Identfication of Brain Tumor using Enhanced Adaptive Fuzzy K Means and SVM Classifier

Jayamani.K
Assistant professor, Department of ECE
Rajalakshmi Institute of Technology, Chennai

Monisha.P
Student, Department of ECE
Rajalakshmi Institute of Technology, Chennai

Manopriya.M
Student, Department of ECE
Rajalakshmi Institute of Technology, Chennai

Haritha.D
Student, Department of ECE
Rajalakshmi Institute of Technology, Chennai

Abstract -*In the modern world, detection of tumor is the challenging one. Tumors are abnormal cells which leads to cancer . Brain tumor is the major challenging one in tumor detection. Segmentation is one of the technique which is used for radiographer to detect any abnormality present in the MR brain Images. Segmenting an image is the process of dividing a digital image into N number of regions. Accurate detection of size and location of the brain tumor plays a vital role in diagnosis. There are various methods for image segmentation in which Threshold based and Clustering techniques - KM, FCM, AFCM are efficient, faster and accurate. This paper deals about the efficient image segmentation algorithms i.e., Threshold and Fuzzy c means techniques for brain MRI images. And also, to improve the accuracy and quality rate of the segmented tissue by SVM based classifier, relevant features are extracted.*

Keywords – *Image Segmentation, MR Image, FCM, Thresholding, SVM.*

I.INTRODUCTION

The parts of human body images are created by using the technique of Medical Imaging for clinical purposes, which are followed by medical procedures seeking to diagnose, or examine disease or medical science including the study of normal anatomy and physiology. There are many techniques developed for medical imaging have scientific and industrial applications[1]. In the medical imaging , to classify the set of techniques that are non-invasively produce images of the internal aspect of the body.The term non invasive is a term based on the fact that following medical imaging process do not penetrate the skin physically. But on the electromagnetic and radiation level, they are quite invasive. In other ways, medical imaging can be seen as the solution of mathematical inverse problems that cause (the properties of living tissue) is conclude from effect. Segmentation describes process of splitting observe image data to a serial of non-overlapping important homogeneous regions [2].

Segmentation is done by several methods: Thresholding, Clustering, Compressed-based, Region-based and so on. Clustering algorithm is one of the process in segmentation. In the analysis of medical images for computer-aided diagnosis and therapy, a preliminary processing task often required is segmentation [3].

A brain tumor is an abnormal growth of cells within the brain. It is also termed as an intracranial solid neoplasm. Brain tumors include all abnormal tissues inside the human skull or in the central spinal canal [4]. They are produced by an erratic and excessive cell division, typically in the brain itself, but also in lymphatic tissue, in blood vessels, in the cranial nerves, in the meninges, skull, pituitary gland, or pineal gland. Within the brain , there are number of cells which include neurons or glial cells (which include astrocytes, oligodendrocytes, and ependymal cells).Though, brain tumors are not consistently incurable, particularly lipomas which are intrinsically kind. Brain tumors or intracranial neoplasms may be cancerous or non-cancerous[5]. Brain tumors are classified based on the type of tissue involved, the location of the tumor, whether it is benign or malignant. Benign brain tumor does not consists of cancer cells and can be removed and have an obvious border or edge. They don't spread to other parts of the body. However, this tumors of low-grade I and II glioma are considered to be corrective under complete surgical excursion. Malignant brain tumor consists of cancer cells and hence also called as brain cancer. They are likely to grow rapidly and can affect nearby healthy brain tissues. This type of tumor can be a threat for life. This tumor is of both grade III and IV gliomas, which is also referred to as anaplastic astrocytomas and glioblastoma . An anaplastic astrocytoma is a mid-grade tumor that demonstrates abnormal or irregular growth and an increased growth index compared to other low-grade tumors. Grade IV tumor class is glioblastoma which grows rapidly .It is highly malignant when compared to other grades of tumors. The abnormal growth of blood vessels and the presence of the necrosis around the tumor are distinguished glioblastoma from all the other grades of the tumor class. [6,7].

