



ENHANCING BRAIN TUMOR SEGMENTATION AND AREA ESTIMATION WITH HYBRID SHAFT CLUSTERING

Pitta Satya Surekha¹, Kandukuri Srinivas², Medipally Nagasri³ Assistant Professor, Department of ECE, Malla Reddy Engineering College and Management Sciences, Hyderabad, Telangana.

Abstract

Accurate segmentation of brain tumors in magnetic resonance (MR) images is a crucial step in quantitative brain image analysis and has garnered significant research interest. Traditional methods for MR brain image segmentation, such as K-means and Fuzzy C-Means (FCM) clustering algorithms, treat each pixel as independent and lack spatial information integration among neighboring pixels. Consequently, these segmentation algorithms face limitations in accuracy due to noise and intensity inhomogeneity in brain MR images. To address this issue, we introduce an advanced approach to segmentation using the hybrid shaft clustering method, which combines adaptive K-means clustering and Fuzzy Kernel C-Means (FKCM). Additionally, we estimate the area by calculating the segmented tumor area and the number of cells it occupies. Simulation results demonstrate that our proposed method offers superior segmentation and area estimation accuracy compared to conventional approaches.

Keywords: Brain Tumor Segmentation, Magnetic Resonance Images, Hybrid Shaft Clustering, Area Estimation, K-means, Fuzzy C-Means, Spatial Information, Quantitative Analysis.

1. Introduction:

The tumor is a standout amongst the most widely recognized as relatable point Brain illnesses. The World Health Organization (WHO) estimates that Analysis and medicine would vital to more than five lakh persons would endure from tumor for every year in the globe. Developments in restorative imaging systems permit using them inside few domains of medicine, for instance, workstation helped pathologies diagnosis, surgical arranging and guidance, longitudinal dissection. Around every last one of restorative image modalities, Magnetic Resonance Imaging (MRI) also Computed Tomography (CT) need aid the mossycup oak intermittently used imaging strategies clinched alongside neuroscience Furthermore neurosurgery. Segmentation of objects, primarily anatomic structures and more Pathologies starting with MRI images may be a crucial task, since the outcomes every now and again turned the foundation to different requisitions. Systems for performing segmentation shift comprehensively contingent upon those specific provisions and image modality. Additionally, the segmentation from claiming medicinal images will be a was troublesome task, Since they for the most part incorporate an expansive amount of data, Furthermore here and there a couple artifacts due to patient's restricted securing run through Furthermore fragile tissue boundaries, typically not great defined. At managing brain tumors, separate issues arise, which make their segmentation troublesome. There may be a limitless population about tumor sorts which bring a mixture of shapes also sizes. It might develop at whatever range also done divergent image intensities. Some about them misshape those encompassing structures or might make identified with edema that transforms those intensities from claiming images around those tumors. Additionally, those presences from claiming a couple MRI procurement conventions provides for divergent majority of the data on the brain. Each image generally highlights a specific region of the tumor. The Robotized segmentation with former models alternately using the former information will be challenged with executes. The flawed segmentation for interior structures of the Brain is from claiming great energy should contemplate also for those medications from claiming tumors. It dives during



diminishing those mortal sins also upgrading the surgical or radio restorative. Over saw economy for tumors. To brain oncology, it is also alluring with bringing a reminiscent human brain model that could coordinate tumor data concentrated from MRI and CT information, for example, such that localization, type, shape, utilitarian positioning, and additionally influence with respect to other brain structures. Despite Different efforts also guaranteeing brings about the therapeutic Imaging community, exact Also proliferation segmentation and abnormalities, Characterization need aid even now difficult assignments. Existing strategies clear out significant Space to expanded automation, materialness Furthermore accuracy. In the requisition for image processing, smoothening of the image will be, Crucial should aggravate the characteristic extraction also segmentation steps simpler. Hence impeccable sifting method is compulsory over biomedical image transforming. The Suitableness denoising calculation to MRI brain images may be vital to finish secondary execution. Those unwanted parcel in the MRI images might Make evacuated by correct segmentation algorithm. Characteristic extraction is the following Stage after preprocessing also segmentation which may be took after Eventually Tom's perusing characteristic Determination. The point when those required offers are chose it may be subjected of the Segmentation transform. Those issues for picking those proper chK-meansels to Denoising, segmentation algorithms, characteristic extraction Furthermore prediction, Calculation for those orders about brain MRI images still remains as a real.

