Qualitative Evaluation of Fuzzy Clustering Methods in Segmentation of Fundoscope Diabetic Retinopathy Images

R.Ravindraiah,

Assistant Professor, ECE Dept., MITS, Madanapalle,Chittor dt, Andhra pradesh
Email: ravindra.ranga@gmail.com

Dr.S.Chandra Mohan Reddy,

Professor & Head of ECE Dept.,JNTU College of Engineering, Pulivendula,
Kadapa dt, Andhra Pradesh,
Email:cmr.ece@jntua.ac.in

Abstract:

At present there is an increased interest in Therapeutic image processing in most fields of engineering. Imaging modality provides detailed information about anatomy. It is also helpful in the finding of the disease and its progressive treatment. The Prime signs of Diabetic Retinopathy (DR) are Exudates which leads to severe vision loss in chronic condition. Exudates are the remnant of exuded blood and protein based particles from the damaged blood vessels of retina. Laser healing requires accurate location of exudates for faithful removal through Laser burns. Segmentation of fundoscope image will help the ophthalmologist in diagnosis, classification and to determine the severity. Multiple methods are designed and developed for Medical Image segmentation based on Thresholding, Region Growing, Markov Random Model, Clustering, Deformable Model, Classifier, Neural Networks, Expectation Maximization and Support vector machines etc. Out of these the fuzzy clustering methods are less complex and are robust in operation. This paper aims in performance evaluation of Fuzzy C means clustering (FCM) algorithm, Kernel induced FCM (KFCM) and Spatial FCM (SFCM) algorithms are done.

Keywords: *Fuzzy C means Clustering (FCM), Kernel based Fuzzy C Means (KFCM), Spatial Fuzzy C means(SFCM), Image segmentation, Diabetic Retinopathy.*

I. Introduction

Diabetes is a chronic condition which requires lifelong therapeutic care and edification of self-management by the patient to put off initial complications and to lessen the threat of continuing problems. Beyond glycemic control, its care is quite complex as it poses the root cause of many issues.

Diabetes is classified into 4 medical classes:

- Type 1 diabetes (outcome of islet cell damage of pancreas, results in absolute insulin deficiency)
- Type 2 diabetes (results in lack of insulin secretion)
- Other explicit causes includes hereditary issues in – functioning of organic cells , hereditary causes in action of insulin, issues with exocrine pancreas (like cystic fibrosis), and chemical actions during treatment of certain complications like AIDS or after organ replacement
- Gestational diabetes mellitus (GDM) (found in pregnant women)

The patients with diabetes will have common symptoms like increased urination (polyuria), thirst (polydipsia), hunger (polyphagia) and weight loss. Vision blurriness, peripheral neuropathy, repeated vaginal infections, and tiredness are analyzed in clinical diagnosis.

Out of the above classification DR is more prevalent in type 2 diabetes. DR is a results mainly due vascular changes in the retina. Exudates are one of the primary sign of Diabetic Retinopathy [1][10. Retina is a light sensitive muscular layer which is nourished by a network of blood vessels. Any vascular change results in difficulties of light perception. Depending upon these vascular changes in the eye fundus, DR is classified into nonproliferative and proliferative. Proliferative refers to uncertainty of neovascularization (abnormal blood vessel growth) in the retina. DR with no neovascularization is called nonproliferative diabetic retinopathy (NPDR). As the pathos progresses into proliferative diabetic retinopathy (PDR) [Presence of neovascularization] serious visual consequences are resulted. The vascular changes include weak and fragile blood vessels that exude blood, and protein particles into the eye fundus [Exudates]. Before they actually break they got swollen and ballooned out. These are called micro-aneurysms. In time these micro-aneurysms burst and leads to accumulation of Exudates. Also the blood vessels may block making the retina cells suffer from nourishment. Hence new blood vessels are stimulated which are abnormal, weak, transparent and probably ruptured. This makes situation chronic. It would be useful to have an automated method of detecting exudates in digital retinal images produced from DR screening programs [2].Early detection and diagnosis of DR help the patient in preventing severe vision loss. Retinal images obtained by the fundus camera are used to diagnose DR. Automated methods of DR screening help to save time, cost and vision of patients, compared to the manual methods of diagnosis [3-4].

