Brain Tumor Detection Using Machine Learning

Asha K S[1], Medha Lakshmi S[2],Goutam R[3]
[1, 2, 3] Dept of Computer Science
[1, 2, 3] Atria Institute of Technology

Abstract- *Brain cancer is caused by the population of abnormal cells called glial cells that takes place in the brain. Over the years, the number of patients who have brain cancer is increasing with respect to the aging population, is a worldwide health problem. The objective of this paper is to develop a method to detect the brain tissues which are affected by cancer especially for grade-4 tumor, Glioblastoma multiforme (GBM). GBM is one of the most malignant cancerous brain tumors as they are fast growing and more likely to spread to other parts of the brain. In this paper, Naïve Bayes classification is utilized for recognition of a tumor region accurately that contains all spreading cancerous tissues. Brain MRI database, preprocessing, morphological operations, pixel subtraction, maximum entropy threshold, statistical features extraction, and Naïve Bayes classifier based prediction algorithm are used in this research. The goal of this method is to detect the tumor area from different brain MRI images and to predict that detected area whether it is a tumor or not. When compared to other methods, this method can properly detect the tumor located in different regions of the brain including the middle region (aligned with eye level) which is the significant advantage of this method. When tested on 50 MRI images, this method develops 81.25% detection rate on tumor images and 100% detection rate on non-tumor images with the overall accuracy 94%.*

Keywords- Glioblastoma Multiforme (GBM), Magnetic resonance imaging (MRI), Naïve Bayes classification, maximum entropy threshold, statistical features extraction.

I. INTRODUCTION

Tumor is an uncontrolled growth of cancer cells in any part of the body. Tumors are of different types and have different characteristics and different treatments. At present, brain tumors are classified as primary brain tumors and metastatic brain tumors. The former begin in the brain and tend to stay in the brain, the latter begin as a cancer elsewhere in the body and spreading to the brain. Brain tumor segmentation is one of the crucial procedures in surgical and treatment planning. Brain tumor segmentation using MRI has been an intense research area. Brain tumors can have various sizes and shapes and may appear at different locations. Varying intensity of tumors in brain magnetic resonance images (MRI) makes the automatic segmentation of tumors extremely challenging. There are various intensity based techniques which have been proposed to segment tumors on magnetic resonance images. Texture is one of most popular feature for image classification and retrieval. The multifractal texture estimation methods are more time consuming. A texture based image segmentation using GLCM (Gray-Level Co-occurrence Matrix) combined with AdaBoost classifier is proposed here. From the MRI images of brain, the optimal texture features of brain tumor are extracted by utilizing GLCM. Then using these features AdaBoost classifier algorithm classifies the tumor and non-tumor tissues and tumor is segmented. This method provides more efficient brain tumor segmentation compared to the segmentation technique based on mBm and will provide more accurate result. Tumor is the abnormal growth of the tissues. A brain tumor is a mass of unnecessary cells growing in the brain or central spine canal. Brain cancer can be counted among the most deadly and intractable diseases. Today, tools and methods to analyse tumors and their behaviour are becoming more prevalent. Clearly, efforts over the past century have yielded real advances. However, we have also come to realize that gains in survival must be enhanced by better diagnosis tools. Although we have yet to cure brain tumours, clear steps forward have been taken toward reaching this ultimate goal, more and more researchers have incorporated measures into clinical trials each advance injects hope to the team of caregivers and more importantly, to those who live with this diagnosis. Magnetic Resonance Imaging (MRI) has become a widely-used method of high-quality medical imaging, especially in brain imaging where MRI's soft tissue contrast and non-invasiveness are clear advantages. An important use of MRI data is tracking the size of brain tumor as it responds treatment. Therefore, an automatic and reliable method for segmenting tumor would be a useful tool. MRI provides a digital representation of tissue characteristics that can be obtained in any tissue plane. The images produced by an MRI scanner are best described as slices through the brain. MRI has the added advantage of being able to produce images which slice through the brain in both horizontal and vertical planes. This makes the MRI-scan images an ideal source for detecting, identifying and classifying the right infected regions of the brain. Most of the current conventional diagnosis techniques are based on human experience in interpreting the MRI-scan for judgment; certainly this increases the possibility to false detection and identification of the brain tumor. On the other hand, applying digital image processing ensures the quick and precise detection of the tumor. One of the most effective techniques to extract information from complex medical images that has wide application in medical field is the segmentation process. The main objective of the image segmentation is to partition an image into mutually exclusive and exhausted regions such that each region of interest is spatially contiguous and the pixels within the region are homogenous with respect to a predefined criterion. The cause of most cases is unknown. Risk factors that may occasionally be involved include: a number of genetic syndrome such as neurofibromatosis as well as exposure to the chemical vinyl chloride, Epstein-Barr virus, and ionizing radiation. Magnetic resonance imaging (MRI) is the prime technique to diagnose brain tumors and monitor their treatment. Different MRI modalities of each patient are acquired and these images are interpreted by computer-based image analysis methods in order to handle the complexity as well as constraints on time and objectiveness. In this thesis, two major novel approaches for analyzing tumor-bearing brain images in an automatic way are presented: Multi-modal tissue classification with integrated regularization can segment healthy and pathologic brain tissues including their sub-compartments to provide quantitative volumetric information. The method has been evaluated with good results on a large number of clinical and synthetic images. The fast run-time of the algorithm allows for an easy integration into the clinical work flow. An extension has been proposed for integrated segmentation of longitudinal patient studies, which has been assessed on a small dataset from a multi-center clinical trial with promising results. Atlas-based segmentation with integrated tumor-growth modeling has been shown to be a suitable means for segmenting the healthy brain structures surrounding the tumor. While a simplistic tumor growth model offered advantages in computation speed, a more sophisticated multi-scale tumor growth model showed better potential to provide a more realistic and meaningful prior for atlas-based segmentation. Both approaches have been combined into a generic framework for analyzing tumor-bearing brain images, which makes use of all the image information generally available in clinics. This segmentation framework paves the way for better diagnosis, treatment planning and monitoring in radiotherapy and neurosurgery of brain tumors.

