A Novel Two-Stage Approach For Automatic Detection of Brain Tumor

Dr. A. Murthi, S. Shameer
Associate professor, Department of EEE, Government College of Engineering, Salem(TN),India.
a.murthi2007@yahoo.com
PG Scholar, Department of EEE, Government College of Engineering, Salem(TN),India.
shameerresearch@gmail.com

ABSTRACT

Brain tumor is one of the most life-threatening diseases, and it is the most common type of cancer that occurs among those in the age group belonging to 0-19. It is also a major cause of cancer-related deaths in children (males and females) under age 20 hence its detection should be fast and accurate. Manual detection of brain tumors using MRI scan images is effective but time-consuming. Many automation techniques and algorithms for detection of brain tumors are being proposed recently. In this paper, we propose an integrated two-step approach combining modified K-means clustering algorithm and Hierarchical Centroid Shape Descriptor (HCSD). The images are clustered using modified K-means based on pixel intensity, and then HCSD helps to select those having a specific shape thus making this approach more effective and reliable. Simulation of the proposed work is done in MATLAB R2013a. Tests are carried out on T1 weighted MRI scan images.

Indexing terms/Keywords

Brain tumor detection, Hierarchical Centroid Shape Descriptor, Modified K-means, MRI, Dice index.

Academic Discipline And Sub-Disciplines

Provide examples of relevant academic disciplines for this journal: E.g., History; Education; Sociology; Psychology; Cultural Studies;

SUBJECT CLASSIFICATION

E.g., Mathematics Subject Classification; Library of Congress Classification

TYPE (METHOD/APPROACH)

Different kinds of tumors lead to an inhomogeneity of their sizes, shapes, textures, locations and intensities, where the automatic detection of abnormal tissues is a challenging task. In [3], the authors show that the automatic tumor detection can be achieved by using some features like texture, shape, intensity, and symmetry. In [4], methods of modified texture based region growing and cellular automata edge detection were used for brain tumor detection. A stochastic model to extract tumor texture was proposed in [5]. The variety of tumor texture makes it difficult to detect them based on shapes and produces unsatisfactory results. In [6] for classifying brain tumor images the Fuzzy C- means algorithm is used, and it works well for tumor detection. A mean shift algorithm is proposed in [7] similar to [6] but both have a high degree of complexity but are best suited when the number of clusters are unknown. Segmentation based on brain symmetry was done by [8] and [9] where the former used texture and intensity for detection of tumors. The application of this kind of feature is limited on axial and coronal planes because there are no symmetric structures in the sagittal plane.The authors in [10] propose to classify brain image as normal or abnormal by using neural network. A computer-aided detection system for detecting tumors based on histogram equalization and morphological mathematical operations was work realized by [11]. The Watershed segmentation method is presented in [12] for brain tumor detection. In [13] the thresholding method of Otsu associated with the Particle Swarm Optimization algorithm for maximizing the optimal threshold values was applied on medical images for tumors detection.

INTRODUCTION

Tumor also known as neoplasm is the condition where a tissue grows unconditionally and forms an abnormal mass which can either be solid or liquid filled. All tumors are necessarily not cancerous in nature. Cancer cells are dead cells that do not enter the phase of regeneration, and they replicate very quickly affecting other healthy cells in our body which may lead to the death of a human being [1]. A brain tumor is an abnormal growth of tissue in the brain or central spine that can disrupt proper brain function.Doctors classify tumor based on where the tumor cells originate and replicate and also based on whether they are cancerous in nature or not. Benign are non-cancerous and least aggressive type of brain tumor, and are referred as benign brain tumor. They are clearly discriminated from other tissues as they have visual borders, and also their growth is very slow in nature. Malignant brain tumors are cancerous in nature and often they are very hard to discriminate as they don’t have clear borders. They are invasive in nature and can quickly grow and spread to surrounding tissues, so they are considered to be life threatening.Primary brain tumors originate in brain cells, and they spread to other parts of the brain or spine but rarely to other organs of the body. Metastatic also called as secondary brain tumors originate in other organs of the body and later spread to brain. They are named based on their origination in human body. Various types of imaging techniques like magnetic resonance imaging (MRI), computed tomography (CT), etc. exist for the diagnosis of brain tumor. MRI is the most prominent one since it does not use ionizing radiation and it exhibits a greater contrast between different soft tissues of the human body.Brain tumors can be detected manually by experts from the MRI images. But manual segmentation leads to difficulties such as over-time consumption, chances of variation of results from expert to expert and chances of getting computational error [2]. Our work focuses mainly on automatic detection of brain tumors with irregular shapes and also in images with low pixel intensity. The tests of our work are carried out mainly in T1 weighted MRI images by a using a robust method that takes into account the shape variations, texture, size, pixel intensity and tumor location by placing a bounding box on it.The anatomy of a human brain can be classified based on its intensity into three groups. The group number increases to four if some pathological tissues appear. The Gray Matter (GM), White Matter (WM), Cerebrospinal Fluid (CSF) and the tumor are the four classified groups. Since the intensity of CSF is very low in T1 weighted images, they are clustered along with black background image.

