Volume 2, No. 2, Mar-Apr 2011

ISSN No. 0976-5697

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Investigation on Dermoscopic Image Segmentation using Fuzzy Clustering Techniques

R.Sowmya Devi*
P.G. Student,
Department of Electrical and Electronics Engineering,
NIU, Kumaracoil.
sow_tulip@yahoo.com

Dr. L.Padma Suresh
HOD,
Department of Electrical and Electronics Engineering,
NIU, Kumaracoil.
suresh_lps@yahoo.co.in

De K.L.Shunmuganathan
Professor,
Department of Computer Science Engineering,
RMK Engineering College.
kls_nathan@yahoo.com

Abstract: Medical image segmentation is the most essential and crucial process in order to facilitate the characterization and visualization of the structure of interest in medical images. This paper explains the task of segmenting skin lesions in Dermoscopy images using various Fuzzy clustering techniques for the early diagnosis of Malignant Melanoma. The various Fuzzy clustering techniques used are Fuzzy C Means Algorithm (FCM), Possibilistic C Means Algorithm and Hierarchical C Means Algorithm. The segmented images are compared with the ground truth image using various parameters such as False Positive Error (FPE), False Negative Error (FNE) Coefficient of similarity, Spatial overlap and their performance is evaluated.

Keywords: Fuzzy C Means clustering, Possibilistic C clustering, Hierarchical C Means, False Positive Error, False Negative Error, Coefficient of similarity, Spatial overlap

I. INTRODUCTION

Malignant melanoma is the most frequent type of skin cancer and its incidence has been rapidly increasing over the last few decades. Nevertheless, it is also the most treatable kind of skin cancer, if diagnosed at an early stage. The clinical diagnosis of melanoma is commonly based on the ABCD rule [3], an analysis of four parameters (asymmetry, border irregularity, color, and dimension), or the 7-points checklist which is a scoring method for a set of different characteristics depending on color, shape, and texture.

Dermoscopy is a non-invasive diagnosis technique for the in vivo observation of pigmented skin lesions used in dermatology .Dermoscopic images have great potential in the early diagnosis of malignant melanoma, but their interpretation is time consuming and subjective, even for trained dermatologists. Therefore, there is currently a great interest in the development of computer-aided diagnosis systems that can assist the clinical evaluation of dermatologists. The standard approach in automatic dermoscopic image analysis has usually three stages: 1) image segmentation; 2) feature extraction and feature selection; and 3) lesion classification. The segmentation stage is one of the most important since it affects the accuracy of the subsequent steps. However, segmentation is difficult because of the great variety of lesion shapes, sizes, and colors along with different skin types and textures. In addition, some lesions have irregular boundaries and in some cases there is a smooth transition between the lesion and the skin. Other difficulties are related to the presence of dark hair covering the lesions and the existence of specular reflections. Some of these difficulties are illustrated below.

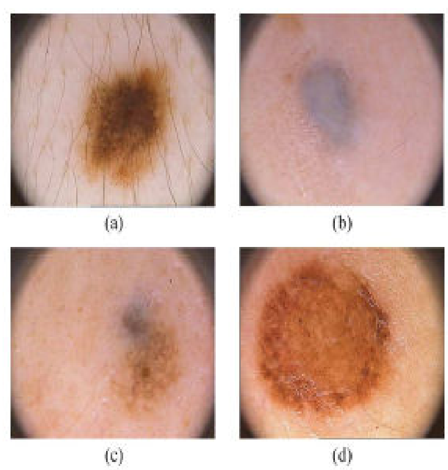

Figure.1 Difficulties of dermoscopic images; (a) presence of hair (b) smooth transition between lesion and skin; (c) multiple colored lesions; and(d) specular reflections.

To address this problem, several algorithms have been proposed. They can be broadly classified as thresholding, edge-based or region-based methods. An example of thresholding can be found in [2]. Thresholding methods achieve good results when there is a good contrast between the lesion and the skin. Edge-based approaches were used where the segmentation is based on the zero-crossings of the Laplacian-of-Gaussian. Edge-based approaches [7] perform poorly when the boundaries are not well defined, for instance when the transition between the skin and the lesion is smooth. In these situations, the edges have gaps and the contour may leak through them. Another difficulty is the presence of spurious edge points that do not belong to the lesion boundary. They are the result of artifacts such as hair, specular reflections or even irregularities in the skin texture and they may stop the contour preventing it to converge to the lesion boundary. Region-based approaches have difficulties when the lesion or the skin region are textured or have different colors presents which lead to oversegmentation.