The remaining part of the paper is systematized as follows. Section 2 describes the related works for brain tumor detection and segmentation with their drawbacks, section 3 deals with proposed method detection and segmentation of brain tumor with detailed operation. Section 4 deals with experimental results of proposed method and comparison with respect to the various state of art approaches using quantitative evaluation and finally section 5 the conclusion and scope for future enhancements.

2. LITERATURE SURVEY

Despite the fact that it may appear to be intelligent that a bigger number of features would be more useful than fewer features, this isn't the case in genuine applications, because of the accompanying three primary reasons as depicted[2].Right off the bat the multifaceted nature and computational cost of the classifier increments profoundly. Also, despite the fact that the quantity of misclassified information may diminish, when more features are included for preparing the classifier, it has been demonstrated that the speculation blunder will inevitably increment [3]. Thirdly, on account of a predetermined number of accessible information and huge number of accessible features, it is more probable that features with little to no separation power will initiate clamor, debasing the speculation of the classifier to obscure information [4]. Thusly, feature determination is an imperative advance for drawing out the more instructive features and for ideal tuning the classifier's ability to dependably characterize obscure data.The utilization of different K-means[5] for image order breaks down. The absence of quicker union rate of the traditional neural systems is additionally clarified. This lays accentuation on the prerequisite of adjusted neural systems with unrivaled meeting rate for image order applications. In[6]authors have arranged four unique types of tumor utilizing LDA method. Be that as it may, the segmentation precision revealed in the work is at the request of 80% which is generally low. This additionally proposes the different purposes behind missegmentations. FCM based characterization of different levels of MR glioma images was performed by authors in [7-8], which is guaranteed to be superior to anything principle based frameworks yet the precision revealed in the FCM based segmentation is low. This FCM based segmentation managed just glioma images and subsequent absence of summing up capacity is another downside of this framework.

In [8-9] authors need to utilize those Kohonen neural networks to image segmentation. A percentage adjustment of the accepted Kohonen neural system is also actualized here, which demonstrated on a chance to be a great deal better than those accepted neural networks.

In [9-10] authors need to utilize a mixture methodology, for example, mix from claiming wavelets and FCM to classifying those abnormal and the typical images. This FCM technique uncovers the prevalence of the mixture FCM of the Kohonen neural networks as far as execution measures. Yet the real detriment for this framework will be the little measure of the data set utilized to usage. The segmentation precision outcomes might lessen when the span of the dataset will be expanded. An change about customary FCM for example, any rate as square FCM (LS-FCM) for brain tumor distinguishment is suggested by authors in [11-12]. Both bi-level segmentation also multiclass order need aid performed with hint at those unrivaled way of the suggested system again the traditional classifiers. This likewise specified a critical note that these contrasts between different calculations build when those amounts about classes' increment. Thus, this approach suggested those needs to multiclass segmentation strategies over bi-level segmentation systems. In turn adaptation for LS-FCM is recommended and successfully actualized by by authors in [13]. A far reaching similar examination will be performed the middle of those FCM, neural classifier and the factual classifiers. 63 Effects proposed the preferences from claiming FCM as far as order exactness. Bi-level order alone is performed, which will be insufficient for judging those natures of the robotized framework

In [14-15] authors utilize those altered PKFCM for tumor image order. Abnormal images, for example, such that metastasis, glioma and meningioma would separated utilizing the any rate as square characteristic conversion based PKFCM. A similar examination may be Additionally, performed for FCM. This methodology inferred that the convert based PKFCM may be better than the FCM as far as segmentation exactness.

3. PROPOSED METHODOLOGY

The brain image segmentation is the essential process in segmentation of tumors and analysis of tumors. The figure 1 presents the detailed operation of proposed tumor segmentation using hybrid shaft clustering approach. It is combination of both adaptive K-means clustering and Fuzzy Kernel C-Means (FKCM) methods.

