# A Comparative Study on Clustering Algorithms using Image Data

Vikas Tondar
Department of Computer
Science and Engineering
MITM Indore

Pramod S. Nair
Department of Computer Science
and Engineering
MITM Indore

## ABSTRACT

Analyzing of image called Segmentation .It is an important concept to viewing and analyzing different type's images and solving a wide range of problems in image. Clustering algorithm and technique for classifying usage image data and the process of analyze image data from dissimilar perception and abbreviation it into valuable information, this information can be use to increase proceeds, cuts costs, or Time complexity. There is different type of algorithms for image data and clustering such as (FCM) fuzzy c-means clustering algorithms, SFCM (Spatial fuzzy c-means clustering), K-Means, and PSOFCM (particle swarm optimization incorporative fuzzy c-means clustering) .The selection between the predictive classifier is extremely important. Fuzzy algorithms based on initial cluster selection without noise data. PSOFCM and SFCM approaches shows better segmentation results can be obtained in noise.

PSOFCM and SFCM approaches shows how better image segmentation of results can be obtained. Image clustering and its applications are used in human image i.e. Medical image segmentation used for detection of Brain images, tumor and more. The result obtained through Particle swarm optimization (PSO), yields better detected image and time complexity compared to FCM and SFCM.

## General Terms

Image segmentation, Clustering algorithm, Time complexity

## Keywords

FCM, Particle swarm optimization based FCM, spatial information based FCM.

## 1. INTRODUCTION

Data mining and clustering has been studied different approach for a long time by researchers. A significant approach of clustering should produce max no clusters without loses. The Accuracy of Segmentation of image method depends on 3 components: how to distance measure, the clustering algorithms used for find the hidden pattern. Clustering method in data mining can be dividing into hierarchical based clustering, partition based clustering. Density-based clustering approach, frequent pattern approach. The database is categorized hierarchical and decomposition of the database called hierarchical clustering .It merge some cluster in order to make a bigger cluster or divide a cluster into some cluster to make small cluster. When database is divided into predefined no of cluster it technique called clustering of partition .They used function with creation criterion to attempt to determine 'K' partitions.

In this paper, the data mining clustering approaches, Fuzzy C-Mean based on initial cluster, SFCM with Spatial information and PSO (PSOFCM) is compare. The set of real data sets are used to establish the functional and compare of the PSOFCM

[2] algorithm is enhanced than the conservative FCM algorithm and SFCM algorithm. Cluster classification in image Data mining is utilize of automated image data analysis method to uncover before undetected relationships between image data segmentation. Many of the main image data segmentation technique is classification and clustering. In this research we are working simply with the clustering since it is nearly all significant process, if we have an extremely image segmentation and clustering discovery object. Clustering is a analysis of explorative data mining, and a frequent method for statistical data analysis use in a lot of fields, counting machine learning, , image analysis, information retrieval, pattern recognition, and bioinformatics. The number of algorithms has been performing for image data classification, but they limitations. A huge scale data set affects the effect of classification and algorithms. Need concentrated computing power for training procedure and image data classification. In addition, based on new work description in the previous work, mainly of algorithms mention beyond worked on small image set.

This paper compares technique fuzzy c-means clustering algorithms of data mining to assist retailer to categorization for image. The aim is to reviewer the accuracy of fuzzy c-means clustering algorithms, SFCM [3], PSOFCM algorithm. Fuzzy clustering, algorithm on various data sets. The performance of image data cluster classification depends on various factors around test mode, size of data set and dissimilar nature of data sets. In this paper we represent first section introduction, second section related work, third methodology and last section we represent the comparative study analysis and conclusion.

## 2. METHODOLOGY
## 2.1 Fuzzy C Means Clustering Alogorithms

Unsupervised Fuzzy clustering is used for the analysis of data and image models .It is an little enhancement of K means clustering algorithms. The main objective of these algorithms is to be defining boundaries between 0 and 1. There are several classes with membership function assigned degree between 0 and 1.Fuzzy c means (FCM) algorithms is used for uncertain data or where there is no boundaries have been calculated.

Fuzzy algorithms membership have different grade between 0 and 1 for different data point in partitioning. This method applied analysis of different image, shape, and medical. The working of fuzzy c mean algorithms is discussed using following steps.

Step of FCM Algorithms:

Fuzzy c Mean algorithm have set of finite elements C= {c1,c2,…..c n } into set cluster of fuzzy according to

predefined method and logic . A predefine set of finite data given such as d.

And a matrix F (partition of initial elements) such that

F =Fmn, m=1, c, n =1

Where Fmn  define the value of function in between point [0, 1]. The result   shows the degree element where Cn have in the m-th cluster.

Fuzzy logic implemented in following steps.

