}{\sum_{i=1}^{r} \left[\mu_{Cj}(X_i)\right]^m}$$

New coordinates of the centroids V_1 to V_k of the clusters are calculated. The value of variable m is chosen by the user for the first iteration and remains the same for the subsequent iterations.

Step 4: For $j = 1,2,..., k$, calculate the Euclidean distances between the current centroid V_j and the centroid V_j' at the beginning of the iteration, as recorded in Step 2. Let the largest of these distances be *D*.

Step 5: If $D > \varepsilon$, go to Step 2. Otherwise, return from the procedure with

$$\mu_{Cj}(X_1), \ \mu_{Cj}(X_2),..., \ \mu_{Cj}(X_r)$$

For $1 \le j \le k$.

Thus, the membership value of every pattern in each cluster is obtained. The variable ε is the convergence criterion of the procedure; it is a small value, say 0.01, chosen by the user.

3.5. MODIFIED FUZZY *C*-MEANS CLUSTERING ALGORITHM

In this proposed modified fuzzy *C*-means clustering algorithm we use trapezoidal membership function instead of the membership function used in the conventional fuzzy *C*-means clustering algorithm. The trapezoidal membership function decides the membership value of belongingness of those elements which lies within a class interval. The value of membership lies between 0 and 1. Those elements or pixel values of an image which are completely belongs to a particular class assign the value 1, while those who do not belong to that class assign a value 0, and those pixels which are mixed assign a partial membership between 0 and 1.

The conventional fuzzy *C*-means clustering algorithm is an iterative clustering method that produces an optimal *c* partition by minimizing the weighted within group sum of squared error new modified objective function MJ_{FCM}:

$$MJ_{FCM} = \sum_{k=1}^{n}\sum_{i=1}^{c}(\mu_{Cj}(X_i)^q d^2(x_k, v_i)) \tag{3}$$

where, $X = \{x_1,\ x_{2,}...,x_n\} \subseteq R^p$ is the p-dimensional vector space, n is the number of data

elements, c is the number of clusters with $2 \le c < n$, $\mu_{Cj}(X_i)$ is the membership of x_k in the

i^{th} cluster defined by the trapezoidal membership function, q is the weighting exponent on every

fuzzy membership, v_i is the prototype of the centre of cluster i, $d^2(x_k, v_i)$ is a distance

measure between object x_k and centre of the v_i. The new objective function MJ_{FCM} gives the

solution by iterative process, which is explained as follows:

1. Give the values of c, ε, and q.

2. Obtain partition matrix (fuzzy), $U = [u_{ik}]$.

3. Set counter loop as, $b = 0$.

4. Calculate c, centres of cluster $\left\{v_i^{(b)}\right\}$ with $U^{(b)}$

$$v_i^{(b)} = \frac{\displaystyle\sum_{k=1}^{n}(\mu_{Cj}(X_i)^{(b)})^q x_k}{\displaystyle\sum_{k=1}^{n}(\mu_{Cj}(X_i)^{(b)})^q} \tag{4}$$

5. Calculate the value of membership $\mu^{(b+1)}$. For $k = 1$ to n, calculate $I_k = \{i\,|\,1 \le i \le c$, and

$d_{ik} = \|x_k - v_i\| = 0\}$, / I; for the kth column of the matrix, now compute new modified

membership values:

(a) If $I_k = \phi$, then

(b) Else $\mu^{(b+1)} = 0 \ \forall \ i \notin I$ and $\sum_{i \in I_k} \mu^{(b+1)} = 1$; then to next k.

6. If the value of $\left\|\mu^{(b)} - \mu^{(b+1)}\right\| < \varepsilon$, stop; or else, set $b = b+1$ and move to step 4.

4. DESCRIPTION OF FRAMEWORK FOR THE STUDY

To validate the effectiveness of the proposed fuzzy unsupervised classification methods, we have conducted classification on high-resolution remotely sensed images using the proposed modified fuzzy clustering methods and the traditional fuzzy clustering classification methods.

