International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2017 IJSRCSEIT | Volume 2 | Issue 1 | ISSN : 2456-3307

Brain Tumor Image Segmentation using K-means Clustering Algorithm

Kanchana Devi E, Dr. D. Devi Aruna
Department of Computer Applications Dr. N. G. P. Arts and Science College, Kalapatti, Coimbatore, Tamil Nadu, India

ABSTRACT

Brain tumor segmentation aims to separate the different tumor tissues such as active cells, necrotic core, and edema from normal brain tissues of White Matter (WM), Gray Matter (GM), and Cerebrospinal Fluid (CSF). MRI based brain tumor segmentation studies are attracting more and more attention in recent years due to non-invasive imaging and good soft tissue contrast of Magnetic Resonance Imaging (MRI) images. With the development of almost two decades, the innovative approaches applying computer-aided techniques for segmenting brain tumor are becoming more and more mature and coming closer to routine clinical applications. The purpose of this paper is to provide a K-means clustering algorithm for MRI-based brain tumor segmentation. K-means clustering algorithm is an unsupervised algorithm and it is used to segment the interest area from the background. However, before applying K-means algorithm, first partial stretching enhancement is applied to the image to improve the quality of the image.
Keywords : Image Segmentation, K-Means Clustering, Magnetic Resonance Imaging (MRI), Segmentation

I. INTRODUCTION

Tumor is an uncontrolled growth of cancer cells in any part of the body. Tumors are of different types and have different characteristics and different treatments[1]. At present, brain tumors are classified as primary brain tumors and metastatic brain tumors. The former begin in the brain and tend to stay in the brain, the latter begin as a cancer elsewhere in the body and spreading to the brain. Brain tumors are divided into two types: benign and malignant. In fact, the most widely used grading scheme has been issued by the World Health Organization (WHO)[2]. It classifies brain tumors into grade I to IV under the microscope. In general, grade I and grade II are benign brain tumor (low-grade); grade III and grade IV are malignant brain tumor (high-grade). Usually, if low-grade brain tumor is not treated, it is likely to deteriorate to high-grade brain tumor. The 2012 CBTRUS (Central Brain Tumor Registry of the United States) Statistical Report has also showed that brain tumors are the second leading cause of cancer-related deaths in children under age 20 and in males ages 20- 39 (leukemia is the first) and the fifth leading cause of cancer-related deaths in females ages 20-39. An estimated 69 720 new cases of primary brain tumors were expected to be diagnosed in 2013 and included both malignant (24 620) and non-malignant (45 100) brain tumors. This estimate is based on an application of age-sex-race-specific incidence rates from the 2013 CBTRUS Statistical Report using SEER and NPCR data to project respective age-sex-race groups (www.abta.org/aboutus/news/brain-tumor-statistics/). Therefore, brain tumor are seriously endangering people's lives and early discovery and treatment have become a necessity. In the clinical aspect, treatment options for brain tumor include surgery, radiation therapy or chemotherapy. Along with the advance of medical imaging, imaging modalities play an important role in the evaluation of patients with brain tumors and have a significant impact on patient care. Recent years, the emerging new imaging modalities, such as XRay, Ultrasonography, Computed Tomography (CT), Magneto Encephalo Graphy (MEG), Electro Encephalo Graphy (EEG), Positron Emission Tomography (PET), Single-Photon Emission Computed Tomography (SPECT), and Magnetic

Resonance Imaging (MRI), not only show the detailed and complete aspects of brain tumors, but also improve clinical doctors to study the mechanism of brain tumors at the aim of better treatment. Clinical doctors play an important role in brain tumor assessment and therapy. Once a brain tumor is clinically suspected, radiologic evaluation is required to determine the location, the extent of the tumor, and its relationship to the surrounding structures. This information is very important and critical in deciding between the different forms of therapy such as surgery, radiation, and chemotherapy. Therefore, the evaluation of brain tumors with imaging modalities is now one of the key issues of radiology departments. MRI is a non-invasive and good soft tissue contrast imaging modality, which provides invaluable information about shape, size, and localization of brain tumors without exposing the patient to a high ionization radiation[3]. MRI is attracting more and more attentions for the brain tumor diagnosis in the clinical[4].