WORKING METHODOLOGY

The presence of pathological structures makes it difficult to obtain accurate results. So we must apply a pre-processing step as done in various works [14] – [16]. The pre-processing involves thresholding technique and morphological operations. The proposed working methodology involves two-step procedure. The first step invloves grouping the images using modified k-means clustering algorithm based on the initial cluster number 'k.' The second step is to use an Hierarchical Centroid Shape Descriptor that helps to locate the tumor and produce accurate results. The basic steps involved in the proposed algorithm is shown in Figure 1. The modified k-means method and its usage is motivated as the number of clusters is known and also the computational complexity is very less. In [18] k-means algorithm is used for clustering which produces almost very similar results but it is slower compared to modified k-means and also the problem of empty vectors are eliminated in modified k–means method. In [17] they have tested the clustering stage with Otsu algorithm which also produces similar results like k-means method. The clustering is based on pixel intensity, and the tumors belong to high intensity clusters. In real world data, along with abnormal tissues, healthy tissues are also grouped in the cluster which makes it necessary to have a second stage, in order to find healthy tissues that are misclassified and are found in the cluster of interest.

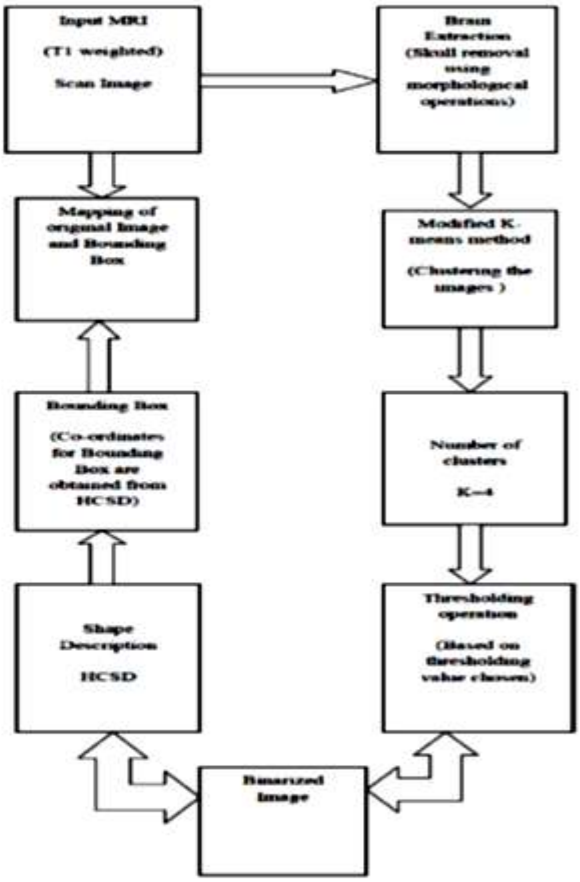


Figure 1: Proposed automatic tumor detection algorithm.

Modified K- Means

The clustering techniques involving the basic *k*-means algorithm have a prominent disadvantage, the generation of empty clusters at different instances due to bad initialization. A scheme to overcome the generation of empty clusters is Modified K-means algorithm. In this algorithm, the center vector updation procedure is modified so that it does not lead to the generation of empty clusters as in basic K-means method. In K-means algorithm, each iteration starts with a set of center vectors which are denoted as old center vectors $\mathbf{V}_k^{(old)}$. After each iteration the center vectors are updated to $\mathbf{V}_k^{(new)}$ by mathematicalformula described in the equation (1), where $n_k$is the number of elements in cluster C_k.

$$\frac{1}{n_k}\{\textstyle\sum_{x_j \in C_k}(x_j)\} \rightarrow \boldsymbol{V}_k^{(new)} \qquad (1)$$

