Brain Tumor Segmentation and Dimension Determination Using Clustering Approach

Harendra Singh1, Dr. Rajeev Ratan2

1,2 MVN University Palwal, India

Abstract— To diagnose the brain tumor in early stage in Magnetic Resonance Imaging (MRI) scans, clustering approach is implemented. Every year a large number of peoples have died due to cancer. Therefore, it becomes a serious and life-threatening issue in this fast-changing modern era. After a data collection a pre-processing is performed to make the MRI images compactible to the algorithm, which includes image resizing, filtering and converting to grayscale. Fuzzy C-means (FCM) required a lot of time for convergence of pixels in final clusters but it provides a much better result as compared to k-means clustering, when pixels belong to more than one cluster. The performance of FCM is improved by selecting an optimal cluster and membership function in terms of reducing the computation time for tumor segmentation. This implemented clustering technique are able to detect the tumor area (no. of pixels) and its stage. Some performance measure features like sensitivity, specificity, accuracy, and error rate are considered for model performance and using Discrete Wavelet transformation (DWT) brain tumor features are measured.

Keywords— MRI Scan , Pre-processing, KCM & FCM Clustering, Discrete Wavelet Transformation.

1. Introduction

For imaging the internal structure of human brain many medical machines are avilabe but MRI scan[3] is used because it can provide more accurate information for clinical purpose without any radioactive effect [1]. In this era of modern technology, There is a need of a fully automatic system in which there is no manual interference to get the result. Different techniques were developed for tumor segmentation such as edge detection, region growing, watershed, but there is always scope for improvement for tumor detection and extraction. In this paper, K-means and the FCM technique has been used to segment the brain tumor, which are gives a much better result as compared to existing tumor segmentation techniques[2]. The specific application in biomedical imaging of this approach is that after detecting the tumor region and its size radiologists can easily plan about patient treatment.

For the existence of any living organ, the cell is the smallest and functional unit. The generation of new as well as replacement of dead cells carries out in a controlled manner. Due to malfunctioning of this cell generation system extra cells are generated instead of one and when these cells gain some mass called tumor[9]. The tumor at the initial stage known as benign and when the tumorous cells start to spread to another part of the body through bloodstream, then known as malignant. Normally, brain tumors are located in the tissue present inside the skull, which is the backbone of the central nervous system (CNS) and it affects the Cerebral Spinal Fluid (CSF) of human beings that causes the strokes [4]. Therefore, at this time, all physicians prescribe the medicines to control this sudden coming stroke rather than the treatment related to the tumor. Nowadays, clustering is the most widely used technique for abnormal region detection in MRI images because due to more flexibility, which allows the pixels to be a large number of a cluster so the degree of membership is varied, but fuzzy c-means (FCM) is a time-consuming technique. Therefore, there is always a scope for further improvement in the detection of tumors in the early stages of a healthy life.

2. Existing Method

In this modern era, there are lot of methods to diagnose the tumor in an MRI scan or CT scan but most of the techniques are based on resolving and region growing method. Now a day's thresholding method is avoided because it is not able to diagnose the spatial characteristics of an MRI scan, which is the most important factor for diagnosing the malignant tumor. The threshold segmented image [P(x)] is based on two values either zero (Black) or one (white) according to the assigned threshold level (T). Therefore, pixel value '0' is assigned to the background of the image and '1' foreground of an image [4]. Therefore, sometimes it is not able to tumor cells. The typical thresholding operation on an input image shown below.

$$p(x) = \left\{\begin{matrix} 1 & if\ x \geq T \\ 0 & if\ x > T \end{matrix}\right\} \qquad \ldots (1)$$

For multiple thresholding according to assigned pixels, values the input image is segmented in n segments. The values of [p(x)] may be shown as;

$$p(x) = \left\{\begin{matrix} 0 & if\ x < t_1 \\ 1 & if\ t_1 \leq x < t_2 \\ 3 & if\ t_2 \leq x < t_3 \end{matrix}\right\} \qquad \ldots (2)$$

In this approach center of tumorous tissues is randomly selected i.e. also known as a seed, it is an iterative process, when neighboring pixels follow the predefined criteria of threshold, then the region is growing otherwise pixel is not counted[7]. The whole process is terminated, when all the pixels belong to the same region. Therefore, large number of the interaction of user is required and time consuming.it may be cause an intensity for homogeneity related issues and not may provide acceptable results for all types of images.