The image under study is the LISS-III Sentinel image of Jawan, Tehsil Koil, District Aligarh, U.P., India, terrain with Latitude: 27° 54' 1.37" and Longitude:78° 42' 20.21". The ortho-rectified image has 4 spectral bands 2, 3, 4 and 5 with a resolution of 10m $\times$ 10m. The image was captured on September 22, 2016, and was procured from United States Geological Survey (USGS) [https://www.usgs.gov]. The dimensions of the image are 249×48 pixels.
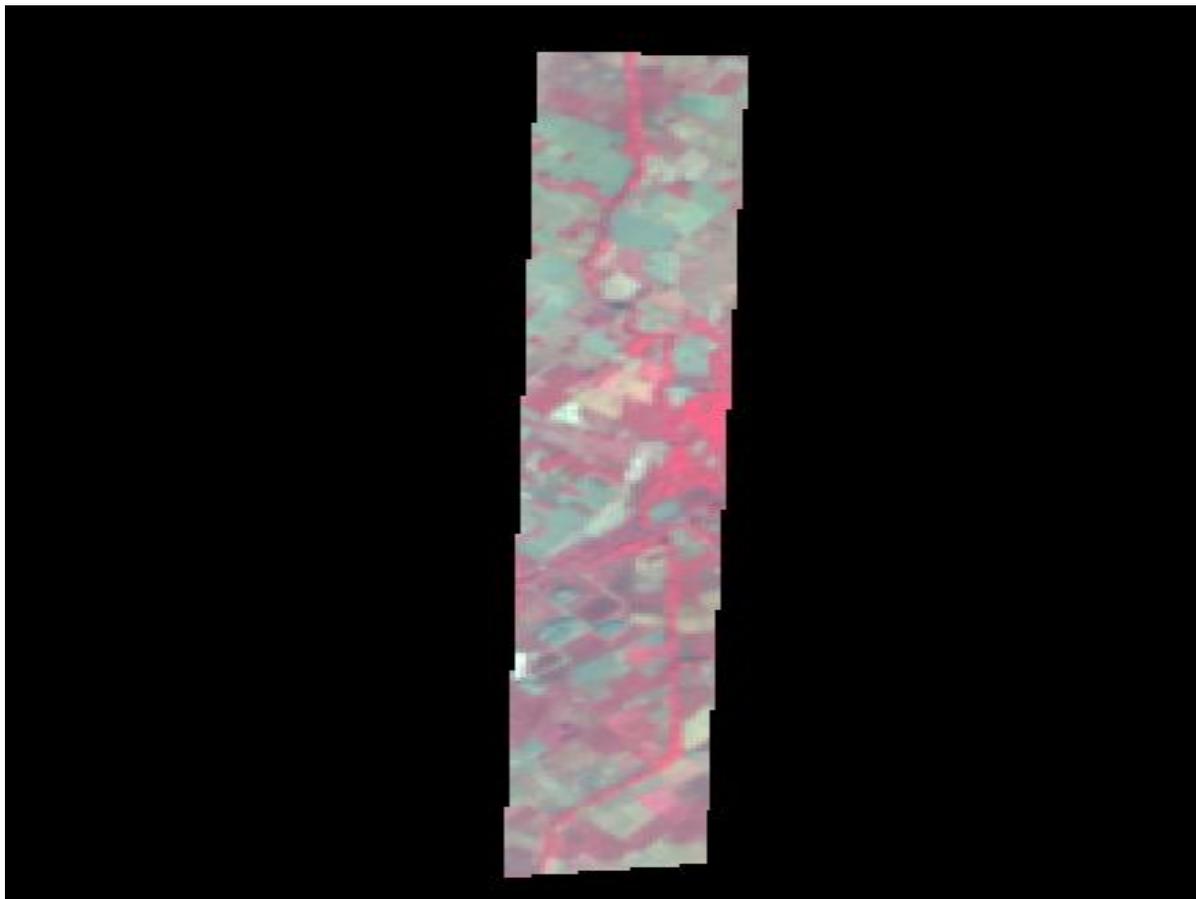
Fig 4: LISS-III Sentinel image of Jawan, Tehsil Koil., District Aligarh, U.P., India, terrain with Latitude: 27° 54' 1.37" and Longitude:78° 42' 20.21".