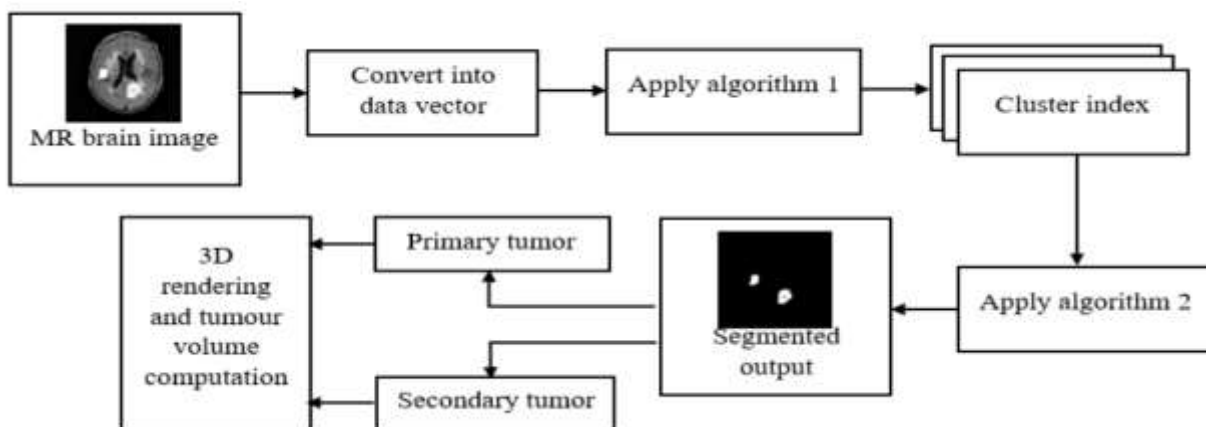


Fig 1. Block diagram of proposed model

The constraints presented in the conventional K-means clustering solved and optimized by utilizing the localization of cluster indexes in adaptive manner, this will effectively introduces the mean property of cluster centers and similarity matching. This adaptive approach of clustering will introduce the new



adaptive K-means clustering respectively. Each cluster will be further applied to the FKCM clustering process for effective segmentation of brain tumor. And Area estimation also carried out by calculating segmented tumor area and number of cells it occupied with respect to the characterization of primary tumor and secondary tumors separately.

Algorithm 1: adaptive K-means clustering

There are two phases in the adaptive K-means clustering algorithm. In the first phase, adaptive K-means clustering algorithm is used for the selection of initial centroids. This way of fixing initial centroids leads to initial centroids remaining unchanged for all the execution. It leads to consistency in grouping of similar objects with reduced iteration. Ultimately, through the unique initial centroids selection the K-means algorithm is able to provide local optimum. That in turn leads global optimum solution for K-means clustering algorithm. In the second phase, Euclidean distance based weighted adaptive K-means clustering algorithm is used for clustering the similar objects. In both the phases, weights associated with every range of attributes values in the data set are considered for processing. First step of the proposed enhanced weighted adaptive K-means clustering algorithm is weighted ranking approach. In the weighted ranking algorithm, the summation of the distance from origin to each weighted attributes of a data point in a data set is calculated. Weighted data points are calculated with the equation 1.

$$U_{j=1}^n = \sum_{i=1}^n W_i X_i \quad (1)$$

where, W_i represents the weightage of X_i th attribute. For n data points, n numbers of U values are calculated and the format is given in Equation 2.

$$U_{j=1}^n = \begin{pmatrix} U_1 \\ U_2 \\ \vdots \\ U_n \end{pmatrix} \quad (2)$$

Then the weighted data points (U_i) are sorted in accordance with the sorted distances. After sorting, the sorted data points are partitioned into k equal sets where k is the predefined number of clusters. In each set, the middle points or mean value of the middle points are taken as the initial centroids. Initial centroids selected through this algorithm leads to the consistency in cluster members. As K-means algorithm makes use of distance measures for finalizing group members, the proposed weighted ranking algorithm based initial centroids selection also works on the basis of the distance formula. Hence, the principle of the proposed weighted ranking is positively relevant to the optimization of initial centroids selection. Two significant features of the first phase of the algorithm are the consistency in group members due to the unique initial centroids selection and the attributes value based weights considered for processing. Adaptive K-means clustering is a partitioned algorithm where the given image is broken into clusters. It works with centroids which represent its cluster, that are artificially created entity and re-estimates it. The first object of a cluster is known as cluster seed and it is compared with the initiator. Considering I as input image and S as clustered image.

