INDIAN INSTIT UTE OF REMOTE SENSING

Indian Sp ace Research Organization

Departmen t of space, Government of India

DOCUMENTATION

“Implementation of Unsupervised FCM and PCM Clustering Algorithms. ”

June-July, 2018

Under the guidance of

Dr. Anil Kumar
Head PRSD, Scientist/Engineer ‘SG’
PRSD/IIRS/ISRO

Submitted By:

Yash Dubey (Roll no: 15103044)

Abhinav Kumar Singh (Roll no: 15103051)

INTRODUCTION:

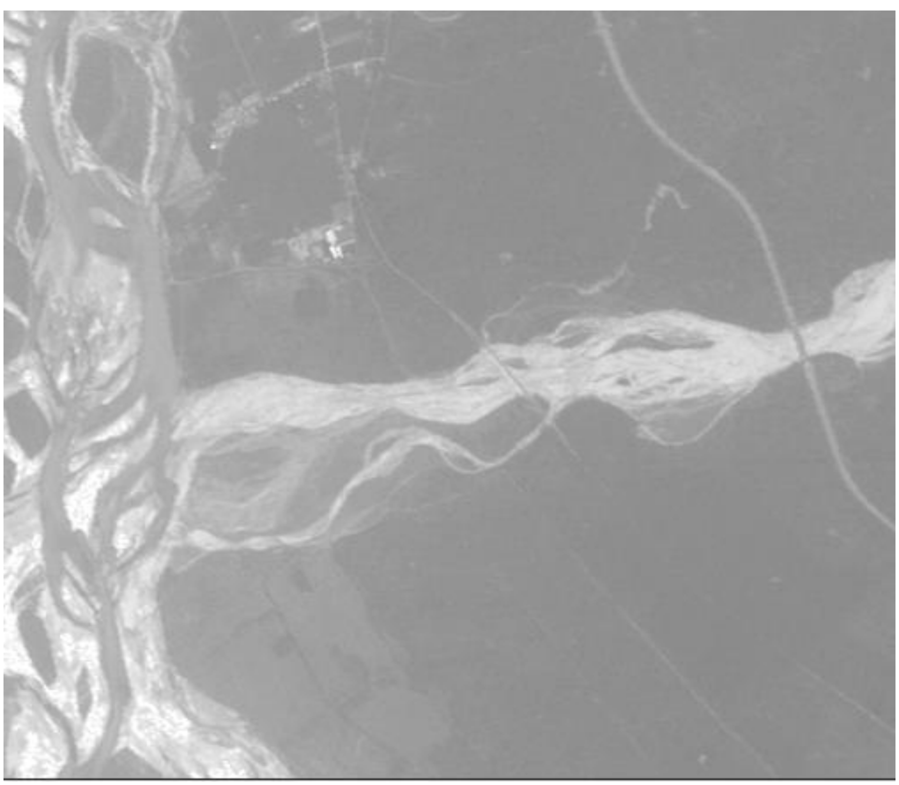
Fuzzy C-means (FCM) is one of the most widely used clustering algorithms and assigns memberships to which are inversely related to the relative distance to the point prototypes that are cluster centers in the FCM model. In order to overcome the problem of outliers in data, models such as possibilistic C-means (PCM) models have been proposed. In Unsupervised FCM and PCM no Training Data is provided instead initial values of variables are given by the user and the program runs for a specified number of iterations.

REQUIREMENTS:

Knowledge of algorithms of FCM and PCM and coding skills in Java.
SOFTWARE USED: Eclipse IDE.

EXPLANATION:

The image classification is done on the dataset provided and has been tested against 4 bands. The input image is and image consisting of various rivers and vegetation:



Sequence of Implementation:

FCM unsupervised:

The FCM algorithm (El-Aziz, 2004) is as follows:

1) Fixing values for m, c (number of classes to be extracted), A and maximum number of iterations.