In this paper, modified fuzzy K-means clustering algorithm and modified fuzzy C-means clustering algorithm has been suggested as the classifier of choice when dealing with mixed

pixels in imagery. Their performance in terms of confusion matrix and misclassification error has been compared with that of the traditional fuzzy *K*-means clustering algorithm and fuzzy *C*-means clustering algorithm.

On the parameter of the LISS-III digital image, a total of 1049 points were under study. These shape files of the surface covers were created using the software package ERDAS Imagine and computational work of this study, has been carried out by using the software R. The data under study consists of a total of 1049 pixels from different classes.

5. RESULTS

It is very obvious that any classification method results in some misclassification probabilities and these misclassification probabilities play a vital role in assessing the performance of the classification method in future samples Johnson & Wichern 2007 [16].

Error Rate = Total no of wrongly classified pixels / Total no. of pixels

	Methods Used	Correctly Classified Pixels	Wrongly Classified Pixels	Error Rates
A	Fuzzy *K*-means clustering Algorithm	758	291	0.2774
B	Modified Fuzzy *K*-means clustering Algorithm	769	280	0.2669
C	Fuzzy *C*-means clustering Algorithm	793	256	0.2440
D	Modified Fuzzy *C*-means clustering Algorithm	811	238	0.2268

Table 1: Probability of Misclassification
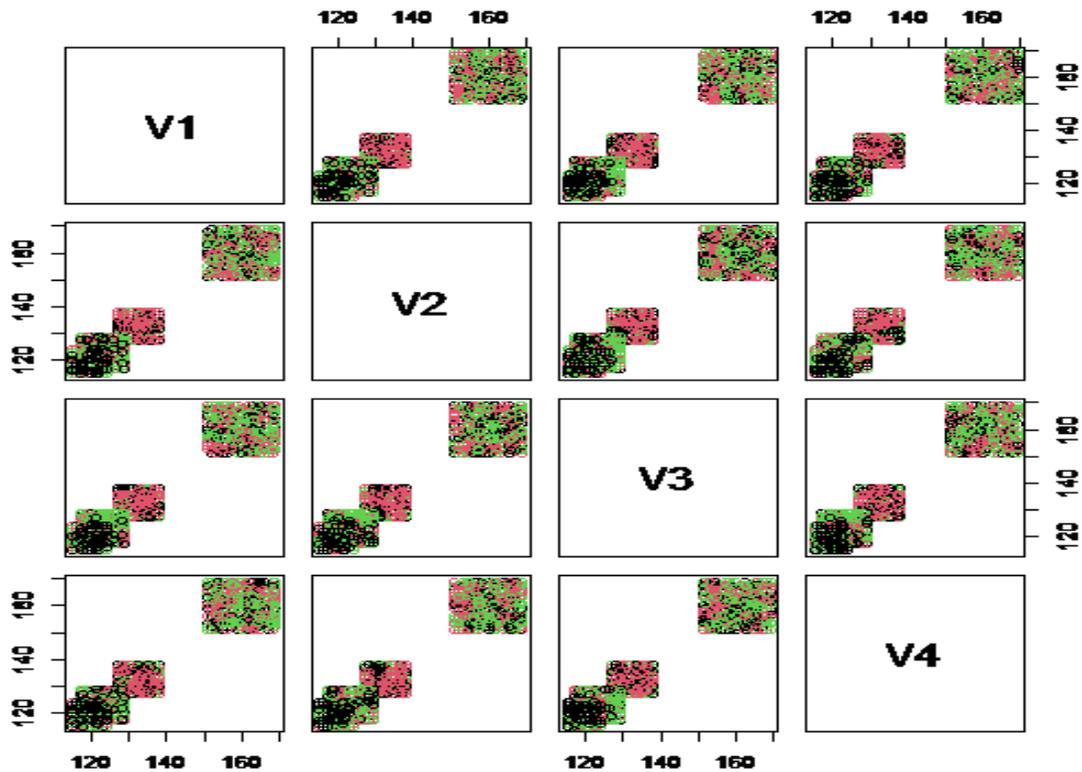
Fig 3. Plot shows the data points of four different classes of imagery.