Table 1: proposed adaptive K-means clustering

Input: I; Output: S
Step 1: convert the input image I into data vectors
Step 2: The initial centroids are identified by the k data points known as cluster centers
Step 3: Label the clusters based on Weighted data centroids as presented using eq 1
Step 4: perform the weighted sorting operation using eq 2
Step 5: The distance is computed from 'U' to the centroid.
Step 6: The closest centroid is assigned as segmented cluster.



Step 7: repeat the process step 3 to step 6 until all clusters covered

Step 8: combine all the segmented clusters and concatenate output S
--

Algorithm 2: FKCM**Proposed Kernel Function**

The kernel function is used to make non-linear projections that map data to a higher feature space. The kernel trick is one of the additional methods that can overcome non-linear problems that generalize well in combination with linear models. Suppose $x_1, x_2, \dots, x_n \in R^n$ is the original dataset in R^n . Then, there is a function ϕ that maps data to a new feature space that is higher F namely:

$$\phi : R^n \rightarrow F \quad (3)$$

So the kernel function is defined as follows:

$$K(x, y) = \langle \phi(x), \phi(y) \rangle \quad (4)$$

and the distance function is defined as follows:

$$\begin{aligned} d^2(x, y) &= \|\phi(x) - \phi(y)\|^2 \quad (5) \\ &= \langle \phi(x), \phi(x) \rangle - 2\langle \phi(x), \phi(y) \rangle \\ &\quad + \langle \phi(y), \phi(y) \rangle \\ &= K(x, x) - 2K(x, y) + K(y, y) \end{aligned}$$

There are several types of kernels that are commonly used including:

Linear kernel function : $K(x, y) = x \cdot y \quad (6)$

RBF kernel function : $K(x, y) = \exp\left(-\frac{\|x-y\|^2}{\sigma^2}\right) \quad (7)$

Polynomial Kernel Function : $K(x, y) = (x \cdot y + 1)^d \quad (8)$

Fuzzy Kernel C-Means method is a modification of the Fuzzy C-Means method. This method is used to group incomplete data using non-linear functions into a much higher dimensional space using kernel functions. The kernel function is proven to be very powerful in extracting information available in the database by mapping actual data from small dimensions to higher dimensional spaces. This function is bridging from linear to nonlinear. A mathematical model of Fuzzy C-Means as follows :

$$J(V, U, X, c, m) = \sum_{i=1}^c \sum_{k=1}^N (u_{ik})^m \|x_k - v_i\|^2 \quad (9)$$

Subject to

$$0 \leq u_{ik} \leq 1 \quad (10)$$

$$\sum_{i=1}^c u_{ik} = 1 \quad (11)$$

$$0 \leq \sum_{i=1}^c u_{ik} \leq N \quad (12)$$

$$i = 1, 2, \dots, c \text{ and } k = 1, 2, \dots, N$$

with c is the number of cluster ($c \geq 2$), N is the amount of data, m is the degree of the fuzziness with $m > 1$, $\|x_k - v_i\|$ is the distance between data x_k and cluster center v_i , dataset that will be in the cluster $X = \{x_1, x_2, \dots, x_n\}$, $V = \{v_1, v_2, \dots, v_c\}$ is a cluster center set and $U = [u_{ik}]$ is a membership function with U containing a membership value that maps the membership level from the center of the cluster i to the k data. The optimum condition of the objective function will be achieved if u_{ik} and v_i are optimum. So find the optimum u_{ik} and v_i can be found by calculating the derivative of the Lagrange function. Based on the Lagrange multiplier theory when a Lagrange multiplier λ . Then the Lagrange function for Fuzzy C-Means can be formed as follows:

$$L = \sum_{i=1}^c \sum_{k=1}^N (u_{ik})^m \|x_k - v_i\|^2 + \lambda (1 - \sum_{i=1}^c u_{ik}) \quad (13)$$

