INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Brain Tumor Extraction Using Matlab

P. Chandra Sandeep[1], K. Sai Manas[2], K. Sujay[3], Dr. Krishna Samalla[4]

[1, 2, 3]B. Tech IV year, Department of ECE, Sreenidhi Institute of Science and Technology, Hyderabad, Telangana, India
[4]Professor, Department of ECE, Sreenidhi Institute of Science and Technology, Hyderabad, Telangana, India

Abstract: The brain is the most crucial part of our human body which acts as central coordinating system for all the controlling and all regular functions of our body. The continuous growth of abnormal cells which creates certain mass of tissue is called as tumor. Tumor in the brain can be either formed inside the brain or gets into brain after formed at other part. But there is no clear information regarding the formation of brain tumor till date. Though the formation tumor in brain is not common or regular but the mortality rate of the infected people is very high because the brain is major part of body. So, it is very important get the treatment at the early stages of brain tumor but there is no direct procedure for detection and classification of tumor in the very first step of diagnosis. In actual medical diagnosis, mri images alone can't be able to determine the detected tumor as either the cancerous or non-cancerous. But the tumor may be sometimes danger to life or may not be danger to life. Tumor inside the brain can be of either the benign(non- cancerous) or the malignant(cancerous). So, we need to detect the tumor from the MRI images through image processing and then to classify the detected tumor as it belongs to either the benign or malignant tumor. We are going to get the brain mri images as our dataset for our proposed method but the images we got may have the noise. So, we need to preprocess the image using the image preprocessing techniques. We are going to use several algorithms like thresholding, clustering to make the detection of tumor by using the image processing and image segmentation and after the detection of tumor we are going do feature extraction. This step involves the extraction of detected objects features using DWT. This extracted features are given as input to classifier algorithms like SVM's and CNN after reduction of features using the PCA.

I.   INTRODUCTION

We know that brain is the most important part of our body for doing all the instinct and controlling operations. The growth of harmful cells inside side the brain called as the tumor cells. This growth makes brain to do malfunctions. These cells may be harmful or not depends on the type of cells that are being growing inside the brain. This abnormal growth of harmful cells rapidly is called the tumor tissue. This tissue grows rapidly and increases uncontrollably where the normal diagnosis mechanisms are not useful for stopping its growth. This tumor cells can get into brain from other parts of the body it means that actually the tumor gets started at other part and gets into brain. Another type is tumor gets started at brain. Tumor is nothing but a cancer cell, but all tumor tissues in brain may be cancerous(malignant) sometimes they may be non-cancerous(benign). Brain tumors of human body of type primary or metastatic. Metastatic brain tumors are tumors which are originated other part and gets into the brain. Another harmful disorder of brain is the epilepsy where the nerves(neurons) coordinating the body forms into cluster and gives wrong abnormal signal.

Magnetic Resonance Imaging (MRI) is a very crucial imaging technique in the medical imaging processing where we can study the anatomical structure of the brain. Through the mri scanning we get the overview of the soft tissues which are inside any strong structures. Brain is covered by the strong bone called skull, and through the mri scanning we can get the anatomical structure of our brain.

In next step of our project, we are going to use these MRI images as our input data for finding the tumor detected in our brain is either the benign or malignant. These get processed using several image processing techniques and finally gets classified using another set of algorithms. In present days of machine learning era, there are several types of techniques and methods which can be used for classification are like supervised learning algorithms, unsupervised algorithms, convolutional neural network, fuzzy methods. Before the classification process we are going to do several methods like image segmentation, feature extraction, feature reduction.

Before get into medical imaging analysis of image of brain MRI, we need to do the image pre-processing like skull striping, enhancement of image, noise reduction using the several image pre-processing techniques like grayscale matching, thresholding. We can also add the morphological operations on the image to get the features of tumor. Here tumor is considered as the object in mri image. Here we extract the features of the objects of image based on the assumptions we made at the shape and size of the tumor of the brain. And finally extracted tumor is mapped into the actual greyscale mri image to male it visible.

Tumors are basically divided into two types:

A. Benign Brain

Benign tumor cells rarely invade surrounding tissue and cause health Problems. Benign tumors can be removed but repeat themselves. It is a rare thing. But this tumor cells won't be affecting other parts or never get into other parts. It is non-cancerous one.