2) An initial membership matrix, , is selected and its elements are assigned membership values ranging from 0 to 1 to for fuzzy classification.

3) Cluster centre is calculated as given in equation (4).

4) The distance is computed based on the selected A norm using equation (3).

5) The U matrix is updated for the next iteration until the user defined error limit is reached.

6) The final U matrix will represent the class proportions.

$$J_m(U,V) = \sum_{i=1}^{N}\sum_{j=1}^{c}(\mu_{ij})^m \parallel X_i - V_j \parallel_A^2 + \sum_{j=1}^{c}\eta_j \sum_{i=1}^{N}\left(1-\mu_{ij}\right)^m \qquad (6)$$

The equation (6) is subject to constraints,

For all i $\max_j \mu_{ij} > 0$

For all j $\sum_{i=1}^{N}\mu_{ij} > 0$

For all i, j $0 \le \mu_{ij} \le 1$

$d_{ij}^2 = \parallel X_i - V_j \parallel_A^2$ and the distance in feature space, μ_{ij} is the membership of pixel i in class j, N is the total number of pixels, m is the weighted constant (1< m < ∞), V_j is the cluster center for class j, X_i is the feature vector for pixel i, A is the weight matrix and the Euclidean norm used here.

η_j is dependent on the shape and average size of cluster j and is computed as in Eq. (7):

$$\eta_j = K\,\frac{\sum_{i=1}^{N}\mu_{ij}^m d_{ij}^2}{\sum_{i=1}^{N}\mu_{ij}^m} \qquad (7)$$

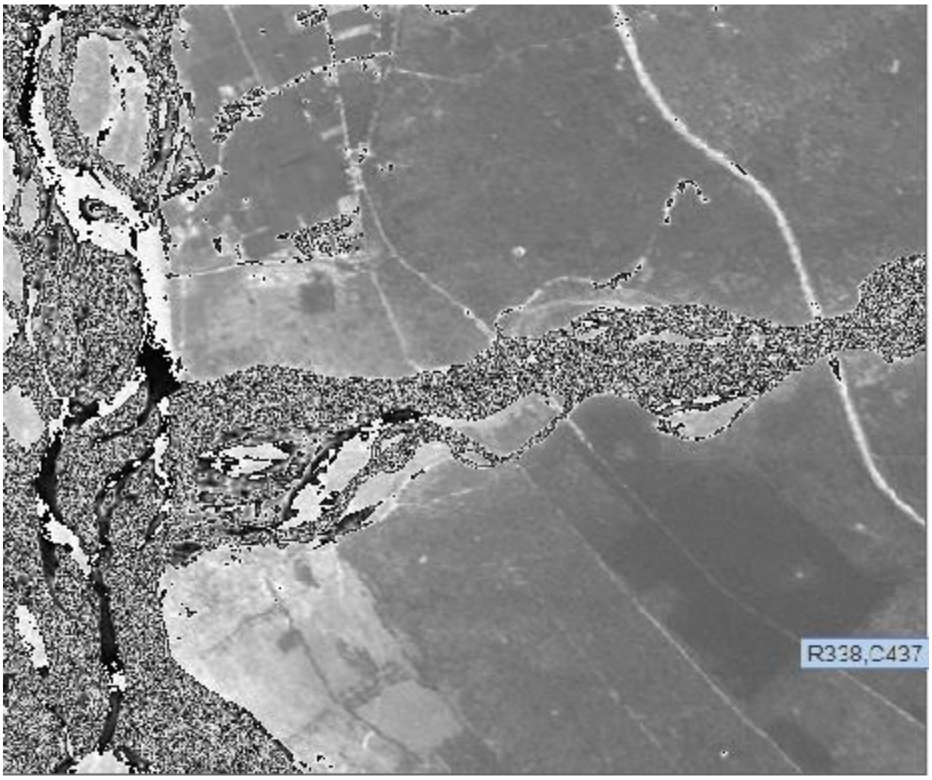
K is a constant generally kept as unity.