So that the Lagrange function derivative of each parameter is zero.

$$\frac{dL}{du_{ik}} = 0 \quad (14)$$

So that the optimum conditions of the membership value and cluster center can be obtained:

$$u_{ik} = \frac{1}{\sum_{j=1}^c \frac{(\|x_k - v_i\|^2)^{\frac{1}{m-1}}}{(\|x_k - v_j\|^2)^{\frac{1}{m-1}}}} \quad (15)$$

$$v_i = \frac{\sum_{k=1}^N (u_{ik})^m x_k}{\sum_{k=1}^N (u_{ik})^m} \quad (16)$$

The accuracy of the FCM method depends on the type of data. When data cannot be linearly separated, the convergence will be slow and inaccurate. So that another method is needed as a link to overcome problems like this. The method that can be used is the kernel method. This study will apply the kernel method to the FCM to complete the segmentation of thalassemia data using Fuzzy Kernel C-Means (FKCM). Suppose $x_1, x_2, \dots, x_n \in R^p$ as the set of data in R^p . Then the function ϕ which maps data on R^p to H as a new feature space has a higher dimension.

$$\phi = R^n \rightarrow F \quad (17)$$

So that the objective functions of Fuzzy Kernel C-Means are as follow :

$$J(V, X, c, m) = \sum_{i=1}^c \sum_{k=1}^N (u_{ik})^m \|\phi(x_k) - \phi(v_i)\|^2 \quad (18)$$

subject to 0

$$0 \leq u_{ik} \leq 1 \quad (19)$$

$$\sum_{i=1}^c u_{ik} = 1 \quad (20)$$

$$0 \leq \sum_{i=1}^c u_{ik} \leq N \quad (21)$$

Where c is a lot of clusters ($c \geq 2$), N is the number of data, m is the degree of fuzziness with $m > 1$, $\|\phi(x_k) - \phi(v_i)\|$ is kernel mapping distance between x_k data and cluster center v_i , $X = \{x_1, x_2, \dots, x_n\}$ is the set of data to be clustered, $V = \{v_1, v_2, v_c\}$ is the cluster center set, and $U = [u_{ik}]$ is membership function. As if Fuzzy C-Means, to find the optimal u_{ik} and v_i , we will give a Lagrange multiplier λ , so the Lagrange function will be applied to find the optimal solution for Fuzzy Kernel C-Means :

$$L = \sum_{i=1}^c \sum_{k=1}^N (u_{ik})^m \|\phi(x_k) - \phi(v_i)\|^2 + \lambda (1 - \sum_{i=1}^c u_{ik}) \quad (22)$$

with the Lagrange function derivative of each parameter equal to zero, it can be written as follows:

$$\frac{dL}{du_{ik}} = 0 \quad (23)$$

Based on the RBF kernel function form to be used, then $K(x_i, x_i) = K(v_i, v_i) = 1$

So that the optimum condition of the membership value and cluster center is as follows:

$$u_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{1 - k(x_k, v_i)}{1 - k(x_k, v_j)} \right)^{\frac{1}{m-1}}} \quad (24)$$

$$v_i = \frac{\sum_{k=1}^N (u_{ik})^m K(x_k, v_i) x_k}{\sum_{k=1}^N (u_{ik})^m K(x_k, v_i)} \quad (25)$$

$$m = m_i + \frac{t}{T} (mf - m_i) \quad (26)$$

With m_i and mf are each initial value and final value of m . When the mf value is small and m_i is larger enough, it is expected that m will decrease and otherwise.