B. Malignant Brain

A malignant tumor is also known as a brain cancer tumor. They grow rapidly and can congregate or invade nearby healthy brain tissue. This tumor cells are cancerous tumor cells where all the affected cells will be combined into a cluster and spreads to reach to other parts so that each and every visited path will be affected.

Mainly the capturing the anatomy of a brain can be done MRI scanning. Magnetic resonance imaging is a medical imaging technique used in radiology to create images of the body's anatomical and physiological processes. The MRI Scanner uses a strong magnetic field, radio waves, and magnetic field gradients to create images of the body. This software applies digital image processing technology to MRI images performed by MATLAB.

C. AIM

The purpose of this article is to identify tumors on CT images of the brain. The main reason for detecting brain tumors is to help with the clinical diagnosis. It provides an algorithm that ensures the presence of a tumor by combining several steps to provide a reliable method of detecting tumors on CT images of the brain. The methods used were filtering, contrast adjustment, image rejection, image subtraction, erosion, extension, thresholding, and tumor contouring. The goal of this project is to extract a tumor from a CT image of the brain and represent it in a simple way that anyone can understand. Sufficiently so that patients and healthcare professionals understand.

D. Objective

The objective of this task is to provide useful information in a user-friendly manner, especially to healthcare professionals who treat patients. Define an algorithm that produces tumor images extracted from CT brain imaging. The resulting image provides information such as the size, size and location of the tumor, follows the outline and c- marker of the tumor, and its boundaries provide information about the mass. This can be useful in various situations. It provides a better basis for staff decisions regarding the healing process. Plotting f-contour plots and c-label histograms of tumors and their boundaries allows medical personnel to better understand images using different colors at different levels of intensity for 3D images from 2D images. It will be easier to understand.

II. PROPOSED METHOD

In this paper, we used following methods for implementation. We used few algorithms and methods form the following list:
1) Data/Image processing.
2) Image segmentation.
3) Feature extraction and feature reduction.
4) Classification.

First two steps of our methodology belong to *Image processing* steps. Another two steps regarding the getting features and applying to *machine learning algorithm*.

A. Image Pre-Processing

The main task of pre-processing is to improve the quality of MR images and make them suitable for further processing by mechanical or human visual systems. In addition, pre-processing includes improving the noise ratio of the input signal, improving the visual appearance of the MR image, removing unwanted noise and unwanted parts of the background, smoothing the interior of the area, preserving the edges, etc. It helps to improve certain parameters. To improve the signal-to-noise ratio and thus the sharpness of the raw MR image, we applied adaptive contrast enhancement based on modified sigmoid function. Head resection is the process of removing all non-brain tissue from a brain image. Skull removal allows you to remove additional brain tissue such as fat, skin, and skull from the brain image. There are several techniques for removing the skull. Some common techniques include automatic skull extraction using image contours, segmentation-based skull extraction and morphometry operations, histogram analysis, or skull-based extraction. single-valued threshold. An extraction technique based on threshold operation to remove tissue from the skull.

1) *Conversion to grayscale/ grayscale Matching:* Any image can be created on the basis of combination of three colors. Those three colors are red, blue, green. Any image which has many colors can be formed by mixing those three colors in required proportion. As we aimed at processing of image, it will be easy for work on many problems like image segmentation and morphological operation when we turn the actual rgb image to grayscale image. As we know that rgb image contains three colors, then image is a three-layered image where each layer stacked upon another. But the conversion of rgb to grayscale makes the resultant image as the single layered because grayscale is done by mixing the equal proportion of those three primary colors. But the formed single layered grayscale image will be having different intensity. And another important advantage of grayscale image is that storage capacity for gray scale image is very less than the rgb image because for storing the single color of rgb image we may be need byte *3 but for grayscale only one byte. And it will be very easy for us to catch out the features of an image when it is single layered. We can use direct function which is available in Matlab for conversion of rgb image to grayscale image. Rgb2gray () is function which we are using. Here this function will be converting the triplet pixel values of those colors to single pixel array of gray scale.

