International Journal of Scientific & Engineering Research, Volume 4, Issue 6, June 2013 1415
ISSN 2229-5518

A Novel Based Fuzzy Clustering Algorithms for Classification Remote Sensing Images

Lakshmana Phaneendra Maguluri, Shaik Salma Begum, T Venkata Mohan Rao

Abstract— Clustering is an unsupervised classification method widely used for classification of remote sensing images. Remote sensing can be defined as any process whereby information is gathered about an object, area or phenomenon without making physical contact with the object [1]. The remote sensing technology is used to classify objects on the Earth by means of propagated signals. In this paper, six different clustering algorithms such as K-means, Moving K-means, Fuzzy K-means and Fuzzy Moving K-means, Adaptive Moving k-means, Adaptive Fuzzy Moving k-means are used for classification of remote sensing images. The Adaptive Fuzzy Moving K-means clustering algorithm avoids the problems such as, the occurrence of dead centers, center redundancy and trapped center at local minima. The final center values obtained are located with their respective group of data. The experimental results show that Adaptive Fuzzy Moving K-means has classified the remote sensing image more accurately than remaining algorithms.

Keywords — Image Clustering, Remote Sensing, Image Processing.

---

1 Introduction

Remote sensing image classification is a key technology in remote sensing applications [2]. Rapid and high accuracy remote sensing image classification algorithm is the precondition of kinds of practical applications. In remote sensing, sensors are available that can generate multispectral data, involving five to more than hundred bands. As of now, there is different image classification methods used for different purposes by various researches. These techniques are distinguished in two main ways as supervised and unsupervised classifications [3]. Supervised classification has different sub-classification methods which are named as parallelepiped, maximum likelihood, minimum distances and Fisher classifier methods. Unsupervised classification has evolved in two basic strategies [4], Iterative and Sequential. In an iterative procedure such as K-Means or ISODATA, an initial number of desired clusters are selected, and the centroid locations are then moved around until a statistically optimal fit is obtained. In a sequential algorithm such as Classification by Progressive Generalization, the large number of spectral combinations is gradually reduced through a series of steps using various proximity measures [5][10]. Clustering is an unsupervised classification and widely classified into hierarchical and partitional clustering**.** Hierarchical clustering which seeks to build a hierarchy of clusters. Hierarchical clustering generally falls into two types: bottom up (Agglomerative) and top down (Divisive).

In partitioned based clustering pixels are partitioned into several parts.

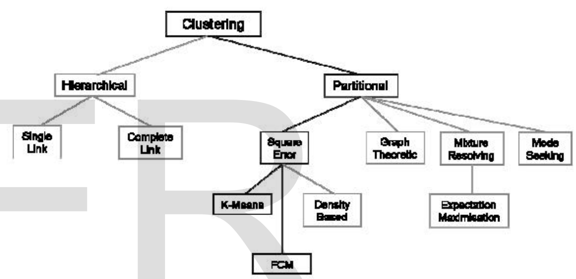


Fig 1 : Types of Clustering.

In fuzzy clustering the elements are assigned not only to one cluster, but to all the clusters with certain degree of membership. This membership to groups is not hard/crisp, rather soft and gradual and is represented by a numeric value between 0 to 1 (called, "membership grade"). Amongst various fuzzy clustering algorithms, Fuzzy C-Means (FCM) is the basic one. As it has some limitations, several algorithms have been developed further to improve its performance .

In this paper, we used versions of conventional k-means clustering algorithm for classification of remote sensing image. The qualitative and quantitative results show that Adaptive Fuzzy Moving K-means clustering algorithm has classified the image better than other clustering algorithms. The paper is organized as follows: Section 2 presents the K-means clustering algorithm, Section 3 presents Moving K-means clustering algorithm, Section 4 presents Fuzzy K-means clustering algorithm, Section 5 presents Fuzzy Moving K-means Clustering algorithm, Section 6 presents Adaptive moving K-means Clustering algorithm, Section 7 presents Adaptive Fuzzy moving K-means clustering algorithms 8 presents Experimental results and finally Section 9 report conclusions.