Area Estimation

We can represent the segmented output image $f_{x,y}$ as a summation of total number of white and black pixels.

$$M = \sum_{x=1}^L \sum_{y=1}^L [f_{x,y}(0) + f_{x,y}(1)] \quad (27)$$

where $L=1, 2, 3 \dots 256$

$f_{x,y}(0)$ = black pixel having the value of zero,

$f_{x,y}(1)$ = white pixels having the value of one

$$P = \sum_{i=1}^L \sum_{j=1}^L f_{x,y}(1) \tag{28}$$

Where,

P = number of white pixels

Now, by using the above equation, we can calculate the area of the segmented tumor based on the typography and digital imaging units [20], where one pixel is equal to 0.264583 millimeters. i.e., 1 pixel = 0.264583 mm

Then the area of tumor can be expressed as follows:

$$A_{Tumor} = (\sqrt{P}) * 0.2646mm^2 \tag{29}$$

4. SIMULATION RESULTS

This section gives the detailed simulation analysis of proposed method with respect to the various test datasets of brain images using Matlab software tool. Both qualitative and quantitative evaluation is performed and comparison analysis with respect to the various approaches has been analyzed.

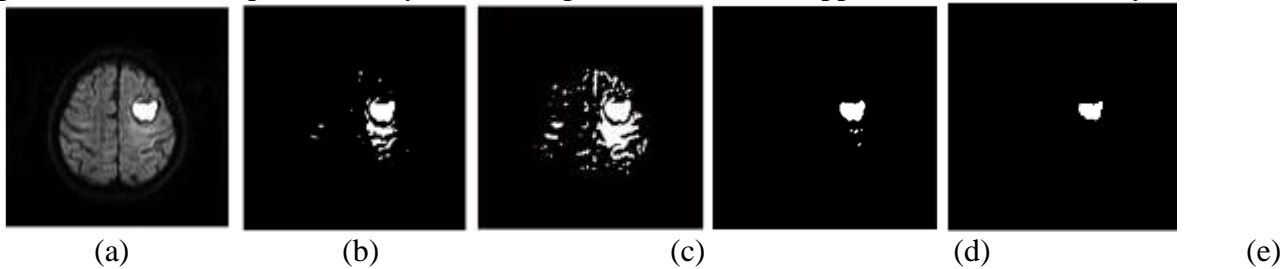


Fig. 2. Single tumor case (a) Original Image (b) manual segmentation (c) FCM (d) K-means (e) proposed clustering

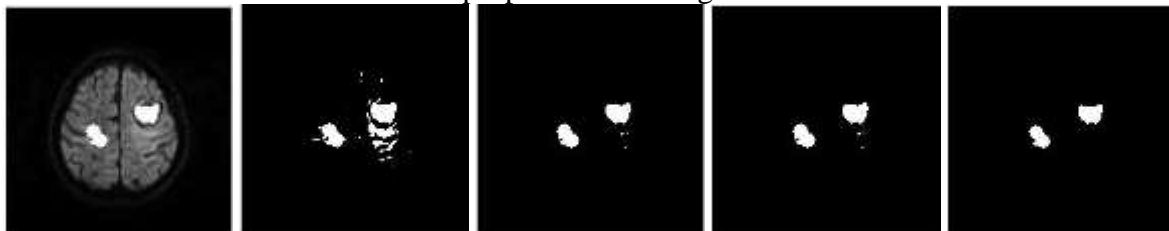


Fig. 3. two tumor case (a) Original Image (b) manual segmentation (c) FCM (d) K-means (e) proposed clustering

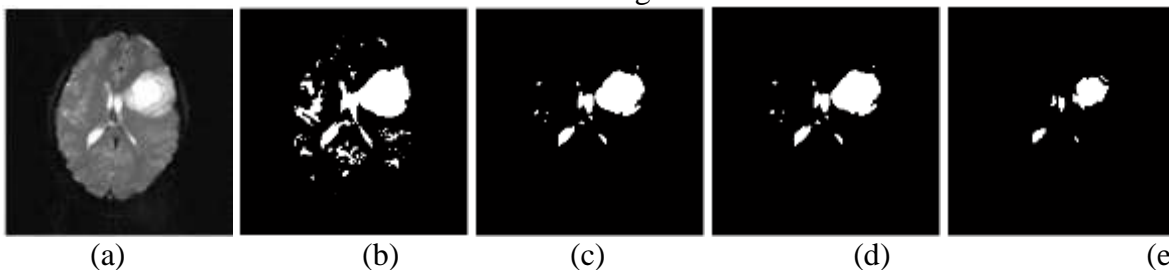


Fig. 4. Multiple tumor s (a) Original Image (b) manual segmentation (c) FCM (d) K-means (e) proposed clustering



Fig 2 and 4 shows that the segmented outputs of single tumor of MR brain images clustering, K-means clustering and proposed hybrid clustering algorithms. From the obtained outputs, we can observe that the proposed hybrid clustering algorithm has detected the tumor more effectively with higher accuracy. Although, our proposed algorithm running time will be quite bit of more than the k-means clustering but however the accuracy of segmented output will be more i.e., tumor area will be estimated more precisely to diagnosis further.

