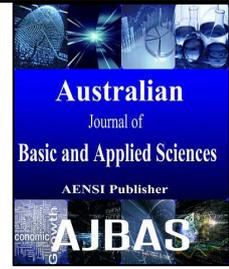




AUSTRALIAN JOURNAL OF BASIC AND APPLIED SCIENCES

ISSN:1991-8178 EISSN: 2309-8414
Journal home page: www.ajbasweb.com



A Gui Based Hybrid Clustering Technique for Brain Tumor Segmentation and Validation in Mri Images

¹Karnam Gopi and ²Thirumala Ramashri

¹Department of ECE, Sri Venkateswara University College (SVU) of Engineering, Tirupati, Andhra Pradesh, India

²Professor, Department of ECE, Sri Venkateswara University College (SVU) of Engineering, Tirupati, Andhra Pradesh, India.

Address For Correspondence:

Karnam Gopi, Department of ECE, Sri Venkateswara University College (SVU) of Engineering, Tirupati, Andhra Pradesh, India
E-mail: gopiradha47@gmail.com

ARTICLE INFO

Article history:

Received 12 February 2016

Accepted 18 March 2016

Available online 20 April 2016

Keywords:

Medical Imaging, Brain tumor segmentation, Magnetic Resonance Image(MRI), KMeans algorithm, EM algorithm, Intelligent CAD system, Image quality Metrics

ABSTRACT

Background: Image segmentation is considered as the most vital and preliminary process in image analysis for Computer aided diagnosis (CAD). Segmentation has proved its applicability in various areas like medical image processing, satellite image processing and many more. CAD systems can enhance the diagnostic capabilities of physicians and reduce the time need for precise diagnosis. In this paper a fusion technique for segmentation of brain tumor through MRI is proposed. MRI is a superior medical imaging technique providing rich information about the human soft-tissue anatomy. **Objectives:** The proposed work presents an efficient image segmentation approach using K-means clustering technique integrated with Expectation Maximization (EM) algorithm to provide accurate brain tumor detection. The proposed technique can get benefits of the K-means clustering in the aspects of minimal computation time. In addition, it can get advantages of the Expectation Maximization (EM) in the aspects of accuracy. The performance of the proposed image segmentation approach was evaluated by comparing it with some state of the segmentation algorithms in case of accuracy, and performance. The accuracy was evaluated by comparing the results with the ground truth of each processed image. The exploratory results illuminate the effectiveness of our proposed approach to deal with higher number of segmentation issues by means of enhancing the segmentation quality and precision. In the proposed work a multifunctional Graphical user interface tool (GUI) is designed that performs the processing of hybrid brain tumor segmentation, evaluation, quantification and validation. **Results:** The performance measures defined as probability random index (PRI), global consistency error (GCE), structural similarity (SSIM), variation of information (VOI), tumor area and Peak signal to noise ratio (PSNR) integrated in GUI-CAD is explained. **Conclusion:** At the end of the process significant results were obtained in case of proposed algorithm when the performance measures were compared with existing work.

INTRODUCTION

The use of automatic computerized algorithm to assist the image interpretation process is commonly referred to as CAD. The aim of this study is to develop Computer Aided Diagnosis (CAD) system for the detection of brain tumor .Computer aided diagnosis. CAD has generally been defined by diagnosis made by a physician who takes into description the computer output based on quantitative examination of radiological images. The goal of CAD is to improve the quality and productivity of radiologists. Improving the effectiveness of CAD could improve the detection of breast cancer, and could improve the survival rate by detecting the cancer earlier.

Open Access Journal

Published BY AENSI Publication

© 2016 AENSI Publisher All rights reserved

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

To Cite This Article: Karnam Gopi and Thirumala Ramashri., A Gui Based Hybrid Clustering Technique for Brain Tumor Segmentation and Validation in Mri Images. *Aust. J. Basic & Appl. Sci.*, 10(8): 84-96, 2016

Therefore for advancement of a successful CAD scheme it is essential to create computer algorithm, as well as to research how valuable the computer output would be for radiologists in their diagnosis, how to evaluate the benefits of the computer output for radiologists. Brain tissue has a complex structure and its segmentation is an important step. Automated brain disorder diagnosis with MR images is one of the challenging medical image analysis methodologies. The impact of image segmentation happens as an arrangement of locales that jointly covers the entire image (Janani, V. and P. Meena, 2013). Medical image segmentation plays a significant role in clinical diagnosis and considered as a complicated problem since medical images normally have poor contrasts, different sorts of noise, and absence or diffusive limits (Dong, B. and A. Chien, 2010). Normally, the anatomy of brain tumor can be examined by MRI scan or CT scan. The main benefit of MRI over CT scan is, it does not include any radiation. MRI provides accurate visualization of anatomical structure of tissues. The anatomy of the brain can be scanned by Magnetic Resonance Imaging (MRI) scan or computed tomography (CT) scan. The MRI scan is more agreeable than CT scan for diagnosis. It does not influence the human body because it does not utilize any radiation. It depends on the magnetic field and radio waves (Patel, J. and K. Doshi, 2014). The lifetime of the person who is affected by the brain tumor will amplify if it is detected early (Rohit, M., 2013). Thus, there is a requirement for an efficient medical image segmentation method with some preferred properties such as minimum user interaction, quick computation, precise, and strong segmentation results (Aslam, H.A., T. Ramashri, 2013). On the other hand, image segmentation algorithms are based on one of the two primary properties of image strength values: discontinuity and similarity (Acharya, J., 2013). In the formal classification, the segmentation method depends on partitioning the processed image based on changes in intensity, such as edges and corners.

One perspective of image segmentation is a grouping problem that concerns how to determine which pixels in an image belong together most appropriately. There is a numerous literature on the methods that perform image segmentation based on clustering techniques. These methods typically demonstrate clustering in one of the two different ways, either by partitioning or by grouping pixels. Grouping, the pixels are collected together based on some assumption that determines how to group preferably (Kaur, J., 2012). The clustering algorithms used in image segmentation such as k means clusters, hard clustering and fuzzy clustering. For classification the analysis of cluster serves as pre-processing for other algorithms in order to calculate and detect the brain tumor we taken image segmentation approach called k means integrated with expectation maximization.

