

Classification of Abnormalities in Brain MRI Images Using PCA and SVM

Ajala Funmilola A^a., Fenwa Olusayo D^b., Amusan Elizabeth A. ^c

^{a b c}Department of Computer Science and Engineering, LAUTECH, P.M.B 4000, Ogbomoso, Nigeria

(* Corresponding authors: fenwadeborah@yahoo.com, eaadewusi@lautech.edu.ng, faajala@lautech.edu.ng)

Abstract

The impact of digital image processing is increasing by the day for its use in the medical and research areas. Medical image classification scheme has been on the increase in order to help physicians and medical practitioners in their evaluation and analysis of diseases. Several classification schemes such as Artificial Neural Network (ANN), Bayes Classification, Support Vector Machine (SVM) and K-Means Nearest Neighbor have been used. In this paper, we evaluate and compared the performance of SVM and PCA by analyzing diseased image of the brain (Alzheimer) and normal (MRI) brain. The results show that Principal Components Analysis outperforms the Support Vector Machine in terms of training time and recognition time.

Keywords: Alzheimer, Support Vector Machine, Principal Components Analysis, Medical Image

I. Introduction

In recent decades, the need for computers in assisting the processing and analysis of medical images has become inevitable with the mounting size and number of medical images [1]. Brain tumors are abnormal and uncontrolled proliferations of cells. Some originate in the brain itself, in which case they are termed primary. Others spread to this location from somewhere else in the body through metastasis, and are termed secondary. Primary brain tumors do not spread to other body sites, and can be malignant. Secondary brain tumors are always malignant. Both types are potentially disabling and life threatening, this is because the space inside the skull is limited, their growth increases intracranial pressure, and may cause edema, reduced blood flow, displacement with consequent degeneration of healthy tissue that controls vital functions [2]. Brain tumors are, in fact, the second leading cause of cancer related deaths in children and young adults. According to the Central Brain Tumor Registry of the United States (CBTRUS), there were 64,530 new cases of primary brain and central nervous system tumors diagnosed by the end of 2011. In general, more than 600,000 people currently live with the disease (www.cewebsource.com).

A major challenge facing MRI brain image segmentation is that the brain tissue is often divided into white matter (WM), gray matter (GM) and cerebrospinal fluid (CSF). The precise measurement of WM, GM and CSF is important for quantitative pathological analyses and so it becomes a goal of lots of methods for segmenting MRI brain image data. Because of the imperfections of imaging scanner and imaging techniques, obtained medical images will inevitably be affected by some

corruption factors such as random additive noises, partial volume effect and intensity bias field. For improving segmentation performance, many different strategies, for example, Mercer Kernel techniques, Filtering techniques and so on, can be adopted [3]. However, Alzheimer's Disease (AD) is a condition where the brain slowly goes down, as well as a serious loss of thinking ability in a person and cognitive impairment. PET scan image is mainly used for the Alzheimer's neurological disorder treatment and other kind of dementia. In 2006, there were 26.6 million sufferers worldwide. Alzheimer's is predicted to affect 1 in 85 people globally by 2050.

II. Related Works

A number of approaches have been used to segment and predict the grade and volume of the brain tumor. [4], in their work proposed a Fuzzy Cognitive Map (FCM) to find the grade value of tumor. Authors used the soft computing method of FCM to represent and model expert's knowledge. FCM grading model achieved a diagnostic output accuracy of 90.26% and 93.22 % of brain tumors of low grade and high grade respectively. They proposed the technique only for characterization and accurate determination of grade. [5] in their work proposed an implementation of evaluation method known as Image Mosaicking in evaluating the MRI brain abnormalities segmentation study. 57 mosaic images are formed by cutting various shapes and size of abnormalities and pasting it onto normal brain tissue. PSO, ANFIS, FCM are used to segment the mosaic images formed. Statistical analysis method of Receiver Operating Characteristic (ROC) was used to calculate the accuracy.

[6], in their work proposed detection of tumor growth by advanced diameter technique using MRI data. To find the volume of brain tumor they proposed diameter and graph based methods. The result showed tumor growth and volume.[8] proposed a system that automatically segments and labels tumor in MRI of the human brain. They proposed a system which integrates knowledge based techniques with multispectral analysis. The results of the system generally correspond well to ground truth, both on a per state basis and more importantly in tracking total volume during treatment over time.

