Brain Tumor Segmentation from Brain Magnetic Resonance Images using Clustering Algorithm

V.Malathy, S.M.Kamali

Abstract: The main purpose of this article is to ascertain the presence of the tumor as well as quantify it. This is achieved by using clustering-based methods of segmentation. Both clustering methods for K-means and FCM are used to segment the tumor. FCM clustering is better than K-means clustering due to its fuzzy assigning of the pixels to each cluster. This helped to isolate the tumor portion more effectively as pixels, previously not enlisted as tumor region, are detected to be tumor in FCM clustering. The paper's aim is also to find out if the tumor has reached critical stage or not. This is done by taking the scale of the MRI and using the number of pixels isolated. The other important aim of the paper is to remove noise effectively to get the best possible segmented output. Though median and mean filters are highly effective noise removal filters they do not aide in the removal of multiplicative noise. These filters only remove the additive noise present in the image and are not suitable for removing the multiplicative noise present. Therefore, discrete wavelet transformation is used. This helps to get better detection of multiplicative noise from the image as it works in the frequency domain. This fact is ascertained by the comparisons made with detection of the tumor from the original image and filtered images. For this paper, noise was added for it to be more visual though unnecessary. This fact is driven by the results obtained. The results clearly indicate that the filtered image gives better results.

Keywords: Magnetic resonance Imaging, K-means clustering, Fuzzy C-means clustering, Brain tumor

I. INTRODUCTION

A digital image is a binary representation of a two-dimensional image. A digital image can be a gray image composed of only black (or white) shades. A digital image is represented by pixels. A pixel is the smallest addressable element in an image which is generally represented by an 8 bit binary number. So, a pixel value ranges from 0 to 255.

Digital image processing is using computer algorithms to develop process, interact and show digital images.

The brain tumor is an abnormal tissue growth in the brain. Brain tumors spread through local extension and rarely metastasize (spread) outside of the brain. A benign brain tumor consists of non-cancerous cells and does not metastasize beyond the part of the brain from which it originates.

A brain tumor is considered to be malignant if it contains cancer cells or if it consists of harmless cells located in an area where it suppresses one or more vital functions. About half of all main brain tumors are benign but at life-threatening places. The rest of them are malignant and invasive.

Benign brain tumors, composed of harmless cells, have clearly defined boundaries that can be permanently removed. They're unlikely to recur. Benign brain tumors do not infiltrate nearby tissues but may cause severe pain, permanent brain damage, and death.

Malignant brain tumors have no distinct boundaries. They tend to develop rapidly, growing pressure within the brain, and may spread beyond the point where they originate in the brain or spinal cord. It is highly unusual for malignant brain tumors to spread across the central nervous system.

Primary tumors of the brain originate in the brain. They represent about 1% of all cancers and 2.5% of all deaths from cancer.

Approximately 25% of all cancer patients develop secondary or metastatic brain tumors when cancer cells spread to the brain from another part of the body. These metastatic brain tumors can develop on any part of the brain or spinal cord. Secondary brain tumors are most apt to occur in patients with breast cancer, colon cancer, kidney cancer, lung cancer, melanoma (skin cancer) and cancer within the nasal passages and/or throat that follow the nerve pathways into the skull.

Manisha et al. (2010) referred to the grouping of image elements exhibiting similar characteristics. Grouping means splitting an image into its constituent regions or objects. Watershed transform generates closed contours for each region in the original image. But the watershed transformation may lead to over segmentation. Therefore, an improved watershed transformation based on morphological reconstruction is proposed in this paper.

Suman (2008) used three segmentation algorithms, namely k-means clustering, EM (Expectation maximization) and Normalized Cuts. K-means clustering is a clustering algorithm wherein the given image is segmented into K number of clusters. This is done by considering a centroid for each cluster and mapping all points to a cluster based on how close the specific point is to the cluster. This is an iterative process. Expectation maximization algorithm works to estimate the density of data points in an unsupervised environment.

Revised Manuscript Received on May 22, 2019.