We used image segmentation techniques based on clustering to detect the brain tumor and calculating the tumor area and other quality metrics. We developed a novel image segmentation approach, called K-means integrated with expectation Maximization (EM) for abnormal MRI images. We incorporated K-means clustering algorithm with the EM algorithm to overcome the limitations and get benefits of them. Following clustering stage, the removal of the tumor is done automatically. Therefore, we get benefits from integrating the algorithms to decrease the number of iterations, which affect implementation time and gives an exact result in tumor detection. Clustering is also popular method for medical image segmentation. The distinctive issue with medical image segmentation is failure of information is improper and may result in losing potential diagnostic information. In this paper a concise study of the various segmentation techniques for the MR image segmentation are reviewed and talk about advantage and disadvantage regarding particular methods. This work presents a tool that is designed to automatically segment brain tumor in MRI images. The projected tool is developed in the form of a graphical user interface that is able for reading the image, performing necessary explanation and presents the processed results.

This paper is organized as follows. In Section 2, the current research in medical image segmentation is introduced. Section 3 shows the current brain tumor methods, pre-processing and proposed medical image segmentation system based on clustering. In Section 4&5, we analyze the evaluation and validation of the current brain tumor segmentation methods, experimental results and section 6 conclusions is discussed.

Ning Li and Miamian Liu (2007) has proposed a system Image Segmentation Algorithm using Watershed Transform and Level Set Method. It combines the watershed transform and region Based level set method. The watershed is initially used to pre segment the image. The section based level set strategy is then associated for separating the boundaries of objects. The addicted time does not rely on upon the span of the image but rather the quantity of pre fragmented regions. This strategy is computationally efficient.

Ruoyu Du and Hyo Jong Lee (2009) has proposed A modified FCM segmentation for brain MR images. An enhanced segmentation technique has been proposed in this paper on the basis of FCM clustering algorithm. The neighbour pixels of targets are varied by applying the Sigma filter principle. The proposed algorithm is compared with FCM algorithm in visual evaluation and quantitative evaluation thereby the efficacy of the proposed method was demonstrated.

Yongyue Zhang and Michael Brady (2001) has proposed a Segmentation of Brain MR Images through a Hidden Markov Random Field Model and the probability Maximization. HMRF model is a stochastic procedure produced by a MRF whose state grouping can't be watched specifically yet can be in a roundabout way evaluated through perceptions. HMRF EM structure, an exact and strong segmentation can be accomplish

Christe *et al.* (2010) has proposed a improved hybrid segmentation of brain mri tissue and tumor using statistical feature. In this paper the fuzzy c means and K means segmentation has been proposed In addition to this, new hybrid segmentation technique, namely, Fuzzy Kohonen means of image segmentation is proposed and implemented along with standard pixel value clustering method. It is found that the feature based hybrid segmentation gives improved performance metric and improved classification accuracy rather than pixel based segmentation

Funmilola *et al.* (2012) made the Fuzzy K-C-means method. Here the calculation peruses the image, determines the iterations, reduces the iterations by distance checker, gets the extent of the image, concatenates the measurement, generates vast data items with distance calculation, and reduces repetition when possible distance has been attained. The iteration begins by distinguishing critical segment of data then it stops when possible identification elapses.

Gopi karnam, Ramashri Thirumala (2015) has proposed a Design and development of a computer aided diagnosis (CAD) system for segmentation of brain tumor. In this paper a segmentation approach namely k means segmentation is implemented and this tool has provision for quantification of the segmented region.

Rajiv kumar and A.M.Arthanarjee (2014) has proposed a concept on comparative analysis of proposed image segmentation algorithm. In this paper a novel approach has been proposed. The parameter such as Global consistency error, variation of information was evaluated and the analysis was made with the existing algorithms.

Data Collection and Analysis:

The system is thought to be able to analyze double-blind studies of radiologist techniques. It can be designed so that various radiologists can comment the same image under various conditions without surveying the biopsy-based ground truth or the other explanations. Their data can then be seen by a user which can only view the doctor information after it has been automatically anonymized to maintain the integrity of the study. The outcomes can then be stored, and the study can be distributed by permitting user access. The graphical client interface for the segmentation is shown in figure1. The radiologists can store and compose medical images such as x-rays, CAT, MRI and other image that is stored in a DICOM format. The radiologist's diagnoses are captured and the data can be accessed remotely thereby allowing telemedicine applications. The CAD framework helps store and recover information and images for the radiologists who are the primary data collection agents in the development of medical image CAD work.

Pre-processing stage:

This phase is actualized by applying a series of initial processing procedures on the image before any unique purposes processing. It enhances the image quality and removes the noise. Since, the brain images are more sensitive than other medical images; they should be of minimum noise and extreme quality. Therefore, this stage consists of the accompanying two sub-stages: (a) De-noising: MRI images are usually contaminated by disturbances like Gaussian, poisson noise and impulse noise. The majority of the de-noising algorithms presume additive white Gaussian noise. In this paper, we used a hybrid filter which is a combination of adaptive median filter and wiener filter. The adaptive median filter has been applied as a superior method compared with standard median filtering. This filter classifies pixels as noise by comparing each pixel in the image to its surrounding neighbour pixels. The size of neighbourhood is adaptable as well as threshold for the similarity. The adaptive filter solves the dual purpose of removing the impulse noise from the image and reducing the distortion in the image. The output of this sub-step in pre-processing is the free noising MRI image. (b) Skull removal: Image background does not usually contain any helpful information but increase the processing time. Therefore, removing background, skull, scalp, eyes, and all structures that are not in the interest decrease the amount of the memory used and improved the processing speed. Skull removed is done by using BSE algorithm. The BSE algorithm is used only with MRI images. It filters the image to remove irregularities, distinguish edges in the image, and performs morphological erosions and brain isolation. It also performs surface cleanup and image masking. The output of this sub step is the free noising MRI image contains only the human brain.

The BSE calculation as follows:

This algorithm attempts to confine the brain from the rest of a T1-weighted MRI using a series of image manipulations. Basically, it depends on the fact that the brain is the largest area surrounded by a strong edge within an MRI of a patient's head. There are essentially four stages to the BSE algorithm:

- Step 1. Filtering the image to remove abnormalities.
- Step 2. Distinguishing edges in the image.
- Step 3. Performing morphological operations and brain isolation.
- Step 4. Perform image masking.