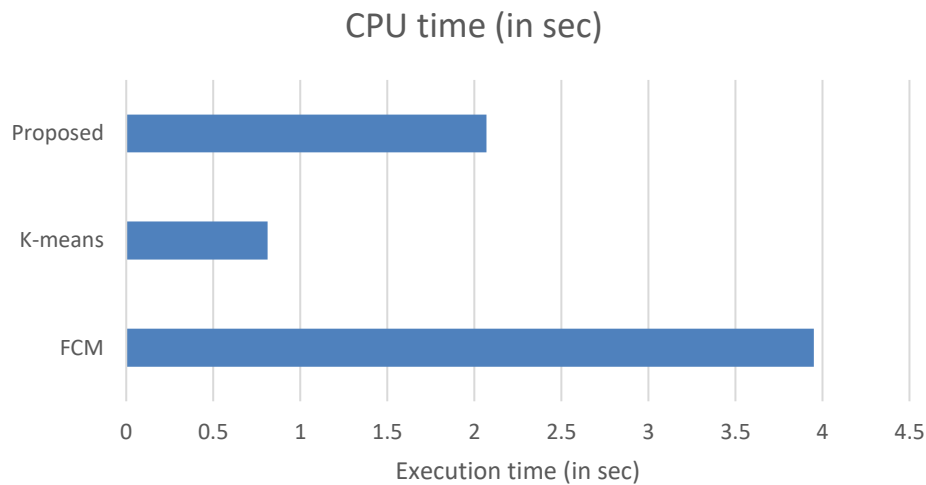


Fig. 5. Performance evaluation with CPU running time for multi tumors detection

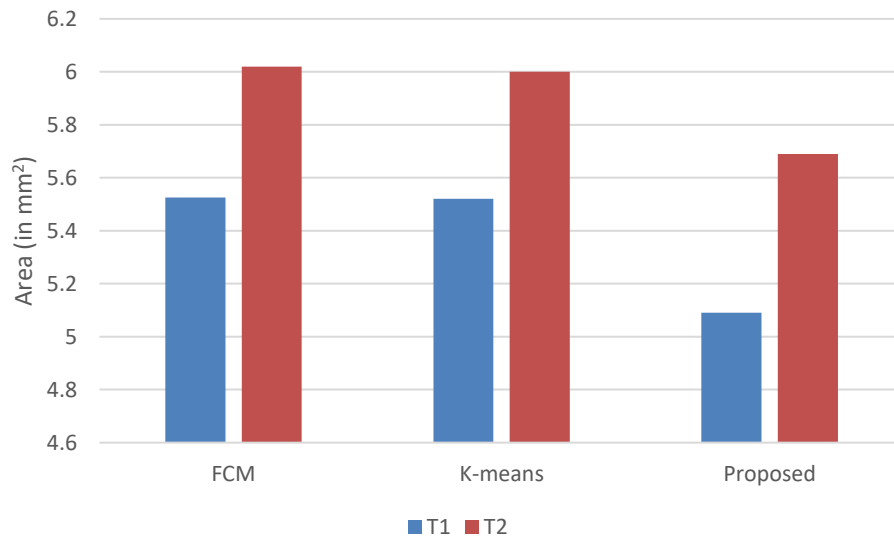


Fig. 6. Estimated area of segmented tumors T1 = tumor 1 and T2 = tumor 2

Above figures demonstrates that the performance evaluation of proposed hybrid clustering algorithm with comparison to the conventional clustering algorithms presented in the literature. We calculated execution time in seconds and tumors area in mm^2 . From the figure 5 it is observed that the proposed method consumes very low time for segmentation process and from the figure 6 it is observed that the proposed method provides maximum accuracy in area of the segmented area compared to FCM and K-means clustering approaches.