In this paper we propose and evaluate several Fuzzy based clustering techniques: Fuzzy C Means Algorithm (FCM), Possibilistic C Means Algorithm, and Hierarchical C Means Algorithm. These algorithms are applied to the dermoscopic image and are compared with the expected lesion segmentation (ground truth). The evaluation is based on different parameters and quality metrices that take into account different types of error.

II. FUZZY CLUSTERING TECHINQUES

Cluster analysis is a technique for classifying data, i.e., to divide a given dataset into a set of classes or clusters. The goal is to divide the dataset in such a way that two cases from the same cluster are as similar as possible and two cases from different clusters are as dissimilar as possible. Thus one tries to model the human ability to group similar objects or cases into classes and categories. In classical cluster analysis each datum must be assigned to exactly one cluster. Fuzzy cluster analysis relaxes this requirement by allowing gradual memberships, thus offering the opportunity to deal with data that belong to more than one cluster at the same time. The general philosophy of clustering is to divide the initial set into homogeneous groups and to reduce the data. Clustering methods can be of two types: Crisp and Fuzzy clustering. Crisp clustering assigns each data to a single cluster but in fuzzy the membership function measures the degree of belonging of each feature in a cluster. Most fuzzy clustering algorithms are objective function based: They determine an optimal classification by **minimizing** an **objective function**.

The degrees of membership to which a given data point belongs to the different clusters are computed from the distances of the data point to the cluster centers. These distances depend on the size and the shape of the cluster as stated by the additional prototype information. The closer a data point lies to the center of a cluster (i.e. size and shape), the higher is its degree of membership to this cluster. Several fuzzy clustering algorithms can be distinguished depending on the additional size and shape information contained in the cluster prototypes, the way in which the distances are determined, and the restrictions that are placed on the membership degrees [8], [9]. Here we focus on the fuzzy c-means algorithm [10], which uses only cluster centers and a Euclidean distance function, and the Gustafson{Kessel algorithm, which uses cluster centers, covariance matrices and a Mahalanobis distance function.

We propose and compare various Fuzzy clustering [4] techniques. The various Fuzzy clustering methods are:

- Fuzzy C Means Algorithm (FCM)
- Possibilistic C Means Algorithm (PCM)
- Hierarchical C Means Algorithm (HCM)

The fuzzy clustering differs from the conventional hard computing in that, unlike the later, it is tolerant of imprecision, uncertainty, partial truth, and approximation. The guiding principle is that it exploits the tolerance for imprecision, uncertainty, partial truth, and approximation to achieve tractability, robustness and low solution cost. As it resembles human brain, the results are fast and accurate.

A. Fuzzy C Means Algorithm (FCM)

The most prominent algorithm is the FCM or Fuzzy C Means algorithm. The Fuzzy C Means algorithm was proposed as an improvement of the classic Hard C-Means clustering algorithm. It is a method where large data is grouped into clusters in which each point has a degree of belonging completely to just one cluster. The data points that are nearer to the centre have high degree of membership rather than belonging completely to just one cluster. The data points that are nearer to the centre have high degree of membership than the points on the edge of a cluster have a lesser degree. The FCM algorithm receives the data or sample space in matrix format. The number of clusters, the assumption partitioning matrix, the convergence value all must be given to the algorithm. The FCM algorithm assigns pixels to each category by using fuzzy memberships N. The algorithm is an iterative optimization that minimizes the cost function.