Clustering stage:

Now a days, image segmentation assume fundamental part in medical image processing. The segmentation of brain tumor from magnetic resonance images is an important assignment. Manual segmentation is one of the procedures for discovering tumor from the MRI. This strategy is time consuming but also generates errors. Several automated techniques have been developed for MRI segmentation.

Clustering the process of accumulation of objects which are comparative between them and are dissimilar objects belonging to other clusters. Clustering is suitable in biomedical image segmentation when the number of cluster is known for particular clustering of human anatomy. Clustering algorithm are classified two types: Exclusive clustering and overlapping clustering. In exclusive clustering, one data (pixel) is belonging only one cluster then it could not belong to another cluster. Kmeans is sample of select grouping algorithm. In overlapping clustering, one data (pixel) is belonging two or more clusters

K-Means Clustering:

K-Means calculation is an unsupervised algorithm that characterizes the input information into various classes in view of their inherent separation from one another. The calculation accept that the information highlights frame a vector space and tries to discover normal grouping in them. The points are clustered around centroids.

A cluster is a gathering of objects which are similar between them and are not at all like the objects having a place with different groups. clustering is an unsupervised learning system which manages discovering a structure in a gathering of unlabeled information. K-means clustering is a algorithm to gathering objects in view of attributes/features into k number of gatherings where k is a positive integer. The gathering (clustering) is finished by minimizing the Euclidean separation between the information and the comparing cluster centroid.

Finally, this algorithm aims at minimizing an *objective function*, in this case a squared error function. The objective function

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (1)$$

Where $\|x_i^{(j)} - c_j\|^2$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster centre c_j , is an indicator of the distance of the n data points from their respective cluster centers.

1. In the image let D be the data points
2. Partition the data points into k equal sets.
3. Take the middle point in each set as initial centroid.
4. Calculate the distance between each data point ($1 \leq i \leq n$) to all initial centroids ($1 \leq j \leq k$).
5. For each data point di , find the closest centroid cj and assign di to cluster j.
6. Set $[i] = j$.
7. Set $[i] = (di, cj)$.
8. For each cluster ($1 \leq j \leq k$), recalculate the centroids.
9. for each data point di , (i) compute its distance from the centroid of the present nearest cluster. (ii) The data point stays in the same cluster If this distance is less than or equal to the present nearest distance, Otherwise compute the distance (di, cj) for every centroid ($1 \leq j \leq k$).
10. Repeat from steps 5 to 9 until convergence is meet.

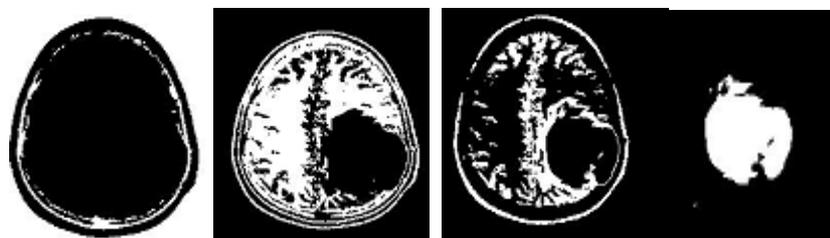


Fig. 1: Results of K means segmentation of Brain MRI image.

The Kmean algorithm clusters the image according to some characteristics(Figure 1) is the output for K-Means algorithm with four clusters. At the fourth cluster the tumor is extracted.The GUI tool for K-means segmentation shown in (Figure 2).

Watershed Transform:

Watershed segmentation is a most efficient segmentation originating from the field of numerical morphology. The instinctive thought of this transform is quite simple: if we consider the image as a scene or topographic relief, where the height of every point is directly identified with its gray level, and consider rain gradually falling on the territory, then the watersheds are the lines that separate the "lakes" really called catchment basins that form. The watershed transform is computed on the gradient of the original image, so that the catchment bowl limits are situated at high inclination points. This transform has been generally utilized as a part of numerous fields of image processing, including medicinal MR image segmentation, because of the advantages that it possesses. It is very simple, instinctive, quick, parallelized strategy and produces a complete division of the image in separated regions regardless of the fact that the complexity is poor, along these lines staying away from the requirement for any sort of contour joining. Some essential disadvantages related to the watershed transform are the over segmentation and poor detection areas with low contrast that commonly results in MRI brain images. Marker Controlled Watershed Segmentation Touching objects detachment in an image is the one of the most difficult works in image processing. The watershed transform is often applied to this issue.

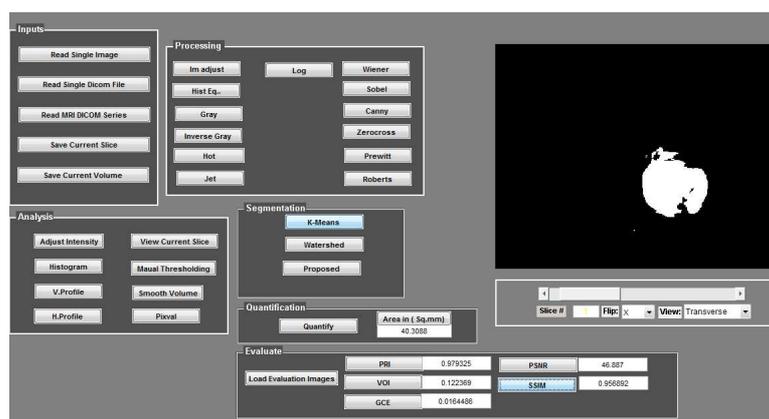


Fig. 2: GUI for K-Means segmentation of brain MRI image.

Marker Controlled Watershed Segmentation Touching items partition in a picture is a standout amongst the most troublesome works in medicinal picture handling. The watershed change is frequently connected to this issue. The watershed transform discovers catchment basins and watershed edge lines in a image by treating it as a surface where light pixels are high and dark pixels are low. Segmentation utilizing the watershed transforms works better, if you can identify or mark foreground objects and background locations.