Pass1: Initialization on initial value of cluster C. Exponential factor of weight  $\delta$ ($1 < \delta < \infty$), partition matrix of initial values  $F^0$, and criteria of termination.  .Also;

Pass2: Calculate the center of fuzzy by using  $F^1$.

{  $Y_m^2$| m=1, 2, ..., n}

Pass3: Update  $F^{m+1}$    matrix by using  { $Y_m^2$| m=1, 2, ..., n}

 Pass4.  If  $\| F(m+1) - F(m) \| < \delta$   is true then stop.

This algorithm first allocates the membership function to dataset point according to distance between the cluster and its center and the data point. If data is have more function value then we say cluster have near to center. When we add all function for summarization then we find it is equal to one. When we calculate membership function and iteration we need to update according to condition occurs in algorithms.

Advantage:

1. Unsupervised learning algorithm
2. Converge and give the better result than K-Means algorithms
3. Easy to access and implement

## 2.2   Spatial FCM Algorithms

Fuzzy C-means clustering with additional information of neighbor called   spatial information. This spatial is implemented for analyzing of image to which are not FCM due to noise. One of the major facts of an image is that the nearest pixels are closely related to each other and there are high chances that they belong to the same class, group or cluster. Spatial information is ignored in conventional FCM which are an important parameter for clustering. In spatial FCM information is fused in the membership function to obtain better image segmentation results and is defined as follows:

The SFCM clustering technique is a two iteration procedure at each iteration. Steps which are involved in SFCM clustering as follows:

Step1. In first pass iteration the membership function of SFCM is computed initial cluster as similar as in conventional FCM.

Step2.  The spatial information of each nearest calculated pixel is delineated to the function of spatial information, and the new summation function is determined by spatial old information.

Step3. Next iteration procedure continues until some result reached at point where the difference between two successive iteration of cluster is same.

Step4. Relocated to each pixel according to cluster when membership function is calculated maximum.

Advantages of spatial fuzzy c mean are:

1. Eliminates noisy spots between pixel

2. Reduces false blobs between initial pixel

3. Less sensitive to noise

4. More homogeneous regions pixel  are obtained

## 2.3  Pso Fcm Clustering Alogorithms

Particles swarm optimization clustering approach works in the form of Population and Candidate solution. Later population called swarm and candidate solutions called particles.

These calculated particles are moved around the space area by pre define formula and function. This space called the space area. The moments of particles in search space by function are guided by their own best position which is optimal in search space .This position guide by best position of space it is optimal position in space area. If any point position particle is improved his position then we calculates his position and guide to swarm to movement in space

This process completed and repeated again and hopes for best position but not sure that solution is satisfactory discovered.

The steps of PSOFCM are given below:

Pass1: Define a multi dimension search space M, Particle C for consisting swarm. The ith effect of    particles shows in Vector M .Q is the velocity of M-dimensional vector of particles.

Pass2:   Pc = (Pc1, Pc2, P cM) , Time t

Velocity, Qc = (Qc1, Q c2… QcM), time t

The last best particle position discovered by particle in ith iteration Y is denoted by

Rc = (Rc1, R c2… RcM) * Time t

The position of particle and velocity are initialization by randomly. Let we take some assumption B* to be best global position among particles and take a program function iteration z. calculate the each position and velocity in each iteration and update all swarm and position of particles .

Now we will trying calculate and minimize a function

Further we are trying to minimize a function F =Fmn using unsupervised clustering algorithm. The main objective is used to function of FCM is to maximize global minima of cluster data set. The following step try to minimize the function value of FCM and get the good position of particle achieved.

 Pass 3:

$$R_c(t+1) = AR_c(t)+c_1 Z1(F_c(t) - P_i(t)) + c_2 Z2(F^* - P_c(t))$$

$$\text{And}\quad Pc(t+1) = P_c(t) + R_c(t+1)$$

Where         $F_c = P_c$, if f $(F_c) \geq$ f $(P_c)$

= $P_c$, if f $(P_c) <$ f $(F_c)$

$Y^*$      $Y_0, Y_1, Y_i$} as

$\pounds(F^*) = $ minimum ($\pounds$ $(f_0)$, $\pounds$ $(f_1)$,……., $\pounds$ $(f_c)$)

Z1 and Z2 define two random uniform range sequence between (0, 1) , A is matrix those are containing weight   of Rc(t) same as calculation of $R_c(t+1)$.

It is cleared from equation of last step F* is Best Global Position to be considered.

Advantages of PSO FCM are:

1. Maximum no of edges marks in image
2. Each edges closed to each other .so localization is to be high
3. It is working in noise so image marks one ,it take less response time

Result analysis we perform the implementation and experiment using core to duo 2.9 GH processor ,2GB RAM , 25GB Hard Disk, we used the tool Visual studio 2010 ,do the coding in C# we used for data base implementation and result generation SQL SERVER -2008.