- *Lakshmana phaneendra maguluri is currently pursuing masters degree program in Gitam Institute of Technology in Gitam University, INDIA PH-9666966077. E-mail: phanendra51@gmail.com.*
- *Shaik Salma Begum, Assistant Professor in DEPT of CSE,* Gudlavalleru Engineering college, INDIA, *PH-9948053635. E-mail: salmashaik.86@gmail.com.*
- *T Venkata Mohan Rao,Assistant Professor in DEPT of CSE, Amrita Sai Institute of Science and Technology, INDIA. PH:9848055835 E-mail: tallapudimohan16@gmail.com.*

2 K-MEANS CLUSTERING ALGORITHM

K-means is one of the basic clustering methods introduced by Hartigan [6]. This method is applied to segment the remote sensing image in recent years. The K-means clustering algorithm for classification of remote sensing image is summarized as follows:

Algorithm K-means(x,n,c)
Input:

N: number of pixels to be clustered; $x=\{x_1,x_2,x_3,\ldots\ldots,x_N\}$: pixels of remote sensing image
$c=\{c_1,c_2,c_3,\ldots.,c_j\}$: clusters respectively.

Output:

cl: cluster of pixels

Begin

Step 1: cluster centroids are initialized.

Step 2: compute the closest cluster for each pixel and classify it to that cluster, ie: the objective is to minimize the sum of squares of the distances given by the following:

$$\Delta_{ij} = || x_i - c_j ||. \ \ arg\ min \sum_{i=1}^{N} \sum_{j=1}^{C} \Delta_{ij}^2 \qquad (1)$$

Step 3: Compute new centroids after all the pixels are clustered. The new centroids of a cluster is calculated by the following

$$c_j = \frac{}{} \sum x_i \ \ where\ x_i\ belongs\ to\ c_j\ . \qquad (2)$$

Step 4: Repeat steps 2-3 till the sum of squares given in equation is minimized.

End.

The K-means clustering algorithm has many weaknesses which are as follows:

1. The number of clusters must be determined before the algorithm is executed.
2. The algorithm is sensitive to initial conditions. It produces different results for different initial conditions.
3. The K-means algorithm may be trapped in the local optimum. As a result, the trapped clusters would represent wrong group of data.
4. Data which are far away from the centers may pull the centers away from the optimum location, leading to poor representation of data.

3. MOVING K-MEANS CLUSTERING ALGORITHM

The Moving K-means clustering algorithm is the modified version of K-means proposed in [7]. It introduces the concept of fitness to ensure that each cluster should have a significant number of members and final fitness values before the new position of cluster is calculated. The Moving K-means clustering algorithm for classification of remote sensing image is summarized as follows:

Algorithm Moving K-means(x,n,c) Input:

N: number of pixels to be clustered; $x=\{x_1,x_2,x_3,\ldots\ldots,x_N\}$: pixels of remote sensing image.
$c=\{c_1,c_2,c_3,\ldots.,c_j\}$: clusters respectively.

Output:

cl: cluster of pixels

Begin

Step 1: cluster centroids are initialized

Step 2: compute the closest cluster for each pixel and classify it to that cluster, ie: the objective is to minimize the sum of squares of the distances given by the following:

$$\Delta_{ij} = || x_i - c_j ||. \ \ arg\ min \sum_{i=1}^{N} \sum_{j=1}^{C} \Delta_{ij}^2 \qquad (3)$$

Step 3: The fitness for each cluster is calculated Using:

$$f(c_k) = \sum_{t \in c_k} (|| x_t - c_k ||)^2 \qquad (4)$$

All centers must satisfy the following condition:

$$f(c_s) \geq \alpha_a f(c_l) \qquad (5)$$

where α_a is small constant value initially with value in range $0< \alpha_a <1/3$, c_s and c_l are the centers that have the smallest and the largest fitness values. If (5) is not fulfilled, the members of c_l are assigned as members of c_s, while the rest are maintained as the members of c_l. The positions of c_s and c_l are recalculated according to:

$$C_s = 1/n_{cs} \left(\sum_{t \in c_s} x_t \right) \qquad (6)$$

$$C_l = 1/n_{cl} \left(\sum_{t \in c_l} x_t \right) \qquad (7)$$

The value of α_a is then updated according to:

$$\alpha_a = \alpha_a - \alpha_a / n_c \qquad (8)$$

The above process are repeated until (5) is fulfilled. Next all data are reassigned to their nearst center and the new center positions are recalculated using (2).

Step 4: The iteration process is repeated until the following condition is satisfied.

$$f(c_s) \geq \alpha_a f(c_l) \qquad (9)$$

End

The Moving K-means algorithm has the following drawbacks:

1. The Moving K-means algorithm is sensitive to noise.
2. For some cases of Moving k-means, the clusters or centers are not located in the middle or centroid of group of data, leading to imprecise results.
3. The fitness concept in the Moving k-means algorithm lead to a problem where some members of centers with the

largest fitness are enforced to be assigned as a members of a center with the smallest fitness.

