An Efficient Brain Tumor Detection Methodology Using K-Means Clustering Algorithm

J.Vijay,J.Subhashini

***Abstract*-** **Segmentation of images holds an important position in the area of image processing. It becomes more important while typically dealing with medical images where pre-surgery and post surgery decisions are required for the purpose of initiating and speeding up the recovery process. Computer aided detection of abnormal growth of tissues is primarily motivated by the necessity of achieving maximum possible accuracy. Manual segmentation of these abnormal tissues cannot be compared with modern day's high speed computing machines which enable us to visually observe the volume and location of unwanted tissues. A well known segmentation problem within MRI is the task of labeling the tissue type which include White Matter (WM), Grey Matter (GM),Cerebrospinal Fluid (CSF) and sometimes pathological tissues like tumor etc. This paper describes an efficient method for automatic brain tumor segmentation for the extraction of tumor tissues from MR images. In this method segmentation is carried out using K-means clustering algorithm for better performance. This enhances the tumor boundaries more and is very fast when compared to many other clustering algorithms. The proposed technique produce appreciative results.**

***Index Terms*-** **Magnetic Resonance Imaging (MRI), Image segmentation, K- means, White Matter (WM), Grey Matter(GM) , Cerebrospinal Fluid (CSF)**

I.INTRODUCTION

The influence and impact of digital images on modern society is tremendous, and image processing is now a critical component in science and technology. With the computer techniques, multidimensional digital images of physiological structures can be processed and manipulated to help visualize hidden diagnostic features that are otherwise difficult or impossible to identify using planar imaging methods.

J.Vijay[1] is with Electronics And Communication Engineering Department, SRM University,Chennai(email:ramanvijay89@gmail.com).
J.Subhashini[2] is with Electronics And Communication Engineering Department, SRM University,Chennai(email:subhasivaj@gmail.com).

Image segmentation may be defined as a technique, which partitions a given image into a finite number of non-overlapping regions with respect to some characteristics, such as gray value distribution, texture Distribution, etc. Segmentation of medical images is required for many medical diagnoses like radiation treatment, planning volume visualization of regions of interest (ROI) defining boundary of brain tumor and intra cerebral brain hemorrhage, etc[1]. Many approaches are based on fuzzy logic means and Neural Networks (NN) distribution, etc. segmentation of medical images is required for many medical diagnoses like radiation treatment, planning volume visualization of regions of interest (ROI) defining boundary of brain tumor and intra cerebral brain hemorrhage, etc [2]. Basically, image segmentation methods can be classified into three categories: edge-based methods, region based methods and pixel-based methods[3].

K-means clustering is a key technique in pixel-based methods. In which pixel-based methods based on K-means clustering are simple and the computational complexity is relatively low compared with other region-based or edge-based methods, the application is more practicable[4]. It is an unsupervised clustering algorithm that classifies the input data points into multiple classes based on their inherent distance from each other[5]. This paper proposes automatic method to find the characteristics of Tumor cells using Morphological technique. Morphological operations are applied on the segmented image for smoothening the brain tumor part[6-9] .

The rest of this article is organized as follows. Section II describes the literature review which we have done to implement our idea, section III describes the proposed methodology, section IV describes the results and discussion and in section V we draw the conclusions.

II.LITERATURE REVIEW

A. Image Enhancement

Image enhancement is one of the major research fields in image processing. In any applications such as medical application, military application, media etc., the image enhancement plays an important role. Recently, neural networks turn to be a very effective tool to support the image enhancement. Neural network is applied in image enhancement because it provides many advantages over the other techniques [1]. Also, neural network can be suitable for removal of all kinds of noises based on its training data. This paper provides survey about some of the techniques applied for image enhancement.

978-1-4673-4866-9/13/$31.00

B. Morphology

Morphological processing is constructed with operations onsets of pixels. Binary morphology uses only set membership and is indifferent to the value, such as gray level or color, of a pixel. Morphological image processing relies on the ordering of pixels in an image and many times is applied to binary and gray scale images. Through processes such as erosion, dilation, opening and closing, binary images can be modified to the user's specifications [2,3]. Binary images are images whose pixels have only two possible intensity values. They are normally displayed as black and white. Numerically, the two values are often 0 for black, and either 1 or 255 for white. Binary images are often produced by thresholding a gray scale or color image, in order to separate an object in the image from the background. The color of the object (usually white) is referred to as the foreground color. The rest (usually black) is referred to as the background color. However, depending on the image which is to be threshold, this polarity might be inverted, and in which case the object is displayed with 0 and the background is with a non-zero value.