Due to the large amount of brain tumor images that are currently being generated in the clinics, it is not possible for clinicians to manually annotate and segment these images in a reasonable time. Hence, the automatic segmentation has become inevitable. Brain tumor segmentation is to segment abnormal tissues such as active cells, necrotic core, and edema (Fig. 2) from normal brain tissues including GM, WM, and CSF[6]. In recent years, medical imaging and soft computing have made significant advancements in the field of brain tumor segmentation. In general, most of abnormal brain tumor tissues may be easily detected by brain tumor segmentation methods. But accurate and reproducible segmentation results and representation of abnormalities have not been solved all the way. Since brain tumor segmentation has great impact on diagnosis, monitoring, treatment planning for patients, and clinical trials, this paper focuses on MRI-based brain tumor segmentation and presents a relatively detailed overview for the current existing methods of MRI-based brain tumor segmentation. The rest of this paper is organized as follows: In Section 2, we briefly introduce the preprocessing methods of MRI images. In Section 3, we discuss the current different brain tumor segmentation algorithms including conventional methods, classification and clustering methods, and deformable model methods. In Section 4, we analyze the evaluation and validation of the current brain tumor segmentation methods. Finally, in Section 5, an objective assessment is presented and future developments and trends are addressed for MRIbased brain tumor segmentation methods.

Image segmentation is one of the mostly used methods to classify the pixels of an image correctly in a decision oriented application. It divides an image into a number of discrete regions such that the pixels have high similarity in each region and high contrast between regions. It is a valuable tool in many field including health care, image processing, traffic image, pattern recognition etc. There are different techniques for image segmentation like threshold based, edge based, cluster based, neural network based1. From the different technique one of the most efficient methods is the clustering method. Again there are different types of clustering: K-means clustering, Fuzzy C-means clustering, mountain clustering method and subtractive clustering method. One of most used clustering algorithm is k-means clustering. It is simple and computationally faster than the hierarchical clustering. And it can also work for large number of variable. But it produces different cluster result for different number of number of cluster. So it is required to initialize the proper number of number of cluster, k2. Again, it is required to initialize the k number of centroid. Different value of initial centroid would result different cluster. So selection of proper initial centroid is also an important task. Nowadays image segmentation becomes one of important tool in medical area where it is used to extract or region of interest from the background. So medical images are segmented using different technique and process outputs are used for the further analysis in medical. But medical images in their raw form are represented by the arrays of numbers in the computer3, with the number indicating the values of relevant physical quantities that show contrast between different types of body parts. Processing and analysis of medical images are useful in transforming raw images into a quantifiable symbolic form, in extracting meaningful qualitative information to aid diagnosis and in integrating complementary data from multiple imaging modalities. And one of the fundamental problems in medical analysis is the image segmentation which identifies the boundaries of objects such as organs or abnormal region in images. Results from the segmentation make it possible for shape analysis, detecting volume change, and making a precise radiation therapy treatment plant.