4. Fuzzy K-Means Clustering Algorithm

The Fuzzy K-means [8] is an unsupervised clustering algorithm. The main idea of introducing fuzzy concept in the Fuzzy K-means algorithm is that an object can belong simultaneously to more than one class and does so by varying degrees called memberships. It distributes the membership values in a normalized fashion. It does not require prior knowledge about the data to be segmented. It can be used with any number of features and number of classes. The fuzzy K-means is an iterative method which tries to separate the set of data into a number of compact clusters. The Fuzzy K-means algorithm is summarized as follows:

Algorithm Fuzzy K-Means (x,n,c,m)
Input:
N=number of pixels to be clustered;

x = {x1, x2 ,..., xN}: pixels of remote sensing image;

c ={c1,c2,c3,….,cj} : clusters respectively. m=2: the fuzziness parameter;

Output:

u: membership values of pixels and clustered Image

Begin

Step_1: Initialize the membership matrix u_{ij} is a value in (0,1) and the fuzziness parameter m (m=2). The sum of all membership values of a pixel belonging to clusters should satisfy the constraint expressed in the following.

$$\sum_{j=1}^{C} u_{ij} = 1 \qquad (10)$$

for all i= 1,2,…….N, where c (=2) is the number of clusters and N is the number of pixels in remote sensing image.

Step_2: Compute the centroid values for each cluster c_j. Each pixel should have a degree of membership to those designated clusters. So the goal is to find the membership values of pixels belonging to each cluster. The algorithm is an iterative optimization that minimizes the cost function defined as follows:

$$F = \sum_{j=1}^{N} \sum_{i=1}^{C} u_{ijm} \, || x_j - c_i ||_2 \qquad (11)$$

where u_{ij} represents the membership of pixel x_j in the ith cluster and m is the fuzziness parameter. Step_3: Compute the updated membership values u_{ij} belonging to clusters for each pixel and cluster centroids according to the given formula.

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{\|x_j - v_i\|}{\|x_j - v_k\|} \right)^{2/(m-1)}},$$

and

$$v_i = \frac{\sum_{j=1}^{N} u_{ij}^m x_j}{\sum_{j=1}^{N} u_{ij}^m}. \qquad (12)$$

Step_4: Repeat steps 2-3 until the cost function is minimized.

End.

5. Fuzzy Moving K-Means Clusteing Algorithm

In the Fuzzy Moving K-means clustering algorithm [9], the membership function is used in addition to the Euclidian distance to control the assignment of the members to the proper center. The algorithm minimizes the sensitivity to the noisy data by updating the moving member function. It is not obligatory for the members of the center with the largest fitness value to follow the center with the smallest fitness value. The Fuzzy Moving K-means clustering algorithm is summarized as follows:

Algorithm Fuzzy Moving K-Means (x,n,c,m)
Input:

N: number of pixels to be clustered; x={x_1,x_2,x_3,……,x_N}: pixels of remote sensing image

c ={c_1,c_2,c_3,….,c_j} : clusters respectively m=2: the fuzziness parameter;

Output:

u: membership values of pixels and clustered Image

Begin

Step_1: Initialize the membership matrix u_{ij} is a value in (0,1) and the fuzziness parameter m (m=2).
The sum of all membership values of a pixel belonging to clusters should satisfy the constraint expressed in the following.

$$\sum_{j=1}^{c} u_{ij} = 1 \qquad (13)$$

for all i= 1,2,…….N, where c (=2) is the number of clusters and N is the number of pixels in remote sensing image.
Step_2: Compute the centroid values for each cluster c_j. Each pixel should have a degree of membership to those designated clusters. So the goal is to find the membership values of pixels belonging to each cluster.
The algorithm is an iterative optimization that minimizes the cost function defined as follows:

$$f = \sum_{j=1}^{N} \sum_{I=1}^{C} {U_{IJ}}^{M} \left\| x_j - c_i \right\|^2 \qquad (14)$$

ISSN 2229-5518

where u_{ij} represents the membership of pixel x_j in the i^{th} cluster and m is the fuzziness parameter.