## 3. EXPERIMENTAL RESULTS

Particle swarm optimization and Spatial Fuzzy c mean approaches shows how better image segmentation results can be obtained. PSO and its applications are used in image i.e. Graphic image, abnormalities condition of brains. The result obtained through Particle swarm optimization using FCM yields better detected image and time complexity compared to FCM and SFCM.
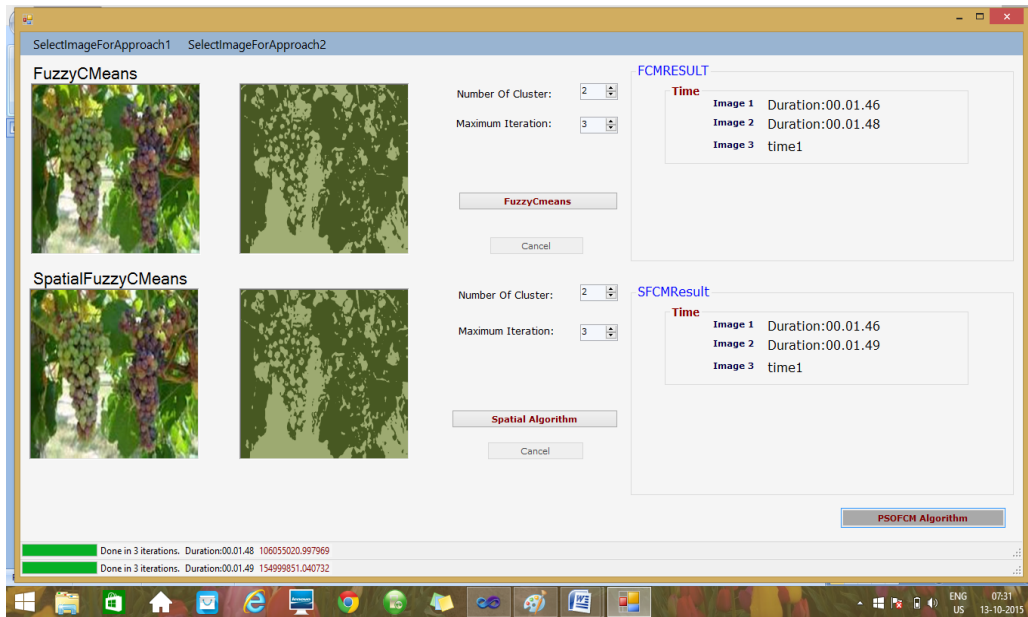


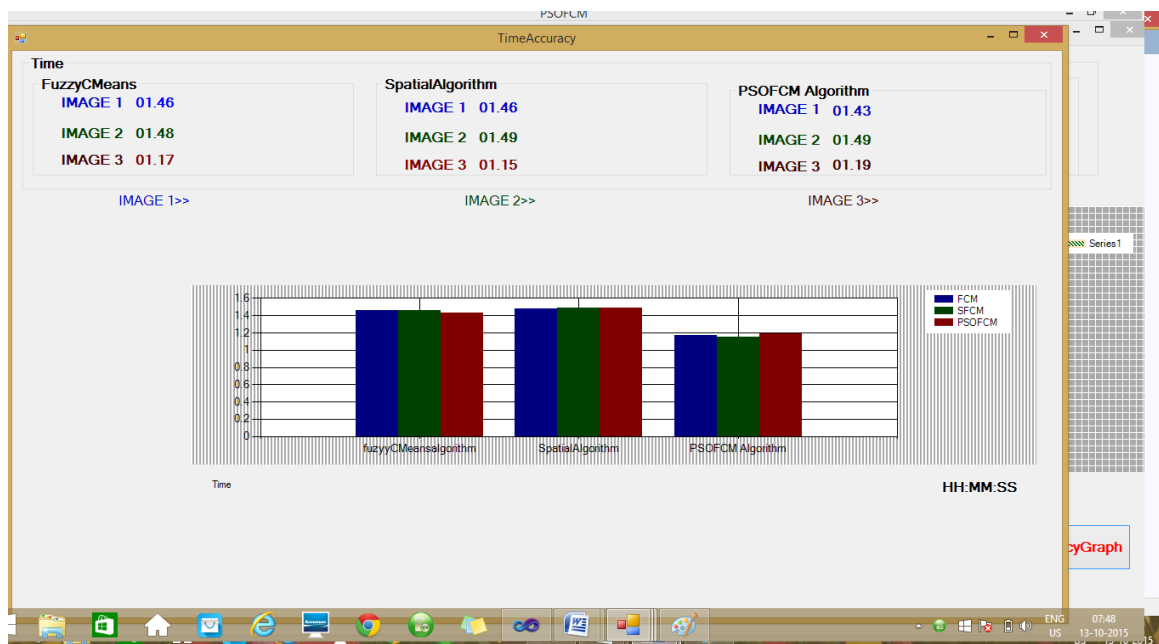**Figure 1: Comparison of FCM, SFCM, and PSOFCM clustering Algorithms for selection of image.**