3. Proposed Methodology

The proposed method to diagnose the tumor in Brain MRI scan mainly consists of four modules; MRI database, pre-processing module, image segmentation module (K-Means & FCM), and feature extraction with approximate reasoning. Figure 1, shows a block diagram of proposed model.

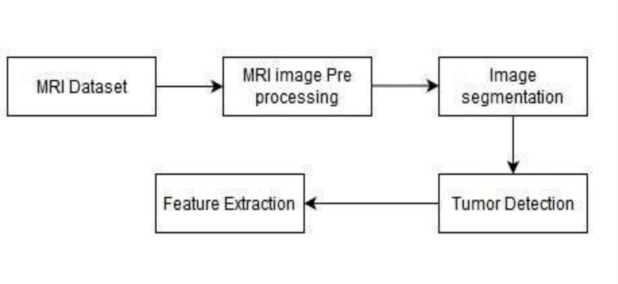


Figure 1. Block diagram Proposed Model

3.1 MRI Dataset

Brain tumor MRI dataset consisting of four types of tumors images for brain tumor detection with 256*256 gray levels having intensity range from ‘0’ to ‘255’ are considered for practical implementation. First of all, collected MRI images are normalized to gray level values from ‘0’ to ‘1’ and due to normalization, the dynamic range of images is reduced so that feature extraction becomes much simpler. Figure 2, shows a samples of the dataset of different type of tumor in MRI scans.

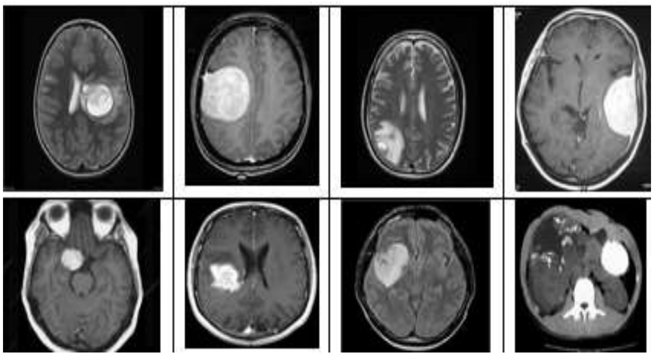


Figure 2. Sample data set: Meningioma, Metastate, Glioma, and Astrocytoma

3.2 MRI Image Pre-Processing

Image pre-processing is a technique, in which by performing some operations like Image-filtering, histogram equalization, histogram adjustment, image enhancement, and image quality is improved so that fine features of an image can easily be extracted. Figure 3, shows a samples of pre-processed MRI images with a histogram.

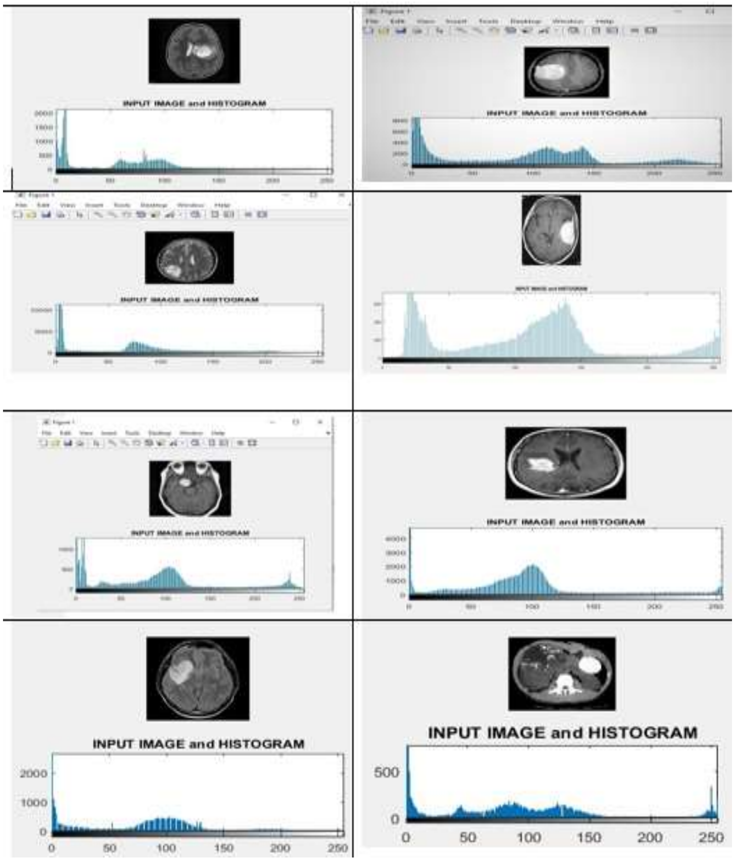


Figure 3. Pre-processed Image with histogram

3.3 Image Segmentation

Image segmentation is a process used to extract the most important features required brain tumor area detection and on the basis of area tumor the stage of tumor is also predict. In modern MRI scans, the possibility for the acquisition of noise is very less but due to thermal effect, some noise is added during image acquisition. Therefore, a median filter gives a better result as compared to others for removing the error. There are many image techniques which are used for tumor segmentation in MRI images but in the proposed system, KCM & FCM clustering approach is used to segment the tumor in minimum number of iteration with high level of accuracy.

3.4 K- Means clustering

It is an un-supervised clustering approach in which randomly the number of clusters is taken then pixels are moved to that cluster which has less Euclidian distance to that particular pixel [5]. Then again, new cluster centers has been calculated. This process is terminated when all pixels belong to the clusters. K-Means clustering is not suitable for overlapping clusters [11]. The Figure 4, shows a flow chart of the k-means (KCM) algorithm.

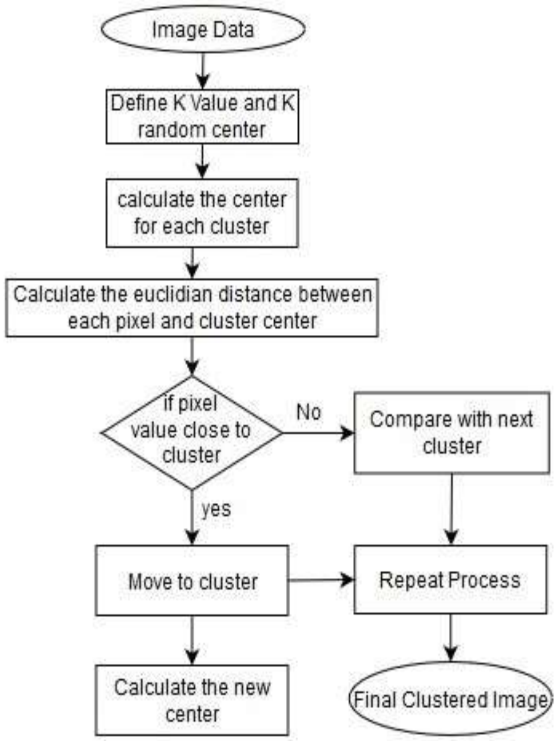


Figure 4. Flowchart of the k-means algorithm

Basic steps of the k-means algorithm

- Determine the Euclidian distance using equation 3.

$$D(i) = arg\ min\|x_i - M_k\|^2 \quad , i = 1,2 \ldots N \qquad \ldots (3)$$

- Compute the new mean values of each cluster

$$M_k = \frac{\sum_{i:c(i)=k} x_i}{n_k} \quad , k = 1,2 \ldots., k \qquad \ldots (4)$$

- Repeat the above steps until the all the pixels are convergence

3.5 Fuzzy C- Means (FCM) clustering

When image pixels of more than two clusters belong to each other than FCM can give better results to diagnose the object [5], [8]. A membership value (‘0’ or ‘1’) are to are assigned to each data point based on the distance of cluster center and data point

Basic steps of Fuzzy c-means algorithm

- Choose ‘c’ (cluster) between ‘2’ to ‘n’ and ‘m’ (fuzziness parameter) greater than ‘1’.
- Calculate the center of each cluster;

$$c_j = \frac{\sum_{k=1}^{n} M_{ik}^{m} p(r) r}{\sum_{k=1}^{n} M_{ik}^{m}} \qquad \ldots (5)$$

- Calculate the convergence matrix of each cluster

$$G_i = \frac{\sum_{k=1}^{n} M_{ik}^{m} (x_k - c_i)(x_k - c_i)^T}{\sum_{k=1}^{n} M_{ik}^{m}} \qquad \ldots (6)$$

- Calculate the new membership values using

$$M_{ik} = \frac{1}{\sum_{j=1}^{n} \left(\frac{\left| (x_k - c_i)^T G_i (x_k - c_i) - ln|G_i| \right|}{\left| (x_k - c_j)^T G_i (x_k - c_j) - ln|G_i| \right|} \right)^{\frac{2}{m-1}}} \qquad \ldots (7)$$

- Above process ends when

$$max_{ij}\left\{\left|M_{ij}^{(k+1)} - M_{ij}^{k}\right|\right\} < \delta \qquad \ldots (8)$$

Where δ a termination value i.e. lie in the range = $0 < \delta < 1$

4. Feature Extraction And Appropriate Reasoning

The feature extraction process is based on the dynamic intensity variation in the segmented MRI scan. The overall extracted cluster is processed through thresholding operation. In the threshold process, the intensities of segmented image are redistributed according to threshold values so that after threshold operation it become as binary image. Tumor area is easily calculated with the help of approximate reasoning, which makes the dark pixels to be darker and white pixels to be brighter. The intensity values [19] according to threshold are set to ‘0’ or’1’ as shown by equation 9.

$$p(x) = \begin{Bmatrix} 1 & if\ x \geq T \\ 0 & if\ x > T \end{Bmatrix} \qquad \ldots (9)$$

The binary image having only two levels of intensity like as ‘0’ for black and’1’ for white.[13] The area of tumor from the binary image calculated by using equation 10 as shown below.

$$Image\ (P(x,y) = \sum_{w=0}^{255} \sum_{H=0}^{255} [P(0) + P(1)] \qquad ... (10)$$

Where

Size of Pixels = width*Height =255*255
P (0) = white pixel
P (1) = Black pixels

Total number of white pixels in binary image

$$P_{total} = \sum_{w=0}^{255} \sum_{H=0}^{255} [f(0)] \qquad (11)$$

Where

P (0) = (width*height) = 0.264 mm
$Area\ of\ tumor$ = [($\sqrt{P_{total}}$) *0.264] mm^2

4.1 Performance measures

In this paper, to measure the performance of clustering approach for brain tumor segmentation is evaluated based on following parameters as shown below [5].

Mean Square Error

$$MSE = \frac{1}{rc} \sum_{i=0}^{r-1} \sum_{j=0}^{c-1} (O(i,j) - F(i,j)^2 \qquad \ldots (12)$$

Where;

O (i, j) = MRI input database
F (i, j) = Output MRI database
‘r’ & ‘c’ = No. of rows and columns.

Peak Signal to Noise Ratio

$$PSNR = 10 log_{10} \left(\frac{max_I^2}{MSE} \right) \qquad \ldots (13)$$

Sensitivity, Specificity, Accuracy, and Error rate are considered for performance scale

$$\%\ Sensitivity = \frac{(T^{+VE})}{(T^{+VE}) + (F^{-VE})} * 100 \qquad \ldots (14)$$

$$\%\ Specificity = \frac{(T^{-VE})}{(T^{-VE} + F^{+VE})} * 100 \qquad ... (15)$$

$$\% \, Accuracy = \frac{(T^{+VE}+T^{-VE})}{(T^{+VE}+T^{-VE}+F^{+VE}+F^{-VE})} * 100 \qquad \ldots (16)$$

$$\% Error \ Rate = \frac{(F^{+VE}+F^{-VE})}{(T^{+VE}+T^{-VE}+F^{+VE}+F^{-VE})} * 100 \qquad \ldots (17)$$

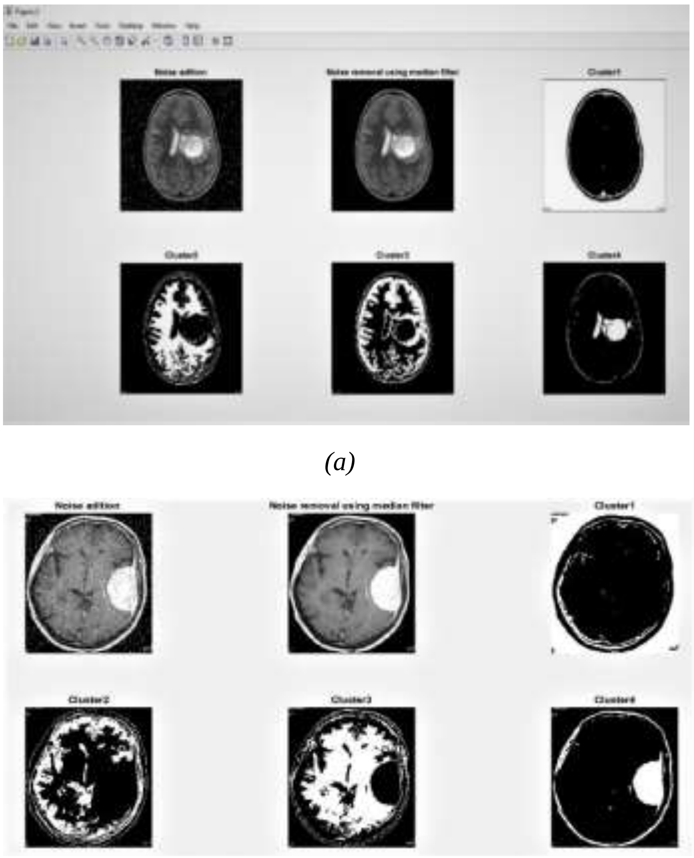
Where;

T^{+VE} = Malignant classified as Malignant tumor is correct
T^{-VE} = Benign classified as Benign tumor is correct
F^{+VE} = Benign classified as Malignant tumor is not correct
F^{-VE} = Malignant classified as Benign tumor is not correct

5. Result

MATLAB 2015b platform with Core i5 with hybrid 2.5 GHz speed is used to implement the proposed model for brain tumor diagnosis and classification. The MRI dataset is downloaded from an open source to everyone i.e. https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection/data.

Screenshot for pre-processing and k-means algorithm



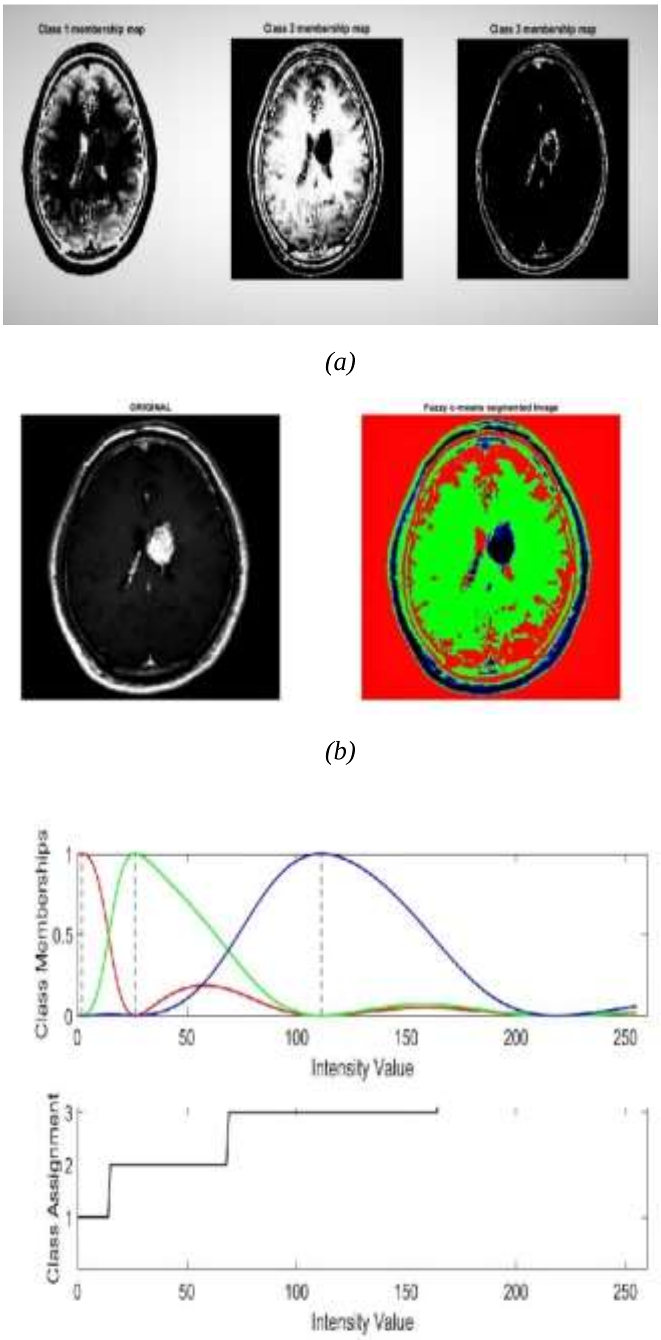
(a)

(b)

Figure 5(a) and (b). Output Image form pre-processing to tumor detection using k-map, for k=4

In figure 5 (a) & (b) first MRI image is first pre-processed and tumor is segmented by k-means algorithm in 4th cluster.

Screenshot for pre-processing and Fuzzy c-means algorithm



(a)

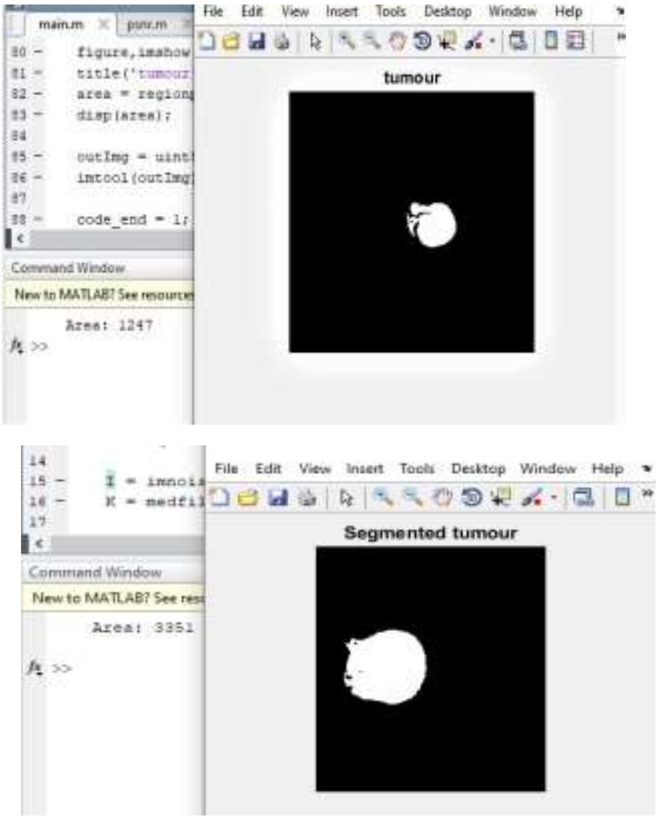
(b)

(c)

Figure 6. Output Images of FCM with membership function

The output image of the fuzzy c-means (FCM) algorithm, which gives better segmentation of tumorous cells in an image as compared to k- means due to the overlapping effect of pixels.

Screenshot for area calculation of segmented image using a clustering approach



```
figure,imshow
title('tumour
area = regiong
disp(area);

outImg = uint
imtool(outImg

code_end = 1;
```

```
Area: 1247
```

```
I = imnois
K = medfil
```

```
Area: 3351
```

Fig. 7 (a) & (b) Calculated tumor area in MRI Image

For a small sample of 10 mixed benign and malignant brain tumors images performance measures are calculated as shown in table 1.

Table 1. performance measures of brain tumor detection

Sr. No.	Mean Square Error	Peak Signal to Noise Ration	Processing Time (s)	Segmented Tumor size (Pixels)
1	0.2337	55.2641	15.215	1247
2	0.0967	57.2410	8.012	3351
3	0.1983	56.1246	25.954	6259
4	0.1015	58.1245	18.012	3118
5	0.3468	52.1942	8.999	1289
6	0.0537	60.7625	9.019	3824
7	0.0658	61.7470	18.235	1905
8	0.0574	59.0165	12.985	768
9	0.0684	66.3124	7.019	3027
10	0.0522	60.9499	29.216	15148

In figure 8, the last column shows the segmented brain tumor size (pixels) in MRI images using clustering approach.

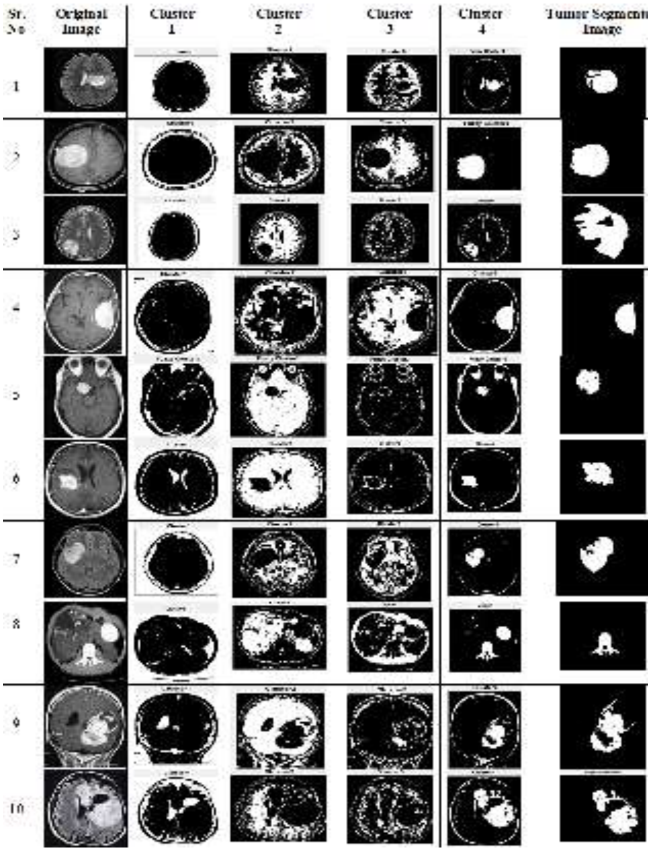


Fig. 8 (i) MRI Image (ii) custer-1 (iii) cluster-2 (iv) cluster-3 (v) cluster-4 (vi) Tumor segmented Image

The proposed model is able to detect the tumor size in MRI scan with good accuracy [5] and predict the class of tumor like benign or malignant according to area size, some features are calculated like as: mean, standard deviation, entropy, root mean square, variance, smoothness, kurtosis, skewness, IDM, contrast, correlation, energy, and homogeneity shown in Table 2.

Table 2. All extracted features for a brain MR Image [5]

S. No.	**Extracted features**	**Corresponding values**
1	Mean	.0047
2	Standard Deviation	.0856
3	Entropy	1.7554
4	Root Mean Square	.0857
5	Variance	.0073
6	Smoothness	.9546
7	Kurtosis	24.7162
8	Skewness	2.1895
9	IDM	1.6994
10	Contrast	.3557
11	Correlations	.1266
12	Energy	.8499
13	Homogeneity	.9555

All (13) features are extracted for 10 benign and malignant images from MRI database and used to train the classifier. In the testing phase, 90 images (50 Benign & 40 Malignant) are used. 49 images are classified as benign correctly and 1 images as wrongly benign out of 40 malignant images. The performance of the classifier is shown in table 3.

Table 3. Performance measures of classifier

Sr. No.	Performance parameter	Benign tumor (%)	Malignant Tumor (%)
1	Sensitivity	98	97.5
2	Specificity	97.5	98
3	Accuracy	97.77	97.77
4	Error Rate	2.23	2.23

6. Result Analysis

This modified clustering technique is able to segment the size of brain tumor tissues in MRI images with a good accuracy in minimal processing time. When a tumor gains some mass then size and exact location of tumor is detected using KCM. The results of FCM technique is much better than K-means clustering (KCM) because it is soft clustering approach in which pixels of one cluster may be a part of other cluster is also considered. Therefore, it is a time-consuming process but gives a more accurate segmented area of tumor in MRI images. When the area of the tumor is determined then the classification of types of tumors is easy with Accuracy 97.77%, Sensitivity 98%, 97.77% of Specificity, and 2.23 % of Error rate in experimental result.

7. References

[1] M.Shasidhar, V.S raja, and B. V. Kumar, "MRI Brain Image Segmentation using modified Fuzzy C- Means Clustering Algorithm," *in IEEE Int'l Conf. on Communication System and Network Technology,* 2011 pp. 473-478.

[2] F. Liao and C.Zhao, " Improved Fuzzy C- Means Clustering for Automatic Detection of Lung Nodules," *in IEEE Int'l Conf. on Image and Signal Processing, China,* 2015 pp. 464-469.

[3] Praveen and A. Singh, "Detection of Brain Tumor in MRI Images, Using Fuzzy C- Means and SVM, " *in IEEE Int'l Conf. on Signal Processing and Integrated Networks,* 2015 pp. 98-102.

[4] J. Selvankumar, A. Lakshmi, and T. Arivoli, "Brain Tumor Segmentation and Its Area Calculation in Brain MR Images using K- Means Clustering and Fuzzy C- Means Clustering," *in IEEE Int'l Conf. on Advances in Engineering, Science and Management (ICAESM-2012),* March 2012, pp. 186-190.

[5] B. Srinivas and G. S. Rao, " Performance Evaluation of Fuzzy C- Means Segmentation and Support Vector Machine Classification for MRI Brain Tumor, " *in Springer Nature Singapore Pte Ltd. on Soft Computing for Problem Solving,Advances in Intelligent System and Computing,* March 2019, pp. 355-367.

[6] R. Ratan, P. G. Kholi, S. K. Sharma, and A. Kholi, "Unsupervised segmentation and quantization of malignancy from breast MRI images," *in Journal of the National Science Foundation of Sri Lanka,* vol. 44, no. 4, pp. 437-442, December 2016.

[7] A.R. kavita, Dr.C. chellamuthu, and K. Rupa, "An Efficient Approach for Brain Tumor Detection Based on Modified Region Growing and Network in MRI Images," *in IEEE Trans. on Information Forensics and Security,* vol. 9(2), May 2012.

[8] Md. S. Alam, Md. M. rahman,Md. A. Hossain,Md. K. Islam,K.M. Ahmed,K.T. Ahmed,B.C.singh, and md. S. Miah, "Automatic Human Brain Tumor Detection in MRI Image Using Template-Based K Means and Improved Fuzzy C Means Clustering, *in Big Data Cogn. Comput. 2019, 3, 27; doi:10.3390/bdcc3020027.*

[9] S. Kaushal, "An efficient brain tumor detection system based on segmentation technique for MRI brain images," *Int'l Journal of advanced research in computer science,* vol. 8, no. 7, pp. 1131-1136, March 2017.

[10] M. P. Arakeri and G. R. M. Reddy, "Efficient Fuzzy Clustering Based Approach to Brain Tumor Segmentation on MR Images," CIIT 2011, CCIS 250, pp. 790–795, 2011. © Springer-Verlag Berlin Heidelberg 2011.

[11] J.Vijay and J.Subhashini , "An efficient brain tumor detection methodology using k-means clustering algorithm," IEEE Int'l Conference on Communication and Signal Processing, April 3-5, 2013.

[12] A. Das and M. Bhattacharya, "GA based neuro fuzzy techniques for breast cancer identification," *IEEE Int'l Conf. on machine vision and image processing,* vol. 08, 2008, pp. 136-141.

[13] G.M. Xian, "An identification method of malignant and benign liver tumors from Ultrasonography based on GLCM texture features and fuzzy SVM*," Journal of Elsevier Expert Systems with Applications*, vol. 37, Issue 10, pp. 6737–6741, February 2007.

[14] P. N. H. Tra, T. H. Nguyen, and T.T. Mai, "Image Segmentation for Detection of Benign and Malignant Tumors," *in Proc. IEEE Int'l Conf. on Biomedical Engineering (BME-HUST),vol.3 ,oct 2016*, pp. 51-54.

[15] R.S.P. Saranya and C. Lokanayanki, "A Survey Paper on Fuzzy Image Segmentation Techniques," *Int. Journal of Engineering Research and Applications*, vol. 4, Issue 3(Version 1), pp. 429-434, 2014.

[16] B. Fatima and C.M. Amine, "A neuro-fuzzy inference model for breast cancer recognition," *Int'l Journal of Computer Science & Information Technology (IJCSIT),* vol. 4, no. 5, pp.163-173, 2012.

[17] N.M. Ahmed, S.M. Yamany, N. Mohamed, A.A. Farag ,and T. Moriarty , "A Modified Fuzzy C-Means Algorithm for Bias Field Estimation and Segmentation of MRI Data," *IEEE Trans. on medical imaging,* vol. 21, no. 3, march 2002,pp.193-199.

[18] T. Manikandan and N. Bharathi, "Lung Cancer Diagnosis from CT Images Using Fuzzy Inference System," *in journal of Communications in Computer and Information Science*, vol. 5, pp. 474-479, July 2015.

[19] F.S. Ahadi, M.R. Desai, C. Lei, Y. Li, and R. Jia, "Feature-Based classification and diagnosis of breast cancer using fuzzy inference system," *in Proc. IEEE Int'l Conf. on Information and Automation (ICIA) Macau SAR, China,* October 2017 pp. 517-522.

[20] A. Moghaddamzadeh and N. Bourbakis, "A Fuzzy Region Growing Approach for Segmentation of Color Images," *in journal Elsevier science Ltd on Pattern recognition society,* vol. 30, no. 6, pp. 867 -881, 1997.