V.Malathy, Assistant Professor, Department of Electronics and Communication Engineering, S R Engineering College, Warangal.

S.M.Kamali, Professor, Department of Electrical and Electronics Engineering, Mahendra College of Engineering, Salem.

Here, the expectation and maximization steps are performed alternatively. This is continued iteratively until both of them converge. Normalized cuts are also a segmentation algorithm which works to attain optimal partitioning of the image. Image segmentation methods can be separated by ideas like pixel oriented, contour oriented, region-oriented, model oriented, color oriented and hybrid. Color image segmentation is a crucial operation in image analysis and in many systems for computer vision, image interpretation and pattern recognition.

Anil (2010) used color-based image segmentation with K-means clustering unsupervised algorithms. There is no data on training. In the first stage process, satellite image color separation is enhanced using de-correlation stretching. The regions are grouped into a set of five classes using the K-means clustering algorithm in the second stage process. De-correlation stretching enhances image color separation with significant band-to-band correlation. The exaggerated colors improve visual interpretation and make discrimination of features easier. Mary et al. (2010) developed a cluster of K-means to obtain a non-overlapping and reliable image. Using a genetic algorithm, the optimal range for the clustering value of k-means is obtained. These images are to get a proper segmentation output that is non-overlapping. The fusion of these different images results in a much more reliable output. Van et al. (2003) introduced the concept of noise removal by the use of fuzzy logic. The filter averages a pixel using other pixel values from its neighborhood, but at the same time to take care of important image structures such as edges. Due to noise and image structure, the filter distinguishes between local variations.

II. MAGNETIC RESONANCE IMAGING

Magnetic resonance imaging (MRI) uses a magnetic field and radio wave energy pulses to take pictures of organs and structures inside the body. MRI provides different information about the body's structures. MRI can show problems that can't be seen with other imaging methods. Figure 1 shows a human brain's MRI scan.

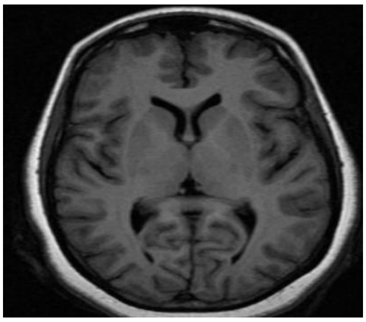


Fig. 1 MRI scan of a brain

III. WAVELET TRANFORM

Wavelet transform converts time domain to frequency domain. In images, the pixel variations are mapped in terms of high frequency and low frequency components. Wavelet transform segments the image based on edge detection. Wavelet transform performs noise reduction on the image. The two-dimensional discrete wavelet transform functional block diagram is shown in Figure 2.

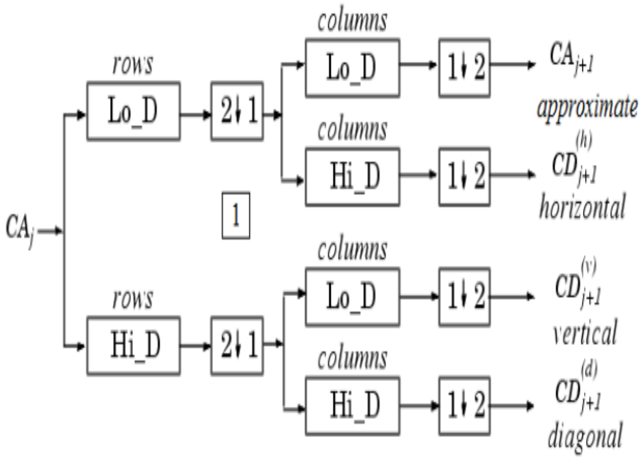


Fig. 2 Two dimensional discrete wavelet transform

The image is divided into 4 components, namely horizontal, vertical, diagonal and approximate, in two dimensional discrete wavelet transform. The diagonal component is done by taking the high frequency components in the rows and columns. Noise in an image is in the form of components of high frequency. So, by analyzing the diagonal coefficient and reducing noise on it can significantly improve the quality of the image. Also, this process can be applied through many more steps thus enabling us to identify the noise component more distinctly.