II. METHODS AND MATERIAL

1. Related Work

There have been many works done in the area of image segmentation by using different methods. And many are done based on different application of image segmentation. K-means algorithm is the one of the simplest clustering algorithm and there are many methods implemented so far with different method to initialize the centre. And many researchers are also trying to produce new methods which are more efficient than the existing methods, and shows better segmented result. Some of the existing recent works are discussed here. Pallavi Purohit and Ritesh Joshi4 introduced a new efficient approach towards K-means clustering algorithm. They proposed a new method for generating the cluster center by reducing the mean square error of the final cluster without large increment in the execution time. It reduced the means square error without sacrificing the execution time. Many comparisons have been done and it can conclude that accuracy is more for dense dataset rather than sparse dataset. Alan Jose, S. Ravi and M. Sambath5 proposed Brain Tumor Segmentation using K-means Clustering and Fuzzy C-means Algorithm and its area calculation. In the paper, they divide the process into three parts, pre-processing of the image, advanced k-means and fuzzy c-means and lastly the feature extraction. First pre-processing is implemented by using the filter where it improves the quality of the image. Then the proposed advance K-means algorithm is used, followed by Fuzzy c-means to cluster the image. Then the resulted segment image is used for the feature extraction for the region of interest. They used MRI image for the analysis and calculate the size of the extracted tumor region in the image. Madhu Yedla, Srinivasa Rao Pathakota, T. M. Srinivasa6 proposed Enhancing K-means clustering algorithm with improved initial center. A new method for finding the initial centroid is introduced and it provides an effective way of assigning the data points to suitable clusters with reduced time complexity. They proved their proposed algorithm has more accuracy with less computational time comparatively original k-means clustering algorithm. This algorithm does not require any additional input like threshold value. But this algorithm still initializes the number of cluster k and suggested determination of value of k as one of the future work. K. A. Abdul Nazeer, M. P. Sebastian7 proposed an enhanced algorithm to improve the accuracy and efficiency of the k-means clustering algorithm. They present an enhanced k-means algorithm which combines a systematic method consisting two approaches. First one is finding the initial centroid and another is assigning the data point to the clusters. They have taken different initial centroid and tested execution time and accuracy. From the result it can be conclude that the proposed algorithm reduced the time complexity without sacrificing the accuracy of clusters.

2. K-Means Clustering Algorithm

Clustering is a method to divide a set of data into a specific number of groups. It’s one of the popular method is k-means clustering. In k-means clustering, it partitions a collection of data into a k number group of data11, 12. It classifies a given set of data into k number of disjoint cluster. K-means algorithm consists of two separate phases. In the first phase it calculates the k centroid and in the second phase it takes each point to the cluster which has nearest centroid from the respective data point. There are different methods to define the distance of the nearest centroid and one of the most used methods is Euclidean distance. Once the grouping is done it recalculate the new centroid of each cluster and based on that centroid, a new Euclidean distance is calculated between each center and each data point and assigns the points in the cluster which have minimum Euclidean distance. Each cluster in the partition is defined by its member objects and by its centroid. The centroid for each cluster is the point to which the sum of distances from all the objects in that cluster is minimized. So K-means is an iterative algorithm in which it minimizes the sum of distances from each object to its cluster centroid, over all clusters. Let us consider an image with resolution of $x \times y$ and the image has to be cluster into k number of cluster. Let $p(x, y)$ be an input pixels to be cluster and c_k be the cluster centers.

The algorithm for k-means13 clustering is following as:

1. Initialize number of cluster k and centre.
2. For each pixel of an image, calculate the Euclidean distance d, between the center and each pixel of an image using the relation given below. $d = p(x, y) - c_k$ (3)
3. Assign all the pixels to the nearest centre based on distance d.

4. After all pixels have been assigned, recalculate new position of the centre using the relation given below. $ck = 1\ k\ y \in ck\ x \in ck\ p(x, y)$ (4)
5. Repeat the process until it satisfies the tolerance or error value. 6. Reshape the cluster pixels into image. Although k-means has the great advantage of being easy to implement.