The fuzzy c-means algorithm steps are

1. Random initialization of inputs to the cluster C.
2. Calculate centroid V_i for each cluster

$$V_i = \frac{\sum_{j=1}^{N} (u_{ij})^m X_j}{\sum_{j=1}^{N} (u_{ij})^m} \qquad i=1,2....C \quad (1)$$

3. Using objective function find the coefficients of the cluster for each point.
4. Repeat steps 2 and 3 until the algorithm has converged (that is the coefficients change between two iterations is no more than ε, the threshold)

The objective function is minimized when pixels close to the centroids are assigned high membership values and low membership values assigned to pixel far from centroid. The standard FCM objective function is given by

$$J_m(U, V) = \sum_{j=1}^{N} \sum_{i=1}^{C} u_{ij}{}^{m} d^2(X_j, V_i) \quad (2)$$

Where $X = \{X1, X_2, \ldots\ldots X_j, \ldots.. X_N\}$ is a p x N input data matrix, where p represents the dimension of each feature vector, and N represents the number of feature vectors.

C is the number of clusters.

Uij represents the membership function of the j^{th} data in i^{th} cluster C_i.

d is the distance between input and centroid.

Vi is the ith cluster center

m is a constant.

The parameter m controls the fuzziness of the resulting partition, and m=2 is used in this study. The cost function is minimized when pixels close to the centroid of their clusters are assigned high membership values, and low membership values are assigned to pixels with data far from the centroid. The membership function represents the probability that a pixel belongs to a specific cluster. In the FCM algorithm, the probability is dependent solely on the distance between the pixel and each individual cluster center in the feature domain. The membership functions and cluster centers are updated by the following

$$\text{Uij} = \frac{1}{\sum_{K=1}^{C} \left(d(x_j, v_{ij}) \mid d(x_j, v_k) \right)^{2/m-1}} \quad (3)$$

Where m is a weighting factor on each fuzzy membership, which controls the degree of fuzziness. A measure of similarity between X_j and V_i is given as

$$d^2(X_j, V_i) = \| X_j - V_i \| \quad (4)$$

The results of this algorithm depend upon the initial choice of weights but it minimizes intra-cluster variance. Similar to K means it also has local minimum. Starting with an initial guess for each cluster center, the FCM converges to a solution for V_i representing the local minimum or a saddle point of the cost function. Convergence can be detected by comparing the changes in the membership function or the cluster center at two successive iteration steps. This is ineffective in situations, in which the data is contaminated by noise,

B. Possibilistic C Means Algorithm (PCM)

In possibilistic fuzzy [5] clustering one tries to achieve a more intuitive assignment of degrees of membership by dropping the probability constraint of FCM, which is responsible for the undesirable effect. However, this leads to the mathematical problem that the objective function is now minimized by assigning $u_{ij} = 0$ for all i € {1,……c} and j € {1,…..,n}. In order to avoid this trivial solution, a penalty term is introduced, which forces the membership degrees away from zero. That is, the objective function J is modified to

$$J_m(x,\mu,c) = \sum_{i=1}^{c}\sum_{j=1}^{N} \mu_{ij}^{m} d_{ij}^{2} + \sum_{i=1}^{c} \eta_i \sum_{j=1}^{N} (1-\mu_{ij})^{m} \quad (5)$$

Where, d_{ij} is the distance between the j^{th} data and the i^{th} cluster center, μ_{ij} is the degree of belonging of the j^{th} data to the i^{th} cluster, m is the degree of fuzziness, η_i is a suitable positive number, c is the number of the clusters, and N is the number of the data. μ_{ij} can be obtained as

$$\mu = \frac{1}{1 + \left(\frac{d_{ij}^{2}}{\eta_i} \right)^{\frac{1}{m-1}}} \quad (6)$$

Where d_{ij} is the distance between the j^{th} data and the i^{th} cluster center, μ_{ij} is the degree of belonging of the j^{th} data to the i^{th} cluster, m is the degree of fuzziness, η_i is a suitable positive numbers. The value of η_i determines the distance at which the membership value of a point in a cluster becomes 0.5. The value of η_i is obtained as

$$\eta_i = \frac{\sum_{j=1}^{N} \mu_{ij}^{m} d_{ij}^{2}}{\sum_{j=1}^{N} \mu_{ij}^{m}} \quad (7)$$