Marker-controlled watershed division takes after this essential methodology:

1. Compute a segmentation function. This is a image whose dark regions are the items you are attempting to segment.
2. Compute frontal area markers. These are joined blobs of pixels inside of each of the objects.
3. Compute background markers. These are pixels that are not a portion of any object.
4. Change the segmentation function so that it just has maxima at the foreground view and background marker areas.
5. Compute the watershed transform of the altered segmentation capacity

The detailed view of watershed transform with all processing images as shown below in (Figure 4).The GUI for watershed segmentation of brain MRI image shown in (Figure 3).

Methodology:

In the proposed approach an improved K -means algorithm and EM algorithm are combined to formulate a hybrid strategy for better clustering. The proposed approach aims to exploit the capability of providing well distributed cluster of K-means and the compactness of clusters provided by EM. The initial clusters are provided by the improved K-means algorithm. This initial clustering operation results in centers which are widely spread in the given data. These centers form the initial variable for EM, which subsequently uses these variables and iterates to find the local maxima.

EM is an iterative algorithm which employs a model based approach to identify clusters. EM is classified as a suitable optimization algorithm for constructing suitable statistical model of data unlike distance based or hard member ship algorithm such as K-means algorithm. . EM is widely used in applications such as computer vision, speech processing and pattern recognition. The clustering operation of EM differs from the clustering

operation of K means in that EM starts with an initial estimate for the missing variable and iterates to find the maximum likelihood (ML) for these variables. Unlike K-means, in EM the numbers of clusters that are desired are predetermined. It is initialized with values for unknown (hidden) variables. Since EM uses maximum likelihood it most likely converges to local maxima, around the beginning values. Hence selection of initial values is vital for EM. However, the EM algorithm works well on clustering data when the number of clusters is known. The proposed algorithm has described as follows.

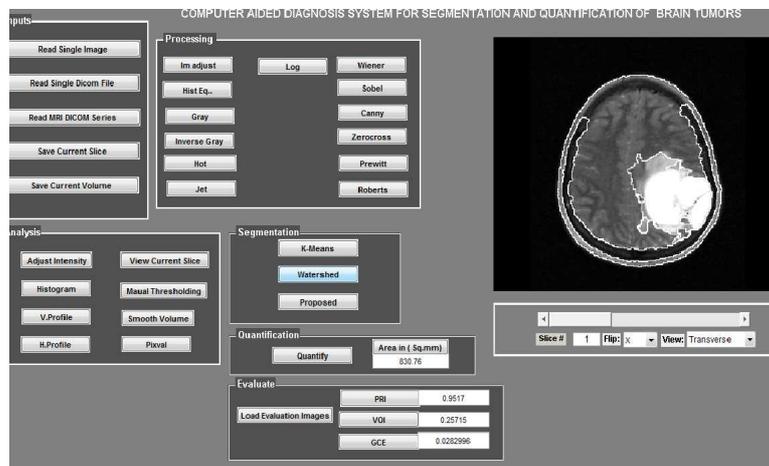


Fig. 3: GUI for watershed segmentation of brain MRI image.

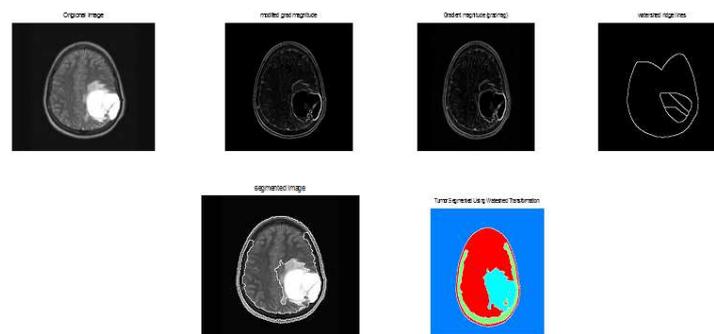


Fig. 4: Results of watershed segmentation of brain MRI image a)original image b)modified gradient image c)watershed ridge line d)gradient magnitude e) segmented image f)tumor segmented using watershed.

Hybrid Improved K-means Clustering and EM Algorithm:

The k-means clustering calculation creates k points as starting centroids subjectively, where k is a user indicated parameter. Every point is then assigned to the cluster with the nearest centroid. At that point the centroid of every group is updated by taking the mean of the data points of every cluster. Euclidean distance is utilized to find separation between data points and centroids.

The steps involved in simple K-means clustering can be illustrated as:

- Step 1. Arbitrarily choosing k data points from D as initial centroids.
- Step 2. Assigning each point d_i to the cluster which has the closest centroid.
- Step 3. Calculate the new mean for each cluster.
- Step 4. If convergence criteria is met then giving the clusters or going back to Step 2.

In the proposed improvement, a computationally less complex approach is suggested to identify better initial clusters thereby enhancing the efficiency and performance of the clustering operation. The steps involved in the implementation of the improved K-means clustering are mentioned below.

- Step 1: Considering middle point in each data set as the initial centroids.
- Step 2: Computing the Euclidean distance for each data point from the origin.
- Step 3: Sorting the obtained data point using the distance computed.
- Step 4: Portioning the sorted data points into K equal sets.
- Step 5: Considering the middle point in each set as the initial centroid.
- Step 6: Computing the distance between each data point to all the initial centroids.

- Step 7:** Finding the closest centroid c_j and assign d_i to cluster j for each data point d_i
- Step 8:** Setting the Cluster $Id[i] = j$. // j : Id of the closest cluster.
- Step 9:** Setting the nearest $Dist[i] = d(d_i, c_j)$.
- Step 10:** Recalculate the centroids for each cluster j ($1 \leq j \leq k$),
- Step 11:** For each data point d_i , its distance from the centroid of the present nearest cluster is calculated. If this distance is less than or equal to the present nearest distance, the data point stays in the same cluster. The operation moves to Step.
- Step 12:** For every centroid c_j ($1 \leq j \leq k$) the distance $d(d_i, c_j)$ is computed.
- Step 13:** If convergence criteria are met then giving the clusters or going back to Step 2.

Expectation maximization clustering evaluates the probability densities of the classes utilizing the Expectation Maximization (EM) algorithm. The algorithm depends on finding the extreme likelihood estimates of parameters when the information model relies on certain idle variables. In this proposed approach the initial clusters are identified with the help of improved K –means algorithm and alternating steps of Expectation (E) and Maximization (M) are performed iteratively till the outcomes join.