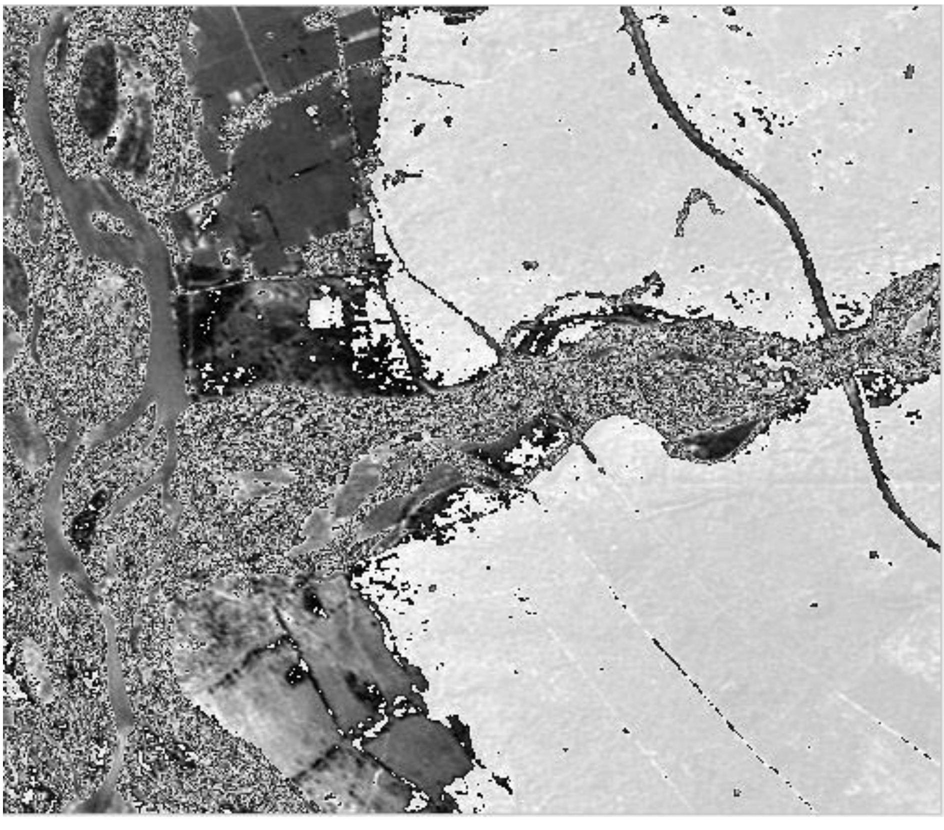
There after class memberships (μ_{ij}) are calculated as in Eq. (8):