The value of η_i can be fixed or changed in each iteration by changing the values of μ_{ij} and d_{ij}. This method is more robust in the presence of noise, in finding valid clusters, and in giving a robust estimate of the centers. At first sight this approach looks very promising. However, if we take a closer look, we discover that the objective function J defined above is, in general, truly minimized only if all cluster centers are identical. The reason is that formula for the membership degree of a datum to a cluster depends only on the distance of the datum to that cluster, but not on its distance to other clusters. Hence, if there is a single optimal point for a cluster center (as it will usually be the case, since multiple optimal points would require a high symmetry in the data), all cluster centers will converge to this point. More formally, consider two cluster centers β_1 and β_2 which are not identical, and let

$$z_i = \sum_{j=1}^{n} u_{ij}^{m} d^2(\beta_i, x_j) + \eta_i \sum_{j=1}^{n} (1-u_{ij})^{m} \quad (8)$$

That is, let z_i be the amount that clusters β_i contributes to the value of the objective function. Except in very rare cases of high data symmetry, it will then either be $z_1 > z_2$ or $z_2 > z_1$. That is, we can improve the value of the objective function by setting both cluster centers to the same value, namely the one which yields the smaller z-value, because the two z-values do not interact.

Note that this behavior is specific to the possibilistic approach. In the probabilistic approach the cluster centers are driven apart, because a cluster, in a way, consumes part of the weight of a datum and thus leaves less that may attract other cluster centers. Hence sharing a datum between clusters is disadvantageous. In the possibilistic approach there is nothing equivalent to this effect. Nevertheless, possibilistic fuzzy clustering [11] usually leads to acceptable results, although it suffers from stability problems if it is not initialized with the corresponding probabilistic algorithm [6]. We assume that other results than all cluster centers being identical are achieved only, because the algorithm gets stuck in a local minimum of the objective function. This, of course, is not a desirable situation. Hence we tried to improve the algorithm by modifying the objective function in such a way that the problematic property examined above is removed.

C. Hierarchical C Means Algorithm (HCM)

Given a set of elements X, a mixed approach is applied to build a fuzzy hierarchical structure. The process starts building a fuzzy partition of X applying fuzzy c-means. This results into a set of fuzzy membership functions μ_i, each one built on the centroid v_i. This fuzzy partition bootstraps the process. Then, the iterative process is applied to build the hierarchical clustering following a bottom-up strategy. Each step of the process starts with a fuzzy partition of X represented by a set of membership functions μ_i. Such set of membership functions is partitioned using a partitive clustering method for fuzzy sets. Such partitive clustering method returns a new fuzzy partition $\mu_i^{/}$ that is used as the starting point of the new step.

In this algorithm, the fuzzy c-means algorithm is used for building the initial fuzzy partition. Such fuzzy partition is obtained by applying the fuzzy c-means algorithm to X. In this case, the algorithm is applied with a large number of clusters (i.e., c is large). This selection of c is to have a large number of leaves in the fuzzy hierarchy. In the iterative process, fuzzy c-means based clustering method is used.

Differences consist on the way the distance $\|x_k\text{-}v_i\|$ is computed. Here, x_k and v_i represent fuzzy sets. More specifically, x_k stands for the k-th fuzzy set to be partitioned and v_i is one of the fuzzy sets in the new partition. Accordingly, $\|x_k\text{-}v_i\|$ is a distance between fuzzy sets. Following the standard approach in fuzzy c-means, the fuzzy membership of a fuzzy set with centroid v is defined considering all other centroids v_i. In our case, the membership of the fuzzy set with centroid x_k is computed for all x taking into account all other centroids x_j as follows:

$$\mu_{x_k}(x) = \left(\sum_{j=1}^{C} \left(\frac{d(x_k, x)^2}{d(x_k, x_j)^2} \right)^{\frac{1}{m-1}} \right)^{-1} \quad (9)$$

Similarly, the membership of the fuzzy set with centroid v_i is computed for all x taking into account all other centroids v_j as follows:

$$\mu_{v_k}(x) = \left(\sum_{j=1}^{C} \left(\frac{d(v_k, x)^2}{d(v_k, v_j)^2} \right)^{\frac{1}{m-1}} \right)^{-1} \quad (10)$$

Note that here, x_j are the centroids of the fuzzy sets being clustered and v_j are the centroids of the clusters we are constructing with the fuzzy c means. Similarly, c is the number of centroids xj_j and c is the number of centroids in v_j .Then, the distance between a fuzzy set with centroid x_k and another with centroid v_i will be computed. Another element to be taken into account is how to compute the new centroid, once the membership is known. This is, how to determine the new v_i.

$$v_i = \frac{\sum_{k=1}^{n} (u_{ik})^m x_k}{\sum_{k=1}^{n} (u_{ik})^m} \quad (11)$$

Note that this approach leads to different membership values. In particular, in the new approach it is possible that the membership of x to a cluster $\mu$is smaller than the membership of x to a sub cluster μ_j of μ. Thus this method builds hierarchies of clusters where membership to clusters is fuzzy.