II. LITERATURE SURVEY

A.Hamamci, N. Kucuk, K. Karaman, K. Engin, G. Unal, "Tumor-cut: segmentation of brain tumors on contrast enhanced MR images for radiosurgery applications", Medical Imaging IEEE Transactions, vol. 31, no. 3, pp. 790, March 2017 : In this paper, we present a fast and robust practical tool for segmentation of solid tumors with minimal user interaction to assist clinicians and researchers in radiosurgery planning and assessment of the response to the therapy. Particularly, a cellular automata (CA) based seeded tumor segmentation method on contrast enhanced T1 weighted magnetic resonance (MR) images, which standardizes the volume of interest (VOI) and seed selection, is proposed. B.A. Makropoulos, I. S. Gousias, C. Ledig, P. Aljabar, A. Serag, J. V. Hajnal, A. D. Edwards, S. J. Counsell, D. Rueckert, "Automatic whole brain MRI segmentation of the developing neonatal brain", Medical Imaging IEEE Transactions, vol. 33, no. 9, pp. 1818-1831, September 2018: Magnetic resonance (MR) imaging is increasingly being used to assess brain growth and development in infants. Such studies are often based on quantitative analysis of anatomical segmentations of brain MR images. However, the large changes in brain shape and appearance associated with development, the lower signal to noise ratio and partial volume effects in the neonatal brain present challenges for automatic segmentation of neonatal MR imaging data.C.M. T. El-Melegy, H. M. Mokhtar, "Tumor segmentation in brain MRI using a fuzzy approach with class center priors", EURASIP Journal on Image and Video Processing, January 2017: This paper proposes a new fuzzy approach for automatic segmentation of normal and pathological brain MRI volumetric data sets. MRI is generally useful for brain tumor deduction because it provide more detailed information about its type, position, size. Brain tumor segmentation is the separation of different tumor tissues from normal brain tissue. In automatic brain segmentation MRI is a sophisticated tool for medical imaging. Fuzzy is used to segment all the tissues at the same time.D.J. Selvakumar, A. Lakshmi, T. Arivoli, "Brain tumor segmentation and its area calculation in brain MR images using K-mean clustering and Fuzzy C-mean algorithm", Advances in Engineering Science and Management (ICAESM) 2012 International Conference, vol. 7, no. 7, pp. 186-190, October 2018. : This paper deals with the implementation of Simple Algorithm for detection of range and shape of tumor in brain MR images. Tumor is an uncontrolled growth of tissues in any part of the body. Tumors are of different types and they have different Characteristics and different treatment. As it is known, brain tumor is inherently serious and life threatening because of its character in the limited space of the intracranial cavity (space formed inside the skull). E.Rajesh C. Patil, A. S. Bhalchandra, "Brain Tumour Extraction from MRI Images Using MATLAB", International Journal of Electronics Communication & Soft Computing Science and Engineering, vol. 2, no. 1 Medical image processing is the most challenging and emerging field now a days. Processing of MRI images is one of the part of this field:

This paper describes the proposed strategy to detect & extraction of brain tumour from patient's MRI scan images of the brain. This method incorporates with some noise removal functions, segmentation and morphological operations which are the basic concepts of image processing. Detection and bextraction of tumour from MRI scan images of the brain is done by using MATLAB software.