C. Morphological Segmentation

This section details the segmentation of mammograms for identifying the tumor in brain. The proposed approach utilizes mathematical morphology operations for the segmentation. The morphological operations are applied on the gray scale images to segment the abnormal regions [4,5]. Erosion and dilation are the two elementary operations in Mathematical Morphology. An aggregation of these two represents the rest of the operations of opening and closing. And all of these together forms the four fundamental binary morphological operations: dilation, erosion, opening, and closing.

III. PROPOSED METHODOLOGY

Clustering refers to the process of grouping pixels of an image such that pixels which are in the same group (cluster) are similar among them and are dissimilar to the pixels which belong to the other groups (clusters). Let the feature vectors derived from I clustered data. The generalized algorithm initiates k cluster centroids by randomly selecting k feature vectors from X. Given a set of observations $(x_1, x_{2,} \dots x_n)$, where each observation is a d-dimensional real vector, k-means clustering aims to partition the n observations into k sets (k ≤ n) S = $\{s_1, s_{2,} \dots s_k\}$ so as to minimize the within cluster sum of squares (WCSS)

$$arg_s min \sum_{i=1}^{k} \sum_{x_{j \in s_i}} \left\| x_{j - \mu_i} \right\|^2 \quad (1)$$

Where μ_i is the mean of points in s_i .

Given an initial set of k means $m_1^1, \dots m_k^1$ the algorithm proceeds by alternating between two steps:

Assignment step: Assign each observation to the cluster whose mean is closest to it (i.e. partition the observations according to the Voronoi diagram generated by the means).

$$S_i^{(t)} = \{x_p : \left\| x_{p - m_i^{(t)}} \right\| \leq \left\| x_{p - m_j^{(t)}} \right\| \forall 1 \leq j \leq k \quad (2)$$

where each x_p is assigned to exactly one $s^{(t)}$, even if it could be is assigned to two or more of them.

Update step: Calculate the new means to be the centroids of the observations in the new clusters.

$$m_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x_{j \in s_i^{(t)}}} X_j \quad (3)$$

Let the feature vectors derived from I clustered data. The generalized algorithm initiates k cluster centroids by randomly selecting k feature vectors from X. Later, the feature vectors are grouped into k clusters using a selected distance measure such as Euclidean j [7,8].. The clustering procedure stops only when all cluster centroids tend to converge. Similarity is measured by distance and defined by an N dimensional feature space. Feature distance calculation differs from spatial distance calculation. Feature distance calculation is based on features such as color or intensity and texture while spatial distance calculation is based on x, y (width, height) coordinates [9]. Devising an appropriate distance calculation method is an important task since it greatly affects final clustering result. Clustering algorithms may be generally classified into four main categories which are hierarchical, overlapping and exclusive [10,11].

K-means is an instance of exclusive clustering algorithms and is the backbone of this paper's methodology [11,12]. K means algorithm starts clustering by determining k initial central points, either at random or using some heuristic data. It then groups each image pixel under the central point it is closest to. Next, it calculates new central points by averaging the pixels grouped under each central point. The two former algorithmic steps are repeated alternately until convergence (central points no longer change by averaging). The limitations of K-means clustering are many iterative rounds may be required. This work strives to reduce that limitation.

A.K-Means based segmentation

K means clustering is the most widely used and studied method among the clustering formulations that are based on minimizing a formal objective function. Modifications to the K-means clustering method that makes it faster and more efficient are proposed. K-means clustering is a key technique in pixel-based methods. In which pixel-based methods based on K-means clustering are simple and the computational complexity is relatively low compared with other region-based or edge-based methods, the application is more practicable.

