 ISSN: 2321-8169
122 – 127

Illustration of Medical Image Segmentation based on Clustering Algorithms

Harsmeet Singh
Department of Computer Science
CGC-College of Engineering
Mohali, Punjab
er.harsmeet@gmail.com

Jagbir Singh Gill
Department of Computer Science
CGC-College of Engineering
Mohali, Punjab

Abstract- Image segmentation is the most basic and crucial process remembering the true objective to facilitate the characterization and representation of the structure of excitement for medical or basic images. Despite escalated research, segmentation remains a challenging issue because of the differing image content, cluttered objects, occlusion, non-uniform object surface, and different factors. There are numerous calculations and techniques accessible for image segmentation yet at the same time there requirements to build up an efficient, quick technique of medical image segmentation. This paper has focused on K-means and Fuzzy C means clustering algorithm to segment malaria blood samples in more accurate manner.

Keywords: *Image segmentation; K-means algorithms; fuzzy C-means algorithms.*

*****

I. INTRODUCTION

Image processing is a technique to play out a few operations on an image, so as to get an enhanced image or to extract some valuable data from it. It is a kind of flag processing in which information is an image and yield might be image or characteristics/highlights associated with that image. Medical imaging is the representation of body parts, tissues, or organs, for use in clinical conclusion, treatment and ailment observing. Imaging techniques encompass the fields of radiology, nuclear medicine and optical imaging and image-guided mediation, examples Radiology, Nuclear Medicine, Optical Imaging.

Image segmentation alludes to the process of apportioning a digital image into numerous regions. The objective of segmentation is to change the portrayal of an image to be more meaningful and less demanding to break down. It is utilized as a part of a request to locate objects and limits in images. The aftereffect of image segmentation occurs as an arrangement of regions that collectively cover the whole image [1]. In this manner, medical image segmentation assumes a significant part in clinical diagnosis. It can be considered as a difficult issue because medical images commonly have poor contrasts, diverse sorts of noise, and absent or diffusive boundaries [2].

Many approaches have been proposed for image segmentation like boundary based, edge based, cluster based, and neural framework based. From the different technique, a champion among the most efficient procedures is the clustering methodology. Again there are distinctive sorts of clustering: K-means clustering, Fuzzy C-means clustering, mountain clustering methodology and subtractive clustering procedure.

The clustering process discovers meaningful and regular clusters in the datasets. The real obstacle in clustering is that no earlier learning about the given dataset is accessible. There are numerous information clustering calculations accessible in the writing to perform clustering. Among them most broadly utilized categories of the clustering are partitional and hierarchical calculations. The most prevalent class of partitional clustering is FCM clustering calculation.

In any case, FCM calculation relies on upon the underlying seed indicates and converges local optima. FCM clustering calculation is connected to a wide assortment of geostatistical information examination issues [2]. Cannon recommended an efficient execution of FCM clustering calculation. To enhance and tackle the shortcoming of the FCM calculation, a brought together a structure for performing thickness weighted FCM clustering is created. A comparative examination of FCM calculation is performed regarding the nature of clusters and their computational time. They indicate that the significant drawback of FCM calculation is that it generally converges to a local least or a set point [3]. Hence, the researchers have demonstrated their enthusiasm for advancement calculations to overcome the difficulties in FCM clustering technique and to enhance the nature of clusters.

K-means is one of the least difficult unsupervised learning calculations that take care of the outstanding clustering issue. The procedure takes a straightforward and simple approach to classifying a given informational index through a certain number of clusters (accept k clusters) settled from the earlier [4]. The principle thought is to characterize k centroids, one for each cluster. These centroids should be placed cleverly because of various location causes the distinctive outcome. Along these lines, the better choice is to place them however much as could reasonably be expected far from each other. The subsequent stage is to take each direct having a place toward a given informational collection and associate it to the closest centroids [5].