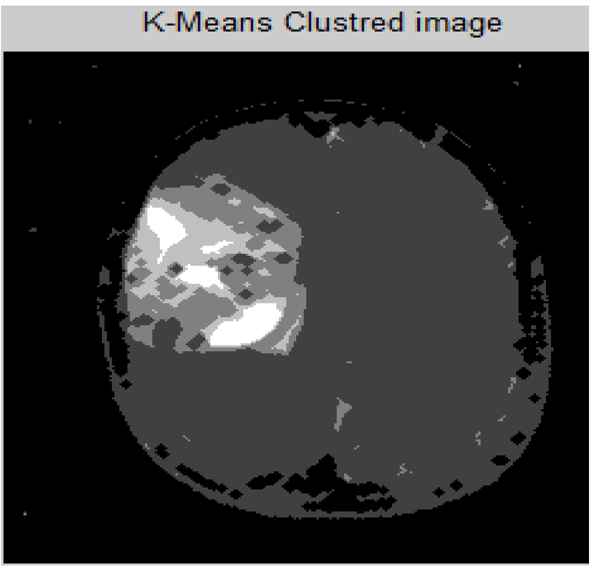
iT has some drawbacks. The quality of the final clustering results is depends on the arbitrary selection of initial centroid. So if the initial centroid is randomly chosen, it will get different result for different initial centers. So the initial center will be carefully chosen so that we get our desire segmentation. And also computational complexity is another term which we need to consider while designing the K-means clustering. It relies on the number of data elements, number of clusters and number of iteration.

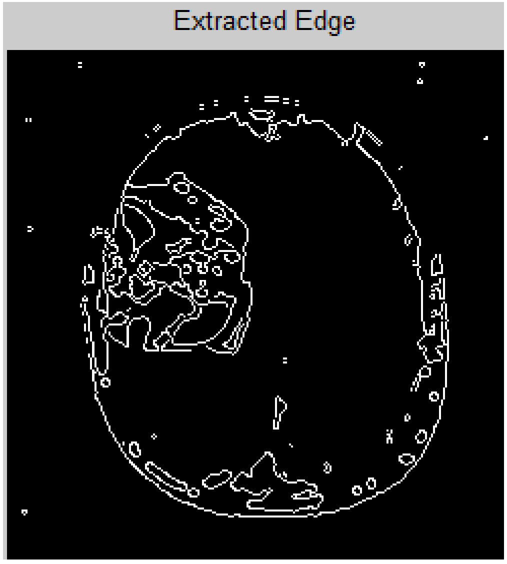
III. RESULTS AND DISCUSSION

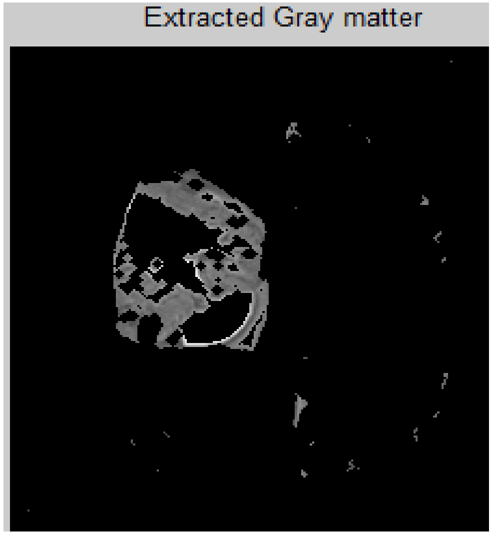
Experimental Results

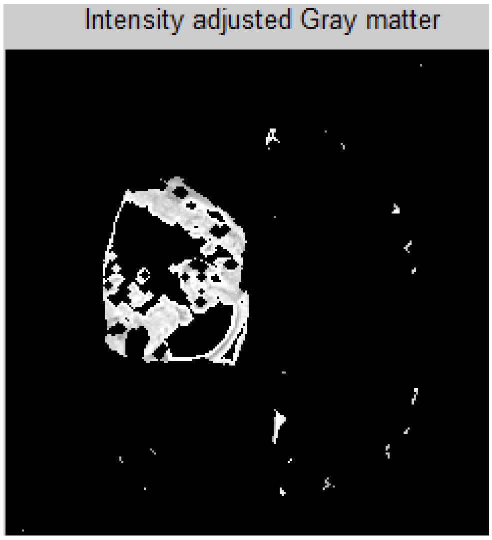
Images are obtained by MRI scan of brain and the output of MRI provides gray level images. A gray scale image is a data matrix whose value represents shades of gray. The elements of gray scale matrix have integer values or intensity values in range [0 255].