The following Tables 2 (a-d) shows the confusion matrix by conventional methods and the confusion matrix by proposed approaches. Fuzzy K-means clustering algorithm (A) and the modified Fuzzy K-means clustering algorithm (B) are compared in terms of area on the ground, the improvement is of 11 pixels or 1100 sq.mt. area is correctly classified by using proposed method B. Similarly, FCM (C) and the modified FCM (D) are compared in terms of area on the ground; the improvement is of 18 pixels or 1800 sq.mt. area is correctly classified by using the proposed method D.

Classes	Water	Vegetation	Bare Soil	Concrete
Water	131	21	16	8
Vegetation	16	242	57	31
Bare Soil	11	37	279	57
Concrete	8	12	21	106

Table 2 (a) Confusion Matrix by Fuzzy *K*-means clustering Algorithm.

Classes	Water	Vegetation	Bare Soil	Concrete
Water	131	21	16	8
Vegetation	14	248	55	29
Bare Soil	10	35	283	56
Concrete	8	12	21	107

Table 2 (b) Confusion Matrix by Modified Fuzzy *K*-means clustering Algorithm.

Classes	Water	Vegetation	Bare Soil	Concrete
Water	134	20	13	9
Vegetation	12	257	52	25
Bare Soil	9	31	288	56
Concrete	6	7	16	114

Table 2 (c) Confusion Matrix by Fuzzy *C*-means clustering Algorithm.

Classes	Water	Vegetation	Bare Soil	Concrete
Water	138	16	15	7
Vegetation	11	259	50	26
Bare Soil	10	26	298	50
Concrete	6	8	13	116

Table 2 (d). Confusion Matrix by Modified Fuzzy *C*-means clustering Algorithm.

Classes	Means				
	Methods	Band1	Band2	Band3	Band4
Water	Fuzzy *K*-means clustering Algorithm	160.10	160.28	159.62	160.03
	Modified Fuzzy *K*-means clustering Algorithm	160.10	160.28	159.62	160.03
	Fuzzy *C*-means clustering Algorithm	160.10	160.28	159.62	160.03
	Modified Fuzzy *C*-means clustering Algorithm	160.10	160.28	159.62	160.03
Vegetation	Fuzzy *K*-means clustering Algorithm	133.26	133.33	133.32	133.34
	Modified Fuzzy *K*-means clustering Algorithm	133.63	132.48	132.08	132.11
	Fuzzy *C*-means clustering Algorithm	132.44	133.98	133.48	133.01
	Modified Fuzzy *C*-means clustering Algorithm	123.95	132.69	132.71	132.63
Bare Soil	Fuzzy *K*-means clustering Algorithm	124.39	123.93	124.04	123.90
	Modified Fuzzy *K*-means clustering Algorithm	123.01	124.54	124.99	124.81
	Fuzzy *C*-means clustering Algorithm	123.09	123.17	123.51	123.37
	Modified Fuzzy *C*-means clustering Algorithm	116.16	123.10	123.25	123.00
Concrete	Fuzzy *K*-means clustering Algorithm	117.42	117.77	118.12	118.07
	Modified Fuzzy *K*-means clustering Algorithm	118.37	117.32	117.97	117.88
	Fuzzy *C*-means clustering Algorithm	119.60	118.32	118.82	119.0
	Modified Fuzzy *C*-means clustering Algorithm	119.92	119.19	119.54	119.51

Table 3: Means by proposed methods and by conventional methods.