B. *Image Segmentation*

In the field of medical image processing, the most important step is the image segmentation. It is more important in the low-level visualization and recognition of pattern. Image segmentation process is most important and difficult task in image processing to find out the end result in analysis of image, extracting textures from a particular image. This process involves the numbering the all the pixels of an image with label number so that all the group of pixels which having the same label number will share the same common image attributes. And it is process of classification of entire image into large number of small segments image so that each segment is homogenous. The process of isolating the tumor from the background of image is main important task in image segmentation of medical image processing. This allows the image to focus on the tumor and separate important information from the background. It becomes even more important that routine medical image processing requires pre and postoperative decisions to initiate and speed up the recovery process. Image segmentation is computer vision for segmenting digital images. Data-driven models and processes are used in. Divided into several segments for easy analysis. It is an integral part of image analysis and is often the first task used to distinguish between different objects in an image. Properties are calculated properties such as color, texture, grayscale, depth, and intensity. The method used to extract and represent information from an image is called segmentation.

1) *Image Segmentation Algorithms:* We are using the Ostu's algorithm in our project. Below we are discussing the matter regarding the algorithm:

a) *Otsu's Algorithm:* It is important algorithm in image segmentation, which is named after the Nobuyuki Ostu, is a popular global thresholding technique. This is used for conversion of gray image to binary image which means only in black and white color. This algorithm creates the bimodal histogram, it means to create the two classes. Two classes are foreground pixels and background pixels. This algorithm calculates the intra-class variance and inter class variance. It maintains the low intra class variance and optimum threshold between the two classes to make the high inter-class variance. This algorithm performs the search operation across image segments to find the threshold value which minimizes intra-class variance of the image that is segmented. The threshold value in this method is determined by the class threshold value which has the larger variance, may that class can be of foreground or background. We will get the good results from this algorithm when the image histogram has two different peaks in which one is from the foreground and another is from the background. This searching operation continues until the pixel values of that class reaches the minimum of variance.

Following steps are of ostu's algorithm:

➢ Finding the histogram for the image and finding the probabilities if each intensity pixel level.
➢ Set the initial class probabilities and mean of the class.
➢ Iterate through all the possible classes thresholds T= 1,2,3,4….max intensity
  • then update the class probabilities and mean accordingly.
  • calculate the Inter class variance.

Then the required threshold value shows the maximum inter class variance.

2) *K-means Algorithm:* It is most preferred hard clustering algorithm because of its high performance and easy implementation. Clustering is an unsupervised model where the different types of statistics are divided into different cluster. This is done simply by considering the two criterions where one step is describes about the data that is belongs to cluster is having high similarity and another step is about the data that belongs to 1-cluster ought to be differs with the other clusters data. This k-means algorithm needs the initial information regarding the image and the centroids of the cluster which are initially formed and about the number of clusters we are going to implement. But there may not be having prior information regarding the image in real- world problems. Whatever the parameters we are going to use for implementation of this algorithms are going to impact the overall result of the performance of the algorithm. So, we are aimed at using the algorithm without any of the parameters. In this scenario, we do the searching of the histogram along vertical and horizontal direction and getting record of the local and global maximum values. Image segmentation which is going to implemented through the k-means will provides the centroids of the clusters and the optimized cluster formed with the help of histogram.

Where K is number of clusters specified. The official steps included in this algorithm are:

a) Step-1: Select the first K centroid.
b) Step-2: then find the distance between each data item and the centroid
c) Step-3: Generate a group k using the method d 'Element data assignment closest Place it in the center.
d) Step-4: Recalculate the center of each group.
e) Step-5: Repeat steps 3 and 4 until the center position does not change.

The error which is formed due to the quadratic error method is reduced by the K-means algorithm. This is the chosen measure of distance between the data point and the center of the long group for the same color or measure of similarity.

C. *Feature Extraction and Feature Reduction using The GLCM*
1) *Feature extraction:* Wavelet transform is most powerful tool for feature extraction. We need to get the features of the objects found after image segmentation so that we can train the dataset easily for doing the classification of tumor cells from brain MRI image using the SVM machine learning algorithm. We are used the wavelet transform because it has multi- dimensional property where it allows the analysis of image at different levels of resolution. But it requires the high storage and computationally expensive. So, we combine the feature reduction with feature reduction. Fourier transform is most common wavelet transform for studying the features of the wave. Fourier transform transforms the time signal into frequency signal. But some signals are bounded to some limit the continuous transform of signal will create the loss of signal at certain time slots. So, we are using the short time Fourier transform also called as the windowing. This will add small window at specific location of wave so that both the time information and the frequency information is preserved. It also adds the dynamic window where the size of window can be changed according to the requirement. The precision of feature extraction depends on the size of window and another advantage is that it provides the time-scale view of data rather the traditional time-frequency view. Here we are using the discrete wavelet transform as for feature extraction. It is most useful method for implementation of wavelet transform. DWT uses the dyadic scale and positions.