Image segmentation is defined as a method of sub dividing an image into group of pels that have homogeneous characteristics with respect to various principles. These are area oriented instead of pixel oriented. Segmentation splits an image into meaningful connected regions. These methods are classified into two: (i) Local segmentation, and (ii) global segmentation.

Local segmentation deals with segmenting sub images which are small windows on a whole image. The number of pixels available to local segmentation is much lower than global segmentation. Local segmentation must be frugal in its demands for pixel data. Global segmentation is concerned with segmenting a whole image. Global segmentation deals mostly with segments consisting of a relatively large number of pixels. This makes estimated parameter values for global segments more robust. Image segmentation can be approached from three different philosophical perspectives. They are (i) region approach, (ii) boundary approach, and (iii) edge approach.

Clustering method tries to attain the common features/characteristics of the pixels in different data sets and organizes the patterns into separate groups based on specific similarities. In the clustering technique, an attempt is made to extract a feature vector from local areas in the image. A standard procedure for clustering is to assign each pixel to the class of nearest cluster mean [5]. Most clustering algorithms do not depend on expectations common to traditional numerical methods, such as the basic statistical distribution of data, and therefore they are useful in situations where little prior knowledge exists. The strength of clustering procedures expose the principal structures in data can be made use in wide-ranging applications, including image processing, pattern recognition, classification, identification and modeling.

Clustering methods can be classified into two - **Hierarchical and Partitional clustering**. Hierarchical clustering methods make use of proximity matrix to determine the resemblance of data points needed to be grouped. They produce cluster trees which signifies the nested set of patterns and likeness levels at which grouping varies. In these methods the resulting clusters are always created as the inner nodes of the tree, while the root node is kept reserved for the whole dataset and leaf nodes are for individual data samples. The clustering procedures vary with respect to the rules by which two small clusters are fused or a large cluster is fragmented. The two key types of procedures used in the hierarchical clustering framework are agglomerative and divisive. Partition

based clustering uses an iterative optimization procedure that aims at minimizing an objective function, which measures the goodness of clustering. Partition based clustering are composed of two learning steps— the partitioning of each pattern to its closest cluster and the computation of the cluster centroids [5]. A common characteristic of the partition based clustering is that the clustering process starts with an initial assumption of known number of clusters. The clustering centroids are typically calculated based on the optimality condition i.e., minimization of objective function. Partitioning procedures are characterized into Partitioning relocation methods and density based partitioning. Partitional clustering methods e.g., K-means clustering and fuzzy clustering are more advantageous than hierarchical clustering methods, where a portion of the data points optimizes some criterion functions.

In this paper a comparative analysis is done on three fuzzy based clustering methods. These includes Fuzzy C Means clustering (FCM) proposed by Dunn J C[6], and Kernel induced FCM proposed by D.Zhang [8] and Spatial FCM proposed by Keh-Shih Chaung et.at., [9]

II. Fuzzy based Image Segmentation methods

Clustering methods are classified as crisp (hard) and fuzzy. In Crisp clustering data is fragmented mutually exclusive subgroups. Whereas Fuzzy methods gives significance on the uncertainty of data points as they poses variable amount of belongingness to multiple clusters concurrently. This belongingness is termed as membership degree whose value varies between 0 and 1. In many applications, fuzzy clustering is more natural and quite beneficiary than crisp methods. Even the data elements on the corners are not forced to merge in any one of the classes, but rather assigned with membership values prior to clustering.
The discrete nature of the hard partitioning also causes difficulties with algorithms based on analytic functionals, since these functionals are not differentiable [5]

II(a) Fuzzy C Means algorithm [FCM]

Joe Dunn [6] reported FCM method in 1974 and improved by Bezdek [7] in 1981. FCM aims at Minimizing the Objective Function J_{FCM} defined as

$$J_{FCM} = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{m} d_{ij}^{2} \qquad [1]$$

where u_{ij} is the Membership function which represents the degree of belongingness of a pixel to a particular Cluster or Region and d_{ij} is the Euclidean distance metric.