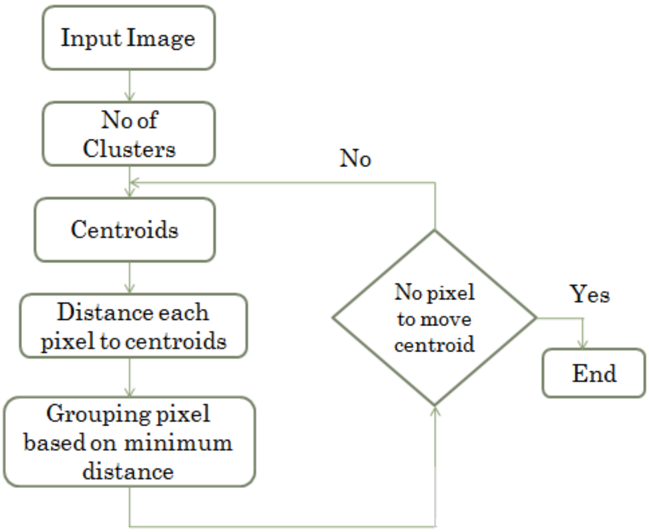


Figure.1-Working methodology of K-Means clustering algorithm

The main argument of the proposed modifications is on the reduction of intensive distance computation that takes place at each run (iteration) of K-means algorithm between each data point and all cluster centers. To reduce the intensive distance computation, a simple mechanism by which, at each iteration, the distance between each data point and the cluster nearest to it is computed and recorded in a data structure is suggested. Thus, on the following iterations the distance between each data point and its previous nearest cluster is recomputed.

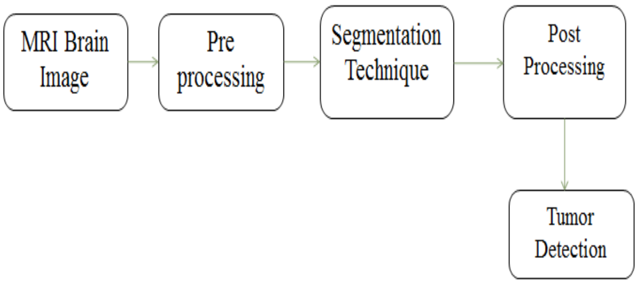


Figure.2- Block diagram of the proposed method

Figure.2 explains the flowchart of the proposed method. In the proposed method, we combine segmentation and the K-means clustering. A brain Image consists of four regions i.e. gray matter (GM), white matter (WM), cerebra spinal fluid (CSF) and background. Therefore, an input image needs to be divided into these four classes. In order to avoid the chances of misclassification, the outer elliptical shaped object should be removed. After the enhancement of image morphological process is carried out to extract the required region. The Next step is by implementing K-means with clusters exact result is produced.

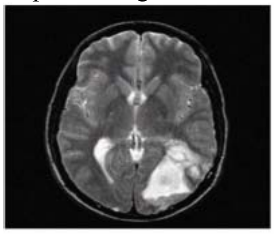


Figure. 3 Image with Outer Ring (Skull)

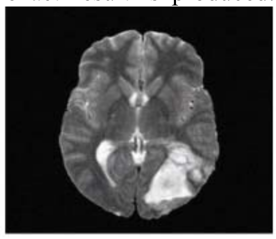


Figure.4 Removing Skull Tissues

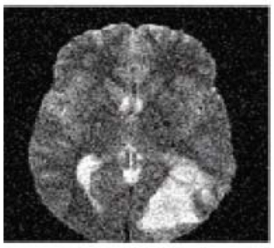


Figure.5 Noisy image

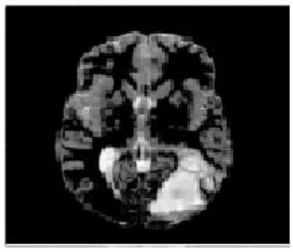


Figure.6 Enhanced image

Figure.3 shows an image of the brain with skull seen as an outer elliptical ring. In figure.4, this elliptical ring is removed and we are left with only soft tissues. This is done by employing the following morphological function, i.e. erosion and dilation.

In Figure.5, the noisy image is shown, which is enhanced by removing noise in the image by a suitable filter. The enhanced image is shown in figure.6. Due to unsupervised nature of the approach, the proposed system is efficient and is less error sensitive. It can be deduced from the results that unsupervised segmentation methods are better than the supervised segmentation methods. Because for using supervised segmentation method a lot of preprocessing is needed. More importantly, the supervised segmentation method requires considerable amount of training and testing data which comparatively complicates the process. Whereas, this study can be applied to the minimal amount of data with reliable results. However, it may be noted that, the use of K-Means clustering method is fairly simple when compared with frequently used fuzzy clustering methods. The Figure.7 shows the segmentation using the proposed work. In this system brain tumors have been segmented with the help of K-means algorithm. The execution time for K-means clustering was less compared to the other clustering methods.