In EM, alternating steps of Expectation (E) and Maximization (M) are performed iteratively till the outcomes merge. The E step processes an expectation of the likelihood by including the dormant variables as if they were observed, and maximization (M) step, which figures the maximum probability evaluations of the parameters by maximizing the expected likelihood found on the last E step. The parameters found on the M step are then used to begin another E step, and the procedure is repeated until convergence .

Precisely for a known training dataset $\{x(1), x(2) \dots \dots x(m)\}$ and model $p(x, z)$

Where z is the latent variable, we have:

$$l(\theta) = \sum_1^m \log p(x; \theta) \tag{2}$$

$$\sum_1^m \log \sum_z p(x; z; \theta) \tag{3}$$

As can be seen from the above equation, the log probability is described in terms of x, z and θ . But since z , the latent variable is not known, we utilize approximations in its place. These approximations obtain the form of E & M steps mentioned above and formulated mathematically below.

E Step, for each i:

$$Q_i(z^{(i)}) := P(z^{(i)} / x^{(i)}; \theta) \tag{4}$$

M Step, for all z:

$$\theta := \operatorname{argmax}_{\theta} \sum_i \sum_{z^{(i)}} Q_i(z^{(i)}) \log \frac{p(x^{(i)}, z^{(i)}; \theta)}{Q_i(z^{(i)})} \tag{5}$$

Where Q_i the posterior distribution of is $z^{(i)}$,s given the $x^{(i)}$ s. Conceptually



Fig. 5: Stages of tumor segmentation using Hybrid Improved K –means Clustering and EM Algorithm.

The detailed view of Hybrid Improved K –means Clustering and EM Algorithm with all processing images as shown below in (Figure 5).The GUI for watershed segmentation of brain MRI image shown in (Figure 6).

RESULTS AND DISCUSSION

Two sorts of experiments are done to assess the execution of the proposed approach qualitatively and quantitatively. The performance measures used to analyze the execution are PRI (Probabilistic Rand Index), VOI (variation of information) &GCE (global consistency error). PRI assesses the pair wise connections between pixels of divided result and numerous ground-truth divisions and takes values in the reach (0, 1). Hence

higher PRI esteem shows a superior match between the sectioned result and the ground-truth information. VOI, GCE are error measures that have to be reduced by the good segmentation algorithm.

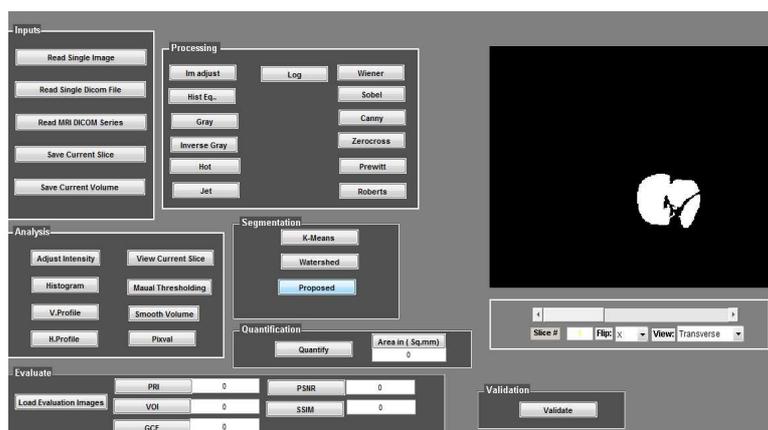


Fig. 6: GUI for Hybrid Improved K-means Clustering and EM Algorithm of brain MRI image.

Rand Index (RI):

The Rand Index (RI) checks the part of pairs of pixels whose marking are reliable between the computed segmentation and the ground truth averaging over numerous ground truth segmentations. The Rand measure is a measure of the similitude between two data clusters. Given a set of n components and two partitions of S to look at, and, we characterize the following:

- The quantity of pairs of elements in S that are in the same set in X and in the same set in Y
- The quantity of pairs of elements in S that are in different sets in X and in distinctive sets in Y
- The quantity of pairs of elements in S that are in the same set in X and in diverse sets in Y
- The quantity of sets of elements in S that are in different sets in X and in the same set in Y

The Rand Index (RI) defined as :

$$R = (a + b) / (a + b + c + d) = (a + b) / ((n/2)) \quad (6)$$

Where $a + b$ as the number of agreements between X and Y and $c + d$ as the number of disagreements between X and Y . The Rand index has a value between 0 and 1, with 0 indicating that the two data clusters do not agree on any pair of points and 1 indicating that the data clusters are exactly the same.

Variation of Information (VOI):

The Variation of Information (VOI) metric characterizes the distance between two segmentations as the average conditional entropy of one segmentation given the other, and in this manner measures the measure of randomness in one segmentation which can't be clarified by the other. Assume we have two clustering (a division of a set into a several subsets) X and Y where

$$X = \{X_1, X_2, \dots, X_k\}, \quad p_i = |X_i| / n, \quad n = \sum_k |X_i| \quad (7)$$

At that point the variety of information between two clustering is:

$$VI(X; Y) = H(X) + H(Y) - 2I(X, Y) \quad (8)$$

Where, $H(X)$ is entropy of X and

$I(X, Y)$ is mutual information between X and Y . The common data of two clustering is the loss of uncertainty of one grouping if the other is given. Consequently, shared positive and limited by $\{H(X), H(Y)\}_{\log_2(n)}$.

Global Consistency Error (GCE):

The Global Consistency Error (GCE) measures the degree to which one segmentation can be seen as a refinement of the other. Segmentations which are connected are thought to be steady, since they could represent to the same image segmented at distinctive scales. Segmentation is essentially a division of the pixels of a picture into sets. The portions are sets of pixels. If one segment is appropriate subset of the other, then the pixel lies in an area of refinement, and the error should be zero. If there is no Subset relationship, then the two regions overlap in an Inconsistent manner.

$$\tau = \sum_{j=1}^K \sum_{i=1}^K \mathbf{1} \parallel X^{(i)} - C_j \parallel^2 \quad \text{where } \parallel X^{(i)} - C_j \parallel^2 \quad (9)$$

Where, segmentation error measure takes two divisions S1 furthermore, S2 as input, and produces a genuine esteemed output in the reach [0::1] where zero signifies no error. For a given pixel pi consider the segments in S1 and S2 that contain that pixel.