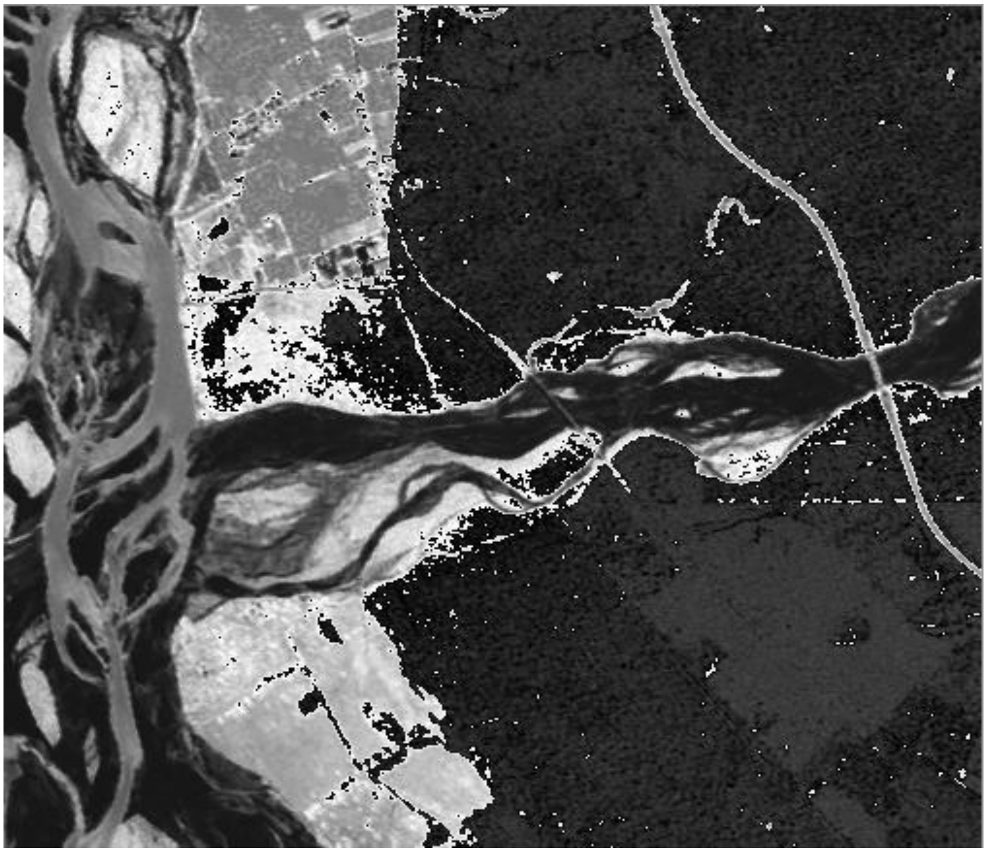
$$\mu_{ij} = \frac{1}{1+\left(d_{ij}^2/\eta_j\right)^{1/(m-1)}} \qquad (8)$$

From equation (5) it can be seen that the membership values generated for pixels are dependent on the number of classes to be extracted, where $d_{ik}^2 = \sum_{j=1}^{C} d_{ij}^2$. While extracting a single class from a remote sensing image the following case of $d_{ik}^2 = d_{ij}^2$ is encountered and the memberships (μ_{ij}) of all pixels in class j become unity. This means that all the pixels in the image belong to the same class j which is not the case. In such a case while using PCM, from Eq. (7) $\eta_j = K\,\frac{\sum_{i=1}^{N}\mu_{ij}^m}{N}$ and the class memberships are calculated as given in Eq. (8).

OUTPUT: For m=2.2, no.of.clusters=4, no.of.iterations=10.







PCM unsupervised:

The PCM algorithm (El-Aziz, 2004) is also the same as that of the FCM algorithm but with some
significant changes.

1) Fixing values for m, c (number of classes to be extracted), A and maximum number of iterations.

2) An initial membership matrix, , is selected and its elements are assigned membership values ranging
from 0 to 1 to for fuzzy classification.

3) Then is estimated from equation (7).

4) The cluster centre is calculated as in equation (4).

5) The distance is computed based on the selected A norm using equation (3).

6) The U matrix is updated for the next iteration dotil the user defined error limit is reached.

7) The final value of is estimated from equation (7) using the updated U matrix.

8) the elements of U matrix is then again computed from the final value.

9) The final U matrix will represent the class proportions.

PCM (*Possibilistic c-Means*) is a modified form of FCM clustering technique that assigns representative feature points the highest possible membership, while unrepresentative points get low memberships (Krishnapuram and Keller, 1993). Thus to satisfy this requirement the objective function of FCM (Eq. (2)) has been modified to as given in Eq. (6):

$$J_m(U,V) = \sum_{i=1}^{N}\sum_{j=1}^{c}(\mu_{ij})^m \parallel X_i - V_j \parallel_A^2 + \sum_{j=1}^{c}\eta_j \sum_{i=1}^{N}(1-\mu_{ij})^m \qquad (6)$$

The equation (6) is subject to constraints,

For all i $\max_j \mu_{ij} > 0$

For all j $\sum_{i=1}^{N} \mu_{ij} > 0$

For all i, j $0 \le \mu_{ij} \le 1$

$d_{ij}^2 = \parallel X_i - V_j \parallel_A^2$ and the distance in feature space, μ_{ij} is the membership of pixel i in class j, N is the total number of pixels, m is the weighted constant (1< m < ∞), V_j is the cluster center for class j, X_i is the feature vector for pixel i, A is the weight matrix and the Euclidean norm used here.