III. PROPOSED SYSTEM

The system proposed a novel semi-automatic segmentation method based on population and individual statistical information to segment brain tumors in magnetic resonance (MR) images. A new cost function for segmentation is constructed through these probabilities and is optimized using graph cuts. It can easily be realized that the full or semi-automatic watersheds-based methods are in fact region growing methods constrained, at the beginning, by the competition between the different seeds and then by the differential ready grown regions. In its present form, the fuzzy-connectedness-based method does not incorporate such a competition paradigm: only one object can be detected at a time. A direct consequence of this non-competitive learning is the need to threshold the connectedness map. A way to improve the fuzzy connectedness-based method is thus to introduce the concept of competition. This is very natural, from a computer vision point of view: an object or a region in an image does not exist intrinsically, but against other (adjacent) objects or against the background. Thus, we modified the fuzzy connectedness method along the following lines: First, not only one object, but the different objects or regions the user wants to differentiate, are designated and labeled. Second, in order to avoid thresholding, the affinity to any seed can be computed and the labeling of the pixel can be done according to the maximum affinity. In practice, instead of computing all the affinities for every pixel, it proves faster to label the pixels in the course of the computation. Before the presentation of the brain tumor segmentation methods, the MRI preprocessing operations are introduced because it is directly related to the qualities of the segmentation results. In general, the raw MRI images need to be preprocessed to realize the segmentation purposes. These pre-processing operations include de-noising, skull-stripping, intensity normalization, etc, and have direct impact on the results of brain tumor segmentation. Image de-noising is a standard preprocessing task for MRI. Noise in MRI image makes it difficult to precisely delineate regions of interest between brain tumor and normal brain tissues. For this reason, it is necessary to preprocess MRI image to reduce noise and to enhance contrast between regions. Many de-noising methods for MRI image have been proposed, such as Anisotropic Diffusion Filtering (ADF),wavelets, Non-Local Means (NLM), and Independent Component Analysis (ICA). ADF is the current most popular method for the de-noising of brain tumor MRI images. A critical review of the effects of de-noising algorithms on MRI brain tumor segmentation was discussed. It concluded that, although the noise of images was reduced, it has always existed and became a negative effect on the brain tumor segmentation. Skull stripping is an important preprocessing step for the analysis of MRI images. For example, the system shows a result of skull stripping. Skull stripping is the process of delineation and removal of non-cerebral tissue region such as skull, scalp, and meninges from the brain soft tissues. The accuracy in skull striping process affects the efficiency in detecting tumor, pre-surgical planning, cortical surface reconstruction, and brain morphometry, and has been considered as an essential step for brain tumor segmentation. Removal of the skull region reduces the chances of misclassifying diseased tissues. The process of skull stripping is faced with many challenges due to the complexity of the human brain, variability in the parameters of MR scanners, and individual characteristics. Poor quality and low contrast images also contribute to difficulties in segmenting the images precisely. Many of robust skull stripping algorithm shave been proposed to reduce these influences. Intensity normalization is a very critical step for the preprocessing of MRI, especially when classification and clustering methods are used for the segmentation. However, due to the confounding effects caused by the differences in brain tumor appearance, the segmentation of tumor-bearing images are more challenging than healthy images. When operating on multi-modal images, pre-processing always includes the registration of all modalities in a common space of reference. In most cases, this is performed using a linear transformation model with the Mutual Information (MI) similarity metric and resampling in order to ensure voxel-to-voxel correspondence across all modalities. Nowadays, brain tumor segmentation methods can be organized into different categories based on different principles. In the clinic, brain tumor segmentation methods are usually classified into three main categories including manual, semi-automatic, and fully automatic segmentations based on the degree of required human interaction. For manual brain tumor segmentation, the experts of brain tumor must master the information presented in the brain tumor images and some additional knowledge such as anatomy because manual brain tumor segmentation aims to manually draw the boundaries of the brain tumor and paint the regions of anatomic structures with different labels. To date, manual segmentation is widely applied to clinical trial. In the clinic, since many of brain tumor images are emerging, the manual segmentation of the different regions of brain tumor will become an error-prone and time-consuming task for the experts and yield poor results in a way. Therefore, more advanced segmentation methods such as semi-automatic and fully automatic segmentation methods are required to address this problem. For semi-automatic brain tumor segmentation, it mainly consists of the user, interaction, and software computing. In the semi-automatic brain tumor methods, the user needs to input some parameters and is responsible for analyzing the visual information and providing feedback response for the software computing. The software computing is targeted at the realization of brain tumor segmentation algorithms. The interaction is in charge of adjusting segmentation information between the user and the software computing. The semi-automatic brain tumor segmentation methods were divided into three main processes: initialization, feedback response, and evaluation. Although brain tumor semi-automatic segmentation methods can obtain better results than manual segmentation, it also comes into being different results from different experts or the same user at different times. Hence, fully automatic brain tumor segmentation methods were proposed. For fully automatic brain tumor segmentation, the computer determines the segmentation of brain tumor without any human interaction. In general, a fully automatic segmentation algorithm combines artificial intelligence and prior knowledge. With the development of machine learning algorithms that can simulate the intelligence of humans to learn effectively, the study of fully automatic brain tumor segmentation has become a popular research issue. The semi-automatic and fully automatic segmentation of tumor brain images are faced with great challenges due to usually exhibiting unclear and irregular boundaries with discontinuities and partial-volume effects for brain tumor images.