2) *Gray-Level Co-Occurrence Matrix Algorithm (GLCM):* GLCM is also refers as the co-occurrence matrix. It is a statistical texture analysis method. It gives us the relationship between the pixels and how many times those combined pixels are present in an image through a given direction and distance. There will be 20 features for each image. It is mainly used for medical image analysis, classification. It gives us the spatial relationship. Matrix is updated by calculating the occurrences of pixel value pairs at a certain distance and through specific degrees. The GLCM values are updated by calculating the f(i,j), at a distance vector d=(x,y). The (i,j)th element of the matrix is defined as the probability of gray values i and j occurrence at distance d and angle θ. Then the values of GLCM matrix is updated in this manner. We need the four distances (1,2,3,4) and four angle (0,45,90,135) needed for calculation of co-occurrence of matrix.

3) *Feature Reduction:* As the extracted features are increasingly, the computational features and storage capacity will also increase. Not only making the increase of storage capacity, but also risk of wrong classification. So, we required to reduce to number of features. Principal component analysis is a powerful tool for reducing the dimension of data while maintaining the variance among huge number of corelated attribute. It can be done by transformation of old dataset into the new dataset by ordering the attributes based on the variance or importance. This technique uses do three things: Normalization of components so that making attributes uncorrelated to each other by using orthogonality property. All the input vectors are orthogonalizes and arranged in such a order that components with high variance will comes first and components with least variance is eliminated in new dataset. Before performing the feature reduction, we need to do the normalization of input vectors such that they should have the zero mean and unity variance.

D. Classification

Classification is process of machine learning to get the results as the either yes or no type. Here in our project, we are using classifier algorithms for extract the whether the brain tumor found in the image after giving to algorithm is either non-cancerous or cancerous. Classification algorithms are of many types. Here we are using the support vector machine classifier algorithm because it is most powerful classifier algorithm in classifier machine learning algorithms. SVM as its own powerful advantages for making it in use for several projects. It is of supervised learning type where the prediction of whether the given tumor is benign or malignant is determined by using trained data which we provided to SVMs. It is very crucial to note that the whatever the examples we are giving that should be true. Algorithms working is based on the training of data. Training data will be contained of both cancerous and non-cancerous images and algorithms will try to get the knowledge from the provided training data. Based on the training data, algorithms are going to work on the testing data. This classifier works like the following steps:

First it will make a model where the training examples are being represented as the points/dots in space. These points/dots are mapped in space into two categories where the one belongs to non-cancerous category will be separated from cancerous category with large gap in between them. When the new example data is fed to classifier-model they will be plotted into space according to the prediction made by the classifier and made to get into any one category based on the which side of gap they fall on. These gaps are the functions which constructs the set of hyper-planes or hyper-plane in space. This can be used for regression or classification problems. Hyper- lanes are constructed as classifier function margin to classify the data. A great gap will be made by hyper plane which means there will be large distance to training data dot/point belongs to any category. This makes low generalization error of the classifier due to larger margin.

Here we provide the SVM with a lot of features as input vectors and makes the training model after scaling and gives the output by training model. Actually, original classifiers are linear classifiers. We are using of kernel SVM's which are non-linear classifiers by using the kernel function in the place of dot product form in the linear SVM's. kernel svm's provide maximum gap hyper- plane. Svms are state of art machine learning algorithm dues to various reasons like high accuracy rate, direct geometric interpretation and mathematical tracebility. We are using cross validation function to kernel svm's to overcome the overfitting of data into the classifiers.

1) *Cross Validation:* As we training the classifier with training data, the classifier may give high accuracy to that dataset. It may not produce accuracy for the other independent datasets. This kind of problem is called the overfitting. In order to overcome over-fitting, we are going to use the cross validation in our proposed method. It may not increase the classification accuracy but can make our proposed model as reliable and generalized for using on other independent datasets.

Cross validation process consists of following three types:
a) Random subsampling.
b) K-fold cross validation.
c) Leave-one-out validation.

Here in our proposed method, we are using K-fold stratified cross validation because of its easy implementation and simple for using on training data and validating data. Here we are going to use the 5-fold cross validation, in this method we are going to partition the whole data into 5 parts and we perform training on any one part of those 5 parts and left out parts are for validation. This k-folds are formed by random partitioning where distribution in one of folds is can be different from the other folds. So, we are using the k-fold stratified where distribution of every fold is same. The way of determining the k value is done by trail and error method by using the varying the value of k value incrementing the value of 3 to 10 by 1 value for each step. And then we will train svm classifier for each value. We will select the k value which show high classification accuracy.

If we set the K value as high value then the computational time will be decrease because of increase of variance of estimator and decrease of bias of true error rate estimator. If we set the K value as low value then the computational time will be increase because of decrease of variance of estimator and increase of bias of true error rate estimator.

Use of kernel svm's due to following reasons:
➢ It is very useful in practical applications and gives excellent results in the field of bio-informatics, computer vision, textual classification. And having of tunable parameters.
➢ Convex quadratic optimization is used in training.

Following functions are the kernel function that we are using in kernel svm of our project:
● Homogenous polynomial.
● Inhomogeneous polynomial.
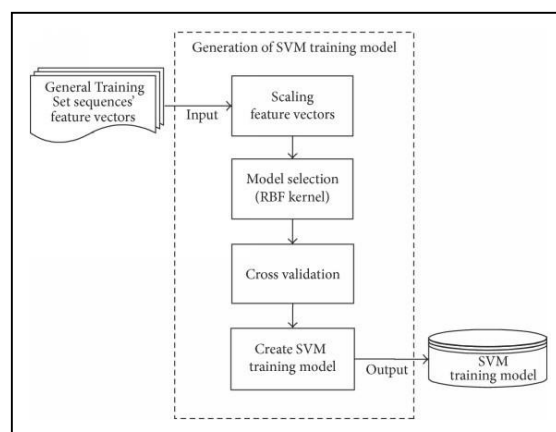● Gaussian radial Basis.

Figure 4.1: flowchart of SVM classifier.

III. FUTURE SCOPE

This detection and classification of brain tumor from the brain mri images helps both the medical practitioner, radiologists and brain specialist doctors called as neurologists. Main diagnosis for brain tumor starts from the MRI scanning but mri scanning alone can't be helps us to find out the tumor that presents in the brain. We can't be able to distinguish the whether the present tumor cells in mri image are of benign or malignant. In order to find out the specific nature of object found in mri image. Additional to MRI scanning doctor may perform the magnetic resonance spectroscopy (MRS) to get the chemical profile and to find out the details of lesions observed on the mri image. To detect recurring brain tumor, we asked to get done the PET scan. In some case, they may recommend us to get done by biopsy for diagnosis of brain tumor.

So, in order to make this so simple we are going to extract the tumor from the MRI scanned images using the machine learning classifiers. This concept very helpful because we are going to build a model using the existing machine learning algorithms. And make model to predict the details about the tumor whether it is benign or malignant. So, this model of making use of image processing and image segmentation for classification of make use of this build model in practical applications by increasing the overall efficiency of classification.

IV. CONCLUSION

The collection of abnormal growing cells in a tissue of brain causes damage to the brain. This affects the normal functioning of brain because this abnormal growth of tissues can't be controlled by body itself. It is very important to identify the abnormal issue at early stage of bad tissue birth and to start the diagnosis. The most crucial goal of medical image processing is to find the accurate, useful, worthful, authentic information using various algorithms with very possible minimum error.

We are mainly used the four steps in our projects to make out the model to get ready. Those four steps are image preprocessing, image segmentation, feature extraction and image classification. The first and foremost important task is to get the image dataset like the MRI images. This MRI images will contain the different illumination intensities and different pixels values each may be correlated to each other or not.

To make obtained image for good to processing we are using the pre-processing to image like skull striping, conversion to gray scale. We used various segmentation algorithms for our project like Ostu's segmentation, k-means clustering algorithm to make image to various regions. After segmentation we did the feature extraction and given those features for the input of classifier. As the feature extraction step is act as input for the classifier then this step impacts on the overall accuracy of the model. As there many classifier-algorithms are available for machine learning we are using SVM algorithm because of its advantages. The main focus of our project should on the accuracy because it is medical application project where the patient lives depend on the level of accuracy the equipment we are using. Then we used to derive different kernel functions to their accuracy values. So that we can use anyone in real time application. We got the desired results from the gaussian radial kernel function gives us the high accuracy.