These days image segmentation becomes one of the vital instrument in medical territory where it is utilized to extract or area of enthusiasm from the background. So medical

ISSN: 2321-8169
122 – 127

images are sectioned utilizing diverse technique and process yields are utilized for the further examination in medical. Be that as it may, medical images in their crude frame are spoken to by the varieties of numbers in the computer, with the number indicating the estimations of important physical amounts that show the contrast between various sorts of body parts. Processing and investigation of medical images are valuable in changing crude images into a quantifiable symbolic frame, in extracting meaningful subjective data to help to find and in coordinating complementary information from various imaging modalities [6].

A. Related work:

Some of the existing recent works are discussed here.

Dhanachandra[7] et al. had proposed an image segmentation utilizing K-means Clustering Algorithm and Subtractive Clustering Algorithm, and the median filter is connected to evacuate undesirable areas and then RMSE and PSNR are checked and watched that they have small and large value respective, which are the condition for good image segmentation quality. Furthermore, comparison for RMSE and PSNR are done for proposed method and classical K-mean algorithms manditis found the proposed strategy has better performance result later on.

Sandeep [8] et al. had implemented Fuzzy C means and Mean Shift based segmentation in MATLAB with the help of image processing tool box. Comparative analysis and results has shown that the Fuzzy C means provides more accurate results than Mean Shift algorithm because it contains highest values for PSNR parameter and almost contains lowest values for MSE parameter.

Cebeci [9] et al. Compared a K-means (KM) and the Fuzzy C-means (FCM) algorithms for their computing performance and clustering accuracy on different shaped cluster structures which are regularly and irregularly scattered in two dimensional space. While the accuracy of the KM with single pass was lower than those of the FCM, the KM with multiple starts showed nearly the same clustering accuracy with the FCM. Moreover the KM with multiple starts was extremely superior to the FCM in computing time in all datasets analyzed.

Maksoud [10] et al. presents an efficient image segmentation approach utilizing K-means clustering technique integrated with Fuzzy C-means algorithm. It is followed by thresholding and level set segmentation stages to provide accurate brain tumor detection. The proposed technique can get benefits of the K-means clustering for image segmentation in the aspects of insignificant computation time. In addition, it can get favorable circumstances of the Fuzzy C-means in the aspects of accuracy. The performance of the proposed image segmentation approach was assessed by comparing it with some state of the art segmentation algorithms in case of accuracy, processing time, and performance.

Patil [11] et al. presented a detailed study and comparison of different clustering based image segmentation algorithms. The traditional clustering algorithms are the hard clustering algorithm and the soft clustering algorithm and compared the hard k-means algorithm with the soft fuzzy c- means (FCM) algorithm. To overcome the limitations of conventional FCM we have also studied Kernel fuzzy c- means (KFCM) algorithm in detail. The K-means algorithm is sensitive to noise and outliers so, an extension of K-means called as Fuzzy c- means (FCM) are introduced.

II. K-MEANS CLUSTERING:

K-means clustering algorithm given by Macqueen is one of the simplest unsupervised learning algorithm that solve the well known clustering problem.

The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a prior.

The idea is to define k centroids, one for each cluster. These centroids should be placed in a cunning way because of different locations causes different results. So, the better choice is to place them as much as possible for away from each other.

The next step is to take each point belonging to a given data set and associate it to the nearest centroids. When no point is pending, the first step is completed and an early group age is done. At this point need to recalculate k new centroids as bary centers of the cluster resulting from the previous steps.

After have these k new centroids, a new binding has to be done between the same data set points and the nearest new centroids. A loop has been generated. As a result of this loop. K-centroids change locations step by step until no more changes are done.

Algorithm:

Step1: Input k, set of points x_1 ... x_n

('k' as a input, 'k' means need to tell it how many cluster you find?)

Step2: Place centroids c_1 ... c_k at random locations in a space.

Step3: Repeat until convergence:

- For each point X_i:
- Find nearest centroid C_j :

(Distance (Euclidean) between Individuals X_i and cluster C_j .)

- Assign the point X_i to cluster j.

(For each cluster j = 1 ….. k)

- New centroids C_j = mean of all points X_i assigned to cluster j in previous step:

$$C_j(a) = \frac{1}{n_{jx_{j-nc_j}}} X_i(a)\ ;\ \text{for } a = 1..d,..(1)$$

Step4: Stop when none of the cluster assignments change.