Nowadays, information technology and e-health care system introduces the medical field helps clinical experts to provide better health care to the patient. This study deals with the problems of segmentation of abnormal brain tissues and normal tissues such as gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) from magnetic resonance (MR) images using feature extraction technique and support vector machine (SVM) classifier [8]

Many Researchers carried out various clustering algorithms. One of the most important and widely used algorithm is the k-means clustering. The k-means method is numerical, unconfirmed, non-deterministic and repetitive. It is commonly used in practice as a form of image segmentation. However, the k-mean clustering algorithm has some of the shortcoming as follows:

1. The number of clusters, K , must be determined before the algorithm is implemented. This process is more time consuming and subjective for different users.
2. The algorithm is sensitive to initial conditions (i.e. different initial conditions may produce different results of cluster). Furthermore, the algorithm may be confined in the best. As a result, the trapped clusters or centers could represent wrong groups of data.
3. Data which are isolated far away from the centers may pull the centers away from their optimum location. This could lead to poor representation of data. The main idea of introducing the fuzzy concept in the FCM algorithm is that an object can belong simultaneously to more than one class and does so by varying degrees called memberships. The FCM algorithm is an iterative method, which tries to separate the set of data into a number of compact clusters. Every cluster is represented by its center. The FCM is an unsupervised clustering algorithm. Thus, no prior knowledge about the tested data is needed. It can be used with any number of features and any number of classes. It distributes the membership values in a normalized fashion[9].

In this study, different magnetic resonance imaging (MRI) sequence images are employed for diagnosis, including T1-weighted MRI, T2]weighted MRI, fluid-attenuated inversion recovery- (FLAIR) weighted MRI, and proton density weighted MRI. The detection of a brain tumor at an early stage is a key issue for providing improved treatment. Once a brain tumor is clinically expected, radiological evaluation is required to determine its location, its size, and impact on the surrounding areas. On the basis of this information the best therapy, surgery, radiation or chemotherapy, is decided. It is fact that the chances of survival of a tumor-infected patient can be increased significantly if the tumor is detected accurately in its early stage [10].

The rest of the paper is organized as follows: Section 2 presents the related works, Section 3 presents the proposed methodology with the steps used in the proposed technique, Section 4 presents the results and discussion, and finally Section 5 contains the conclusions.

II.RELATED WORKS

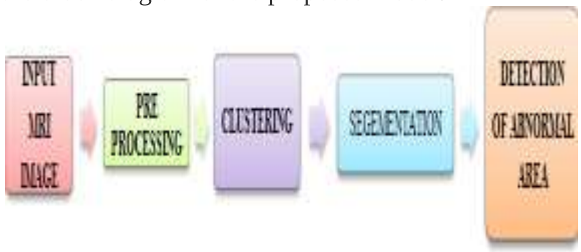
Many techniques have been proposed for detection of brain tumors in MR images. In digital image Processing, several segmentation methods have been developed based on ad hoc. The most common methods are: Amplitude thresholding, texture segmentation, Template matching, and region-growing segmentation. It is very important for detecting tumors, edema and necrotic tissues. Some of the algorithms are used for dividing the brain images into three categories: Pixel based ,Region or Texture Based, Structural based [11].

The extraction of the brain tumor requires MR images in two regions [12]. One region contains the tumor cells of the brain and the second contains the normal brain cells [13]. There have been many researches and study done in the area of image segmentation by using different methods. Most of them are done based on different application of image segmentation. K-means algorithm is one of the simplest and fast clustering algorithm[4]. Zanaty proposed a methodology for brain tumor segmentation based on a hybrid approach, combining FCM, seed region growing, and Jaccard similarity coefficient algorithm to measure segmented gray matter and white matter tissues from MR images[14]. And there are many methods implemented so far to initialize the centre. Researchers are trying to propose new methods which are more efficient than the existing methods, The performance of the technique is systematically evaluated using the MRI brain images received from the public sources.

III.PROPOSED METHODOLOGY

There is no ending in the research on medical image segmentation. Many researches come up with many ideas for efficient segmentation. This paper deals about the efficient algorithms such as

Thresholding with AFCM technique Fig.1 shows the block diagram of the proposed model.