Step 3: The fitness for each cluster is calculated using

$$f(c_k) = \sum_{t \in c_k} (|| x_t - c_k ||)^2 \qquad (15)$$

All centers must satisfy the following condition:

$$f(c_s) \geq \alpha_a f(c_l) \text{ and } m(c_{sk}) > m(c_{lk}) \qquad (16)$$

where α_a is small constant value initially with value in range $0 < \alpha_a < 1/3$, c_s and c_l are the centers that have the smallest and the largest fitness values, $m(c_{sk})$ is the membership value of point k according to the smallest centre and $m(c_{lk})$ is the membership value of point k according to the largest centre. If (5) is not fulfilled, the members of c_l are assigned as members of c_s, while the rest are maintained as the members of c_l. The positions of c_s and c_l are recalculated according to:

$$C_s = 1/n_{cs} \left(\sum_{t \in c_s} x_t \right) \qquad (17)$$

$$C_l = 1/n_{cl} \left(\sum_{t \in c_l} x_t \right) \qquad (18)$$

The value of α_a is then updated according to:

$$\alpha_a = \alpha_a - \alpha_a / n_c \qquad (19)$$

The above process are repeated until (5) is fulfilled. Next all data are reassigned to their nearst center

and the new center positions are recalculated using (2).

Compute the updated membership values u_{ij} belonging to clusters for each pixel according to given formula

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{\|x_j - v_i\|}{\|x_j - v_k\|} \right)^{2/(m-1)}},$$

and

$$v_i = \frac{\sum_{j=1}^{N} u_{ij}^m x_j}{\sum_{j=1}^{N} u_{ij}^m}. \qquad (20)$$

Step 4: The iteration process is repeated until the following condition is satisfied.

$$f(c_s) \geq \alpha_a f(c_l) \text{ and } m(c_{sk}) > m(c_{lk}) \qquad (21)$$

6. Adaptive Moving K-Means Clustering Algorithm

One of the weaknesses of the MKM algorithm is that it obligates the members of the center with the largest fitness value to become a member of the center with the smallest fitness value , even though, the center with the smallest fitness represents a group of noise data. This issue could lead to miss-clustering of data to unwanted noise cluster, which in turn, affects the segmented results. The proposed adaptive moving k-means (AMKM) clustering algorithm provides a solution to this problem. The algorithm is as follows:

Algorithm Adaptive Moving K-means(x,n,c)

Input:

N:number of pixels to be clustered;
$x = \{x_1, x_2, x_3, \ldots, x_N\}$: pixels of microarray image
$c = \{c_1, c_2, c_3, \ldots, c_j\}$: clusters respectively. Here we group the pixels into two clusters, foreground and background, j=2.

Output:

cl: cluster of pixels

Begin:

Step 1: cluster centroids are initialized (minimal and maximal values of pixels in the target area)

Step 2: compute the closest cluster for each pixel and classify it to that cluster, ie: the objective is to minimize the sum of squares of the distances given by the following:

$$\Delta_{ij} = || x_i - c_j ||. \ \arg\min \sum_{i=1}^{N} \sum_{j=1}^{C} \Delta_{ij}^2 \qquad (22)$$

Step 3: The fitness for each cluster is calculated using

$$f(c_k) = \sum_{t \in c_k} (|| x_t - c_k ||)^2 \qquad (23)$$

The members of the center with the largest fitness value must satisfy the following condition:

$$v_i > C_l \qquad (24)$$

Where v_i is the i-th data of the center with the largest fitness value and C_l is the is the center with the largest fitness value. If the members of the center with the largest fitness value does not fulfill condition (24) then they are assigned to the nearest cluster depending upon the minimum Euclidian distance, while the rest are maintained as the members of C_l .

The above process is repeated until (24) is satisfied. Next all data are reassigned to their nearest center and new center positions are recalculated using (2).

Step 4: The iteration process is repeated until the following condition is satisfied.

$$v_i > C_l \qquad (25)$$

7. ADAPTIVE FUZZY MOVING K-MEANS CLUSTERING ALGORITHM:

Adaptive Fuzzy Moving K-means Clustering Algorithm which will combines both the clustering algorithms Fuzzy moving k-means clustering algorithm(FMKM) and Adaptive moving K-means clustering algorithm(AMKM).Thus Adaptive Fuzzy Moving K-means Clustering Algorithm(AFMKM) which will combines the properties of the both clustering algorithms. After assigning each data to the nearest center based on the fuzzy membership function and the fitness function is fulfilled ,all the members will be assigned to the proper center.