IV. SEGMENTATION

Segmentation refers to the process of partitioning a digital image into multiple regions (sets of pixels) with the aim of segmentation to simplify an image to be analyzed more meaningfully. Image segmentation is used to locate objects and boundaries (lines, curves, etc.) in images. The result of image segmentation is a set of regions or contours that cover the entire image collectively. Each of the pixels in a region is similar to some characteristic. With respect to the same characteristic(s), adjacent regions are significantly different. The different methods used for segmentation are threshold-based methods, clustering methods, compression-based methods, histogram-based methods, edge detection, region-based segmentation.

V. K- MEANS CLUSTERING

The K-means algorithm is an iterative technique used to divide an image into K clusters.

K-means algorithm:

1. Select K cluster centers either randomly or on the basis of some heuristic
2. Assign every pixel in the image to the cluster which minimizes the length between the pixel and the cluster middle
3. Recalculate the cluster centers by averaging all the pixels in the cluster
4. Repeat steps 2 and 3 until convergence is attained until no pixels change clusters.

The distance is the square or absolute difference between a pixel and a cluster centre. The variance is typically based on pixel color, intensity, texture and location, or a weighted combination of these factors. This algorithm is guaranteed to converge, but the optimal solution may not be returned. The solution's quality depends on the initial set of clusters and the value of K.

The given pixel belongs close to which cluster is found out by

$$\arg\min_{\mathbf{S}} \sum_{i=1}^{k} \sum_{\mathbf{x}_j \in S_i} \|\mathbf{x}_j - \boldsymbol{\mu}_i\|^2$$

Disadvantages:

1. Difficulty in comparing the quality produced by the clusters. Initial partitions or values of K affect the result.
2. Fixed number of clusters can make it difficult to predict the K value.
3. The algorithm does not work well with non-globular clusters.
4. Different initial partitions can result in different final clusters.

VI. FUZZY C- MEANS CLUSTERING

Fuzzy C-means clustering (FCM) allows the objects to belong simultaneously to several clusters with different membership degrees. FCM is a non-supervised technique (Malathy et al. 2016). Objects on multiple class boundaries are not forced to fully belong to one of the classes. Pixels are assigned degrees of membership between 0 and 1 indicating their partial membership. The algorithm is based on minimizing the objective function.

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^{m} \left\| x_i - c_j \right\|^2$$

where m is real number greater than 1. u_{ij} is the degree of membership of pixel i in cluster j. c_j is the center of the cluster j.

Advantages:

1. Gives a good result for overlapped data set.
2. Data point is assigned to each cluster center as a result of which data point may belong to more than one cluster center.

FCM algorithm:

1. Initialize $U=[u_{ij}]$ matrix, U(0)
2. Calculate the cluster center vector with u^k

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} \cdot x_i}{\sum_{i=1}^{N} u_{ij}^{m}}$$

3. Update u^k, u^{k+1}

$$u_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}$$

4. If $\|u^{j+1} - u^j\| < 1.5$, or the time is more than 100, then stop. Otherwise go back to step 2. 1.5 is a controlling parameter.

VII. METHODOLGY

Figure 3 shows the methodology's block diagram.

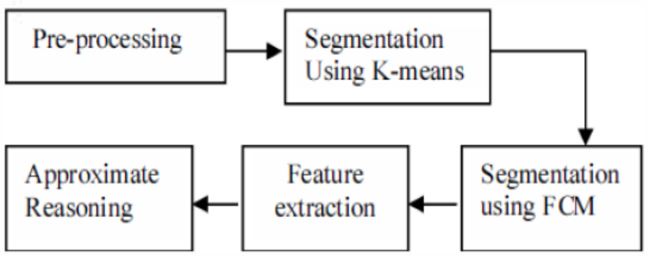


Fig. 3 Block Diagram

Pre-processing

To reduce noise, read the required MRI image. The image needs to be resized so that its resolution is a multiple of 4 for wavelet transformation to give a credible output. The image is converted to grayscale image if it is in RGB since the image is to be processed in only 1 plane and not in all 3. The image is converted to double format because the unit 8 format has a maximum value of 255. Then wavelet transform of five step decomposition and inverse discrete wavelet transform are applied on the image and the noise reduced output image is extracted.