In this work, classification system for abnormalities of MRI brain images using principal component analysis, and support vector machine was proposed.

III. METHODOLOGY

The block diagram of the proposed system showing the stages of the system development is as displayed in figure 3.1. The stages are: Medical Image Acquisition; Image Preprocessing; Feature Extraction; Classification (using SVM and ANN); and Performance Evaluation.

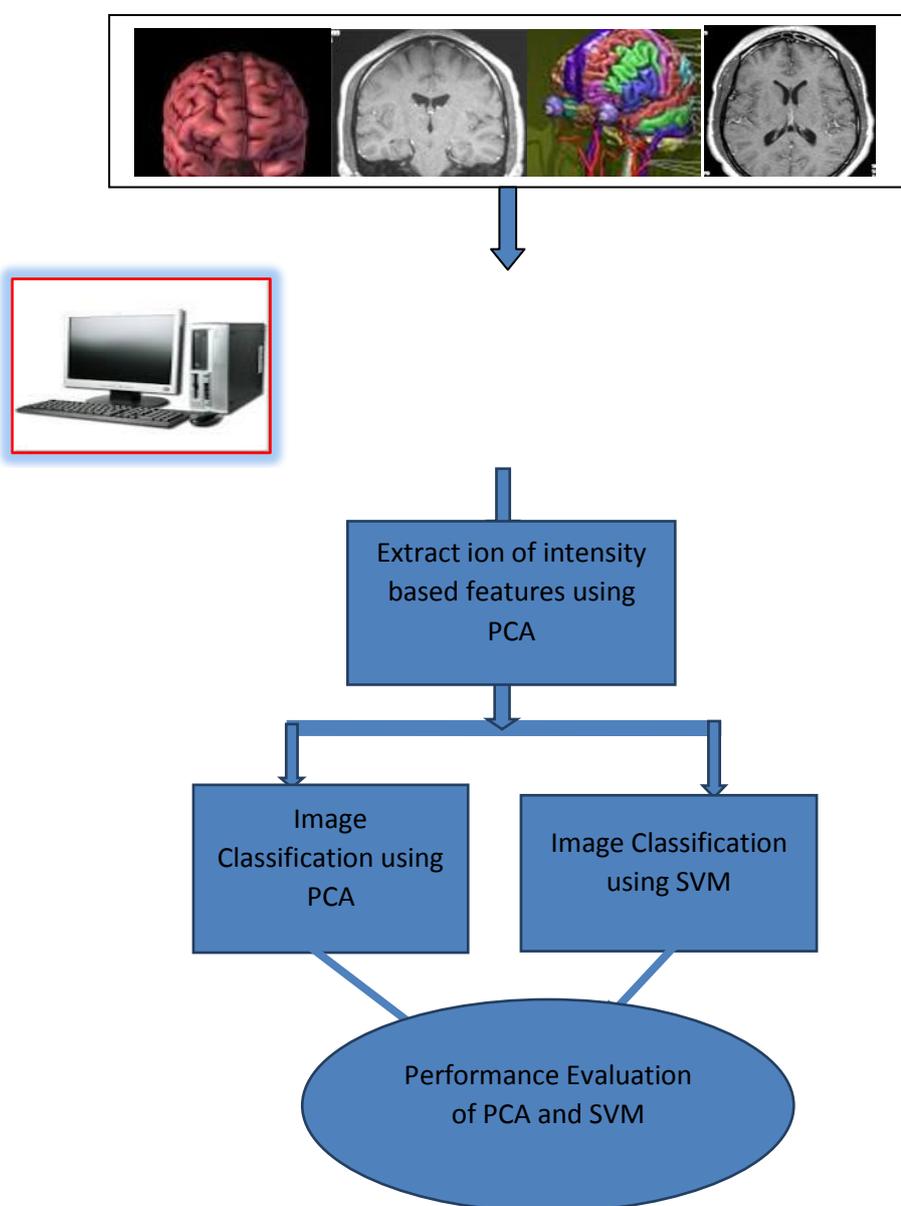


Figure 3.1 Model Architecture of the Proposed System

3.1 Image (Data) Acquisition

The image used for this research work was obtained from an online database of brain MRI images. Typical data samples consisting of Alzheimer diseased brain and normal brain are as shown in figures 3.2. and 3.3. The database provides a repository of these images which can be downloaded and regenerated in the MatLab environment. Some of these images are stored for research purposes, and for other image processing analysis. After the images were gotten from an online source, a database that contains both images was created in the MatLab environment. The image was called from the database using Matlab algorithm.