IV. IMPLEMENTATION AND RESULTS

System Architecture:

The system architecture has many elements some of them are

- MRI Image
- Pre-processing and skull striping
- Segmentation and morphological operations
- Feature Extraction
- Classification

1. MRI Image

Magnetic resonance imaging (MRI), also known as nuclear magnetic resonance imaging, is a scanning technique for creating detailed images of the human body. 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 suspected, 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 evident that the chances of survival of a tumor-infected patient can be increased significantly if the tumor is detected accurately in its early stage. As a result, the study of brain tumors using imaging modalities has gained importance in the radiology department.

2. Pre-Processing

The primary task of preprocessing is to improve the quality of the MR images and make it in a form suited for further processing by human or machine vision system. In addition, preprocessing helps to improve certain parameters of MR images such as improving the signal-to noise ratio, enhancing the visual appearance of MR image, removing the irrelevant noise and undesired parts in the background, smoothing the inner part of the region, and preserving its edges [5]. To improve the signal-to-noise ratio, and thus the clarity of the raw MRimages ,we applied adaptive contrast enhancement based on modified sigmoid function

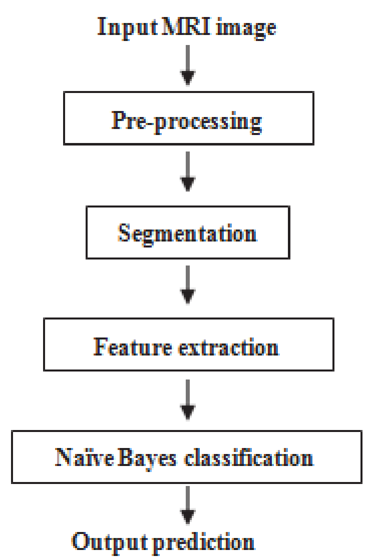


Fig. 1. System Design

3. Skull Striping

Skull stripping is an important process in biomedical image analysis, and it is required for the effective examination of brain tumor from the MR images. Skull stripping is the process of eliminating all nonbrain tissues in the brain images. By skull stripping, it is possible to remove additional cerebral tissues such as fat, skin, and skull in the brain images. There are several techniques available for skull stripping; some of the popular techniques are automatic skull stripping using image contour, skull stripping based on segmentation and morphological operation, and skull stripping based on histogram analysis or a threshold value. Figure 2 provides the stages of the skull stripping algorithm. This study uses the skull stripping technique that is based on a threshold operation to remove skull tissues.