Calculation the tumor region:

The tumor area region is calculated by the following equation:

$$\text{Tumor area} = A \text{ total number of pixel in the tumor region}$$

$$A = V \times H \quad (10)$$

Where, A=the area of each pixel
H=horizontal dimension of the image
V=vertical dimension of the image

Peak signal to noise ratio(PSNR):

The Peak Signal to Noise Ratio (PSNR) is used to find the deviation of segmented image from the ground truth image. Equation (8) represents the PSNR. In this equation mean squared error (MSE) for two M * N monochrome images f and z and it is given by Equation (9). MaxBits gives the maximum possible pixel value (255) of the image.

$$PSNR = 10 \times \log_{10} \frac{MaxBits^2}{MSE} \quad (11)$$

$$MSE = \frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} ((f(x, y) - z(x, y))^2) \quad (12)$$

Structural similarity (SSIM):

Structural Similarity Index (SSIM) is a method for measuring the similarity between two images. The SSIM is measured between two windows X and Y of common size N*N on image using Eq. (3).

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (13)$$

The amount of quantitative information available on medical images is enormous. Computer quantification may hold more potential than computerised detection. The main reason that most radiologic scoring systems are not used routinely in clinical practice that they are too time consuming and cumbersome to apply. Current systems for computer aided detection have been introduced as complementary tools that draw the radiologist attention to certain image areas that need further evaluation. The proposed calculations have been implemented utilizing MATLAB. The execution of different image segmentation methodologies are analyzed and talked about. The measurement of image segmentation is hard to measure. There is no regular algorithm for the image segmentation. The measurable estimations could be utilized to measure the nature of the image segmentation. The rand index (RI), global consistency error (GCE) & VOI are utilized to assess the execution. The detailed description with formulae of RI, GCE, VOI parameters are clarified in point of interest as follows.

Table 1: Performance evaluation.

METHODS	PRI	VOI	GCE	PSNR	SSIM	Tumor area(sq.mm)
K-Means	0.9796	0.126	0.01663	46.88	0.9568	40.3088
WATERSHED	0.5847	1.02	0.0782	27.94	0.3387	830.76
PROPOSED METHOD	0.994645	0.041	0.00496	72.1978	0.992413	29.2913

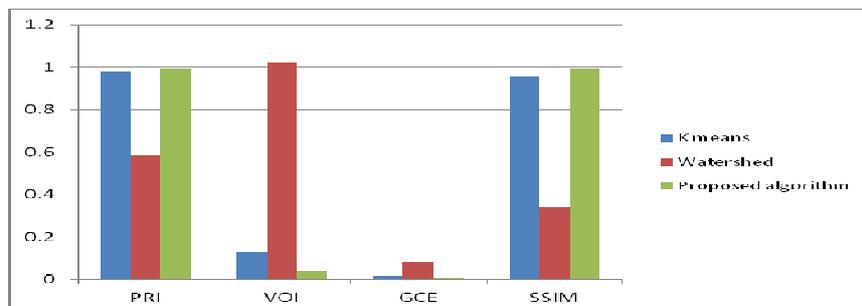


Fig. 7: The performance analysis chart.

The (figure 7) performance analysis chart reveals that the PRI, SSIM values are higher and VOI and GCE values are lower for proposed algorithm compared to k means and watershed algorithm.

The experiment is conducted over the Brain tumor MRI images using the algorithms k means, watershed and the proposed method their outcomes appeared in (Figure2 ,Figure3 and figure6)with required measurable parameters and their outcomes are exhibited in Table 1.If The estimation of PRI,PSNR, SSIM is higher and GCE, VOI are lower than the segmentation methodology is better. The proposed algorithm shows the GCE and VOI Values low and PSNR, PRI values high compare to other methods.

The sequence followed for validating the developed tool is:

- The images obtained from two sources are input to the GUI-CAD tool separately.
- The tool performs required operations and produces output image views that contain segmented Region of Interest of brain tumor, quantification, evaluate based on PRI (Probabilistic Rand Index), VOI (variation of information) &GCE (global consistency error), Structural similarity index(ssim) .These output image views for various Brain MRI images are shown to the expert radiologists and are asked to classify.
- Based on the segmented Region of Interest, quantification and evaluation the clinical experts classified the brain tumor as True Positive, False Positive, True Negative, and False Negative. The experts' classifications are shown in Table. II

Table II: Observed values by Clinical Experts.

Clinical Expert	Image Data Base	True Positive	False Positive	True Negative	False Negative
I	Image Set 1	93% (139)	7% (11)	95% (114)	5% (6)
	Image Set 2	95% (143)	5% (7)	93% (111)	7% (9)
II	Image Set 1	95% (143)	5% (7)	93% (111)	7% (9)
	Image Set 2	91% (137)	9% (13)	90% (108)	10% (12)
III	Image Set 1	96% (144)	4% (6)	98% (117)	2% (3)
	Image Set 2	94% (141)	6% (9)	93% (112)	7% (8)
IV	Image Set 1	90% (135)	10% (15)	94% (113)	6% (7)
	Image Set 2	92% (138)	8% (12)	90% (108)	10% (12)
V	Image Set 1	91% (137)	9% (13)	90% (108)	10% (12)
	Image Set 2	90% (135)	10% (15)	94% (113)	6% (7)

➤ With the above classified values, performance assessment parameters of diagnostic test such as Accuracy, Error, Sensitivity, Specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV), False Discovery Rate (FDR), Matthews Correlation Coefficient (MCC), False Prediction Rate, False Negative Rate, Prediction Conditioned Fallout, Prediction Conditioned Miss, Rate of Positive Prediction , Rate of Negative Prediction, Odds Ratio, Likelihood Ratio positive, Likelihood Ratio negative, Prevalence, Pretest odd, Posttest odds of outcome for given positive test results. The computed values and plots of various radiological diagnosis parameters for the classification values of Table. I of two input image sets and five experts opinion using the add-on user interface performance measure tool integrated in GUI-CAD are directly converted in to Microsoft-Excel sheet for reference and record shown in Figure.8. Each column corresponds to parameters computed for individual experts respectively under image set 1 and 2. All the individual plots of the radiological parameters are saved in separate sheets for easy access by the experts'.