Brain tumor reduces extra time for detection and extraction. It not only reduces the time for detection for doctors but also helps patients to get done diagnosis very less time and with less money.

In future we may also work on the efficiency of classification. The overall increment in the dataset can be done very easily by providing the large diverse dataset as the input to the machine learning model. Prediction power for more output classes can be also increase with help of more data providing to the input.

We also induct the different algorithms for classification of image objects like convolutional neural network.

We can also make use of this model for medical equipment innovation, we can work on it further more to

REFERENCES

[1] R. Mishra, "MRI based brain tumor detection using wavelet packet feature and artificial neural networks," Proceedings of the International Conference and Workshop on Emerging Trends in Technology - ICWET 10, 2010.

[2] Y. K. Dubey and M. M. Mushrif, "Segmentation of brain MR images using intuitionistic fuzzy clustering algorithm," Proceedings of the Eighth Indian Conference on Computer Vision, Graphics and Image Processing - ICVGIP 12, 2012.

[3] S. M. K. Hasan, M. Ahmad, and S. D. Ghosh, "Perceptive Proposition of Combined Boosted Algorithm for Brain Tumor Segmentation," Proceedings of the International Conference on Advances in Information Communication Technology & Computing - AICTC 16, 2016.

[4] J. Han, M. Kamber, and J. Pei, Data mining: concepts and techniques. (2012)

[5] E. Küçükkülahli, P. Erdoğmuş, and K. Polat, "Brain MRI Segmentation based on Different Clustering Algorithms," International Journal of Computer Applications, vol. 155, no. 3, pp. 37– 40, 2016.Mohsen, Fahd, et al. "A new image segmentation method based on particle swarm optimization." Int. Arab J. Inf. Technol. 9.5 (2012): 487-493.

[6] Mohsen, Fahd, et al. "A new image segmentation method based on particle swarm optimization." Int. Arab J. Inf. Technol. 9.5 (2012): 487- 493.

[7] R. V. Rao, "Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems," International Journal of Industrial Engineering Computations, pp. 18–34, 2016.

[8] M. Al-Ayyoub, G. Husari, O. Darwish, and A. Alabed-Alaziz, "Machine learning approach for brain tumor detection," Proceedings of the 3rd International Conference on Information and Communication Systems - ICICS 12, 2012.

[9] Q. Sun and H. Tian, "Interactive image segmentation using power watershed and active contour model," 2012 3rd IEEE International Conference on Network Infrastructure and Digital Content, 2012.

[10] I. Maiti and M. Chakraborty, "A new method for brain tumor segmentation based on watershed and edge detection algorithms in HSV colour model," 2012 National Conference On Computing And Communication Systems, 2012.

[11] Y. Liu and Q. Zhao, "An improved watershed algorithm based on multi- scale gradient and distance transformation," 2010 3rd International Congress on Image and Signal Processing, 2010.

[12] Kaur, Mandhir, and Rinkesh Mittal. "Survey of Intelligent Methods for Brain Tumor Detection." International Journal of Computer Science Issues (IJCSI) 11.5 (2014): 108.

[13] K. Hu, G.-Y. Tang, D.-P. Xiong, and Q. Qiu, "A novel image segmentation algorithm based on Hidden Markov Random Field model and Finite Mixture Model parameter estimation," 2012 International Conference on Wavelet Analysis and Pattern Recognition, 2012.

[14] D. Sridhar and I. V. M. Krishna, "Brain Tumor Classification using Discrete Cosine Transform and Probabilistic Neural Network," 2013 International Conference on Signal Processing , Image Processing & Pattern Recognition, 2013

[15] A. Ladgham, G. Torkhani, A. Sakly, and A. Mtibaa, "Modified support vector machines for MR brain images recognition," 2013 International Conference on Control, Decision and Information Technologies (CoDIT), 2013.

[16] P. Su, Z. Xue, L. Chi, J. Yang, and S. T. Wong, "Support vector machine (SVM) active learning for automated Glioblastoma segmentation," 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI), 2012.

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089   (24*7 Support on Whatsapp)