Algorithm Adaptive Fuzzy Moving K-means(x,n,c)
Input:

N: number of pixels to be clustered;
$x=\{x_1,x_2,x_3,\ldots\ldots,x_N\}$: pixels of microarray image
$c=\{c_1,c_2,c_3,\ldots.,c_j\}$: clusters respectively. Here we group the pixels into two clusters, foreground and background, j=2.

Output:

cl: cluster of pixels

Begin:

Step 1: Initialize the membership matrix u_{ij} is a value in (0,1) and the fuzziness parameter m (m=2). The sum of all membership values of a pixel belonging toclusters should satisfy the constraint expressed in the following.

$$\sum_{j=1}^{c} u_{ij} = 1 \qquad (26)$$

for all i= 1,2,…….N, where c (=2) is the number of clusters and N is the number of pixels in remote sensing image.

Step 2: compute the closest cluster for each pixel and classify it to that cluster, ie: the objective is to minimize the sum of squares of the distances given by the following:

$$\Delta_{ij} = || x_i - c_j ||. \ arg\ min \sum_{i=1}^{N} \sum_{j=1}^{C} \Delta_{ij}^{2} \qquad (27)$$

Step 3: The fitness for each cluster is calculated using the members of centers with largest fitness value must satisfies the following conditions

$$v_i > C_l \qquad (29)$$

where v_i is the ith data of the center with largest fitness value and c_l is the largest center with largest fitness value .if the members of centers with largest fitness value does not satisfiess the above conditions(29) then they are assiged to nearest clusters depending on minimum ecludiean distance while rest are maintained as members of cl .the above process are maintained as upto the above condition satisfied(29).next all data are reassigned to the nearest centers.and positions are recalculated as equation(2)

Step 4:The iteration of the process is repeated until the following condition is satisfied

$$v_i > C_l \qquad (30)$$

8 EXPERIMENTAL RESULTS

Qualitative Analysis:

The proposed four clustering algorithms are performed on a remote sensing image that consists of a total of 38808 pixels. The classification results by the proposed algorithms are shown in figure 2.

From the results, we can visualize that Adaptive Fuzzy Moving K-means have classified the remote sensing image more accurately than other three algorithms. It avoids the problems such as, the occurrence of dead centers, center redundancy and trapped center at local minima. It is also less sensitive to initialization process of clustering value. The finalcenter values obtained are located with their respective group of data, enabling the size and shape of object to be maintained and preserved.

Quantitative Analysis:

Quantitative analysis is a numerically oriented procedure to figure out the performance of algorithms without any human error. The Mean Square Error (MSE) is significant metric to validate the quality of image. It measures the square error between pixels of the original and the resultant images. The MSE is mathematically defined as

$$MSE = \frac{1}{N}\sum_{j=1}^{k}\sum_{i\epsilon c_l}\left\|v_i - c_j\right\|^2 \qquad (31)$$

Where N is the total number of pixels in an image and x_i is the pixel which belongs to the j^{th} cluster. The lower difference between the resultant and the original image reflects that all the data in the region are located near to its centre. Table 1 shows the quantitative evaluations of six clustering algorithms. The results confirm that Addaptive Fuzzy Moving K-means algorithm produces the lowest MSE value for classifying the remote sensing image.

Table 1: MSE values

METHOD	MSE VALUE
K-means	278.671
Moving K-means	206.926
Fuzzy K-means	118.327
Fuzzy Moving K-means	98.223
Addaptive Moving K-means	78.138
Addaptive Fuzzy Moving K-means	62.142

9 CONCLUSION

This paper has presented six clustering algorithms namely K-means, Moving K-means, Fuzzy K-means, Fuzzy Moving K-means, Adaptive Moving K-means and Adaptive Fuzzy moving K-means for the classification of remote sensing image. The qualitative and quantitative analysis done proved that Adaptive Fuzzy Moving K-means has higher classification quality than other clustering algorithms. The occurrence of dead centers, center redundancy and trapped center at local minima problems can be avoided. The algorithm is also less sensitive to initialization process of clustering value.

ACKNOWLEDGMENT

The authors would like to thank J.Hari Kiran Assitant prof. of Gitam University. For their contributions a unconditional support at various stages of the Work.

REFERENCES

[1]. Yi Ma, Jie Zhang, Yufeng Gao," High Resolution Remote Sensing Image Classification of CoastalZone and its Automatic Realization", 2008International Conference on Computer Science and Software Engineering, 2008 IEEE.