Based on the quantification details in the output image views obtained from GUI-CAD tool, the clinical experts classify the tumor as True Positive, False Positive, True Negative, and False Negative. These are the input parameters for the performance measure tool, tabulated in Excel sheet and loaded as input data to measure tool (Figure 8 and 9). The various parameters related to radiological diagnosis mentioned above are computed and corresponding graphs are plotted and can be viewed on the performance measure tool itself for analysis.

Discussion:

The output image views of the GUI-CAD tool for different brain tumor images from two image sets are shown to expert radiologists and are asked to classify the diagnosis test of detected tumor according to their opinion in to True Positive, False Positive, True Negative, and False Negative. The experts' classifications are shown in (Figure10). Based upon these classified values, twenty one (21) performance assessment parameters

are computed for evaluation of the diagnostic test implemented on the tool. All these parameters are computed by the tool for classified values of expert radiologists' and are displayed in Excel sheet. Plot of performance assessment such as Accuracy, Sensitivity parameters are shown in (Figure 11).

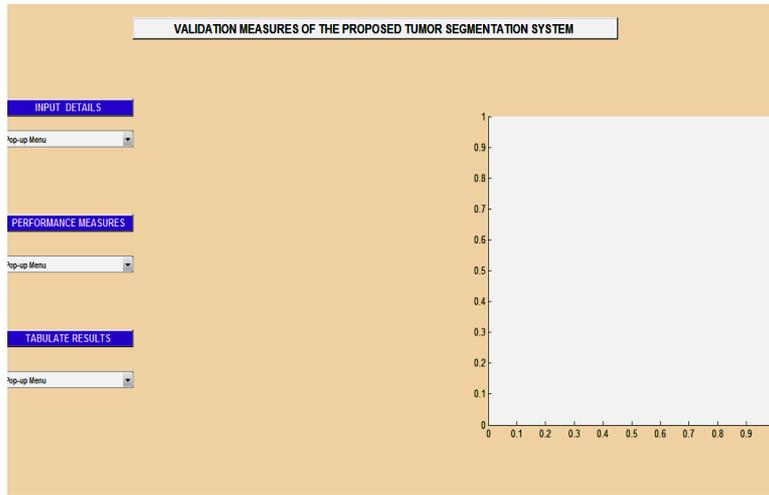


Fig.8: Validation measures of the proposed segmentation.

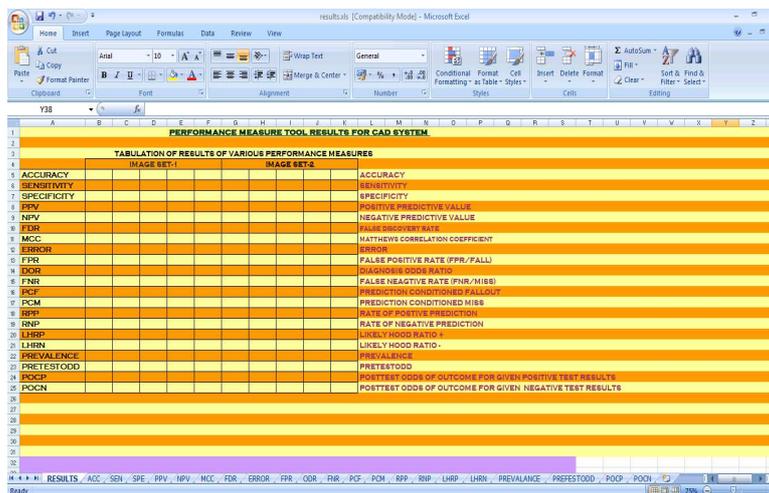


Fig. 9: Performance measure tool for cad system.

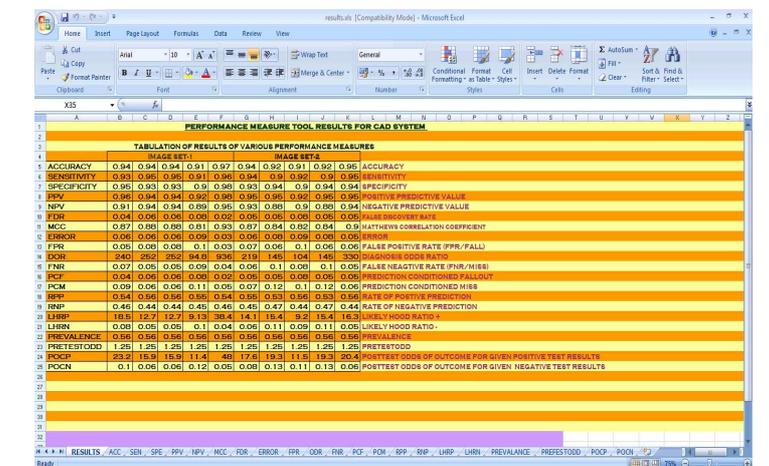


Fig. 10: User Interface performance measure tool for CAD system: Results Excel sheet of various performance assessment parameters and sheets of individual parameters.

The comparison was tested according to the following measures:

True Positive (TP) = number of resulted images having brain tumor/total number of images

True Negative (TN) =number of images that haven't tumor/total number of images

False positive (FP) =number of images that haven't tumor and detected positive/ total number of images

False Negative (FN) =number of images have tumor and not detected/total number of images.

Accuracy= (TP+TN)/(TP+TN+FP+FN)

Recall= (TP/TP+TN)

Precision= (TP/TP+FP)

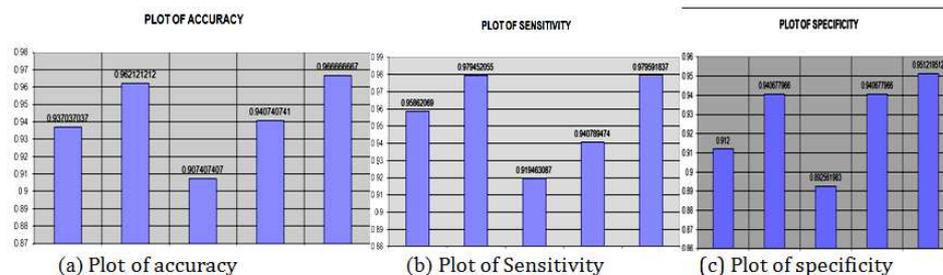


Fig. 11: Plot of performance assessment parameters.