III. EVALUATION OF SEGMENTED RESULTS

Different parameters were used to analyze the performance of various fuzzy clustering algorithms. They are False Positive Error (FPE), False Negative Error (FNE) [1], Coefficient of similarity and spatial overlap. To define the first two types of quality metrics let SR denote the result of an automatic segmentation method and GT denote the ground truth segmentation obtained by the medical expert. Both SR and GT are binary images such that all the pixels inside the curve hove label 1 and all others have label 0. The metrics are calculated as follows:

A. False Positive Error (FPE)

This metric measures the rate of pixels classified as lesions by the automatic segmentation that were not classified as lesion by the medical expert.

$$\text{FPE(SR,GT)} = \frac{\#(SR \cap \overline{GT})}{\#(GT)} \quad (12)$$

B. False Negative Error (FNE)

The FNR measures the rate of pixels classified as lesions by the medical expert that were not classified as lesion by the automatic segmentation.

$$\mathrm{FNE}(SR,GT)=1-\frac{\#(SR \cap GT)}{\#(GT} \quad (13)$$

Clinically, this is worse of two types of error.

Coefficient of Similarity:

The mean and the standard deviation of the coefficient of similarity between the automatic and manual segmentation [12] is given by

$$\in = 1-\frac{|v_{manual} - v_{automatic}|}{v_{manual}} \quad (14)$$

Spatial Overlap:

The measure of spatial overlap between the automatic (algorithmic) and the manual segmentation is given as

$$\in_s = \frac{2 * v_{intersection}}{v_{manual} + v_{algorithm}} \quad (15)$$

IV. EXPERIMENTAL RESULTS

Three various fuzzy clustering algorithmic methods were evaluated, as detailed in section 2. The evaluation was based on the measures described in section 3, using ground truth image manually segmented The proposed Fuzzy clustering algorithms is implemented using MATLAB and tested with ground truth image to explore the segmentation accuracy of the various fuzzy clustering techniques. The effectiveness of the proposed approach is experimentally determined using the ground truth image.

The input malignant melanoma image is as shown in fig (2). The segmented image using the fuzzy c means algorithm is given in fig (3) with number of clusters, c=2. In the output figure the white region indicates the non infected region(portions of the skin free from the malignant melanoma) and the black trace or spots indicates the infected region(portions of skin affected with malignant melanoma). The segmented image using the possibilistic c means algorithm is given in fig (4) and the segmented image using the Hierarchical c means algorithm is given in fig (5). The result of the automatic segmentation method was compared with the ground truth segmented image (Fig 6) using various parameters (FP and FN errors, Coefficient of similarity, and Spatial Overlap) and their performance is being evaluated.

Figure 2: Input Image (Malignant Melanoma image)