The data elements denoted by x_j are distributed based on Euclidean distance among various clusters and then $\mathbf{V}_k^{(new)}$ is updated by averaging data elements as given in equation (1). The K-means algorithm enters next iteration under the assumption $\mathbf{V}_k^{(new)} = \mathbf{V}_k^{(old)}$ if the new center vectors are completely different from the old center vectors.In a Modified K-means algorithm, the old center vectors are also assumed to be members of the clusters whose new centers are to be computed. In this way the computation of new center vectors totally differ from basic algorithm and the mathematical equation for the updation of center vectors in Modified K-means is given as

$$\frac{1}{n_k+1}\{\textstyle\sum_{x_j \in C_k}(x_j) + \boldsymbol{V}_k^{(old)}\} \rightarrow \boldsymbol{V}_k^{(new)} \qquad (2)$$

Here by using this updation procedure, we almost eliminate the formation of empty clusters as equation (2) indicates that every cluster should contain atleast one element always. The inclusion of $\mathbf{V}_k^{(old)}$ in the computation of $\mathbf{V}_k^{(new)}$ does not have any considerable effect on the rate of convergence to final cluster centers as they are insignificant in case of large data sets because the value of $\mathbf{V}_k^{(old)}$ soon becomes well inside the concerned cluster within few iterations. The formation of empty clusters can be completely avoided when random initialization of initial center vectors are done while using a Modified K-means approach. Table 1 shows the performance comparison of K-means and Modified K-means algorithm based on the works done in [20], where both the algorithms are tested out on a large set of IRIS data and compared for generation of empty clusters. Modified k-mean algorithm does not provide a locally optimal solution in most cases, and at the same time it reduces the formation of empty clusters and getting into cluster - error criterion. The basic steps involved in a modified k-means clustering algorithm are discussed below:

Input image

↓

The clusters are decided based on variance of intensity in pixels

↓

The number of clusters is found / decided

↓

The initial centroid is calculated

↓

Distance between each object and centroid is calculated

↓

Grouping based on minimum distance

↓

When there are no more objects to move, terminate.

Table 1. Percentage of empty clusters using IRIS data

Number of Clusters	2	3	4	5	6
***Modified k*-means** **(random initialization)**	0	0	0	0	0
***k*-means** **(random initialization)**	0	0	3	9	12

Hierarchical Centroid Shape Descriptor

Shape descriptor algorithms basically have two approaches namely top down and bottom up approach. The HCSD is a top down approach where the input image is categorized into many levels for proper identification of tumor tissue in the brain. The HCSD is a shape descriptor built based on decomposing of image into sub-images by using Kd – tree decomposition technique.It uses the centroid co-ordinates extracted from the binary images of the brain. The binary images are obtained after the clustering stage by using a threshold value that helps to binarize the image and then the co-ordinates are extracted. The basic decomposition technique such as quad tree decomposition is analyzed before using kd tree technique. In the case of quad tree technique, each level of description produces four sub-images. In terms of general shape discrimination this approximation technique leads to potentially poor results under the circumstances where empty sub-images are yielded by the decomposition, when the co-ordinates at the root level are chosen arbitrarily. In K_d-tree decomposition this problem of empty sub-images are eliminated completely because during each level of decomposition the image is divided equally into two sub- images which is based on Kd tree splitting strategy[21].The strategy of K_d-tree decomposition is explained using the Fig 2 which gives a 3- level normalized feature vectors given in equation (3).