Segmentation using K- means

A user defined function is created to perform the clustering. This function accepts the input image and the number of clusters required. Clustering need not be done based only on position but also on gray level. Number of gray levels equal to the number of required clusters is generated. Then matrices are established to map every gray level‘s distance from each randomly generated centroid. The closest cluster to each gray level is found out figuring out to which cluster the given gray level is associated to.

Segmentation using FCM

Each pixel can either be partially associated or totally associated to a given cluster. So, the points at the center are totally associated into the given cluster and ones on the edges are partially associated. An inbuilt function is used to perform the clustering. The input to this function is a matrix combination of the gray level image and its double counterpart. [2, (m x n)] matrix is obtained, where m and n are the dimensions.

Feature Extraction

The tumor from the clustered image is segmented. Pixels of similar intensity are grouped together. But this will only consist of small blobs and will not be a continuous stretch. Morphological transformation is used wherein the segments below a certain number of pixels of the image are opened.

*Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication*

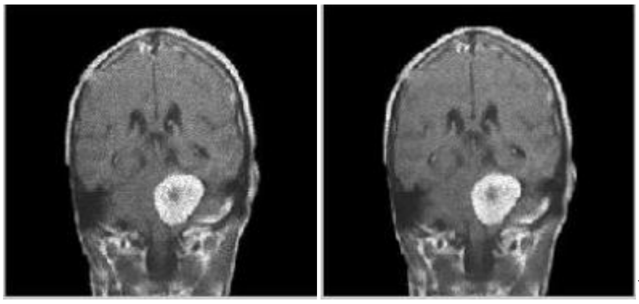
A tumor is there as the threshold is set. The mask obtained is applied onto the original image to view the tumor.

Approximate Reasoning

The tumor is in its critical stage or not is ascertained by calculating the area of the tumor. The mask from the previous step is used to determine the area. The mask consists only of white and black pixels where the white pixels depict the tumor portion and the black pixels represent the normal portion of the brain. The tumor area can be computed by finding the number of white pixels and multiplying the number by the scaling factor.

VIII. RESULTS

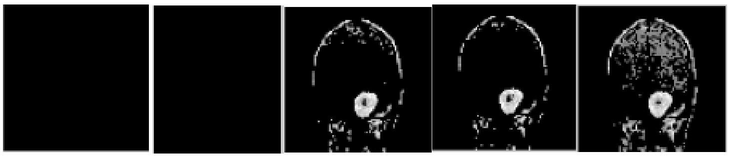
Figure 4 shows the preprocessing output.



Noisy Image Filtered Image

Fig. 4 Preprocessing Output

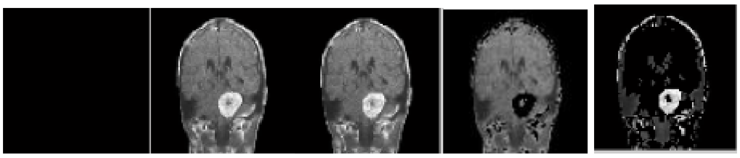
Figure 5 shows the preprocessing output.



K=1 K=2 K=3 K=4 K=5

Fig. 5 K-means segmentation output

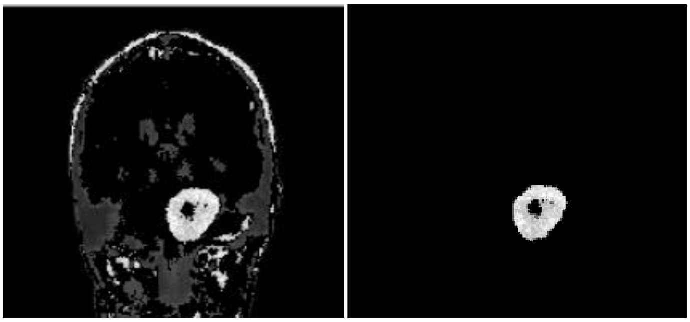
Figure 6 shows the FCM segmentation output.