Algorithm:
Let X=(x1, x2, x3----------- xn) be the set of data points.
C= (c1, c2, c3----------- c3) be the set of centers

(1) Receive the input image in the form of data matrix, X
(2) Fix the number of clusters "C".
(3) Calculate the fuzzy membership by using

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\dfrac{d_{ij}}{d_{kj}} \right)^{\frac{2}{m-1}}} \qquad [2]$$

(4) Compute the fuzzy centers using

$$c_j = \frac{\sum_{i=1}^{n} u_{ij}^m x_i}{\sum_{i=1}^{n} u_{ij}^m} \qquad [3]$$

(5) The iteration will stop when

$$\| U^{(k+1)} - U^{(k)} \| < \epsilon \qquad [4]$$

Based upon the distance metric between the data point and the cluster center the algorithm assigns membership to each data point. Membership value varies inversely with the distance value between them. Clearly, summation of membership of each data point should be equal to one. Cluster centers and Membership are updated according to the formula 2 and 3 after each iteration. The benefit of FCM is its converges speed but suffers with some limits like long computational time, sensitivity to the initial guess (speed, local minima), less immune to noise and one expects low (or even no) membership degree for outliers. For the detailed information of FCM refer [6], [7]

II(b) Kernel Fuzzy C Means algorithm [KFCM]

This is an improved version of FCM in which the basic Euclidian metric in the FCM is modified by inducing a Gaussian kernel. This will improve the clustering efficiency and reduce the outliers' problem present in FCM.

The objective function of KFCM is defined as

$$J_{KFCM} = 2 \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^m \left(1 - k\left(x_j, c_i\right)\right) \qquad [5]$$

Algorithm:

Let X=(x1, x2, x3---------- xn) be the set of data points.
   C= (c1, c2, c3----------- c3) be the set of centers

(1) Receive the input image in the form of data matrix, X
(2) Fix the number of clusters "C".
(3) Calculate the fuzzy membership by using

$$u_{ij} = \frac{\frac{1}{\left(1 - k\left(x_k, c_i\right)\right)^{\frac{1}{(m-1)}}}}{\sum_{i=1}^{c} \left(\frac{1}{\left(1 - k\left(x_k, c_j\right)\right)^{\frac{1}{(m-1)}}}\right)} \qquad [6]$$

(4) Compute the fuzzy centers using

$$c_j = \frac{\sum_{i=1}^{n} u_{ij}^m \left(k\left(x_k, c_i\right)\right) x_j}{\sum_{i=1}^{n} u_{ij}^m \left(k\left(x_k, c_j\right)\right)} \qquad [7]$$

(5) The iteration will stop when

$$\| U^{(k+1)} - U^{(k)} \| < \epsilon \qquad [8]$$

A KFCM algorithm has better noise immunity and reduce outlier's effect present in FCM but has limitations of computational complexity and less performance to large data sets. The detailed information of KFCM is provided in reference [8].

II(c) Spatial Fuzzy C Means algorithm [SFCM]

The earlier versions of Fuzzy clustering methods do not incorporate any spatial information. With the inclusion of spatial information i.e., information of the neighborhood pixels around the Exudates help in achieving optimum Segmentation results.

The objective function of the SFCM is defined as

$$J_{SFCM} = \sum_{i=1}^{c}\sum_{j=1}^{n}u_{ij}^{m}d_{ij}^{2} + \sum_{i=1}^{c}\sum_{j=1}^{n}h_{ij} \quad . \qquad [9]$$

SFCM algorithm has two steps
   1. The first step membership function calculation is done which is identical to standard FCM
   2. In the second step, the mapping of pixel membership information to the spatial domain is done .

A spatial function is computed from the second step.

$$h_{ij} = \sum_{k \in NBx_{j}} u_{ik} \qquad [10]$$

where NB(x$_j$) is a square window centered on pixel x$_j$. A 3*3 square window was used throughout this method; h$_{ij}$ is the probability that pixel x$_j$ belongs to ith cluster.

The Membership function is updated as

$$u_{ij}^{'} = \frac{u_{ij}^{m} h_{ij}}{\sum_{k=1}^{c} u_{ij}^{m} h_{ij}} \qquad [11]$$

Thus the neighbor pixel characteristics are incorporated into the membership function as in equation [11]. The complete information of SFCM is providing in reference [9].

Statistical Measure

The following statistical measures are useful to compare image segmentation methods.

$$Sensitivity = \frac{TP}{TP+FN} \qquad [12]$$

$$Specifity = \frac{TN}{TN+FP} \qquad [13]$$

$$Accuracy = \frac{TP+TN}{TP+FN+FP+FN} \qquad [14]$$

TP=Number of exudates pixels correctly detected,
TN=Number of non-exudates pixels that are correctly detected as non-exudates,
FP= Number of non-exudates pixels that are wrongly detected as exudates,
FN=Number of exudates pixels that are not detected as exudates.

Peak signal-to-noise ratio [PSNR], is defined as the ratio between the maximum possible power of a signal (here the Image) and the power of corrupting noise that affects the fidelity of its representation. PSNR is usually expressed in terms of the logarithmic decibel scale and is given by

$$PSNR = 10\log_{10}[MAX^{2}/MSE] , \qquad [15]$$

where **MAX** is the maximum possible pixel value in an image and MSE is the Mean Square error between the input and processed images, given by the following equation

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^{2}, \qquad [16]$$

where I(i,j) is the input image and K(i,j) is the processed image. m*n refers to size of the image.

III. Results

The figures below show Original Fundoscope images and the results of FCM, KFCM and SFCM methods.

Figure 1(a)   Figure 1(b)   Figure 1(c)   Figure 1(d)

Figure 2(a)   Figure 2(b)   Figure 2(c)   Figure 2(d)

Figure 3(a)   Figure 3(b)   Figure 3(c)   Figure 3(d)