$$\bar{f} = (x_0^0, y_0^1, y_1^1, x_0^2, x_1^2, x_2^2, x_3^2,) \quad (3)$$

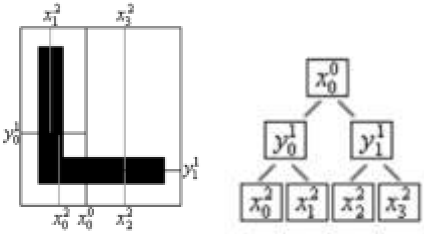


Figure 2: K_d-tree decomposition technique

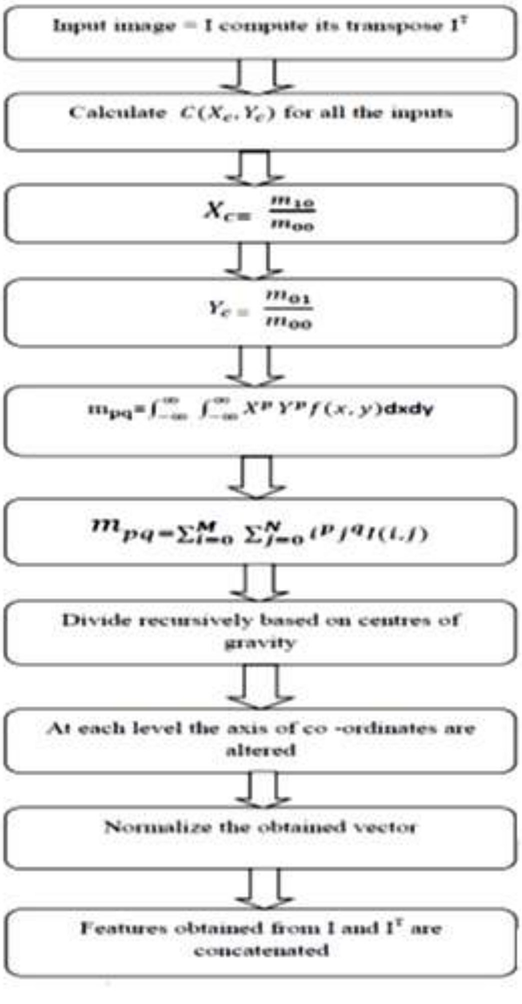


Figure 3: Steps involved in shape descripor (HCSD)

Equation (3) indicates that the process is a recursive decomposition in which x_0^0 is initial point. y_0^1, y_1^1 indicate the co-ordinates of vertical midpoints at 1^{st} level and this partition is continued up to the required depth.A very similar shape descriptor was proposed in [18], where the descriptor length is chosen as 2 x (2^d – 2), where d is defined as the depth of feature extraction. The basic steps involved in the proposed HCSD is explained with the help of a block diagram where various steps are illustrated shown in Fig 3.The image which is input to the shape descriptor and whose features are to be extracted is taken as I. The input image I is taken in the form of a M x N binary image with foreground I_{fg} and background $I_{bg.}$ The centroid at root level is given as C(Xc ,Yc) and m_{10} is the first order moment along x-axis , m_{01} is the first order moment along y-axis and m_{00} is the area of foreground I_{fg}.

RESULTS

The proposed method is analyzed using MATLAB v.R2013a tool. The tests were performed on T1 weighted MRI scan images. These medical images have been provided by Government Mohan Kumaramangalam Medical College Hospital, Salem in Tamil Nadu, India. In order to validate these results Jaccard index and Dice indexes are taken as metrics. The results of Modified K-means HCSD method is compared with delineated detected regions by an expert. The results of Otsu and k – means method are very similar as the segmented structures are almost identical but their shapes are different. The Modified K-means HCSD approach outperforms the results obtained with the one step method such as k-means or Otsu Multilevel thresholding. The score reached with HCSD K- means method is 0.842 Jaccard index and 0.91 based on Dice index[18]. Our approach implements a Modified K- means algorithm which is very similar in nature to K- means the metrics are very much better compared to the previous methods. The major advantage of using Modified K-means over K-means is that it avoids the formation of empty clusters at each iteration, and it is much faster compared to K-means method. In this paper, we have found similarity measure using dice index and compared it with previous works. Different dice indexes have been obtained for different images we have obtained an average value of dice index that can be reached using this method.Many algorithms fail to detect the tumor tissue when it has an irregular shape or when the image has poor contrast. The results of different test input images, obtained using the proposed method is shown in fig Figure 4 where the proposed method is able to locate the tumor in images with poor intensity and difficult shapes.Many research works are performed on simulated database but the main advantage of our work is that it was performed by using real patient data. The metrics obtained for various works done in the past and the proposed method is compared in Table 2. According to the data and analysis shown in Table 2 it is clear that Modified k-means HCSD approach has better similarity measures than the previous methods and it is faster compared to K-means HCSD.