The proposed work also reduces the computational complexity and also provides an accurate method of extracting the Region of Interest (ROI).More importantly, the supervised segmentation method requires considerable amount of training and testing data which comparatively complicates the process. This study can be applied to the minimal amount of data with reliable results.

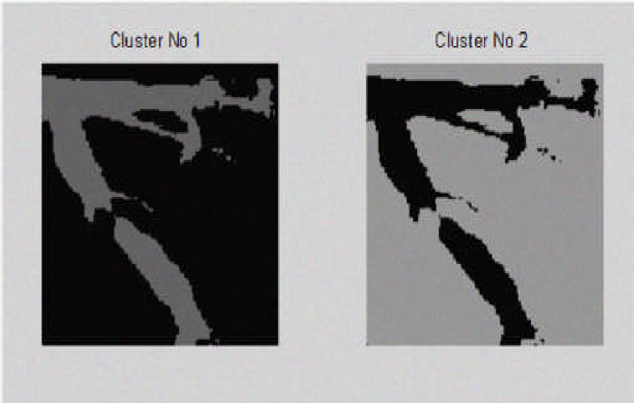


Figure.7 Segmentation using improved K- Means (Proposed work)

Regarding the number of tumor pixels, K-means clustering gave a better result than the other methods. The clustering algorithm was tested with a database of 100 MRI brain images. K-means clustering achieved about 95% result.

IV. RESULTS & DISCUSSION

The tumor affected cells are found out by applying K-means Clustering algorithm using Matlab Simulator. Same process can be applied for brain cell without tumor. The figure.8 shows the brain Image with tumor cells.Figure.9 shows the clustering window in which we specify the number of clusters to be processed. Then we have to pick up a structuring element and then we starts with the K Means clustering methodology and the clustering of Brain tumor takes place in Figure.11 in which the tumor cells are identified in the 4th cluster.