III. FUZZY C-MEANS ALGORITHM:

The most well known fuzzy clustering algorithm is FCM, a modification by ‘Bendex’ of an original CRISP clustering methodology.

‘Bendex’ introduced the idea of a fuzzification parameters (M) in the range [1, N], which determines the degree of fuzziness in the clusters. When M=1 the effect is a crisp clustering of points. When M>1 is the degree of fuzziness among points in the decision space increases.

Algorithm:

Step1: Randomly initializing the cluster center.

Step2: Creating distance matrix from a point X_i to each of the cluster centers to with taking the Euclidean distance between the point and the cluster center.

$$dj_1 = \sqrt{\sum (X_j - C_j)^2} \quad ....(2)$$

Step3: Creating membership matrix takes the fractional distance from the point to the cluster center and makes this a fuzzy measurement by raising the fraction to the inverse fuzzification parameter. This is divided by the sum of all fractional distances, thereby ensuring that the sum of all memberships is 1.

$$\mu_j(X_1) = \frac{\left(\frac{1}{d_1}\right)^{\frac{1}{m-1}}}{\sum_{K-1}^{P}\left(\frac{1}{dj_1}\right)^{\frac{1}{m-1}}} \quad ...(3)$$

Step4: Creating membership matrix FCM imposes a direct constraint of the fuzzy membership function associated with each point, as follows. The total membership for a point in sample or decision space must add to 1.

$$\sum_{J=1}^{P} \mu_j(X_i) = 1 \quad ...(4)$$

Step5: Generating new centroid for each cluster.

$$C_j = \sum\nolimits_1 [\mu_j(X_1)]^m \times \frac{1}{\sum_1[\mu_j(X_1)]^m} \quad ...(5)$$

Step6: Generating new centroid for each cluster with iteration all this step optimize cluster centers will generate.

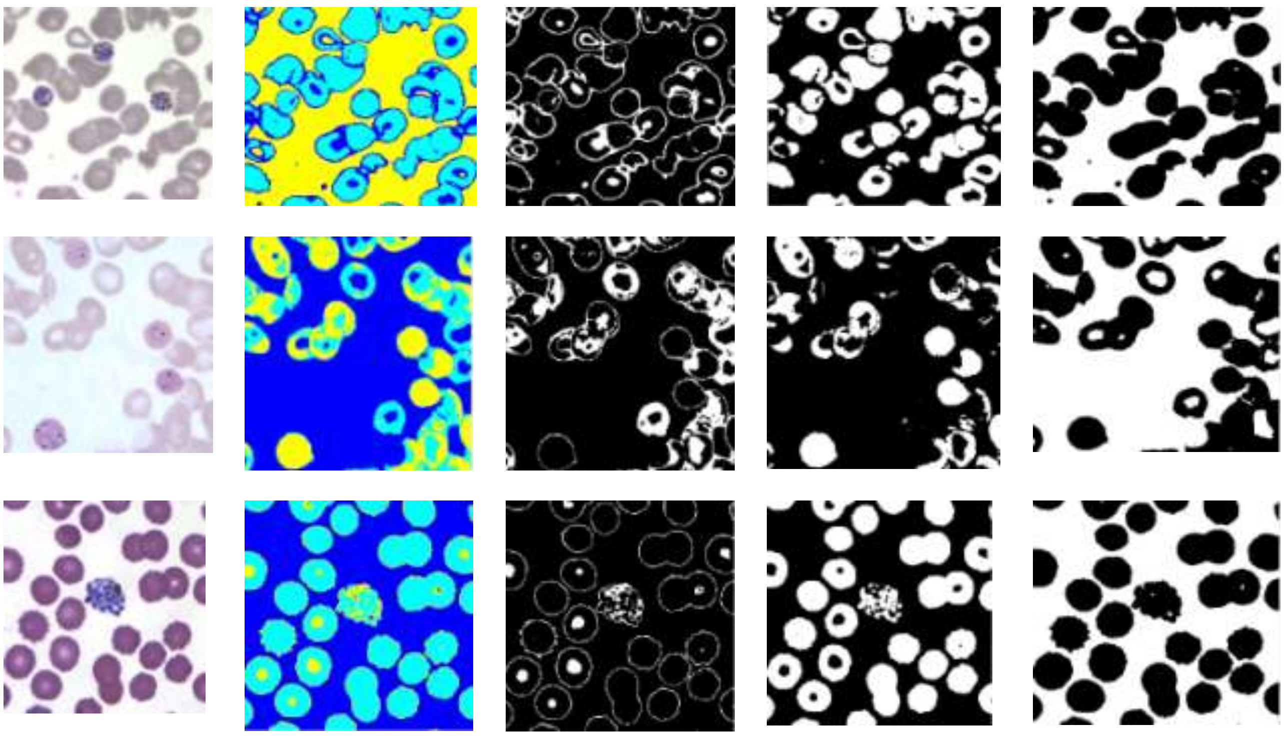
Step7: Weight acceleration cluster assignments.