[2] .Xiaofang Liu, Xiaowen Li, Ying Zhang, CunjianYang, Wenbo Xu, Min Li, Huanmin Luo", RemoteSensing Image Classification Using FunctionWeighted FCM Clustering Algorithm" , 2007 IEEE.

[3] Aykut AKGÜN, A.Hüsnü ERONAT and Necdet TÜRKa, "Compar ing different satellite image classification methods: an application in ayvalik district,western, Turkey.

[4] Josef Cihlar, Rasim Latifovic, Jean Beaubien, "A Comparison of Clustering Strategies for Unsupervised Classification, Canadian Journal of Remote Sensing, February 2000.

[5] D. LU and Q. WENG, "A survey of image classification methods and techniques for improving classification performance", International Journal of Remote Sensing, Vol. 28, No. 5, 10 March 2007, 823–870.

[6] Volkan Uslan, Omur Bucak, " clustering based spot segmentation of microarray cDNA MicroarrayImages ", International Conference of the IEE EMBS, 2010.

[7] Siti Naraini Sulaiman, Nor Ashidi Mat Isa, " Denoising based Clutering Algorithms for Segmentation of Low level of Salt and Pepper Noise Corrupted Images", IEEE Transactions on Consumer Electronics, Vol. 56, No.4, November 2010.

[8] LJun-Hao Zhang, Ming Hu HA , Jing Wu," Implementation of Rough Fuzzy K-means Clustering Algorithm in Matlab", Proceedings of Ninth International Conference on Machine Learning and Cybernetics", July 2010.

[9] Nor Ashidi Mat Isa, Samy A.Salamah, Umi Kalthum Ngah.," Adap tive Fuzzy Moving K-means Clustering Algorithm for Image Segmentation", IEEE 2009.

[10] . N. A. M. Isa et al.: Adaptive Fuzzy Moving K-means Clustering Algorithm for Image Segmentation IEEE 2146 Transactions on Consumer Electronics, Vol. 55, No. 4, NOVEMBER 2009

[11] S. N. Sulaiman and N. A. M. Isa: Adaptive Fuzzy-K-means Cluster ing Algorithm for Image Segmentation IEEE Transactions on Consumer Electronics, Vol. 56, No. 4, November 2010.

[12] . B.Saichandana," De-noising based clustering Algorithm for Classi cation of Remote Sensing Image", Journal of Computing, Vol 4 , Issue 11, November 2012.

AUTHOR'S PROFILE

Lakshmana Phaneendra Maguluri

He has obtained his bachelor's degree from Gudlavalleru Engineering Collag Under JNTUK in Computer Science and Technology in 2011 and Currently doing hisM.Tech (Dept of IT) In GITAM UNIVERSITY, INDIA.

Shaik Salma Begum

She has obtained his masters degree from Acharya Nagarjuna university in 2010 Currently working as Assistant Professor In DEPT of CSE Gudlavalleru Engineering College, Andhra Pradesh INDIA

T Venkata MohanRao

He has obtained his masters degree from Acharya Nagarjuna university in 2010 Currently working as Assistant Professor in DEPT of CSEAmrita Sai Institute of Science and Technology, INDIA

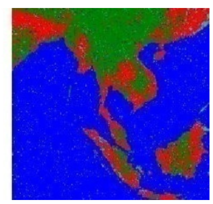
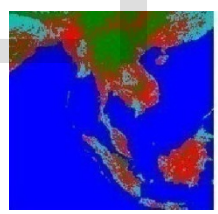
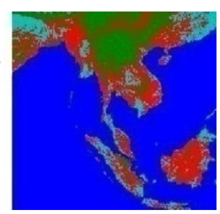
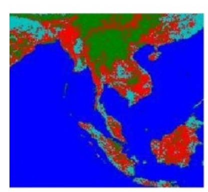
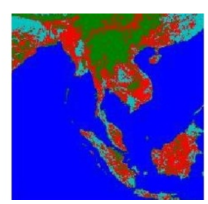
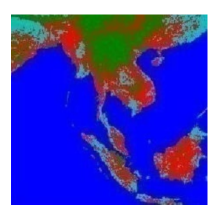
Original Image	K-Means
	
Moving K-Means	Fuzzy K-means
	
Fuzzy Moving K-means	Addaptive Moving K-Means
	
Addaptive Fuzzy Moving K-Means	
	

Fig 2: classification results