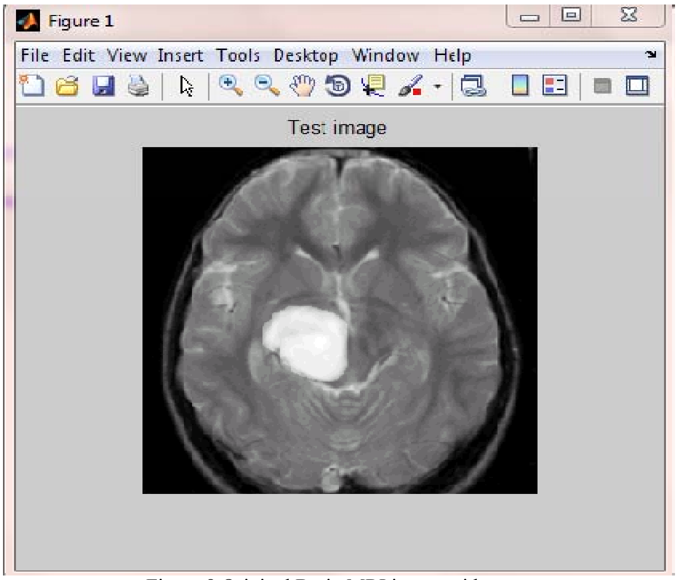


Figure.8 Original Brain MRI image with tumor

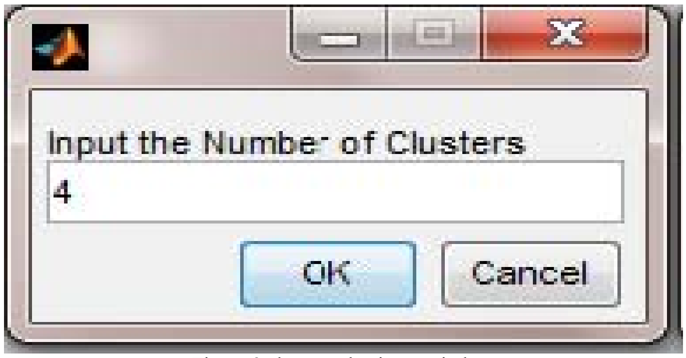


Figure.9 cluster value input window

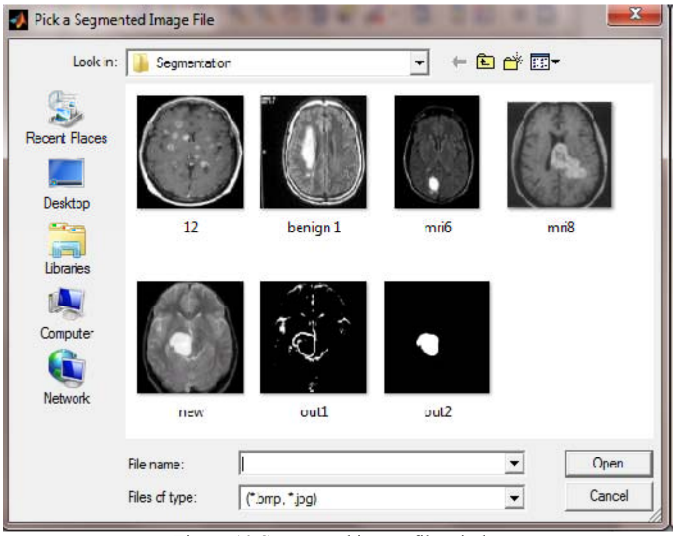


Figure.10 Segmented image file window

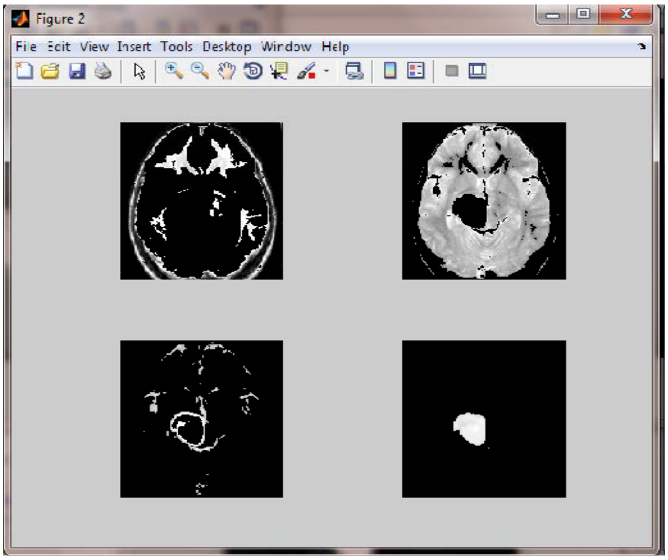


Figure.11 Clustering of Brain tumor

V. CONCLUSION

Thus in our proposed method, we combine segmentation and K-means clustering. It can be deduced from the results that unsupervised segmentation methods are better than the supervised segmentation methods. Because for using supervised segmentation method a lot of preprocessing is needed. More importantly, the supervised segmentation method requires considerable amount of training and testing data which comparatively complicates the process. This study can be applied to the minimal amount of data with reliable results. However, it may be noted that, the use of K-Means clustering method is fairly simple when compared with frequently used fuzzy clustering methods.

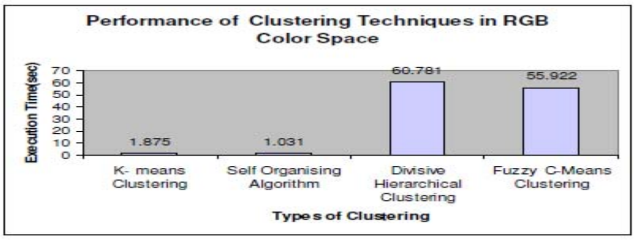
Table 1: Performance of Clustering Techniques in RGB Color Space