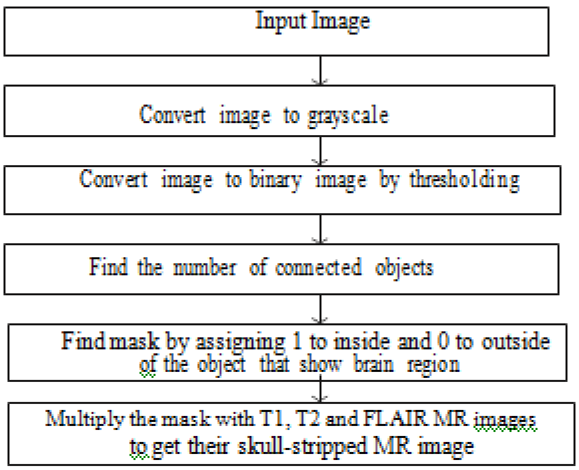


Figure 2: Steps used in the skull stripping algorithm.

4. Segmentation and Morphological operations

The segmentation of the infected brain MR regions is achieved through the following steps: In the first step, the preprocessed brain MR image is converted into a binary image with a threshold for the cut-off of 128 being selected. The pixel values greater than the selected threshold are mapped to white, while others are marked as black; due to this two, different regions are formed around the infected tumor tissues, which is cropped out. In the second step, in order to eliminate white pixel, an erosion operation of morphology is employed. Finally, the eroded region and the original image are both divided into two equal regions and the black pixel region extracted from the erode operation is counted as a brain MR image mask. In this study, Berkeley wavelet transformation is employed for effective segmentation of brain MR image.

5. Feature Extraction

It is the process of collecting higherlevel 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 effectively to improve the accuracy of diagnosis system by selecting prominent features. Haralick et al. [32] introduced one of the most widely used image analysis applications of Gray Level Cooccurrence Matrix (GLCM) and texture feature. This technique follows two steps for feature extraction from the medical images. In the first step, the GLCM is computed, and in the other step, the texture features based on the GLCM are calculated. Due to the intricate structure of diversified tissues such as WM, GM, and CSF in the brain MR images, extraction of relevant features is an essential task. Textural findings and analysis could improve the diagnosis, different stages of the tumor (tumor staging), and therapy response assessment. The statistics feature formula for some of the useful features is listed below.

(1)Mean (M). The mean of an image is calculated by adding all the pixel values of an image divided by the total number of pixels in an image. $M = (f\ (x, y)$.
(2)Standard Deviation (SD). The standard deviation is the second central moment describing probability distribution of an observed population and can serve as a measure of inhomogeneity. A higher value indicates better intensity level and high contrast of edges of an image.
(3)Entropy (E). Entropy is calculated to characterize the randomness of the textural image and is defined as $E = -$.
(4)Contrast (Con): Contrast is a measure of intensity of a pixel and its neighbor over the image, and it is defined as $\mathrm{Con}\quad f\ (x, y)$
(5) Correlation ($Corr$) : Correlation feature describes the spatial dependencies between the pixels.
(6)Skewness: It is the measure of symmetry
(7)Energy (En). Energy can be defined as the quantifiable amount of the extent of pixel pair repetitions. Energy is a parameter to measure the similarity of an image. If energy is defined by Haralicks GLCM feature, then it is also referred to as angular second moment, and it is defined as Energy.

FEATURE EXTRACTION :

In this stage, many morphology features and intensity features are extracted from the grayscale image of the segmented tumor area. Eight region properties (area, perimeter, eccentricity, equivalent diameter, solidity, convex area, major axis length, minor axis length) and three intensity features (maximum, mean, minimum), the total of eleven variables are extracted and saved in a specific matrix to be trained in Naïve Bayes classifier. E.

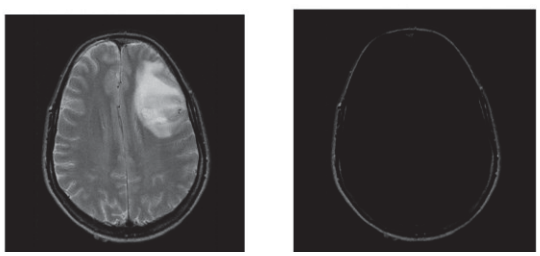
NAÏVE BAYES CLASSIFICATION :

The Naïve Bayes classification is a supervised classification of machine learning, based on a probabilistic approach which uses Bayes' theorem of probability [12]. The Naïve Bayes algorithm is called “naïve” because it makes the assumption that the features occurrences are independent of each other. That is the main reason why we use this algorithm for detecting brain tumor from different locations with

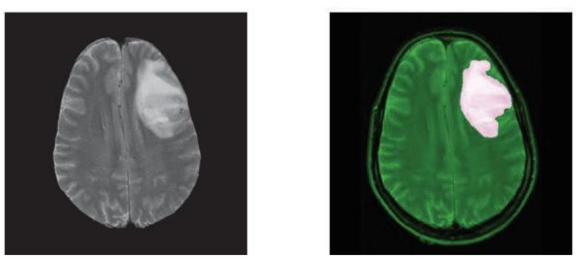
different types of features. As in most fields that deal with events under randomness, probability considerations become significantly effective due to independence on the occurrence of the extracted features. The extracted features matrix is subjected to be trained in the Naïve Bayes classifier so that it could predict the test image whether it is normal or tumor. More false tumor objects are trained than tumor
objects for better performance since the false tumors are detected in different locations.