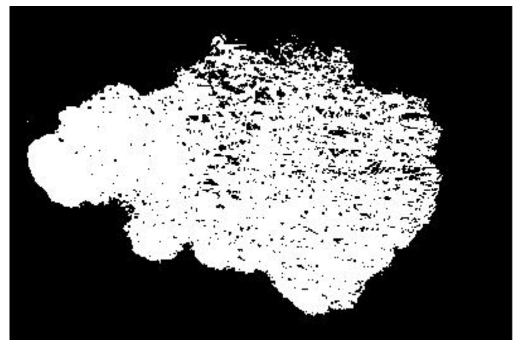

Figure 3: Segmented image using Fuzzy c means

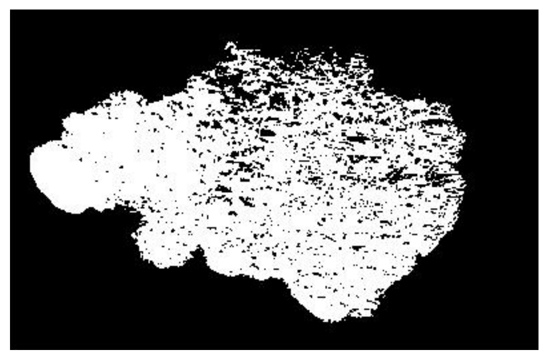

Figure 4: Segmented image using Possibilistic c means

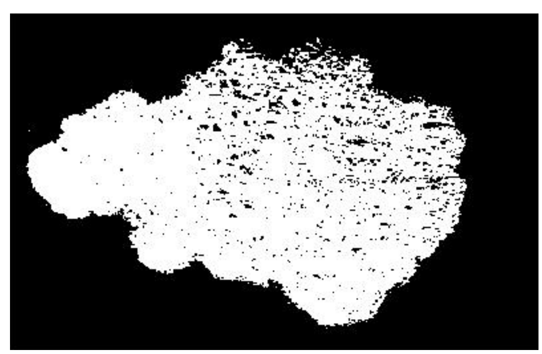

Figure 5: Segmented image using hierarchical c means

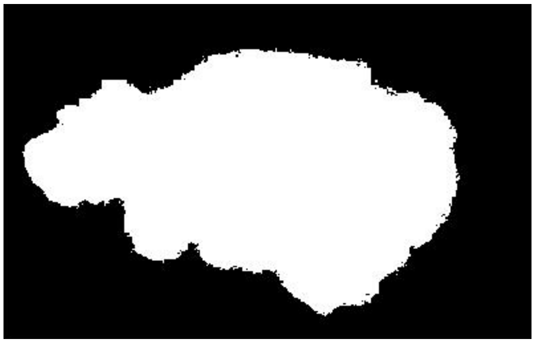

Figure 6: Ground truth image

Fig 3-5 shows the segmentation of malignant melanoma using the various fuzzy clustering techniques. In these cases, all the methods produce segmentation results which are close to the ground truth segmentation. This happens when there is a good contrast between the lesion and the skin, thus the lesion boundaries are well defined. For obtaining the more accurate result in performance analysis, various parameters are to be considered.

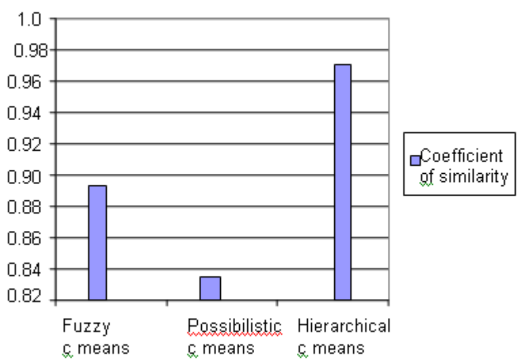

Figure.7 Coefficient of similarity for different clustering methods

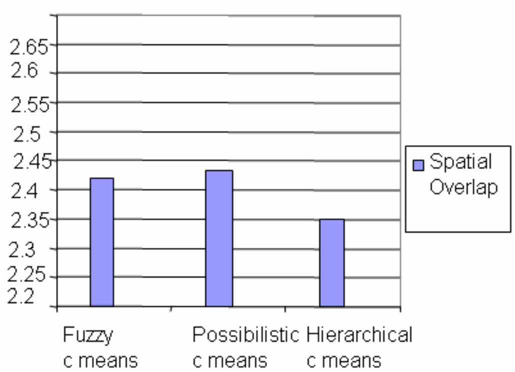

Figure.8 Spatial Overlap for different clustering methods

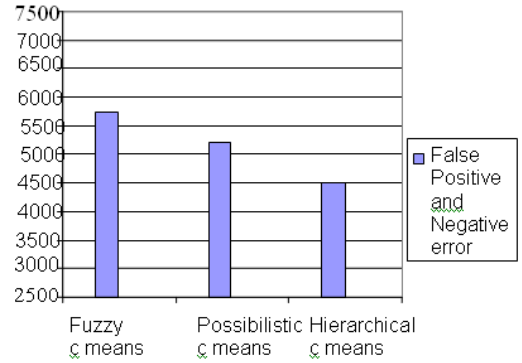

Figure .9 False Negative and Positive errors for different clustering methods

The experimental results obtained by employing the Hierarchical C Means clustering algorithm reveals that it has better performance over the other two clustering techniques (Fuzzy C Means and Possibilistic C Means clustering algorithms). Furthermore, the Hierarchical C Means clustering algorithm eliminates the effect of noise greatly. This in turn increases the segmentation accuracy of the clustering algorithm.