Table 2. Average similarity measure using dice index

Comparative studies	**Similarity Measure**	
	Dice index	**Steps**
k-means	0.677	1
HCSD Multilevel Otsu	0.889	2
HCSD k-means	0.91	2
HCSD Modified K-means	0.93	2

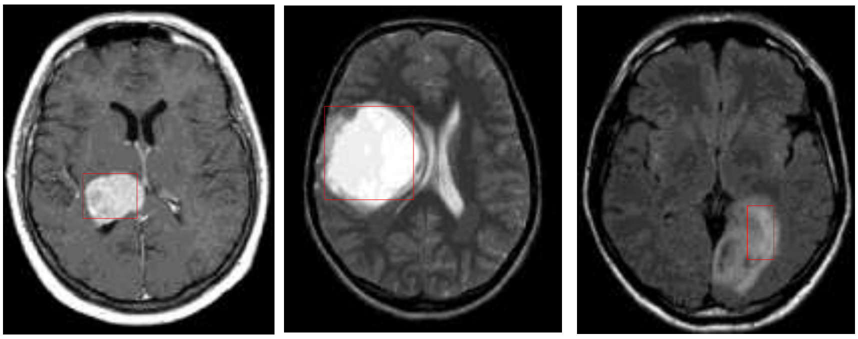


Figure 4: Output of test images showing locations of tumor with bounding box.

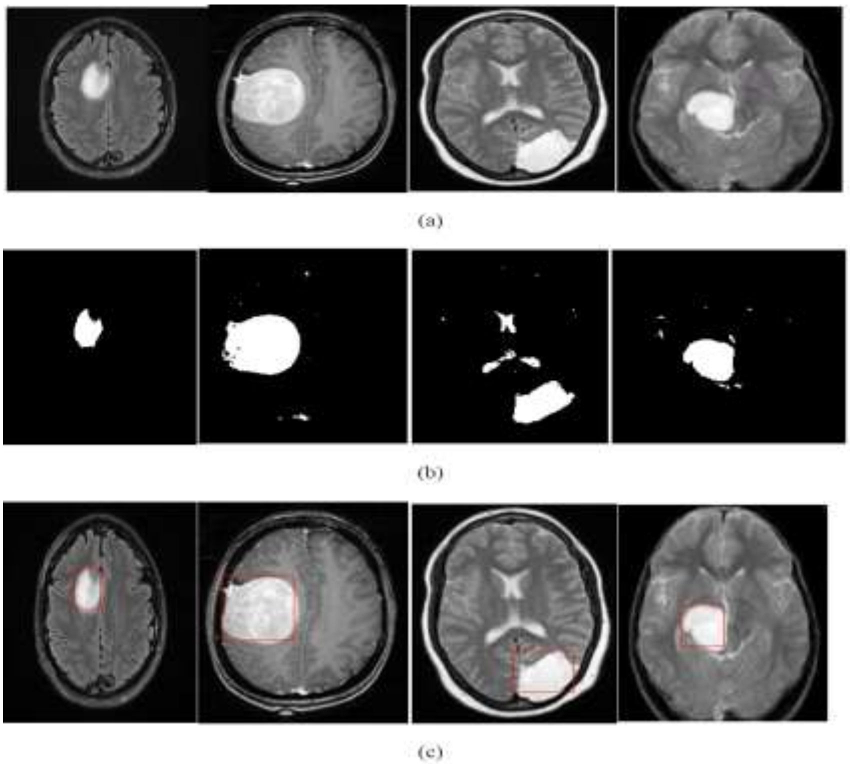
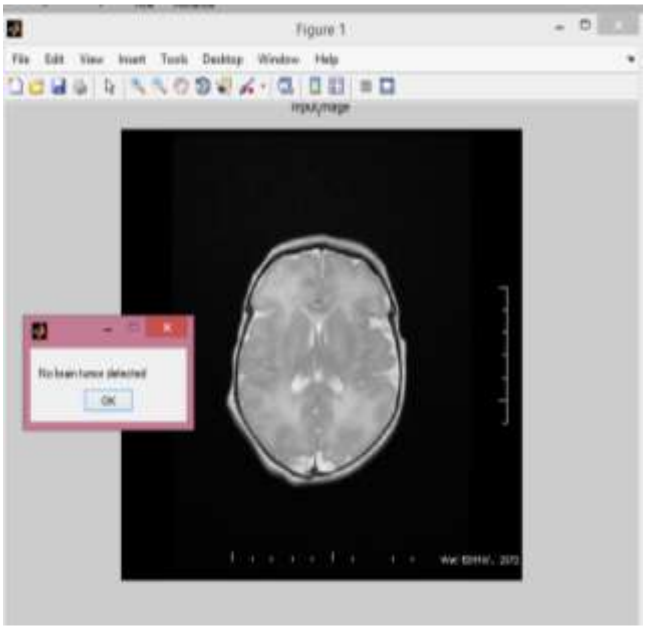