V. EXPERIMENTAL RESULTS

In this section, the simulation of the overall process is shown. Figure 3 shows the result of detecting brain tumors located in the frontal lobe which is the upper level of the brain.



(a) original MRI image (b) skull image



(c) extracted brain image (d) detected object

Fig. 3. Segmentation results with one detected object

VI. CONCLUSION

There are many techniques to detect brain tumor from MRI images but the challenge here is to find the location and size of the tumor and another important challenge is to predict the best therapy for the brain tumor. In the existing system the accuracy of detection of brain tumor depends only on the doctors experience, each doctor will have different perspectives as there is no one standard system. This project will create a standard benchmark, where a user can compare the results with different doctors opinion and make a correct decision. In this study, using MR images of the brain, we segmented brain tissues into normal tissues such as white matter, gray matter, cerebrospinal fluid (background), and tumor-infected tissues. Pre processing is done to improve the signal-to-noise ratio and to eliminate the effect of unwanted noise.From the experimental results performed on the different images, it is clear that the analysis for the brain tumor detection is fast and accurate when compared with the manual detection performed by radiologists or clinical experts.

REFERENCES

[1] Riddhi.S.Kapse1 , Dr. S.S. Salankar2 , Madhuri.Babar3 “Literature Survey on Detection of Brain Tumor from MRI Images”IOSR Journal of Electronics and Communication Engineering (IOSR-JECE) e-ISSN: 22782834,p- ISSN: 2278-8735.Volume 10, Issue 1, Ver. II (Jan - Feb. 2015), PP 80-86

[2] A.Sindhu1 , S.Meera2 “A Survey on Detecting Brain Tumor in mri Images Using Image Processing Techniques ”International Journal of Innovative Research in Computer and Communication EngineeringVol. 3, Issue 1, January 2015

[3] Zeynel Cebeci1 , Figen Yildiz 2 “Comparison of K-Means and Fuzzy C-Means Algorithms on Different Cluster Structures”Journal of Agricultural Informatics (ISSN 2061-862X) 2015 Vol. 6, No. 3:13-23

[4] Soumi Ghosh , Sanjay Kumar Dubey “Comparative Analysis of K-Means and Fuzzy CMeans Algorithms” ((IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 4, No.4, 2013

[5] Russell C. Eberhart and Yuhui Shi “Comparison between Genetic Algorithms and Particle Swarm Optimization ” Indiana University Purdue University Indianapolis 723 W. Michigan St., SL160 Indianapolis, IN 46202

[6] L. O. Hall, A. M. Bensaid, L. P. Clarkeetc, "A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain", IEEE Trans. Neural Networks, vol. 3, no. 5, pp. 672-682, Sep. 1992.

[7] A. Kharrat, N. Benamrane, M. Ben Messaoud and M. Abid, “Detection of Brain Tumor in Medical Images,” 3rd International Conference on Signals, Circuits and Systems (SCS): 1 – 6, 68 November 2009.

[8] J. Zhou, K. L. Chan, V. F. H. Chongand, and S. M. Krishnan, “Extraction of Brain Tumor from MR Images Using One-class Support Vector Machine,” Proc. IEEE Engineering in Medicine and Biology Society (EMBS 05), 2005, pp. 6411-6414.

[9] Gopal,N.N. Karnan, M. , ―Diagnose brain tumor through MRI using image processing clustering algorithms such as Fuzzy C Means along with intelligent optimization

techniques Page(s): 1 – 4, Computational Intelligence and Computing Research (ICCIC), 2010 IEEE International Conference, 28-29 Dec. 2010.

[10] M. Schmidt, I. Levner, R. Greiner, et al. Segmenting brain tumors using alignment-based features[C]. Proceedings of the Fourth International Conference on Machine Learning and Applications, 2005, 6 pp.