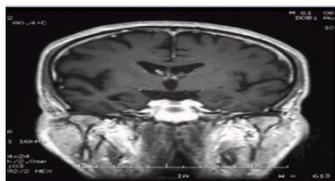


Figure 3.2: A typical image of the Alzheimer Disease
(Source : www.mednet.com)

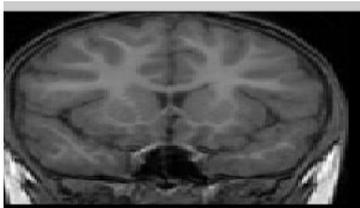


Figure 3.3: Normal MRI Brain (Source:
www.mribrain.com)

3.2 Image Pre-processing

Pre-processing of image is necessary before any image analysis can be carried out. It involves filtering, selecting, randomizing, conversion to gray-scale, resizing and removal of objects that could affect the proper processing of the images. In analysis of the medical images used in this paper, we try to avoid image pre-processing because image preprocessing can decrease image information content.

3.2.1 Pre-processing to gray scale

A major pre-processing is conversion to grayscale. Most images obtained are always in colored form, and the only way to process such image is by conversion to gray scale. An RGB image is a 3 by 3 image matrix consisting of rows, columns, and index type. The value of an image is also identified by its class, such as uint8, uint16, double. For a grayscale image, it is made to be 2 by 2 matrix, depending on the image size, and also retain its class.

3.3 Feature Extraction

The purpose of feature extraction is to reduce the original data set by measuring certain properties of features that distinguish on input pattern from another.

3.3.1 Texture features

Texture is a very useful characterization for a wide range of image. It is generally believed that human visual systems use texture for recognition and interpretation. In general, color is usually a pixel property while texture can only be measured from a group of pixels. A large number of techniques have been proposed to extract texture features. Based on the domain from which the texture feature is extracted, they can be broadly classified into spatial texture feature extraction methods and spectral texture feature extraction methods. For the former approach, texture features are extracted by computing the pixel statistics or finding the local pixel structures in original image domain, whereas the latter transforms an image into frequency domain and then calculates feature from the transformed image.

3.3.2 Shape Features

Shape is known as an important cue for human beings to identify and recognize the real-world objects, whose purpose is to encode simple geometrical forms such as straight lines in different directions. Shape feature extraction techniques can be broadly classified into two groups, viz., contour based and region based methods. The former calculates shape features only from the boundary of the shape, while the latter method extracts features from the entire region.

3.4. Feature Selection

Feature selection (also known as subset selection) is a process commonly used in machine learning, wherein a subset of the features available from the data is selected for application of a learning algorithm. The best subset contains the least number of dimensions that contributes to high accuracy; we discard the remaining, unimportant dimensions.

3.4.1. Forward Selection

This selection process starts with no variables and adds them one by one, at each step adding the one that decreases the error the most, until any further addition does not significantly decrease the error. We use a simple ranking based feature selection criterion, a two-tailed t-test, which measures the significance of a difference of means between two distributions, and therefore evaluates the discriminative power of each individual feature in separating two classes. The features are assumed to

come from normal distributions with unknown, but equal variances. Since the correlation among features has been completely ignored in this feature ranking method, redundant features can be inevitably selected, which ultimately affects the classification results. Therefore, we use this feature ranking method to select the more discriminative feature, e.g. by applying a cut-off ratio (p value < 0.1), and then apply a feature subset selection method on the reduced feature space.

3.4.2. Backward Selection

This selection process starts with all the variables and removes them one by one, at each step removing the one that decreases the error the most (or increases it only slightly), until any further removal increases the error significantly. To reduce over fitting, the error referred to above is the error on a validation set that is distinct from the training set. The support vector machine recursive feature elimination algorithm is applied to find a subset of features that optimizes the performance of the classifier. This algorithm determines the ranking of the features based on a backward sequential selection method that remove one feature at a time. At each time, the removed feature makes the variation of SVM based leave-one-out error bound smallest, compared to removing other features.