.Fig1. Block diagram of the proposed model

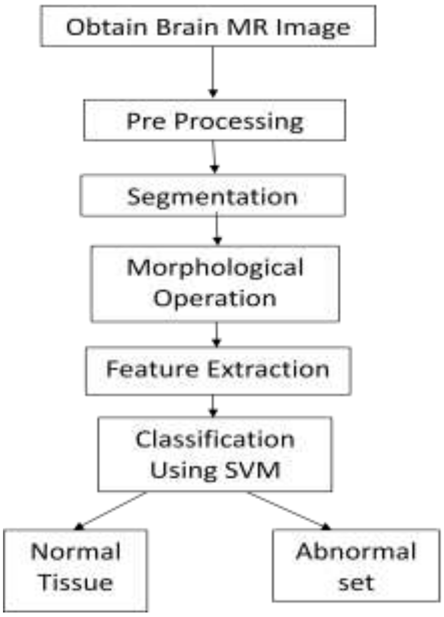


Fig2. Flow diagram of the proposed idea

The following sections discuss the implementation of the algorithm:

1.Preprocessing: The pre-processing step converts the image which is required for further processing . It performs filtering of noise ,sharpening of edges and other artifacts in the image. It is the primary task to improve the quality of the MR brain images and make it ready for the further implementation. In addition to that, it also help in improving the signal to noise ratio, visual appearance of the image is enhanced, smoothening the inner part of the image, removal of irrelevant noise and undesirable part of the image

We have used wiener filter in this paper. Fig.3, shows the block diagram of wiener filter. A wiener filter is an LTI filter which minimizes the mean square error between the estimated random process and the desired process.

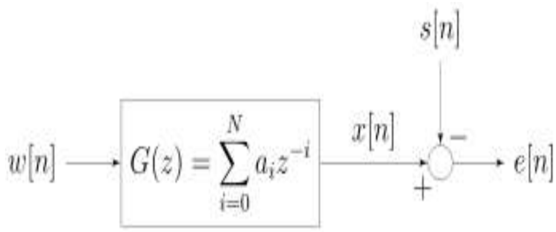


Fig3. Block diagram of Wiener Filter

2.a.Segmentation:

Thresholding*:*

Thresholding is one of the technique used for image segmentation.It is simple and efficient way of partitioning an image into a forefront and backdrop. It isolates objects by converting grayscale images into binary images. This is most effective for images which has high levels of contrast. The common image Thresholding algorithms are Histogram and Multi- band Thresholding. In Histogram Thresholding, an image is subdivided using a threshold which is nothing but a gray level. In multiband Thresholding, the color image is easily thresholded and the approach is to assign the separate threshold for each RBG image.

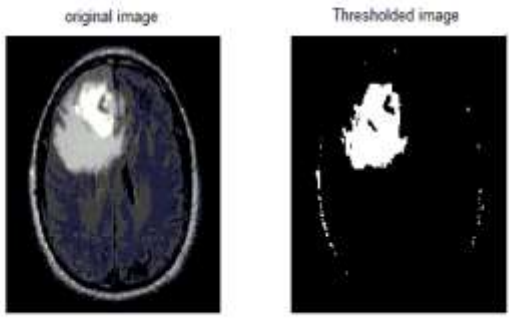


Fig.1 Result of Thresholding for MR brain Image

Fuzzy C Means (FCM) Clustering Algorithm:

Bezdek Fuzzy C Means (FCM) algorithm is commonly used clustering algorithm. The fuzzy logic is used to transform the data by giving the partial membership value to each pixel in the image. The membership value of the fuzzy set is ranges from 0 to 1.It is basically a multi valued logic that allows intermediate values i.e., member of one fuzzy set can also be member of other fuzzy sets in the same image. The membership function defines the fuzziness of an image and the information contained in the image. There are three main basic features involved in characterization of membership function. They are core, support and Boundary. The core is a fully member of the fuzzy set. The support is non membership value of the set and boundary is the intermediate or partial membership with value between 0 and 1.