K=1 K=2 K=3 K=4 K=5

Fig. 6 FCM segmentation output

Figure 7 shows the tumor segmentation output.

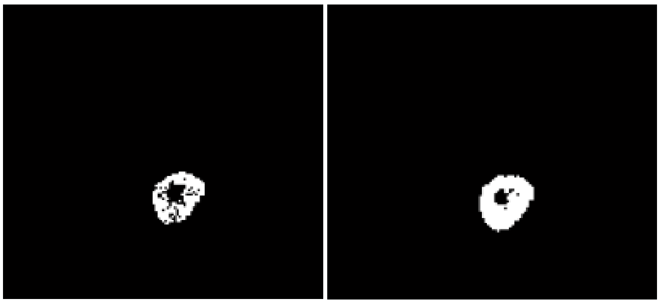


Matrix to gray conversion Brain tumor segmentation

Fig. 7 Tumor segmentation

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Figure 8 shows the feature extraction output.



For threshold level 1 **For threshold level 2**

Fig. 8 Feature extraction output

Figure 9 shows the output of approximate reasoning

Edge of brain tumor **Segmented edge of brain tumor**

Fig. 9 shows the output of approximate reasoning

IX. CONCLUSION

The main purpose of this article is to ascertain the presence of the tumor as well as quantify it. This is achieved by using segmentation methods based on clustering. Both K-means and FCM clustering methods were used to segment the tumor. FCM clustering is better than K-means clustering due to its fuzzy assigning of the pixels to each cluster. This helped to isolate the tumor portion more effectively. The paper‘s aim is also to find out if the tumor has reached critical stage or not. This is done by taking the scale of the MRI and using the number of pixels isolated. The other important aim of the paper is to remove noise effectively to get the best possible segmented output. For better detection, discrete wavelet transform is used to remove multiplicative noise from the image. The detection of the tumor from the original image and filtered images are compared. The results clearly indicate that the filtered image gives better results.

X. FUTURE DEVELOPMENTS

The real world implementation by interfacing with already available technology can be done by running the code right onto the obtained image so as to get the tumor segmented instantly. It can also be implemented into three dimensional images. Though working with three dimensional images is more time consuming, it is far more accurate as well.

REFERENCES

1. Manisha Bhagwat, Krishna R K &Pise V E (2010), ‘Image Segmentation by Improved Watershed Transformation in Programming Environment MATLAB’, in International Journal of Computer Science & Communication Vol. I, No. 2, pp. 17/-/74.
2. Suman Tatiraju A (2008), ‘Image Segmentation using k-means clustering, EM and Normalized Cuts’ in Symposium of Discrete Algorithms
3. Anil Z Chitade (2010), ‘Color based image segmentation using k-means clustering’ in International Journal of Engineering Science and Technology Vol. 2(10), 5319-5325
4. Mary Praveena S & Dr.IlaVennila (2010), ’Optimization Fusion Approach for Image Segmentation Using K-Means Algorithm’ in International Journal of Computer Applications (0975 - 8887) Volume 2 - NO.7
5. Van De Ville D, Nachtegael M, Van Der Weken D, Kerre E E, Philips W &Lemahieu I (2003), ‘Noise reduction by fuzzy image filtering’ in IEEE Transactions on Fuzzy Systems I I
6. Malathy V & Nirmala S 2016, ‘Segmentation of Lumbar Vertebral Bodies and Intervertebral Disc in CT Images’, Asian Journal of Research in Social Sciences and Humanities, vol. 6, no. 6, pp. 360-371.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication