

## **Tumor Stages Detection in Brain MRI Image using Template based K-means and Fuzzy C-means Clustering Algorithm**

Rasel Ahmmed<sup>1</sup> and Md. Foisal Hossain<sup>2</sup>

*In this research work, we have demonstrated a robust method integrating Template based K-means and modified Fuzzy C-means (TKFCM) clustering algorithm incorporated with binary image region properties. This method reduces the inadequacy of human perception for the brain Magnetic Resonance Imaging (MRI) images by the automatic detection of tumor stages based on the tumor area. Firstly, proper segmentation is accomplished based on template selection on the basis of gray level intensity in K-means. The updated membership through automatic clusters and features selection is performed in modified fuzzy c-means (FCM) clustering algorithm. Afterwards, the linearization of membership image from TKFCM provides fruitful region properties like area, eccentricity, perimeters etc. to brain tumor. The tumor image affords characteristic parameters, which is compared to the standard tumor area and automatically recognized the different tumor categories and stages. The results show that one of the tumors is benign and some are malignant with several stages. This method will manipulate experts with identified tumor stages to apply different techniques in further treatment.*

**Keywords:** Brain tumor; tumor stages; magnetic resonance imaging (MRI); template based k-means and modified fuzzy c-means clustering (TKFCM); feature selection.

**Field of Research:** Image Processing

### **I. INTRODUCTION**

Brain is the main portion of our body, where the memory, emotion, speech, and control of whole body are centered. Therefore, the presence of tumor in the brain causes problem for the malfunction of the body, even for the muscle and eyebrow movement. The tumor can paralyze the major areas of brain which is shown in Fig. 1. Among several imaging techniques magnetic resonance imaging (MRI) is gold accepted truthful modality. In MRI technique, brain is imaged on the basis of density of water in soft tissue which is higher compared to other tissues such as bone [1]. Due to the inhomogeneity of brain structure their contrast value differs randomly in MRI, so the proper detection of tumor or tumor size is quite difficult [2].

There are two categories of brain tumor like- (a) Benign and (b) Malignant [3]. The first one behaves like a normal tissue which does not spread. But the last one has four stages such as- i) stage I- tissue area is like normal tumor which just congested in an area, ii) stage II- tumor with or without lymph node but not enough spreading, iii)

---

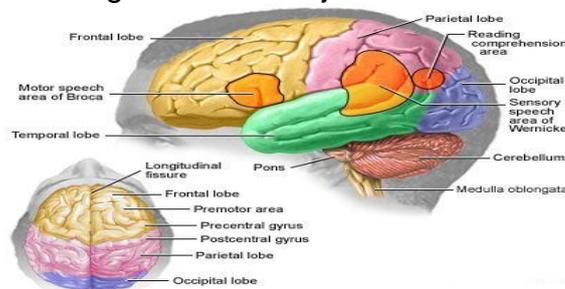
<sup>1,2</sup>Department of Electronics and Communication Engineering  
<sup>1,2</sup>Khulna University of Engineering & Technology, Khulna-9203, Bangladesh  
E-mail:<sup>1</sup> rasel\_kuet2k8@yahoo.com and <sup>2</sup> foisalkuet@yahoo.co

stage III- tumor fixed or spread enough above collarbone, and iv) stage IV- tumors are spread to other parts (metastasis) of body [4].

This spreading tumor causes a hilarious abnormality throughout the body. That is why, proper detection of brain tumor and tumor size with different properties is quite essential for observing the stages of tumor. It helps the experts for further treatment process.

There are some research works in [5-11] used techniques like region growing, thresholding, classifiers, Artificial Neural Networking (ANN), and clustering for segmentation. Based on these aforementioned techniques we can detect only brain tumor. The separation of objects from background is quite difficult for that tissue whose intensity is alike background, is the main drawback of thresholding [8]. The applications of the Expectation Maximization (EM) algorithm to brain MR image segmentation and a common disadvantage of EM algorithms is reported in [9]. In the region growing method, the primary limitation is its seed point finding through manual interaction [10-11]. Furthermore, some works describes only for tumor classification in [12-15]. Due to having several limitations, aforesaid individual method cannot be used properly to detect tumor as well as its stages. Since it is foremost requirement of medical science is to detect tumor size and its stage, it is obligatory to segment proper tumor. As a result, there arises demand of an appropriate method which can be able to segment and detect actual tumor size, simultaneously.

Figure 1: The major areas of brain



In this paper, we have proposed a different technique for tumor size and stages classification from brain MRI image, integrating template based k-means and modified FCM (TKFCM) with some measurement properties in tumor region. This paper describes that, template based k-means extension has applied for the segmentation of tumor based on their gray level intensity. In addition with, FCM algorithm has developed depending on the behavior and number of membership inputs and outputs based on the cluster number and size. Furthermore, the area of the tumor differs from each other, which enable us to make a decision about tumor stages based on the other properties of tumor like eccentricity, perimeters, and bounding box. Comparing these values with the standard tumor area value, we have defined the various tumor stages.

This paper is organized as follows, in section II conventional k-means and fuzzy c-means algorithm is present, the proposed algorithm describes step by step in section III, finally the results and conclusion is shown throughout section IV & V respectively.

## II. K-MEANS AND FUZZY C-MEANS ALGORITHM

### A. K-means algorithm

Historically, the K-mean clustering is normally introduced to group a set of data points  $\{x_1, x_2, \dots, x_N\}$  into K clusters [15]. It has high computational efficiency and can support multidimensional vectors. So it reduces the distortion measurement by minimizing a cost function as:

$$J = \sum_{n=1}^N \sum_{k=1}^K b_{nk} \|x_n - c_k\|^2 \quad (1)$$

$$b_{nk} = \begin{cases} 1 & \text{if } k = \arg \min_a \|x_n - c_a\|^2, a=1, \dots, k \\ 0 & \text{Otherwise} \end{cases}$$

$$c_k = \frac{1}{N_k} \sum_{x \in C_k} x$$

Where,  $\|\cdot\|$  measures distance from the center. The center and number of data points in the cluster  $C_k$  is represented by the variables  $c_k$  and  $N_k$ , respectively.

### B. Fuzzy C-means Algorithm

Historically, the FCM clustering algorithm, introduced by Bezdek, is an improvement of earlier clustering methods [16]. It is based on minimizing an objective function, with respect to fuzzy membership, and set  $U$ , of cluster centroids  $V$ :

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

In (2),  $X = \{x_1, x_2, \dots, x_j, \dots, x_N\}$  is a  $P \times N$  data matrix, where  $P$  represents the dimension of each  $x_j$  'feature' vectors and  $N$  represents the number of feature vectors (pixel numbers in the image).  $C$  is the number of clusters.  $U_{ij} \subseteq U(P \times N \times C)$  is the membership function of vector  $x_j$  to the  $i$ th cluster, which satisfies  $U_{ij} \in [0, 1]$  and  $\sum_{i=1}^C U_{ij} = 1, (j=1, 2, \dots, N)$ . The membership function is expressed as,

$$U_{ij} = \frac{1}{\sum_{k=1}^C \left( \frac{d(x_j, v_i)}{d(x_j, v_k)} \right)^{\frac{2}{m-1}}} \quad (3)$$

$V = \{v_1, v_2, \dots, v_i, \dots, v_C\}$  which is a  $P \times C$  matrix which denotes the cluster feature center.

$$v_i = \frac{\sum_{j=1}^N (U_{ij})^m \times x_j}{\sum_{j=1}^N (U_{ij})^m} \quad (i=1, 2, \dots, C) \quad (4)$$

$m \in (1, \alpha)$  is a weighting exponent on each fuzzy membership, which controls the degree of fuzziness  $d^2(x_j, v_i)$  and that is a measurement of similarity between  $x_j$  and  $v_i$ :

$$d^2(x_j, v_i) = \|x_j - v_i\|^2 \quad (5)$$

$\|\cdot\|$  can be defined as either a straightforward Euclidean distance or its generalization such as Mahalanobis distance [16]. The feature vector  $X$  in MR image represents the

pixel intensity  $P=l$ . The FCM algorithm iteratively optimizes  $J_m(U, V)$  with the continuous update of  $U$  and  $V$ , until  $\|U_{ij}^{(l)} - U_{ij}^{(l+1)}\| \leq \epsilon, \epsilon = \{0 \text{ to } 1\}$ , where  $l$  is the number of iterations.

### III. PROPOSED ALGORITHM

Firstly, we have taken an MRI image from our database which we have made through the brain MRI images collected from [18-20]. This image is incorporated with tumor finding with MRI technique. The k-means algorithm is used to segment this MRI images on the basis of gray level. This gray level is selected depending on the temper of the image. Then the modified fuzzy c-means algorithm which depends on the updated membership is applied to segment the template based k-means segmented image. The membership of modified fuzzy c-means is updated with the cluster distances from centroid defining by the features of the tumor MRI image.

Template based k-means and modified fuzzy c-means clustering algorithm which is used for segmentation of the brain MRI images can be written in equation as below:

$$J_{mn} = B(x_i, y_j) \times \sum_{i=1}^K \sum_{j=1}^C P_{ij} \|x_i - c_j\|^2 \times \sum_{j=1}^R \sum_{i=1}^C (U_{ij})^m d^2(x_j, v_i) \quad (6)$$

Where,  $P_{ij}$ , a binary image matrix. The centroid of the cluster, number of data points in clusters and number of cluster is defined by  $R, K$  and  $C$  respectively. In eqn. (6) the last portion is defined as modified fuzzy c-means whose Euclidian distance is depended on the image features. The middle portion is used as the conventional k-means algorithm, which is defined by the distance from each point to cluster center. Here,  $B(x_i, y_j)$  is the coarse image which is marked in describing desired template could be found through below eqn. (7):

$$B(x_i, y_j) = \sum_{i=1}^M \sum_{j=1}^N P(x_i, y_j) \times T_{mn} \quad (7)$$

Template based window is selected by  $T_{mn}$  which is given as:

$$T_{mn} = \sum_{i=1}^M \sum_{j=1}^N P(x_i, y_j) \oplus \sum_{k=1}^G \sum_{l=1}^S P(x_k, y_l) \quad k \in M, l \in N \quad (8)$$

In this eqn. (8), there is an temper based image matrix with number of gray level intensity,  $G$  and number of bins,  $S$  which is used to detect the temper of the image  $P(x_i, y_j)$ . Here the convolution of temper based image matrix and image cause of obtaining the template for the k-means algorithm.

In the fuzzy C-means,  $U_{ij}^m$  is the membership function whose value updated with Euclidian distance  $d(x, v)$  which relies on the image features  $F = \{F_1, F_2, \dots, F_C\}$ , degree of fuzziness  $m$ , and feature center  $V = \{v_1, v_2, \dots, v_i, \dots, v_C\}$  is expressed as:

$$U_{ij} = \left[ \sum_{k=1}^C \left\{ \frac{d(x_j, v_i)}{d(x_j, v_k)} \right\}^{\frac{2}{m-1}} \right]^{-1} \quad (9)$$

In the previous research works, this Euclidian distance is depended upon only one features for example similarity [18], but in our proposed method this relies on features like contrast, homogeneity, entropy etc.

Clusters center from where the clusters position and tumor are detected can be defined as:

$$V_i = \frac{\sum_{j=1}^N (U_{ij})^m \times j}{\sum_{j=1}^N (U_{ij})^m} \quad \text{where, } (i = 1, 2, \dots, C) \quad (10)$$

We have extracted the following five features for the classifier as:

- a. Energy,  $F_1 = \sum_{i=1}^G \sum_{j=1}^G |P(x_i, y_i)|^2$  Here, G is the gray level co-occurrence matrix.
- b. Contrast,  $F_2 = \sum_{n=1}^G n^2 \sum_{i=1}^G \sum_{j=1}^G |P(x_i, y_j)|$  and  $|i - j| = n$
- c. Homogeneity,  $F_3 = \sum_{i=1}^G \sum_{j=1}^G \frac{P(x_i, y_j)}{1 + |i + j|}$
- d. Entropy,  $F_4 = \sum_{i=1}^G \sum_{j=1}^G P(x_i, y_j) (-\ln(P(x_i, y_j)))$
- e. Dissimilarity,  $F_5 = \sum_{i=1}^G \sum_{j=1}^G P(x_i, y_j) |i - j|$

From the convolution between level of thresholding,  $L_{th}$  and updated membership function we get the binary black and white image,

$$BW = \sum_{i=1}^M \sum_{j=1}^N (U_{ij}^{(l)} \oplus L_{th}) \quad (11)$$

Then applying region properties algorithm to  $BW$ , our desired parameters are obtained.

The TKFCM algorithm can be summarized as follow:

1. Define number of gray level and determine square matrix,  $A = \sum \sum P(x_k, y_l)$   
and image matrix,  $P = \sum \sum P(x_i, y_j)$
  2. Initialize template,  $T_{mn} = P \oplus A$
  3. Determine coarse image,  $B(x_i, y_j)$  from template,  $T_{mn}$
  4. Reshape template based k-means segmented image  

$$P_1 = \sum_{i=1}^K \sum_{j=1}^C P_{ij} \|x_i - c_j\|^2 \times \sum_{i=1}^M \sum_{j=1}^N B(x_i, y_j)$$
  5. Repeat step 2 to 4 until  $T_{mn} \leq \sum_{k=1}^G \sum_{l=1}^S [T_{mn}(k) - T_{mn}(l)]$
  6. Post process the  $P_1$
  7. Determine cluster centroid,  $C$  and degree of fuzziness,  $m$ .
  8. Initialize membership  $U_{ij}^{(0)}$  of FCM
  9. Calculate cluster center,  $v_i^{(l)} \Leftrightarrow U_{ij}^{(l)}, (i = 1, 2, \dots, C)$  and  $(l = 0, 1, 2, \dots)$
  10. Determine image features,  $F(x_j, v_i^{(l)}) \Leftrightarrow v_i^{(l)}$
  11. Update  $U_{ij}^{(l)}$  with  $d(x_j, v_i^{(l)})$  until  $\|U_{ij}^{(l)} - U_{ij}^{(l+1)}\| \leq \varepsilon, \varepsilon = \{0 \text{ to } 1\}$
  12. Obtain the binary black and white image (BW) of updated  $U_{ij}^{(l)}$
  13. From (BW) define area, eccentricity, perimeters and bounding box applying region properties algorithm.
- 

#### **IV. RESULTS AND PERFORMANCE ANALYSIS**

The database of 30 brain tumor images is shown in Fig. 2, which we have made through a collection of brain tumor images from [18]-[20]. We have used these four images numbered as 5, 3, 8 and 29 respectively from database, for the classification of tumor stages and area through linearization of TKFCM algorithm. The output of TKFCM segmentation is finalization with a standard thresholding and using region properties of the binary images cause to produce area and other characteristics parameters whose results are shown in Fig 3.

Figure 2: The database of 30 brain tumor MRI image

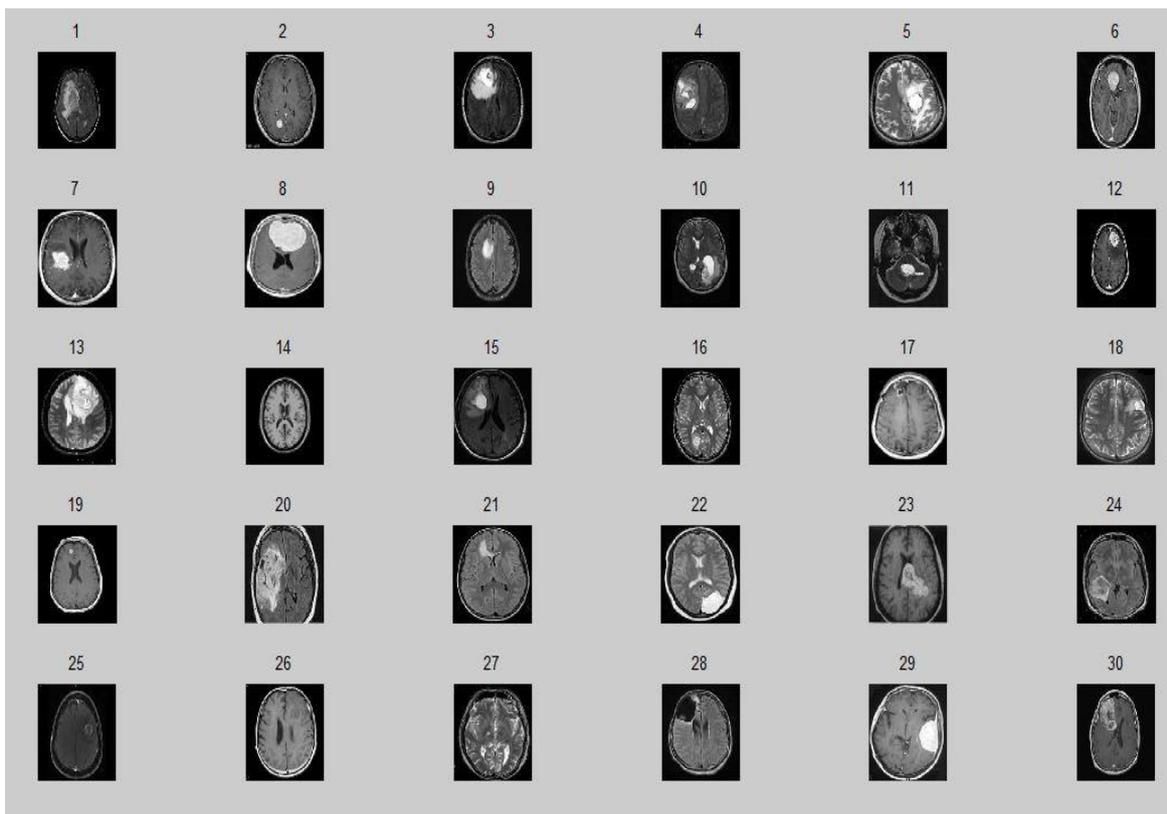


Figure 3. (a) Input images for TKFCM, (b) Detected brain tumor images from the TKFCM, (c) Classified brain tumor from TKFCM

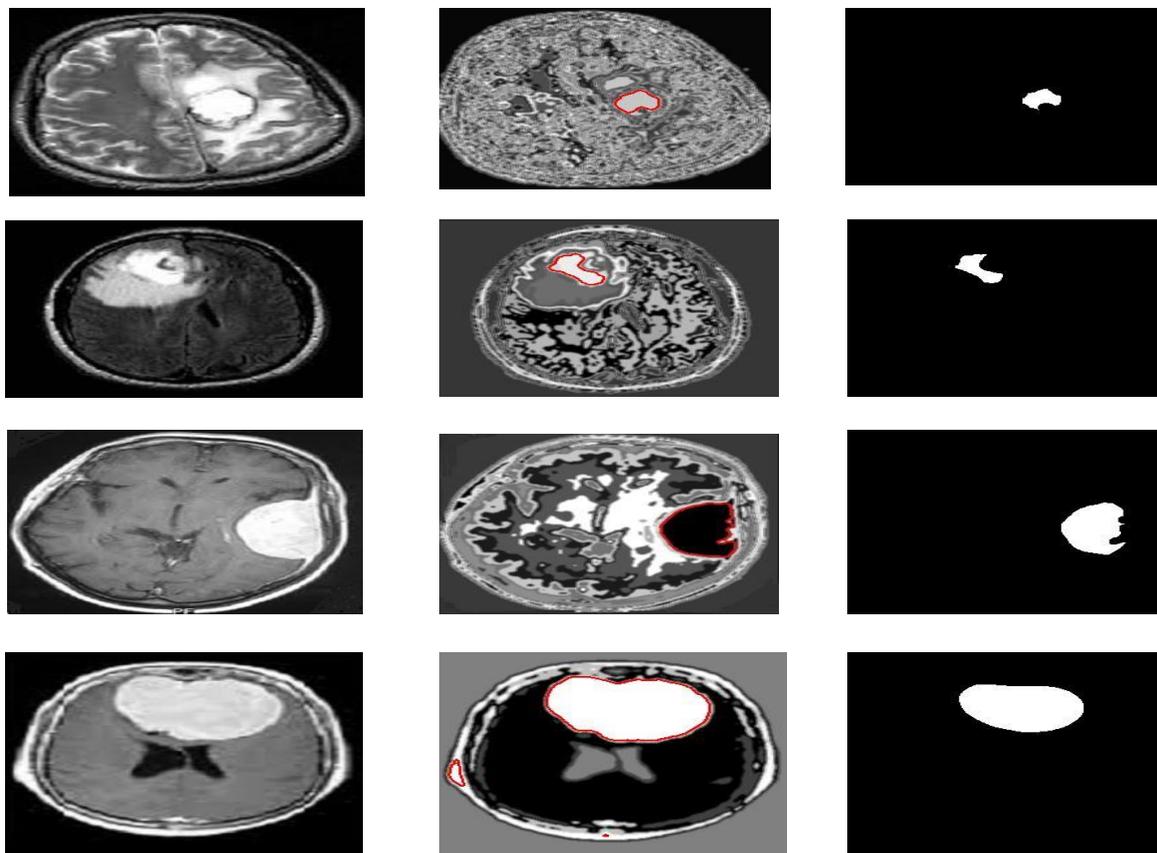
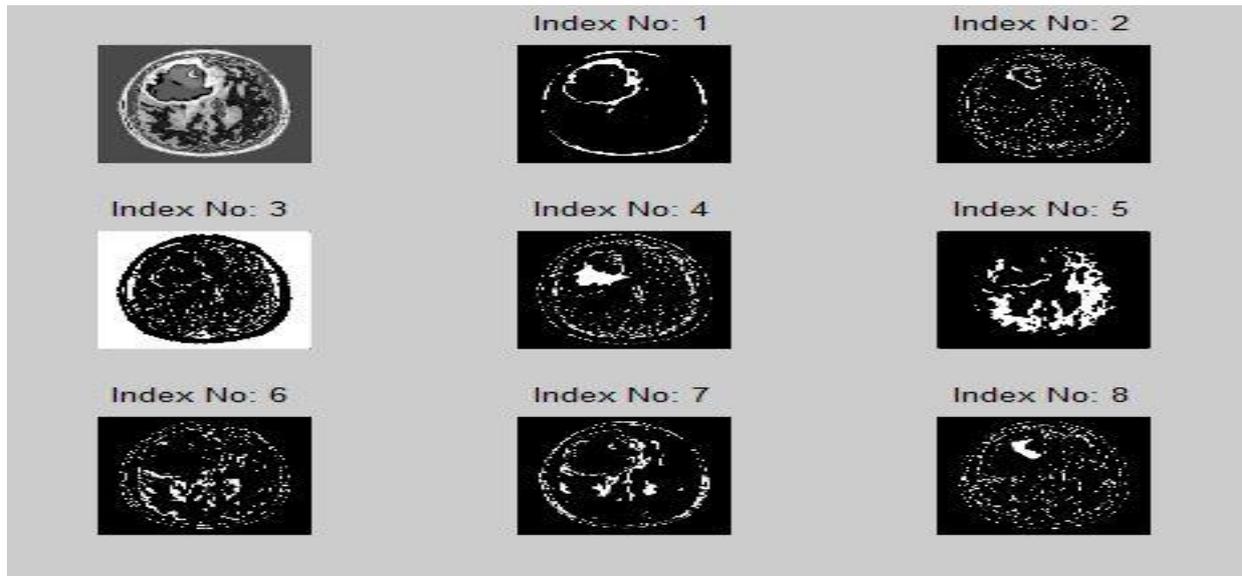


Figure 4: Several clustered images for input image no. 3



On the basis of gray level intensity the modified FCM is performed for 8 clusters. For example, several clustered images for input image no. 3 are shown in Fig 4.

The tumor area is calculated using binary method. This implies that the image having only two values either 0 or 1 as black and white color respectively. The binary image can be calculated using eqn. (12) as a summation of total number of white and black pixels [21]:

$$I = \sum_{W=0}^M \sum_{H=0}^N [f(0) + f(1)] \quad (12)$$

Where black and white pixel are expressed as  $f(0)$  and  $f(1)$  respectively.

The total pixel numbers of the image will be as follow:

$$\text{Pixels} = \text{Width}(W) \times \text{Height}(H) = 256 \times 256$$

The area is calculated through this formula:

$$\text{Size of tumor, } S = [(\sqrt{P}) \times 0.264] \text{ mm}^2 \quad (13)$$

Where, No of white pixel,  $P = \sum \sum [f(0)]$  and 1 Pixel = 0.264 mm.

TABLE I. : CALCULATION OF CHARACTERISTICS PARAMETERS OF TUMORS AND CLASSIFICATION OF TUMOR STAGES

Image No.	Eccentricity	Perimeters	Bounding box	Orientation	Area in pixel	Area in mm <sup>2</sup> (S)	Standard area in mm <sup>2</sup> [21]	Brain Tumor Stages
5.	0.4924	78.1780	[141.5 117.5 23 27]	49.7918	534	6.1	(2-6)	<i>Benign</i> : Just like normal tumor
3.	0.8888	131.5790	[86.5 50.5 38 41]	-53.8841	705	7.009		<i>Malignant</i> : Stage I spreading slowly
29.	0.7489	219.5170	[170.5 102.5 51 72]	84.7361	2592	13.44		<i>Malignant</i> : Stage II ( <b>no lymph</b> ) spread wide
8.	0.8113	249.8840	[89.5 41.5 97 59]	8.6715	4522	17.75		<i>Malignant</i> : Stage III spreading very fast

Table I show different characteristic parameters of brain tumor and classified tumor stages. The first image illustrates the area of 6.1 mm<sup>2</sup> which is like the standard normal tumor area that is why this belongs to benign tumor. On the other hand, other images show deliberately exaggerated area values than normal tumor cause to detect malignant tumors.

TKFCM requires 55 sec for the total computation using *Matlab R2014a* and *core2duo processor*

## **V. CONCLUSION**

This paper presents an effective approach to detect and classify tumor stages from the brain MRI image. The TKFCM is used to segment the tumor and extract features and linearization of TKFCM provides area and characteristics parameters. In TKFCM the template is selected based on convolution between gray level intensity in small portion of brain image and brain tumor image. K-means algorithm is to emphasized initial segmentation through the proper selection of template. Updated membership is obtained from the distance measurement from centroid to clusters, until it reaches to its best. On the basis of updated membership and automatic selected cluster, a sharp segmented image is obtained with tumor from modified FCM technique. The segmented tumor is shown as red marked with their proper detected position. Then taking the thresholding image and region properties of that TJFCM segmented image the area and other parameters are obtained. The classification can be more precise, if we could be able to identify the standard cancerous brain tumor size for all images in our database. In our future research, we will try to reduce these problems and also test this method for other brain imaging modalities like CT, PET, and other neuro image

## **REFERENCE:**

- [1] Ahmed M. N., Yamany S. M., Mohamed N., and Moriarty T. 2002. A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data. *IEEE Transaction on Medical Imaging*. 21(3), 193-199.
- [2] Rana R., Bhadauria H.S., and Singh A. 2013. Study of various methods for brain tumour segmentation from MRI images. *International Journal of Emerging Technology and Advanced Engineering*. 3(2), 338-342.
- [3] Narod S.A. 2012. Tumour size predicts long-term survival among women with lymph node-positive breast cancer. *Current Oncology*. 19(5), 249-253.
- [4] Hemanth D. J., Vijila C. K. S., and Anitha J. 2007. Application of neuro-fuzzy model for mr brain tumor image classification. *Biomedical Soft Computing and Human Sciences*. 16(1), 95-102, 2007.
- [5] Shen S., Sandham W., Granat M., and Sterr A. 2005. MRI fuzzy segmentation of brain tissue using neighborhood attraction with neural-network optimization. *IEEE Transactions on Information Technology in Biomedicine*. 9, 459 – 467.
- [6] Pham D. L., Xu C., and Prince J. L. 2000. Current methods in medical image segmentation. *Annual review on Biomedical Engineering*. 2, 315-337.
- [7] Lemieux L., Hagemann G., Krakow K., and Woermann F. G. 1999. Fast, accurate, and reproducible automatic segmentation of the brain in T1-weighted volume MRI data. *Magnetic Resonance Medical Imaging*. 42, 127–135.
- [8] Pohle R., and Toennies K. D. 2001. Segmentation of medical images using adaptive region growing. *Proceeding International Society for Optics and Photonics (SPIE)— Medical Imaging*. 4322, 1337–1346.
- [9] Wells W. M., Grimson W. E. L., Kikinis R., and Jolesz F. A. 1996. Adaptive segmentation of MRI data. *IEEE Transaction on Medical Imaging*. 15(4), 429–442.

- [10] Law T. Y., and Heng P. A. 2000. Automated extraction of bronchus from 3D CT images of lung based on genetic algorithm and 3D region growing. *Proceeding International Society for Optics and Photonics (SPIE)— Medical Imaging*. 3979, 906–916.
- [11] Leemput K. V., Maes F., Vandermeulen D., and Suetens P. 1999. Automated model-based tissue classification of MR images of the brain. *IEEE Transaction on Medical Imaging*. 18(10), 897–908.
- [12] Li C. L., Goldgof D. B., and Hall L. O. 1993. Knowledge-based classification and tissue labeling of MR images of human brain. *IEEE Transaction on Medical Imaging*. 12(4), 740–750.
- [13] Oweis R., and Sunna M. 2000. A combined neuro–fuzzy approach for classifying image pixels in medical applications. *Journal of Electrical Engineering*. 56(5-6), 146–150.
- [14] Shally H. R., and Chitharanjan K. 2013. Tumor volume calculation of brain from MRI slices. *International Journal of Computer Science & Engineering Technology (IJCSET)*. 4(08), 1126-1132.
- [15] Narkbuakaew W., Nagahashi H., Aoki K., and Kubota Y. 2014. Integration of modified k-means clustering and morphological operations for multi-organ segmentation in ct liver-images. *Recent Advances in Biomedical & Chemical Engineering and Materials Science*. 34-39.
- [16] Bezdek J. C., 1981. *Pattern recognition with fuzzy objective function algorithms*. Kluwer Academic Publishers, USA.
- [17] (2015) PE- Brain tumour website. [Online]. Available: [http:// www.mayfieldclinic.com /](http://www.mayfieldclinic.com/)
- [18] University of Maryland Medical Center (2015) homepage. Primary brain tumor [Online]. Available: <http://umm.edu/health/medical/reports/articles/brain-tumors-primary>
- [19] Oncolgy (2015) homepage. Tumor type [Online]. Available: <http://www.oncolink.org/types/article.cfm?c=52&aid=247&id=9534>
- [20] Selvakumar J., Lakshmi A., and rivoli T. A. 2012. Brain tumor segmentation and its area calculation in brain MR images using K-mean clustering and fuzzy C-mean algorithm," *International Conference on Advances in Engineering, Science and Management (ICAESM)*. 186-190.