IV. RESULTS AND DISCUSSIONS:

In a Result Section Firstly takes a infected blood samples like malaria blood samples for the analysis and for estimate the number of infections in blood using hybrid segmentation. These samples datasets contains the blood samples have been taken from the CDC (center of disease control and prevent)-DPDx (Laboratory Identification of Parasitic Diseases of Public Health Concern) database were sample from 20 different patients are collected. The database contains around 500 blood samples images.

Table1: Segmented images using K-means algorithms.

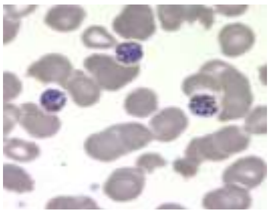
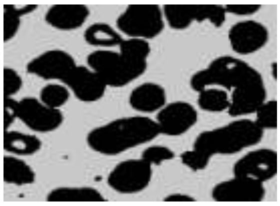
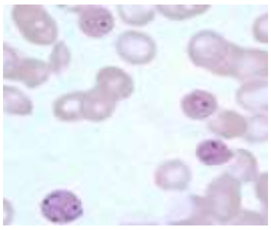
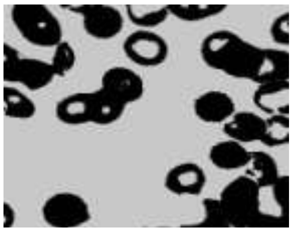
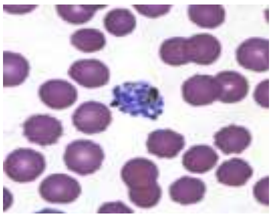
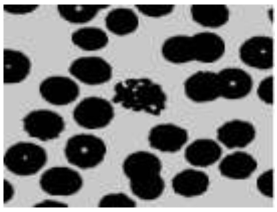
Sample | Original image | All clusters | first cluster | Second cluster | Third cluster



1st
2nd
3rd

Both the clustering algorithms using with edge detection technique shows very effective results when individually applied on database images. From the results it can be seen Fuzzy C-means algorithm matches greater numbers of clusters as compared to K-means algorithm.

Table2: Segmentation using Fuzzy C-means algorithm

Samples	Iteration	Original Image	Final Segmented Image (FCM)
First blood sample	7		
Second blood sample	9		
Third blood sample	5		

A. *Performance Parameters:*

The quality of the segmented image is analyzed using the measurement value of Mean Square Error, Root Mean Square Error and Peak to Signal Noise Ratio.

1. Mean square error (MSE):

The MSE is a measure of the nature of an estimator or a predictor (i.e., a function mapping subjective contributions to a specimen of estimations of some arbitrary variable). it is dependably non-negative, and qualities closer to zero is better. For an image of size M x N the mean square error (MSE) is defined as:

$$\mathrm{MSE} = \frac{1}{MN}\sum_{i=o}^{M-1}\sum_{j=o}^{N-1}[x(i,j) - y(i,j)]^2 \quad (6)$$

where x(i, j) is noise-free grayscale image and y(i, j) is a noisy approximation of x(i, j).