η_j is dependent on the shape and average size of cluster j and is computed as in Eq. (7):

$$\eta_j = K \frac{\sum_{i=1}^{N} \mu_{ij}^{m} d_{ij}^{2}}{\sum_{i=1}^{N} \mu_{ij}^{m}} \qquad (7)$$

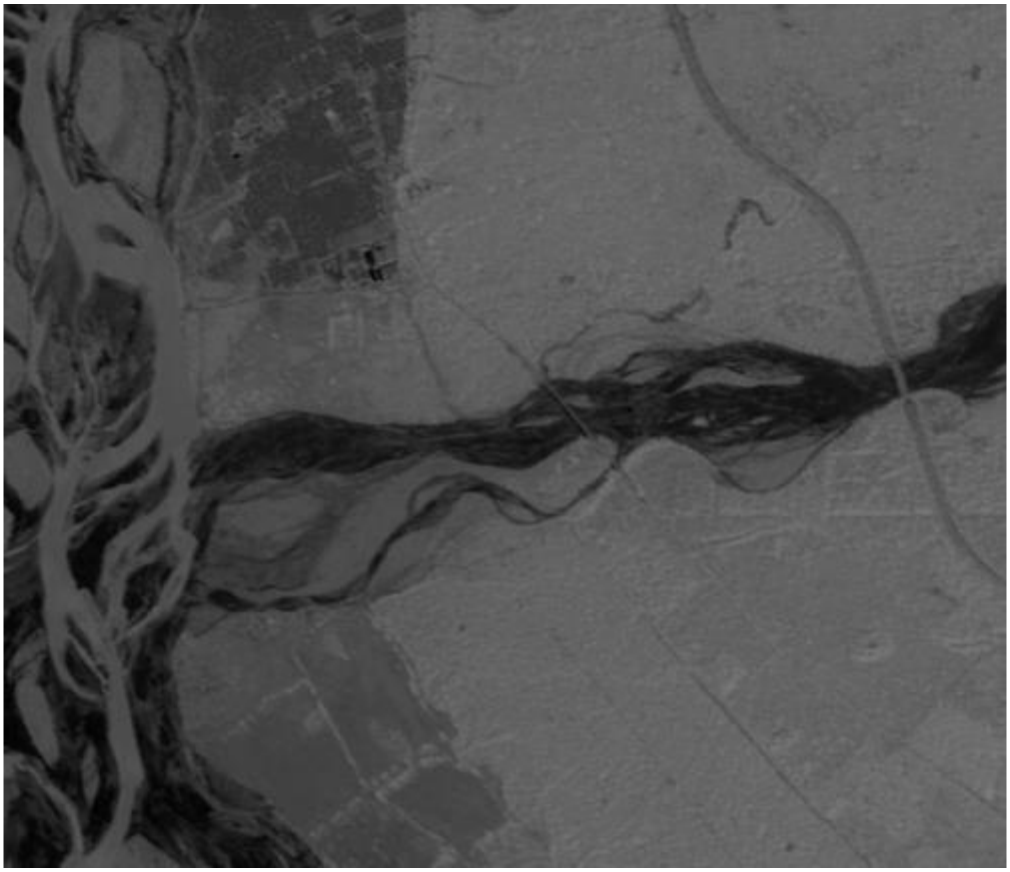
K is a constant generally kept as unity.