5. CONCLUSION: This article proposed an efficient approach for segmentation and detection of tumors from brain medical images using hybrid shaft clustering algorithm. In addition, this approach computes the weight of each picture element based on K-mean shifting process with fuzzy selection clustering. Simulation outcomes shown that, the proposed segmentation approach performed superior to the existing segmentation algorithms in terms of both ocular and quantitative analysis with the low time consumption for segmentation. The robustness of the proposed enhanced weighted hybrid shaft clustering algorithm is proved by comparing with existing K-means and FCM with respect to the accurate area estimation. The work can be extended to develop the segmentation approach and concentrating on the improvement of the segmentation accuracy, developing the classifier for specific application needs consistency at every stage. Furthermore, this can be extended to 3D multi modal medical image segmentation with more effective and accurate clustering algorithms.

References:

1. Saba, Tanzila, et al. "Brain tumor detection using fusion of hand crafted and deep learning features." *Cognitive Systems Research* 59 (2020): 221-230.
2. Sharif, Muhammad, et al. "A unified patch based method for brain tumor detection using features fusion." *Cognitive Systems Research* 59 (2020): 273-286.
3. Toğaçar, Mesut, Burhan Ergen, and Zafer Cömert. "BrainMRNet: Brain tumor detection using magnetic resonance images with a novel convolutional neural network model." *Medical Hypotheses* 134 (2020): 109531.
4. Kumar, P.M.Kumar. "Satellite image denoising using local spayed and optimized center pixel weights." *International Journal of Electrical and Computer Engineering* 4.5 (2014): 751.
5. Wu, Ming-Ni, Chia-Chen Lin, and Chin-Chen Chang. "Brain tumor detection using color-based k-means clustering segmentation." *Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2007)*. Vol. 2. IEEE, 2007.
6. Kumar, P.M.Kumar. "A New Encryption and Decryption for 3D MRT Images." *Asian Journal of Electrical Sciences* 6.1 (2017): 1-6.
7. Kalyanapu Srinivas, P.M.Kumar, Annam Jagadeeswara Rao. "A New Video Watermarking Using Redundant Discrete Wavelet in Singular Value Decomposition Domain." *Journal of Mechanics of Continua and Mathematical Sciences* 14.5 (2019): 476-486.
8. Sharif, Muhammad, et al. "Brain tumor detection based on extreme learning." *Neural Computing and Applications* (2020): 1-13.
9. Amin, Javaria, et al. "Brain Tumor Detection by Using Stacked Autoencoders in Deep Learning." *Journal of Medical Systems* 44.2 (2020): 32.
10. Özyurt, Fatih, Eser Sert, and Derya Avcı. "An expert system for brain tumor detection: Fuzzy C-means with super resolution and convolutional neural network with extreme learning machine." *Medical hypotheses* 134 (2020): 109433.
11. Shivhare, Shiv Naresh, and Nitin Kumar. "Brain tumor detection using manifold ranking in flair mri." *Proceedings of ICETIT 2019*. Springer, Cham, 2020. 292-305.
12. Gumaste, Pratima Purushottam, and Vinayak K. Bairagi. "A Hybrid Method for Brain Tumor Detection using Advanced Textural Feature Extraction." *Biomedical and Pharmacology Journal* 13.1 (2020): 145-157.



13. Gumaste, Pratima Purushottam, and Vinayak K. Bairagi. "A Hybrid Method for Brain Tumor Detection using Advanced Textural Feature Extraction." *Biomedical and Pharmacology Journal* 13.1 (2020): 145-157.
14. Chahal, Prabhjot Kaur, Shreelekha Pandey, and Shivani Goel. "A survey on brain tumor detection techniques for MR images." *MULTIMEDIA TOOLS AND APPLICATIONS* (2020).
15. Chandra, Saroj Kumar, and Manish Kumar Bajpai. "Brain tumor detection and segmentation using mesh-free super-diffusive model." *Multimedia Tools and Applications* 79.3 (2020): 2653-2670.