2. Peak to Signal Noise Ratio (PSNR):

The peak to signal noise ratio is the proportion between most extreme feasible forces and the corrupting noise that influence resemblance of the image. It is utilized to measure the quality of the output image.

$$\mathrm{PSNR} = 10.\log_{10}\left(\frac{MAX_I^2}{MSE}\right) = 20\log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right) \quad (7)$$

Here, MAX_I is the maximum possible pixel value of the image. When samples are represented using linear PCM with B bits per sample, MAX_I is $2^B - 1$.

Table 3: PSNR values of K-means compared with the FCM.

Samples	K-means clustering algorithm	Fuzzy C-means algorithm
	PSNR	PSNR
First blood Sample	34.99	34.74
Second Blood Sample	35.63	35.47
Third Blood Sample	34.86	34.61

The table3 displays the outcomes of evaluation of K-means and Fuzzy C means based segmentation methods on the centre of Peak Signal to Noise Resolution. The PSNR calculates the peak signal to noise ratio, in decibels, among binary images. This proportion is repeatedly castoff as a superiority capacity among the original and a resultant image. The higher the PSNR, the better is the quality of the output image.

ISSN: 2321-8169
122 – 127

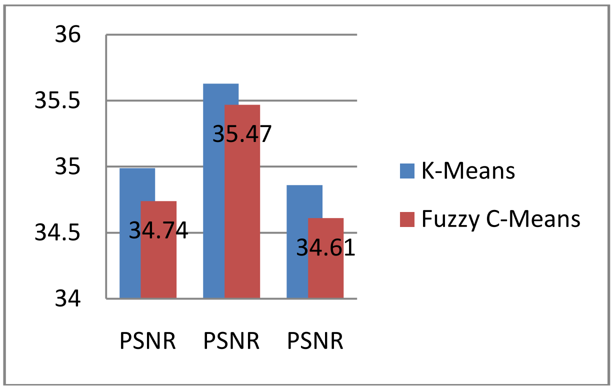


Figure1: Outcomes of evaluation of K-M & FCM based on PSNR.

Table 4: represents the accuracy of FCM and K-Means clustering.