Conclusion:

The results and observations show that the research work is successful in segmenting and identifying brain tumor using Hybrid Improved K –means Clustering and EM Algorithm. The proposed method gives more accurate result compare with the others. The performance of these algorithms is measured using segmentation parameters RI, GCE, VOI, and PSNR, SSIM. The computational results showed that the proposed image segmentation shows accurate results and it performs better than the watershed and k means. Accordingly, we compare the segmentation performance in brain tissue. The GUI-CAD tool designed for aforesaid classification has proved to be computationally less complex and effective in identifying the brain tumors. In addition to primary objective of designing GUI-CAD system to classify brain tumor, research has successfully achieved all objectives initially stated. The resultant images of the GUI-CAD tool provide a second opinion to the radiologists in discriminating lung nodules from blood vessels. The effectiveness of the CAD tool is analysed using a variety of performance assessment parameters. This quantification helps in abstracting clinical opinion to provide a better diagnosis and prognosis. The performance measures of the tool indicate the suitability and reliability of the tool in providing an authenticated secondary opinion. The research work has the potential to serve as an effective tool in managing and early diagnosis of brain tumors. This serves as a basis for the diagnosis, the choice of treatment and the prognosis.

REFERENCES

- Janani, V. and P. Meena, 2013. Image segmentation for tumor detection using fuzzy inference system. Int J Comput Sci Mobile Comput(IJCSMC), 2(5): 244–8.
- Dong, B. and A. Chien, 2010. Frame based segmentation for medical images. Commun Math Sci., 32(4):1724–39.
- Patel, J. and K. Doshi, 2014. A study of segmentation methods for detection of tumor in brain MRI. Adv Electron Electr Eng., 4(3): 279–84.
- Rohit, M., 2013. Segmentation of brain tumour and its area calculation in brain MRI images using K-mean clustering and Fuzzy C-mean algorithm. Int J Comput Sci Eng Technol (IJCSET), 4(5): 524–31.
- Aslam, H.A., T. Ramashri, 2013. A new approach to image segmentation for brain tumor detection using pillar K-means algorithm. Int J Adv Res Comput Commun Eng., 2: 1429–36.
- Acharya, J., 2103. Segmentation techniques for image analysis: a review. Int J Comput Sci Manage Res., 2(4): 1218–21.
- Naik, D. and P. Shah, 2014. A review on image segmentation clustering algorithms. Int J Comput Sci Inform Technol, 5(3): 3289–93.
- Christe, S.A., 2010. A. Improved hybrid segmentation of brain MRI tissue and tumor using statistical features. I CTACT J Image Video Process, 1(1): 34–49.
- Seerha, G.K. and R. Kaur, 2013. Review on recent image segmentation techniques. Int J Comput Sci Eng (IJCSE), 5(2): 109–12.
- Dass, R., Priyanka, S. Devi, 2012. Image segmentation techniques. Int J Electron Commun Technol, 3(1): 66–70.

- Kaur, J., 2012. Integration of clustering, optimization and partial differential equation method for improved image segmentation. *Int J Image Graph Signal Process*, 4(11): 26–33.
- Panda, M. and M.R. Patra 2008. Some clustering algorithms to enhance the performance of the network intrusion detection system. *J TheorAppl Inform Technol* ;4(8):795–801.
- Ning Li, 2007. "Image Segmentation Algorithm using Watershed Transform and Level Set Method". *International Conference on Acoustics, Speech and Signal Processing*", I-613-I-616.
- Ruoyu Du and Hyo Jong Lee, 2009."A modified-FCM segmentation algorithm for brain MR images", *Proceedings of ACM International Conference on Hybrid Information Technology*, 25-27.
- Yongyue Zhang and Michael Brady, 2001. Segmentation of Brain MR Images Through a Hidden Markov Random Field Model and the Expectation-Maximization Algorithm. *IEEE Transactions On Medical Imaging*, 20-1.
- Allin christe, S., 2010. Improved hybrid segmentation of brain MRI tissue and Tumor using statistical features *ICTAT journal of image and video processing*, 01.
- Ajala Funmilola, A., 2012."Fuzzy kmeans Clustering Algorithm for Medical Image Segmentation", *Journal of Information Engineering and Applications*, ISSN 22245782 (print) ISSN 2225-0506 (online), 2-6.
- Meena, A., K. Raja, 2013. Spatial fuzzy c-means PET image segmentation of neurodegenerative disorder. *Indian Journal of Computer Science and Engineering*.
- Gopi Karnam and Ramashri Thirumala, 2015. "design and development of a computer aided diagnosis system for segmentation of brain tumor " in *ARNP Journal of Engineering and Applied Sciences*, 10-22, ISSN 1819-660.
- Rajiv kumar and Arthanarjee, 2014 performance evaluation and comparative analysis of proposed image segmentation algorithm . *Indian Journal of Science and Technology*, 7(1): 39–47.
- Rajendran, P., M. Madheswaran, 2009. An Improved Image Mining Technique For Brain Classification Using Efficient classifier. *International Journal of Computer Science and Information Security*, 6(3): 107-116.
- Ibrahiem, M. and S. Ramakrishnan, 2008. On the application of various probabilistic neural networks in solving different pattern classification problems. *World Applied Sciences Journal*, 4: 772- 780.
- Sandeep, C. and N. Jaganathan, 2006. Classification of MR brain images using wavelets as input to SVM and neural network. *Biomedical signal processing and control*, 1: 86-92.
- Arjan, S., 2004. Combination of feature reduced MR spectroscopic and MR imaging data for improved brain tumor classification. *Nucl.Mag. Res. Biomed*, 18: 34-43.
- Hui, 2006 .Clustering Ensemble Technique Applied in the Discovery and Diagnosis of Brain Lesions. In *Proc: Sixth International Conference on Intelligent Systems Design and Applications (ISDA)*, 2: 512-520.