V. CONCLUSION

This paper proposes and evaluates the performance of various fuzzy clustering techniques for the segmentation of skin lesions in dermoscopic images. The fuzzy technique provides better segmentation when compared to the various existing methods. The various Fuzzy clustering techniques employed in this work are Fuzzy C Means, Possibilistic C Means, and Hierarchical C Means Algorithm. Experiments are conducted on real medical image to evaluate the performance of the proposed algorithm. The output of the automatic segmentation methods was compared with the manually segmented image (the ground truth image) using various parameters. The four most important parameters used to determine the accuracy of the proposed algorithm are False Positive and Negative error, Coefficient of similarity, and Spatial Overlap. The experimental results show that the Hierarchical C Means algorithm provides better performance than other two (Fuzzy C Means and Possibilistic C Means) clustering algorithms.

Thus the Hierarchiacal C Means approach provide better performance and can handle uncertainties that exist in the data efficiently and useful for for the accurate lesion segmentation in a computer aided diagnosis system to assist the clinical diagnosis of melanoma.

. VI. REFERRENCES

[1] Margarida Silvertip, Jacinto C Nascimento, Member IEEE "Comparison of Segmentation Methods for Melanoma Diagnosis In Dermoscopy Images", IEEE Journal Of Selected Topic In Signal Processing, Feb 2009

[2] H.Ganster, P.Pinz, R.Rohhrer, and H. Kittler," Automated melanoma recognition," IEEE Trans.Med.Imag...vol.20, pp.233-239, Mar.2001.

[3] F.NachbarW.Stolz,T. Merckle and G.Plewig," The ABCD rule of dermatoscopy: High prospective value in the diagnosis of doubtful melanocyctic skin lesions,"vol.30, pp.551-559, 1994.

[4] S.R.kannan S.Ramathilagam, R.Pandiyarajan, A.Sathya " Fuzzy Clustering Approach in Segmentation of T1-T2 Brain MRI",International Journal Of Recent Trends in Engineering,Vol 2,No.1,Nov 2009

[5] M.Gomathi,Dr.P.Thangaraj, "A Parameter Based Modified Fuzzy Possibilistic C-Means Clustering Algorithm For Lung Image Segmentation" ,Global Journal Of Computer Science And Technology,vol 10 Issue 4,June 2010

[6] Mahmoud Ramze Rezae, Pieter M.J. Van der Zwet,"A Multiresolution Image Segmentation Technique Based On Pyramidal Segmentation and Fuzzy Clustering" IEEE Transactions On Image Processing,Vol 9,No.7,July 2000

[7] P.Rubegni, A. Ferrari, G. Cevenini, D.Piccolo, M.Biagioli, P. Taddecci, (2001)" Differentiation between pigmented spitz naevus and melanoma by digital dermoscopy and stepwise logistic discriminant analysis", Melanoma Res, vol.11, no.1,pp 37-44.

[8] Bezdek, J.C. and Pal S.K.: "Fuzzy models for pattern recognition-methods that search for structures in data. IEEE Trans.2000.

[9] J.C.Keller, Bezdek, J.Krishnapuram," Fuzzy models and algorithms for pattern recognition and image processing.Kluwer, Boston, London, 1999.

[10] Li Jiang, Zhang," An image segmentation algorithm based on Fuzzy C-Means clustering algorithm" IEEE Trans.2009

[11]Nikhil R Paul, Kuhu Pal,and James C.Bezdek," A possibilistic fuzzy c-means clustering algorithm" IEEE transactions on fuzzy systems, vol.13, No.4, Aug.2005.

[12] T.Kapur, W.E.Grimson and R.Kinkinis," Segmentation of brain tissue from magnetic resonance images," Med. Ima Anal..vol.1, pp.109-127, 1996,