Figure 5: Brain tumor detection results: (a) Input images, (b) Modified K-means result, (c) Final output

In Figure 5 we have shown the results for various MRI T1 weighted scan images which are fed as input data and the results of Modified K-means are shown along with final result.In our proposed method we locate the exact location of tumor tissues and place a bounding box which is very useful for deciding the region of interest for further processing and diagnosis. We have also displayed a message stating that the input image has been detected to have a tumor. If normal MRI brain images are fed to the system the proposed algorithm extracts the brain structure analyses using k means for tumor tissue based on clustering technique further when there is no defective tissues found a message is diplayed on the screen stating that given input image is not affected by brain tumor. The output for a normal MRI T1 weighted scan image is shown in Figure 6.The accuarcy of the system is determined by the positive and negative results that are produced by the proposed method which is compared in Table 3.



No brain tumor detected

Figure 6: Output result for normal MRI scan image

Table 3. Accuracy of HCSD modified k-means

Input images		Output result	
Condition	**Number**	**Positive**	**Negative**
Tumor affected	15	15	00
Not affected by Tumumor	10	09	01

CONCLUSION

In this paper a two-stage approach for automatic detection of brain tumor was proposed, which combines the robustness and accuracy of Modified K-means clustering technique and shape descriptor built based on binarized images. In the first stage the images are clustered based on pixel intensity and then binarized by applying a threshold value. Mostly tumor tissues are surrounded by healthy tissues of the brain. In order to delineate the exact location of tumor we apply a second stage which uses binary shape descriptor to locate the tumor and discard other tissues and places a bounding box around the abnormal cells in MRI scan images. After this the original image is mapped with the result of HCSD to show the exact location of the tumor. The two-stage approach is found to have very good similarity measures when compared with expert opinion. Most of the tumor tissues have irregular shapes due to robust nature of this approach it is less prone to poor contrast in scan images and also complicated shape of tumors. In future the work can be extended to classify the type of tumor whether it is cancerous or not after detecting its location. It can also be analysed to find whether the tumor is primary or secondary and its severity can be analysed.

ACKNOWLEDGMENTS

The authors would like to thank the doctors and management of Government Mohan Kumaramangalam Medical College Hospital, Salem in Tamil Nadu, India for the collection of data and their timely help for validating the reults of our work.

REFERENCES

1. S. Bauer, R. Wiest, L.-P. Nolte, and M. Reyes, 2013. “A survey of MRI-based medical image analysis for brain tumor studies,” Physics in Medicine and Biology, vol. 58, no. 13, pp. 97 – 129.
2. S. Aswathy, G. Glan Deva Dhas, and S. Kumar, 2014. “A survey on detection of brain tumor from MRI brain images,” in Control, Instrumentation, Communication and Computational Technologies (ICCICCT), In-ternational Conference on, July 2014, pp. 871–877.
3. S. Ghanavati, J. Li, T. Liu, P. Babyn, W. Doda, and G. Lampropoulos, 2012. “Automatic brain tumor detection in magnetic resonance images,” in Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on, May 2012, pp. 574–577.
4. S. Charutha and M. Jayashree, 2014. “An efficient brain tumor detection by integrating modified texture based region growing and cellular automata edge detection,” in Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 2014 International Conference on, July 2014, pp. 1193–1199.
5. Islam, S. Reza, and K. Iftekharuddin, 2013. “Multifractal texture estimation for detection and segmentation of brain tumors,” Biomedical Engineer-ing, IEEE Transactions on, vol. 60, no. 11, pp. 3204–3215.
6. R. Preetha and G. Suresh, 2014. “Performance analysis of fuzzy c means algorithm in automated detection of brain tumor,” in Computing and Communication Technologies (WCCCT), 2014 World Congress on, Feb 2014, pp. 30–33.
7. Farmaki, K. Mavrigiannakis, K. Marias, M. Zervakis, and V. Sakkalis, 2010. “Assessment of automated brain structures segmentation based on the mean-shift algorithm: Application in brain tumor,” in Information Technology and Applications in Biomedicine (ITAB), 2010 10th IEEE International Conference on, Nov 2010, pp. 1–5.
8. T. E. M. A. Bianchi A, Miller JV, 2013. “Brain tumor segmentation with symmetric texture and symmetric intensity-based decision forests,” in Proceedings / IEEE International Symposium on Biomedical Imaging: from Nano to macro IEEE International Symposium on Biomedical Imaging., 2013, pp. 748 – 751.
9. K. W. . B. K. Dvok, P., 2013 “Automatic brain tumor detection in t2-weighted magnetic resonance images.” Measurement Science Review, vol. 13, no. 5, pp. 223 – 230.
10. Y. Zhang, Z. Dong, L. Wu, and S. Wang, 2011. “A hybrid method for MRI brain image classification,” Expert Systems with Applications, vol. 38, no. 8, pp. 10 049 – 10 053.
11. E. Ulku and A. Camurcu,2013. “Computer aided brain tumor detection with histogram equalization and morphological image processing techniques,” in Electronics, Computer and Computation (ICECCO), 2013 International Conference on, Nov 2013, pp. 48–51.