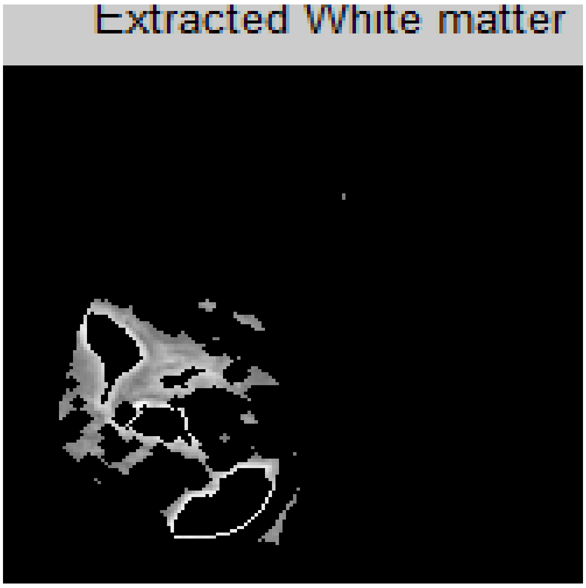
For applying different techniques, the digital images obtained from MRI are stored in matrix form in MATLAB.

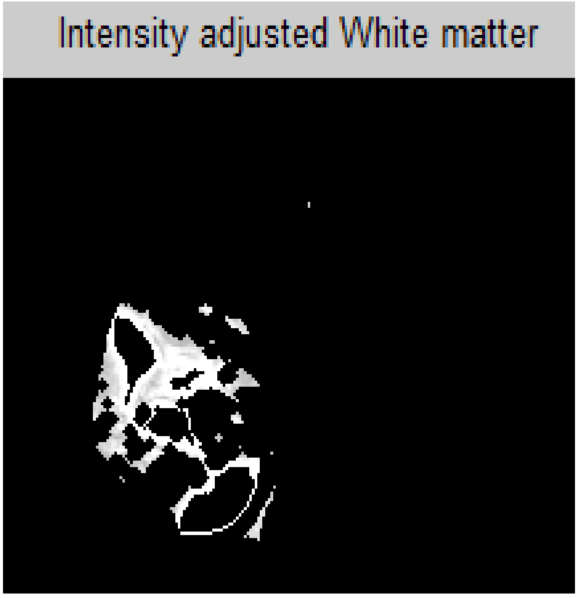


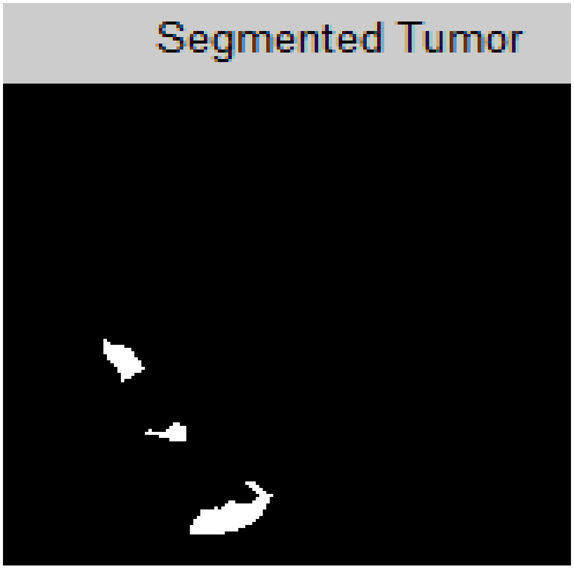


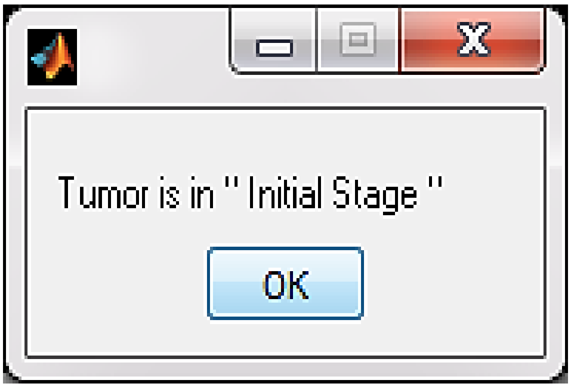












IV.CONCLUSION

We have segmented an image by using k-clustering algorithm, using subtractive cluster to generate the initial centroid. At the same time partial contrast stretching is used to improve the quality of original image and median filter is used to improve segmented image. And the final segmented result is compare with k-means clustering algorithm and we can conclude that the proposed clustering algorithm has better segmentation. The output images are also tune by varying the hyper sphere cluster radius and we can conclude from that result by varying the hyper sphere cluster radius we can get different output. And so we should take the value of hyper sphere cluster very carefully. Finally RMSE and PSNR are checked and observed that they have small and large value respective, which are the condition for good image segmentation quality. And comparison for RMSE and PSNR are done for proposed method and classical K-means algorithm and it is found that the proposed method have better performance result.

In the future, we can improve the quality of the output image more by using the morphological operation and get better performance measurement. We can also implement different clustering method using subtractive clustering algorithm. And lastly we can implement and analyze in different areas of image segmentation.

V. REFERENCES

[1]. K. Atsushi, N. Masayuki, "K-Means Algorithm Using Texture Directionality for Natural Image Segmentation", IEICE technical report. Image engineering, 97 (467), pp.17-22, 1998.