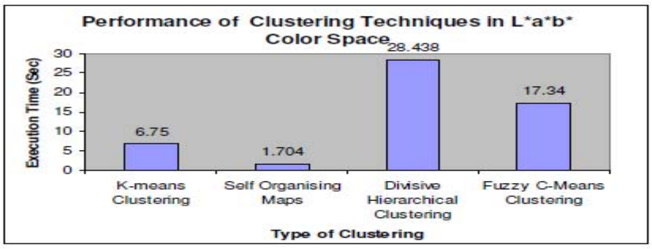
Type of Clustering	Execution Time (sec)
K-means Clustering	1.875
Self Organising Maps	1.031
Divisive Hierarchical Clustering	60.781
Fuzzy C-Means Clustering	55. 922

Table 2: Performance of Clustering Techniques in L*a*b* Color Space

Type of Clustering	Execution Time (sec)
K-means Clustering	6.75
Self Organising Maps	1.704
Divisive Hierarchical Clustering	28.438
Fuzzy C-Means Clustering	17.340

Thus we evaluate the proposed technique with other clustering methods and the results are very appreciable. Less execution time is achieved when compared with other clustering methods due to less number of iterations. Maximum lossless compression is also achieved using new standard and transformations.





Thus from the evaluated results it is clearly shown that the proposed system is efficient and is less error sensitive.

REFERENCES

[1]. M. Masroor Ahmed and Dzulkifli Bin Mohamad "*Segmentation of Brain MR Images for Tumor Extraction by Combining Kmeans Clustering and Perona-Malik Anisotropic Diffusion Model"* International Journal of Image Processing, Volume (2) Issue (l),2010.

[2] Arash Azim Zadeh Irani and Bahari Belaton "*A K-means Based Generic Segmentation System*" 2009 Sixth International Conference on Computer Graphics, Imaging and Visualization.

[3] Ming-Ni Wu, Chia-Chen Lin and Chin-Chen Chang "*Brain Tumor Detection Using Color-Based K-Means Clustering Segmentation*" International Conference on Computer Graphics,2009.

[4] S.Sathish Kumar, M.Moorthi, M.Madhu3and Dr.R.Amutha "*An Improved method of Segmentation using Fuzzy-Neuro Logic"* Second International Conference on Computer Research and development,2009.

[5] S. Murugavalli and V. Rajamani" *An improved Implementation of Brain tumor Detection Using Segmentation Based on Neuro Fuzzy Technique*", Journal of Computer Science 3 ,2007 october.

[6] Kharrat, A. Benamrane, N. Ben Messaoud, M. Abid, Computer. & Embedded Syst. Lab. (CES), Nat. Eng. Sch. of Sfax, Sfax, Tunisia, Nov.2009 .

[8] M. Schmidt, I. Levner, R. Greiner, A. Murtha, and A. Bistritz. "*segmenting Brain Tumors using Alignment-Based Features*".International Conference on Machine Learning and Applications, Los Angeles, December 2005.

[9] S.R. Kannan , Jan. 2005 "*Segmentation of MRl using new Unsupervised Fuzzy C means algorithm"* ICGST-GVIP Journal, Volume 5, Issue2.

[10] S.Saheb Baha and Dr. K.Satya Prasad, "*Morphological image Processing in Bio- Medical Application*", Proceedings of PCEAIFTOMM-International conference - PICA- 2006, held on 11th to 14th of July at Nagpur and NCBME-2006 - National conference on Bio-Medical Engineering held on 28th to 29th of March at Mumbai.

[11] S.Saheb Baha and Dr. K.Satya Prasad, "*Automatic detection of Hard Exudates in Diabetic Retinopathy using Morphological segmentation and Fuzzy Logic*" in [JCSNS International Journal of Computer Science and Network Security, VOL.8 No.12, December 2008.

[12] Juraj Horvath, 2006 *"Image Segmentation using Fuzzy C-means*" SAMI 2006.

[13] Lei jiang, Wenhui Yang , "*A modified Fuzzy C-means algorithm for Magnetic Resonance Images"* Proc. VIIth Digital Image Computing: Techniques and Applications, 2003.

[14] DIGTAL IMAGE PROCESSING,- Rafael C. Gonzalez, -Richard E.Woods, ADDISON-WESLEY, An imprint of Pearson Education, 1st Edition.

[15] J. Goldberger, S. Gordon, H. Greenspan, *"Unsupervised Image-Set Clustering Using an Information Theoretic Framework",* IEEE Transactions on Image Processing, Vol. IS, No. 2, February 2006,