12. P. Dhage, M. Phegade, and S. Shah, “Watershed segmentation brain tumor detection,” in Pervasive Computing (ICPC), 2015 International Conference on, Jan 2015, pp. 1–5.
13. M. Ozic, Y. Ozbay, and O. Baykan, 2014.“Detection of tumor with otsu-pso method on brain mr image,” in Signal Processing and Communications Applications Conference (SIU), 2014 22nd, April 2014, pp. 1999–2002.
14. S. Satheesh, R. Kumar, K. Prasad, and K. Reddy, 2011. “Skull removal of noisy magnetic resonance brain images using contourlet transform and morphological operations,” in Computer Science and Network Technology (ICCSNT), 2011 International Conference on, vol. 4, Dec 2011, pp. 2627–2631.
15. J. Chiverton, K. Wells, E. Lewis, C. Chen, B. Podda, and D. Johnson, “Statistical morphological skull stripping of adult and infant MRI data,” Computers in Biology and Medicine, vol. 37, no. 3, pp. 342 – 357, 2007.
16. Dogdas, D. W. Shattuck, and R. M. Leahy, “Segmentation of skull and scalp in 3-d human MRI using mathematical morphology data,” Human Brain Mapping, vol. 26, no. 4, p. 273285, 2007.
17. N. Otsu, “A Threshold Selection Method from Gray-level Histograms,” IEEE Transactions on Systems, Man and Cybernetics, vol. 9, no. 1, pp. 62–66, 1979.
18. Elisee´ Ilunga-Mbuyamba, Juan Gabriel Avina-Cervantes, Dirk Lindner,Jesus Guerrero-Turrubiates, Claire Chalopin, 2016. “Automatic Brain Tumor Tissue Detection based on Hierarchical Centroid Shape Descriptor in T1-weighted MR images”, 2016 International Conference On Electronics, Communications And Computers (Conielecomp).
19. Shailendra Singh Raghuwanshi , PremNarayanArya ,“Comparison of K-means and Modified K-mean algorithms for Large Data-set”,International Journal of Computing, Communications and Networking,Volume 1, No.3, November – December 2012.
20. Malay K. Pakhira,”A Modified k-means Algorithm to Avoid Empty Clusters”, International Journal of Recent Trends in Engineering, Vol 1, No. 1, May 2009.
21. Alan Sexton, Alison Todman, and Kevin Woodward , “Font Recognition Using Shape-Based Quad-tree and Kd-tree Decomposition”, in Proceedings of The Joint Conference On information sciences , vol.5,no. 2.2000, pp. 212-215.

Author’ biography with Photo

Dr.A.Murthi is presently working as Associate professor in Electrical and Electronics Engineering department of Government college of Engineering, Salem, Tamil Nadu, India. Previously, he has served in various Government Engineering colleges in Tamil Nadu at various levels. He has more than 30-years of rich experience in teaching field. A graduate of Madras university, Chennai, he received his M.E. from Bharathiyar university, Coimbatore and Ph.D from Anna university, Chennai. He is a life member of ISTE. He has published several papers in renowned International and National journals. He has also presented many papers in International and National conferences. His areas of current interest include Electromagnetic theory, Utilization of Electrical energy and Image processing.

S.Shameer received diploma in Electrical and Electronics Engineering from P.S.G Polytechnic college, Tamil Nadu, India, in 2011, and B.E in Electrical and Electronics engineering from Bannari Amman Institute of Technology, Tamil Nadu, India, in 2014. Currently purusing M.E in power electronics and drivesGovernment College of Engineering, Salem, Taminadu, India. His areas of interest include Image processing, AC and DC drives, LED drivers design.