6. CONCLUSION

The remote sensing literature extensively uses cluster analysis and discriminant analysis under the terms of unsupervised and supervised classification methods respectively. The real data application of the methods poses challenges to the end-user in the form of higher error rates of classification of the imageries.

This study identified some of the problems of mixed pixels present in the data and explores fuzzy unsupervised classification methods to overcome the problem of mixed pixels. It is found that the improvement through proposed modified fuzzy K-means clustering algorithm (B) over fuzzy K-means clustering algorithm (A) is improved by 1%. In terms of area on the ground, the improvement is of 11 pixels or 1100 sq.mt. area is correctly classified by using the proposed method B. Similarly, the proposed modified FCM clustering algorithm (D) over traditional FCM clustering algorithm (C) is improved by 2.2%. In terms of area on the ground, the improvement is of 18 pixels or 1800 sq.mt. area is correctly classified by using proposed method D. Tables 2 (a-d) shows the confusion matrix by conventional methods and the confusion matrix by proposed methods.

Therefore, this study concludes that the unsupervised classification based on cluster analysis provides an improvement by modifying the membership function of the available unsupervised classification techniques and provides improved results in identifying the mixed pixels present in an image.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

REFERENCES

[1]   S. K. Fu, J. K. Mui, A survey on image segmentation. Pattern Recognition, 13 (1981), 3–16.

[2]   N. Pal, S. Pal, A review on image segmentation techniques. Pattern Recognition, 26 (1993), 1277–1294.

[3]   L.A. Zadeh, Fuzzy sets, Inform. Control, 8 (1965), 338-353.

[4]   J. C. Bezdek, Pattern recognition with fuzzy objective function algorithms, Plenum Press, New York, (1981).

[5] J.C. Bezdek, L.O. Hall, L.P. Clarke, Review of MR image segmentation techniques using pattern recognition. Med. Phys. 20 (1993), 1033–1048.

[6] S.K. Pal, D.D. Majumdar, Fuzzy mathematical approach to pattern recognition, Wiley, New York, (1986).

[7] P. Bibiloni, M. Gonzalez-Hidalgo, S. Massanet, Soft color morphology: a fuzzy approach for multivariate images, J. Math. Imaging Vision, 61 (2019), 394-410.

[8] A.K. Dwivedi, S. Roy, D. Singh, An adaptive neuro-fuzzy approach for decomposition of mixed pixels to improve crop area estimation using satellite images, in: IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Waikoloa, HI, USA, 2020: pp. 4191–4194.

[9] R.B. Thapa, Y. Murayama, Urban mapping, accuracy, & image classification: A comparison of multiple approaches in Tsukuba City, Japan, Appl. Geography. 29 (2009), 135–144.

[10] P.F. Fisher, S. Pathirana, The evaluation of fuzzy membership of land cover classes in the suburban zone, Remote Sensing Environ. 34 (1990), 121–132.

[11] J.T. Tou, R.C. Gonzalez, Pattern recognition principles. Addison-Wesley Publishing Company, (1974).

[12] R. Bellman, R. Kalaba, L.A. Zadeh, Abstraction and pattern classification, J. Math. Anal. and Appl. 2 (1966), 581-586.

[13] E.H. Ruspini, A new approach to clustering, Inform. Control, 15 (1969), 22-32.

[14] J.C. Dunn, A fuzzy relative of the ISODATA process and its use in detecting compact well separated clusters, J. Cybernetics, 3 (1974), 32–57.

[15] J.C. Bezdek, R. Ehrlich, W. Full, FCM: the fuzzy c-means clustering algorithm. Comput. Geosci. 10 (1984), 191–203.

[16] R.A. Johnson, D.W. Wichern, Applied multivariate statistical analysis, sixth edition, Pearson, (2007).