Figure 4(a)   Figure 4(b)   Figure 4(c)   Figure 4(d)
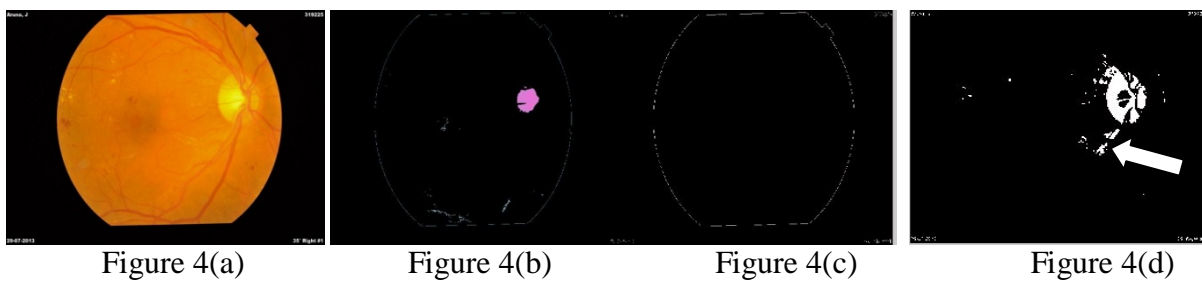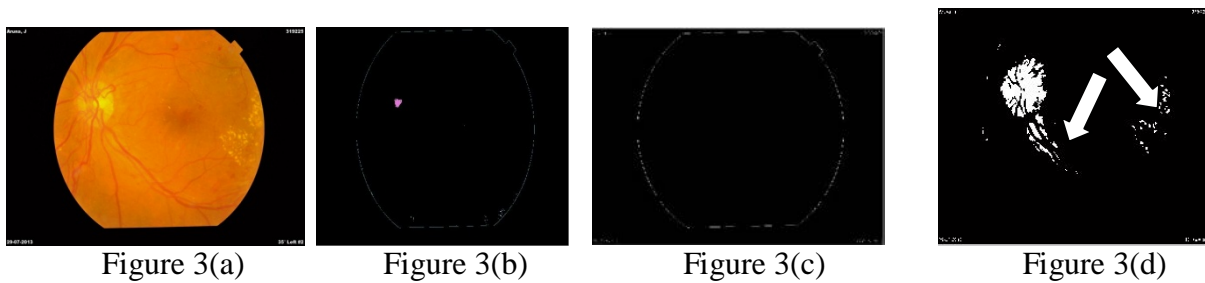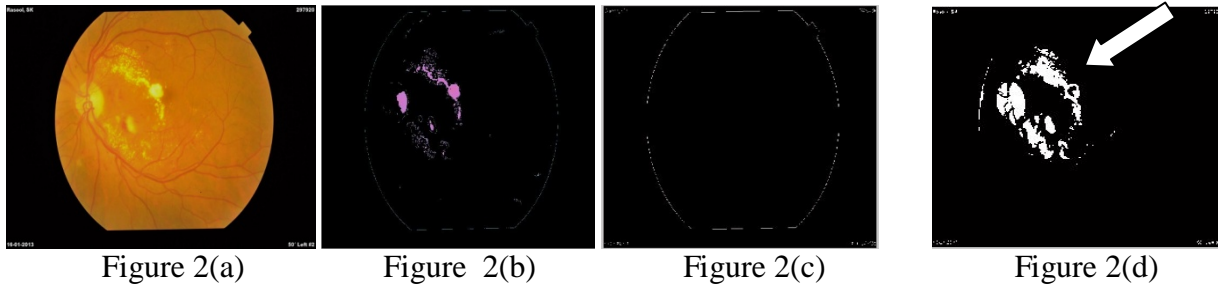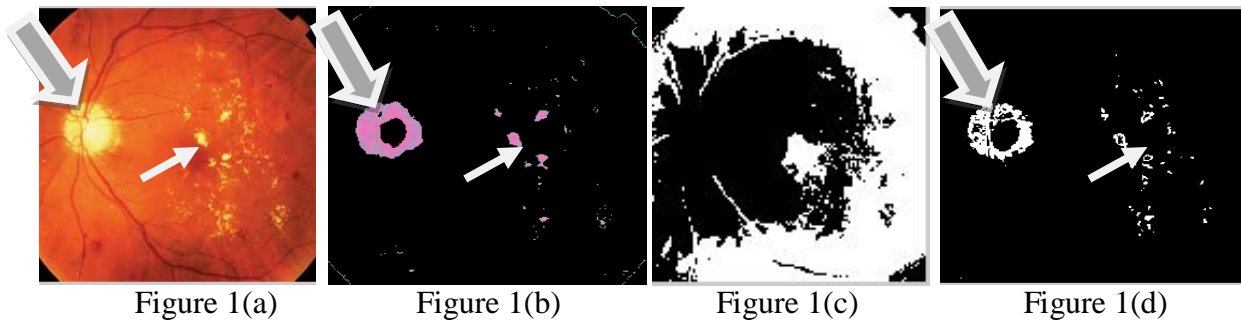
Exudates in Retina Images