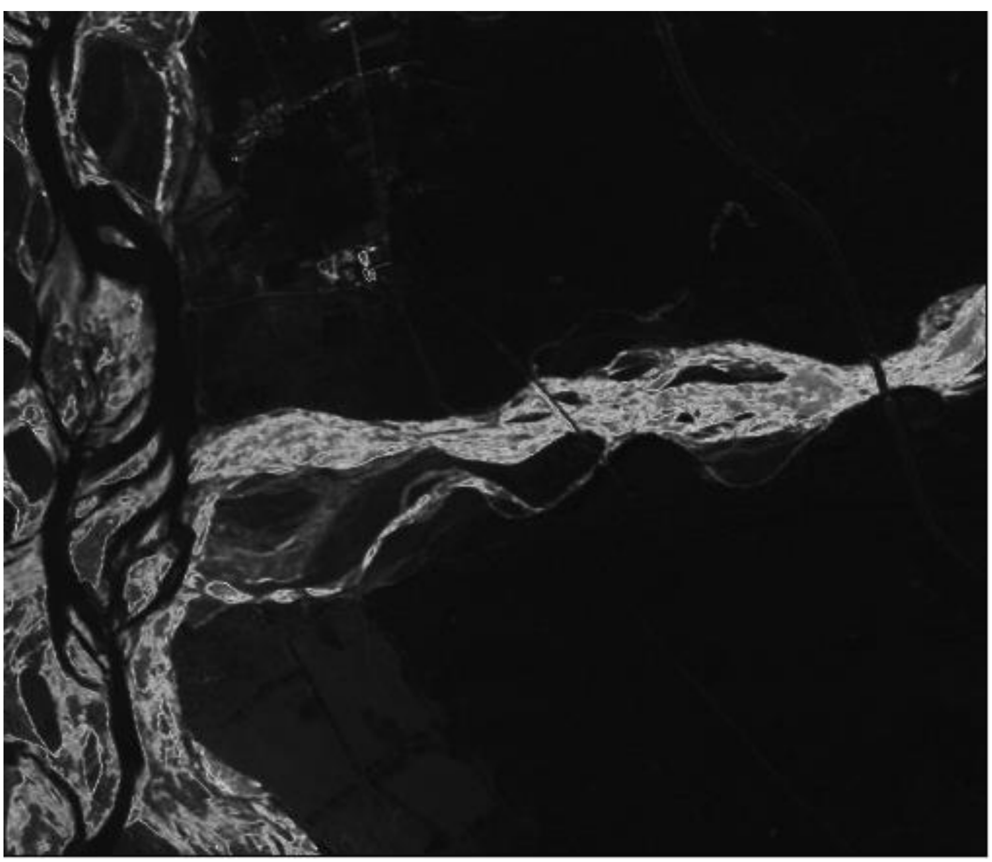
There after class memberships (μ_{ij}) are calculated as in Eq. (8):

$$\mu_{ij} = \frac{1}{1 + \left(d_{ij}^{2}/\eta_j\right)^{1/(m-1)}} \qquad (8)$$

From equation (5) it can be seen that the membership values generated for pixels are dependent on the number of classes to be extracted, where $d_{ik}^2 = \sum_{j=1}^{c} d_{ij}^2$. While extracting a single class from a remote sensing image the following case of $d_{ik}^2 = d_{ij}^2$ is encountered and the memberships (μ_{ij}) of all pixels in class j become unity. This means that all the pixels in the image belong to the same class j which is not the case. In such a case while using PCM, from Eq. (7) $\eta_j = K \frac{\sum_{i=1}^{N} \mu_{ij}^{m}}{N}$ and the class memberships are calculated as given in Eq. (8).

OUTPUT: for m=2.2, no.of.clusters=3, no.of.itr=10 .





CONCLUSION:

On running this project the given i mage dataset of 4 bands, no. of images equal to the no. of clusters are obtained with each image sowing the membership of each pixel in the image to that respective cluster in both FCM and PCM algorithms and membership values of each pixel are also obtained for each image which lies between 0 and 1.