Algorithm	**Accuracy**
FCM	86%
K-Means	78.26%

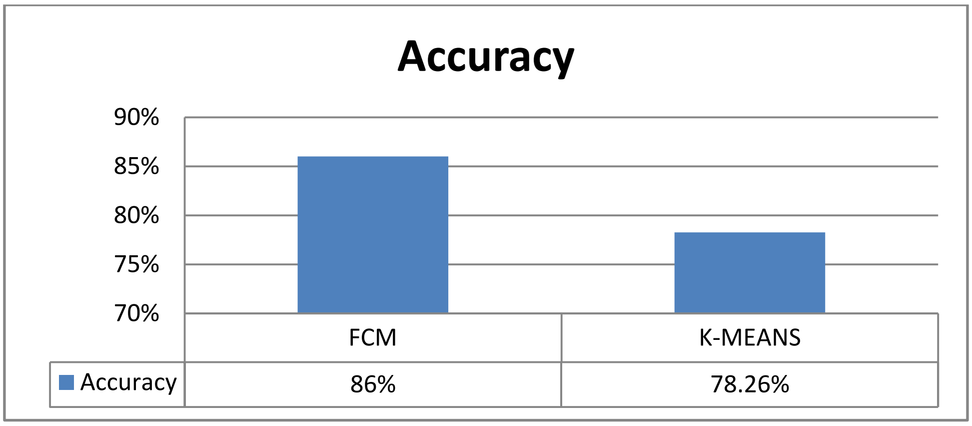


Figure2: accuracy of FCM and K-Means

V. Conclusion:

Fuzzy C-means is slower than K-means in efficiency and lesser PSNR but gives better results in case where data is incomplete or uncertain and has a wider applicability. However the time taken by K-Means algorithm is less as compared to fuzzy C-Means algorithm but after testing on two different datasets which included 400 blood samples images, Fuzzy C-Means clustering shows 86% accuracy and K-means shows 78.26%. As a final conclusion, there is no any algorithm which is the best for all cases. An important factor in choosing an appropriate clustering algorithm is the shape of clusters in datasets to be analyzed. Thus, the datasets should be carefully examined for shapes and scatter of clusters in order to decide for a suitable algorithm. To achieve this, 2D and/or 3D scatter plots of datasets provide good idea to understand the structure of clusters in datasets. When multi-featured objects are analyzed, in order to overcome to plot for multidimensional space, a dimension reduction technique such as multidimensional scaling (MDS) or principal components analysis (PCA) can be applied to reduce dimensions of datasets. Moreover, by using a suitable sampling method this process can be completed in shorter execution times.

In future, medical image segmentation can perform more accurate segmentation results by using both the clustering techniques such as K-means & FCM or by using other classification measures. Combination of different Morphological operations can also be used to gain better Final segmented image.

References:

[1] R. Kaur, N. K. Mavi, and R. Singh, “Comparison of Clustering Techniques for the Detection of Microcalcifications in Mammograms,”International Journal of Advanced Research in Computer Science and Software Engineering, vol. 6, no. 8, pp. 99–104, 2016.

[2] A. Al Malki et al., “Hybrid Genetic Algorithm with K-Means for Clustering Problems,” Open J. Optim., vol. 5, no. 02, p. 71, 2016.

[3] E. Abdel-maksoud, M. Elmogy, and R. Al-awadi, “Brain tumor segmentation based on a hybrid clustering technique,” Egypt. Informatics J., vol. 16, no. 1, pp. 71–81, 2015.

[4] S. Kaur and S. Chawla, “PERFORMANCE EVALUATION OF IMPROVED IMAGE SEGMENTATION USING FUZZY BASED MODIFIED MEAN SHIFT AND MST,” International Journal of Emerging Trends & Technology in Computer Science, vol. 4, no. 4, 2015.

[5] S. A. Christe, K. Malathy, and A. Kandaswamy, “IMPROVED HYBRID SEGMENTATION OF BRAIN MRI

TISSUE AND TUMOR USING STATISTICAL FEATURES," vol. 9102, no. 2, 2010.

[6] H. Singh and J. S. Gill, "Enhancement Methods and Techniques for Brain MR Imaging," vol. 5, no. 2, pp. 38–44, 2017.

[7] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image Segmentation using K-means Clustering Algorithm and Subtractive Clustering Algorithm," Procedia - Procedia Comput. Sci., vol. 54, pp. 764–771, 2015.

[8] S. Kaur, "Evaluation of Performance of Fuzzy C Means and Mean Shift based Segmentation for Multi-Spectral Images," International Journal of Computer Applications, vol. 120, no. 16, pp. 25–28, 2015.

[9] Z. Cebeci and F. Yildiz, "Comparison of K-Means and Fuzzy C-Means Algorithms on Different Cluster Structures,"Journal of AI, vol. 6, no. 3, pp. 13–23, 2015.

[10] E. Abdel-maksoud, M. Elmogy, and R. Al-awadi, "Brain tumor segmentation based on a hybrid clustering technique," Egypt. Informatics J., vol. 16, no. 1, pp. 71–81, 2015.

[11] A. J. Patil, C. S. Patil, R. R. Karhe, and M. A. Aher, "Comparative Study of Different Clustering Algorithms," International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 3, Issue 7, July 2014.

[12] R. D. Pai, P. S. Halvi, and P. B. Hiremath, "Medical Color Image Enhancement using Wavelet Transform and Contrast Stretching Technique," Int. J. Sci. Res. Publ., vol. 5, no. 7, pp. 1–7, 2015.

[13] J. Agarwal and S. S. Bedi, "Implementation of hybrid image fusion technique for feature enhancement in medical diagnosis," Springer Hum. Centric Comput. Inf. Sci., vol. 5, no. 1, pp. 1–17, 2015.

[14] D. M. F. Al-Samaraie, "Medical Colored Image Enhancement Using Wavelet Transform Followed By Image Sharpening," Ubiquitous Comput. Commun. J., vol. 6, no. 5, pp. 1–8, 2014.