Optic Disc

Figure 1,2,3,4(a) Original image; Figure 1,2,3,4(b)-FCM Segmentation results ; Figure 1,2,3,4(c)-KFCM Segmentation results; Figure 1,2,3,4(d)-SFCM Segmentation results;

Table III.1: Performance measures: Statistical analysis of FCM, KFCM, SFCM

Original image	FCM					KFCM					SFCM				
	Se	Sp	Acc	MSE	PSNR	Se	Sp	Acc	MSE	PSNR	Se	Sp	Acc	MSE	PSNR
1	69.39	38.95	51.92	4.3106e+005	8.214	73.87	39.99	54.30	6.5129e+05	10.00	65.13	52.78	58.63	6082.24	34.3556
2	86.59	50.28	65.99	2.8602e+005	6.433	76.93	35.99	52.62	2.2189e+06	15.33	75.68	73.64	74.65	567.771	44.6545
3	75.80	82.29	58.98	5.3869e+005	9.182	52.48	73.40	62.06	1.1604e+07	22.51	53.13	64.17	78.39	240.544	48.3843
4	72.63	86.51	59.26	4.5359e+005	8.435		67.48	58.65	1.1865e+07	2.61	52.30	66.30	78.89	301.281	47.4065

**Se-Sentivity, Sp- Specificity, Acc-Accuracy

Conclusion

From the above results it is quite evident that the results of SFCM are quite optimum compared to FCM and KFCM. The statistical results show that SFCM has highest Accuracy and PSNR compared to the other two clustering methods. Further the robustness of the SFCM towards noise can be improved with the implementation of edge preserving filtration at the front end to it. Even the convergence characteristics of these methods can be improved by weighted distance metric.

References

1. Osareh, M. Mirmehdi, B. Thomas and R. Markham. 2003 Automated Identification of Diabetic Retinal Exudates in Digital Colour Images. British Journal of Ophthalmology, 87(10), 1220-1223.
2. Akara Sopharak , Bunyarit Uyyanonvara and Sarah Barman "Automatic Exudate Detection from Non-dilated Diabetic Retinopathy Retinal Images Using Fuzzy C-means Clustering‖.Journal of sensors/2009. ISSN 1424-8220.www.mdpi.com/journal/sensors
3. C.I. Sanchez, R. Hornero, M.I. Lopez, J. Poza. 2004 Retinal Image Analysis to Detect and Quantify Lesions Associated with Diabetic Retinopathy. Proc. 26th IEEE Annual International Conf. on Engineering in Medicine and Biology Society (EMBC), 3, 1624 – 1627.
4. Selvathi D, N.B.Prakash, and Neethi Balagopal. 2012 Automated Detection of Diabetic Retinopathy for Early Diagnosis using Feature Extraction and Support Vector Machine. International Journal of Emerging Technology and Advanced Engineering (ISSN 2250-2459, Volume 2, Issue 11.

5. "Digital Image processing" by S Jayaraman, S.Esakkirajan and T. Veera Kumar, published by McGraw Hill Education (India) pvt ltd, 2013.
6. Dunn J C, J.Cybernet "A fuzzy relative of the ISODTATA process and its use in detecting compact well separated clusters", pp 32-57, March, 1973
7. Bezdek J C "Pattern Recognition with Fuzzy Objective Function Algorithms", New York: Plenum Press, 1981.
8. Chen, S.C., Zhang, D.Q.: Robust image segmentation using FCM with spatial constrains based on new kernel-induced distance measure. IEEE Trans. Systems Man Cybernet.-Part B 34, 1907–1916 (2004)
9. Keh-Shih Chaung, Hong-Long Tzeng "Fuzzy C Means clustering with spatial information for image segmentation," Computerized medical imaging and graphics, vol. 30, pp. 9-15, 2006
10. J.Ramya, S.Soundarya, "Detection of exudates in color fundus image," International journal of Innovative research in science, engineering and technology, vol. 3, issue 3, Mar 2014, pp. 10659-10665.