[2]. A. Murli, L. D'Amore, V.D. Simone, "The Wiener Filter and Regularization Methods for

Image Restoration Problems", Proc. The 10th International Conference on Image Analysis and Processing, pp. 394-399, 1999.

[3]. S. Ray, R.H. Turi, "Determination of number of clusters in K-means clustering and application in colthe image segmentation", Proc. 4th ICAPRDT, pp. 137-143, 1999.

[4]. T. Adani, H. Ni, B. Wang, "Partial likelihood for estimation of multiclass posterior probabilities", Proc. the IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 2, pp. 1053-1056, 1999.

[5]. B. Kovesi, J.M. Boucher, S. Saoudi, "Stochastic K-means algorithm for vector quantization", Pattern Recognition Letters, Vol. 22, pp. 603-610, 2001.

[6]. J. Z. Wang, J. Li, G. Wiederhold, "Simplicity: Semantics-sensitive integrated matching for picture libraries", IEEE Transactions on Pattern Analysis and Machine Intelligence, 23 (9), pp. 947–963, 2001.

[7]. Y. Gdalyahu, D. Weinshall, M. Wermen, "Self-Organizationin Vision: Stochastic clustering for Image Segmentation, Perceptual Grouping, and Image database Organization", IEEE Transactions on Pattern Analysis and Machine Intelligence,Vol.23, No. 12, pp. 1053- 1074, 2001.

[8]. C. Carson, H. Greenspan, "Blobworld: Image Segmentation Using Expectation-Maximization and Its Application to Image Querying", IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 24, No. 8, pp. 1026-1038, 2002.

[9]. C. J. Veenman, M.J.T. Reinders, E. Backer, "A maximum variance cluster algorithm", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 9, pp. 1273-1280, 2002.

[10]. B. Wei, Y. Liu, Y. Pan, "Using Hybrid Knowledge Engineering and Image Processing in Color Virtual Restoration of Ancient Murals", IEEE Transactions on Knowledge and Data Engineering, Vol. 15, No. 5, 2003.