FCM is based on the minimization of the following objective function

$$J_m = \sum_{i=1}^{D} \sum_{j=1}^{N} \mu_{ij}^{m} \left\| x_i - c_j \right\|^2,$$

Where D is the number of data point, N is the number of clusters, m is fuzzy partition matrix exponent for controlling the degree of fuzzy overlap, with $m > 1$. Fuzzy overlap refers to how fuzzy the boundaries between clusters are, that is the number of data points that have significant membership in more than one cluster, x_i is the ith data point, c_j is the center of the jth cluster, μ_{ij} is the degree of membership of x_i in the jth cluster. For a given data point, x_i, the sum of the membership values for all clusters is 1.

FCM performs the following steps during clustering:

1. Randomly initialize the cluster membership values, μ_{ij}.
2. Calculate the cluster centers:

$$c_j = \frac{\sum_{i=1}^{D} \mu_{ij}^{m} x_i}{\sum_{i=1}^{D} \mu_{ij}^{m}}.$$

3. Update μ_{ij} according to the following:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{N} \left(\frac{\left\| x_i - c_j \right\|}{\left\| x_i - c_k \right\|} \right)^{\frac{2}{m-1}}}$$

4. Calculate the objective function, J_m.
5. Repeat steps 2–4 until J_m improves by less than a specified minimum threshold or until after a specified maximum number of iterations.

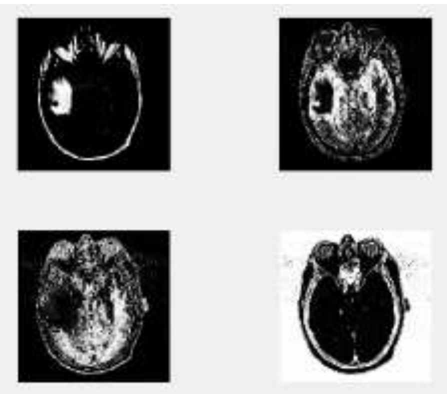


Fig. 2 Output of Fuzzy C Means for 4 clusters

2.b.Morphological Operations:

Morphological operations is assigned to an input image as a structuring element, creating an output image of the same size. In this value of each pixel in the output image is based on a consideration of the corresponding pixel in the input image with its neighbours. The operations are dilation and erosion. Dilation is the process of adding a pixels to the boundaries of objects in an image. Erosion is the process of removal of pixels on object boundaries. The number of pixels added or removed from the objects in an image depends on the size and shape of the structuring element used to process the image. In this paper, in order to eliminate white pixel, an erosion operation of morphology is employed. The erode operation is encountered in the brain MR image mask.

3.Feature Extraction:

Feature extraction is the process of collecting high level information of an image such as shape, texture, color ,and contrast. In fact, texture analysis is an important parameter of human visual perception and machine learning system. It is used to improve the accuracy of diagnosis system by selecting important features. Haralick et al. introduced one of the most widely used image analysis applications of Gray Level Co-occurrence Matrix (GLCM) and texture feature. This technique has two steps. In the first step, the GLCM is computed, and in the following steps texture features are calculated based on the GLCM are computed. Due to the complex structure of diversified tissues such as WM,GM, and CSF in the brain MR images, extraction of relative features is an essential task. The findings and analysis of texture could improve the diagnosis in different stages of the tumor and therapy response assessment. The statistics feature formula for some of the useful features is listed below.

1.Mean(M) : The mean is calculated by adding all the pixels and divided by total number of pixels in theimage.

$$M = \left(\frac{1}{m \times n}\right) \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x, y)$$

2..Standard Deviation(SD) : The standard deviation is describes probability distribution of an observed pixels. A higher value indicates better intensity level and high contrast of edges of an image.

$$SD(\sigma) = \sqrt{\left(\frac{1}{m \times n}\right) \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} \left(f(x, y) - M\right)^2}$$

3Entropy (E). It is calculated to characterize the randomness of the textural image and is defined as

$$E = -\sum_{x=0}^{m-1}\sum_{y=0}^{n-1} f(x,y)\log_2 f(x,y)$$

(4) Skewness (Sk). Skewness is a measure of symmetry or the lack of symmetry. The skewness of a random variable X is denoted as $S_k(X)$ and it is defined as

$$S_k(X) = \left(\frac{1}{m\times n}\right)\frac{\sum (f(x,y)-M)^3}{\mathrm{SD}^3}$$

(5) Kurtosis (Sk).The shape of a random variable's probability distribution is described by the parameter called Kurtosis. For the random variable$X$, the Kurtosis is denoted as $K_{\mathrm{urt}}(X)$ and it is defined as

$$K_{\mathrm{urt}}(X) = \left(\frac{1}{m\times n}\right)\frac{\sum (f(x,y)-M)^4}{\mathrm{SD}^4}$$

(6) Energy (En): The quantifiable amount of the extent of pixel pair repetitions can be termed as Energy. Energy is a parameter to measure the similarity of an image.

$$\mathrm{En} = \sqrt{\sum_{x=0}^{m-1}\sum_{y=0}^{n-1} f^2(x,y)}.$$

(7)Contrast (C_{on}): The intensity of a pixel and its neighbor over the image.

$$C_{\mathrm{on}} = \sum_{x=0}^{m-1}\sum_{y=0}^{n-1} (x-y)^2 f(x,y)$$

(8) Inverse Difference Moment (IDM) or Homogeneity. Inverse Difference Moment is a measure of the local homogeneity of an image. IDM may have a single or a range of values so as to determine whether the image is textured or non-textured.

$$\mathrm{IDM} = \sum_{x=0}^{m-1}\sum_{y=0}^{n-1} \frac{1}{1+(x-y)^2} f(x,y)$$

(9) Directional Moment (DM). Directional moment is a textural property of the image calculated by considering the alignment of the image as a measure in terms of the angle and it is defined as

$$\mathrm{DM} = \sum_{x=0}^{m-1}\sum_{y=0}^{n-1} f(x,y)\,|x-y|$$

(10) Correlation ($Corr$). Correlation feature describes the spatial dependencies between the pixels and it is defined as

$$C_{\mathrm{orr}} = \frac{\sum_{x=0}^{m-1}\sum_{y=0}^{n-1}(x,y)\,f(x,y) - M_x M_y}{\sigma_x\sigma_y}$$

whereMx and σx are the mean and standard deviation in the horizontal spatial domain and My and σy are the mean and standard deviation in the vertical spatial domain

3.SVM Classifier:

The SVM algorithm is based on the study of a superintend learning technique and is applied to one-class classification problem to n-class classification problems. The main objective of the SVM algorithm is to transform a nonlinear dividing objective into a linear transformation. The SVM algorithm defines a hyper plane that is divided into two training classes as defined

$$f(y) = ZT\phi\,(y) + b$$

where Z and T are hyper plane parameters and $\phi(y)$ is a function used to map vector y into a higher-dimensional space. The quality rate parameter accuracy is the proportion of total correctly classified cases that are abnormally classified as abnormal and normally classified as normal from the total number of cases examined

The SVM algorithm's performance can be evaluated in terms of accuracy, sensitivity, and specificity. The confusion matrix defining the terms TP, TN, FP, and FN from the expected outcome and ground truth result for the calculation of accuracy, sensitivity, and specificity are shown in Table 3 and the formulas to calculate accuracy, sensitivity, and specificity are shown in Table 4.

Where TP is the number of true positives,which is used to indicate the total number of abnormal cases correctly classified,TN is the number of true negatives, which is used to indicate normal cases correctly classified; FP is the number of false positive, and it is used to indicate wrongly detected or classified abnormal cases; when they are actually normal cases and FN is the number of false negatives, it is used to indicate wrongly classified or detected normal cases; when they are actually abnormal cases all of these outcome parameters are calculated using the total number of samples examined for the detection of the tumor.

Image	Mean	SD	Skewness	Kurtosis	Entropy
Image1	0.0537	0.0509	3.9578	6.6443	0.3021
Image2	0.6691	0.2214	30.7185	1.5163	0.9159
Image3	0.1357	0.117	2.1270	5.5240	0.5730

		3			
Image4	**0.1373**	**0.118 5**	**2.1074**	**5.4410**	**0.0201**

Table 1. First order Statistical features

Images	Contrast	Homogeneity	Energy	Correlation		
Image1	**0.0106**	**0.9947**	**0.9578**	**0.6443**	**0.8874**	**0.302 1**
Image2	**0.0112**	**0.9944**	**0.5452**	**0.9747**	**0.3021**	
Image3	**0.0208**	**0.9896**	**0.7443**	**0.9118**		
Image4	**0.0475**	**0.9763**	**0.7401**	**0.9788**		

Table 2. Second order Statistical features

Expected outcome	Ground truth		Row total
	positive	negative	
Positive	**TP**	**FP**	**TP+FP**
Negative	**FN**	**TN**	**FN+TN**
Column Total	**TP+FN**	**FP+TN**	**TP+FP+TN+FN**

Table 3. Confusion Matrix

Parameter	Formula
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
Sensitivity	$\frac{TP}{TP+FN}$
Specificity	$\frac{TN}{TN+FP}$
Precision	$\frac{TP}{TP+FP}$
F-score	$\frac{2TP}{2TP+FN+FP}$

Table 4. Formulas for calculation

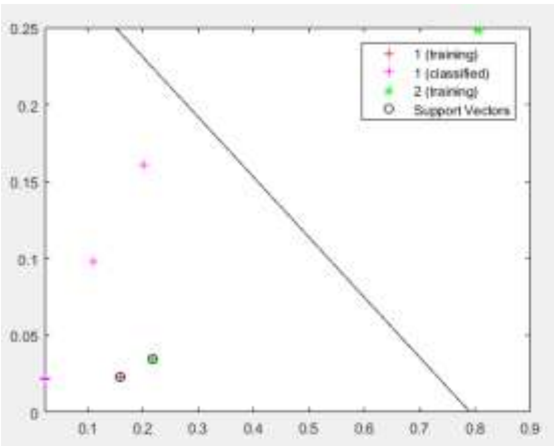


Fig.3 Plot of SVM Classifier

IV.RESULT AND DISCUSSION

To justify the performance of our algorithm datasets are used. The dataset is the Digital Imaging and Communications in Medicine (DICOM) dataset. For the purpose of the images from the DICOM dataset, all of which included are tumor-infected brain tissues. The second dataset is the Brain Web dataset which consists of full three-dimensional simulated brain MR data obtained.

This section presents the results of our proposed image segmentation technique, which are obtained by using real brain MR images. The proposed algorithm was carried out using Matlab 9.2 (R2017a), which runs on the Windows 7 operating system and has an Intel core i3 processor and a 4GB RAM. The proposed algorithm performs segmentation, feature extraction, and classification as is done in human vision perception, which recognizes different objects, different textures, contrast, brightness, and depth of the image. The application of the proposed technique can be applicable to a varied range of tumors and MR modalities.

V. CONCLUSION

In this paper, MR brain images are segmented into normal tissues such as white matter, gray matter, cerebrospinal fluid and tumor-infected tissues. From the infected patients with a glicoplasma tumor, in benign and malignant stages, assisted in this study. We used preprocessing to improve the signal-to-noise ratio and to eliminate the effect of unwanted noise. Thresholding is done to improve performance. Furthermore, SVM is used to classify the tumor stage by analyzing feature vectors and area of the tumor. The texture based features with a commonly recognized classifier for the classification of brain tumor from MR brain images. From the experimental results performed on the different images, it is clear that the analysis for the brain tumor detection is fast and accurate. The various performance factors also indicate that the proposed algorithm provides better result by improving certain parameters such as mean, accuracy, sensitivity, specificity, precision and f-Sscore. Our experimental results show that the proposed algorithm can aid in the accurate and timely detection of brain tumor along with the identification of its exact location. Thus, the proposed algorithm is significant for brain tumor detection from MR images. The experimental results achieved 96.51% accuracy demonstrating the effectiveness of the proposed technique for identifying normal and abnormal tissues from MR images. Our results lead to the conclusion that the proposed method is suitable for integrating clinical decision support systems for primary screening and diagnosis by the radiologists or clinical experts.

VI. REFERENCE

[1] M.Ganesh, M.Naresh & C.Arvind,S "*MRI Brain Image Segmentation using Enhanced Adaptive Fuzzy K-Means Algorithm*", Intelligent Automation & Soft Computing,vol.23,Issue 2,pp-325-330,2017.

[2] Xiao-li Jin, T.W. "*Multi-Spectral MRI brain image segmentationbased on kernel clustering analysis*". International conference on system
Engineering and Modelling, vol. 34, pp. 141–146.,2012.
[3] Saha, S., & Bandyopadhyay, S."*MRI brain image segmentation by fuzzy symmetry based genetic clustering technique*". Evolutionary
Computation, CEC 2007. IEEE Congress on, 2007, pp. 4417–4424.
[4] Siddhi N.Nerurkar, "*Brain Tumor Detection using Image Segmentation*", International Journal of Engineering Research in Computer Science and Engineering , Vol 4, Issue4,April 2017.
[5] Alan Jose1, S.Ravi2, M.Sambath3, *"Brain Tumor Segmentation Using K-Means Clustering And Fuzzy C-Means Algorithms And Its Area Calculation*", International Journal of Innovative Research in Computer Science and Engineering , Vol 2, Issue3,March 2014.
[6] Gauri P.Ananadgaokar, Ganesh.S.Sable, "*Detection and Identification of Brain Tumor in Brain MR Images Using Fuzzy C-Menas Segmentation*", International Journal of Advanced Research in Computer Science and Communication Engineering , Vol 2, Issue10,October 2013.
[7] Nilesh Bhaskarrao Bahadure, Arun Kumar Ray, and Har Pal Thethi, "Image Analysis for MRI Based Brain Tumor Detection andFeature Extraction Using Biologically Inspired BWT and SVM", International Journal of Biomedical Imaging, vol. 47, no. 10, pp3849–3852,2017.
[8] Isa, N.A.M., & Salamah, S.A., "*Adaptive fuzzy moving K-means clustering algorithm for image segmentation".* IEEE Transactions on Consumer Electronics. vol 55, Issue 4, pp. 15–25,2015
[9] G. Coatrieux, H. Huang, H. Shu, L. Luo, andC.Roux, "*A watermarking-based medical image integrity control system and animage moment signature for tampering characterization*," IEEE
Journal of Biomedical and Health Informatics, vol. 17, no. 6, pp.1057–1067, 2013.
[10]T. Logeswari ,M. Karnan, "*An improved implementation of brain tumor detection using segmentation based on soft computing",* Journal of Cancer Research and Experimental Oncology Vol. 2(1) pp. 006-014, March, 2010.
[11] M. Alfonse and A.-B. M. Salem, "*An automatic classification of brain tumors through MRI using support vector machine*,"Egyptian Computer Science Journal, vol. 40, pp. 11–21, 2016.
[12] Q. Ain, M. A. Jaffar, and T.-S. Choi, "*Fuzzy anisotropic diffusionbased segmentation and texture based ensemble classification of brain tumor*," Applied Soft Computing Journal, vol. 21, pp. 330–340, 2014.
[13] E. Abdel-Maksoud, M. Elmogy, and R. Al-Awadi, "*Brain tumor segmentation based on a hybrid clustering technique*," Egyptian Informatics Journal, vol. 16, no. 1, pp. 71–81, 2014.
[14] E. A. Zanaty, "*Determination of gray matter (GM) and white matter (WM) volume in brain magnetic resonance images (MRI),*" International Journal of Computer Applications, vol. 45, pp. 16–22, 2012.
[15] Parveen and A. Singh, "*Detection of brain tumor in MRI images, using combination of fuzzy c-means and SVM*," in Proceedings of the 2nd International Conference on Signal Processing
and Integrated Networks (SPIN '15), pp. 98–102, February 2015.