**Figure 2: Comparison of FCM, SFCM, and PSOFCM clustering Algorithms in term of finding time complexity**

## 4. CONCLUSION

Lastly, we conclude that fuzzy based image clustering algorithms are supposed enhanced of its computation time. We get better the performance of SFCM, PSOFCM than (FCM) fuzzy c-means clustering algorithms, by using image. We process the image file which are free from noisy data and have sequence order data then we can deduct the compilation time and accuracy of fuzzy c means algorithm .These study has been tried utilizing carotid vein ultrasound pictures, ultrasound apparition pictures, and graphic image. Further, its execution has been assessed at diverse grouping quality measures. The concept in this paper shows that PSO based application better result shows than FCM. However results in image mining processing have highly desired result objective. Trial results demonstrate that the PSOFCM procedure offers better grouping contrasted with FCM and SFCM in term of time complexity. We process the image file which are free from noisy data and have sequence order data then we can deduct the compilation time and accuracy of fuzzy c means algorithm.

## 5. REFERENCES

[1.] J. Kennedy and R.C. Eberhart, "Particle swarm optimization" , Proceeding of the 1995 IEEE International Conference on Neural Networks (Perth,Australia), IEEE Service Centre, Piscataway, NI, (1995), Iv: 1942-1948 400

[2]. S. Chen and D. 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, 1998, pp. 1907-1916.

[3] A novel kernelized fuzzy c mean algorithms with application in medical science image segmentation ,Dao-Qiang Zhang,Song,Song –Can Chen,Artificial intelligence in medical Volume 32,Issue 1,page 37-50,September 2004.

[4]. Rafael C. Gonzalez, Richard E.Woods, "Digital Image Processing", Pearson Education, Second Edition, ISBN 81-7758-168-6, 2005

[5]. Yingjie Wang, "Fuzzy Clustering Analysis Using Genetic Algorithm", and ICIC International @ 2008 ISSN 1881-803 X, pp: 331—337.

[6] M. Sonka V Hlavac and R.Boyie ,"Image processing,analsis and machine version",Third edition,Thomson,USA,2008 Iv:208-31

[7] S. Krinidis and V. Chatzis, J. IEEE Transactions on Image Processing, vol. 19, no. 5, (2010)

[8] V. S. Rao and Dr. S. Vidyavathi, "Comparative Investigations and Performance Analysis of FCM and MFPCM Algorithms on Iris data",Indian Journal of Computer Science and Engineering, vol.1, no.2, 2010 pp. 145-151.

[9] SIKKA, K.—SINHA, N.—SINGH, P. K.—MISHRA, A. K.: A Fully Automated Algo-rithm Under Modified FCM Framework for Improved Brain MR Image Segmentation. Magnetic Resonance Imaging, Vol. 27, 2009, No. 7, pp. 994–1004.10] K.Srinivas, P.V.S.Srinivas, A.Govardhan, V.ValliKumari, "Periodic Web Personalization for Meta Search Engine", IJCST, vol. 2, no. 4,December 2011.

[10] Davoud Sedighizadeh and Ellips Masehian, "Particle Swarm Optimization Methods, taxonomy and applications", International Journal of Computer Theory and Engineering (1793-8201) Vol. 1, No. 5, December,2009, pp: 486-502 [12] SheetalChouhan, Manish Shrivastava and KavitaDeshmukh, "A Noble Approach of Web Log Mining", VSRD-IJCSIT, vol. 2, 2012

[11] Romesh Laishram, W.Kanan Kumar Singh, N.Ajit Kumar, Robindro.K, S.Jimriff, "MRI Brain Edge Detection Using GAFCM Segmentation and Canny Algorithm", International Journal of Advances in Electronics Engineering – IJAEE,volume 2 - Issue 3, ISSN:- 2278-215X, pp. 168-171,December 8,2012.

[12] Gaussian smoothing." International Journal of Computational Intelligence in Bioinformatics and Systems Biology 1(3): 316-331. December 3, 2012

[13] Ortiz, A., J. Gorriz, et al. (2012). "Unsupervised Neural Techniques Applied to MR Brain Image Segmentation." Advances in Artificial Neural Systems 2012.