3.5 Principal Components Analysis

Principal components are the projection of the original features onto the eigenvectors and correspond to the largest Eigenvalues of the covariance matrix of the original feature set. PCA provides linear representation of the original data using the least number of components with the mean squared error minimized. PCA can be used to approximate the original data with lower dimensional feature vectors. The basic approach is to compute the eigenvectors of the covariance matrix of the original data, and approximate it by a linear combination of the leading eigenvectors. By using PCA procedure, the test image can be identified by first projecting the image onto the Eigen-space to obtain the corresponding set of weights, and then comparing with the set of weights of the images in the training set.

3.5.1 Algorithm for Principal Component Analysis

- Data: Intensity and label training images
- Result: Intensity, label, and intensity/label coefficient PCA subspaces.
- Step 1: Register the training images group-wise
- Step 2: Apply transformations from the registration to the respective label data
- Step 3: Compute the PCA basis for the registered intensity training data

Step 4: Compute the binary images from the registered training labels

Step 5: Compute the PCA basis for the binary images

Step 6: Project the binary label and intensity training images onto their respective bases

Step 7 Compute the PCA basis of the coefficients of the training data from the projection

3.6 Support Vector Machine

Support Vector Machines are a state of the art pattern recognition technique grown up from statistical learning theory. SVM is capable of producing a hyper plane which linearly separates two distinct classes, or two different types of images. SVM uses an optimum linear separating hyper plane to separate two set of data in feature space as shown in figure 3.4. This optimum hyper plane is produced by maximizing minimum margin between the two sets. Therefore the resulting hyper plane will only be dependent on border training patterns called support vectors. The standard SVM is a linear classifier which is composed of a set of given support vectors \mathbf{z} and a set of weights \mathbf{w} .

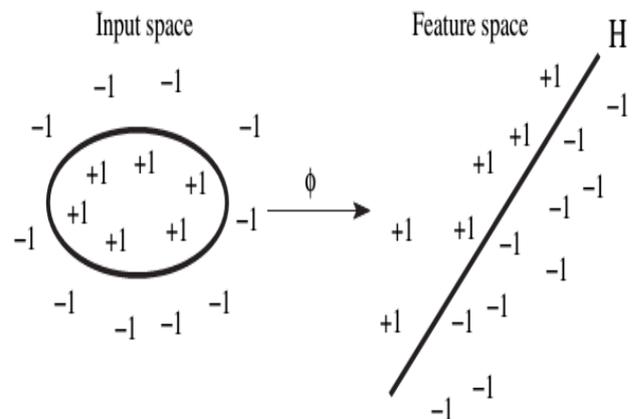


Figure 3.4: Linear separation in feature space

The computation for the output of a given SVM with N support vectors z_1, z_2, \dots, z_N and weights w_1, w_2, \dots, w_N is then given by:

$$F(x) = \sum_{i=1}^N w_i(z_i, x) + b \tag{1}$$

3.6.1 Algorithm for Image Classification using Support Vector Machine

- Step 1: Input image into Mat lab
- Step 2: Perform Pre-Processing, and Normalization
- Step 3: Perform feature extraction by extracting surface texture and roughness
- Step 4: For 100 image datasets, label first 50 normal image as 1, and the last 50 images as 2.

- Step 5: Create train datasets and test datasets
- Step 6: In test datasets, for another 100 image datasets, label first 50 normal images as 1, and the last 50 abnormal images as 2.
- Step 7: Perform SVM classification by calling the command svmstruct in Mat lab.
- Step 8: Check classification by calling the variable group

interface was developed which was used to carry out the process.

4.2 Mat lab Implementation

Any image which comes into the Mat lab environment has a definite value, which is identified by Mat lab. In image processing Mat lab enhances the image and generates its value using the im-read function. This enables further processing. Image components are also gotten in the work environment, after its value has been gotten. This is very important because it helps to enhance image processing. These components are identifiable and are attached to each image, this implies that for every image coming into the Mat lab environment. It has a definite component, which makes it very useful in analysis.

IV. Results and Discussion

4.1. System Design and Implementation

This involves loading the image anywhere in the computer to get the image. The first stage of any image processing is getting the image. After the image has been gotten, the various tasks which involve processing the image, analyzing the image was now performed. For this purpose, an interactive

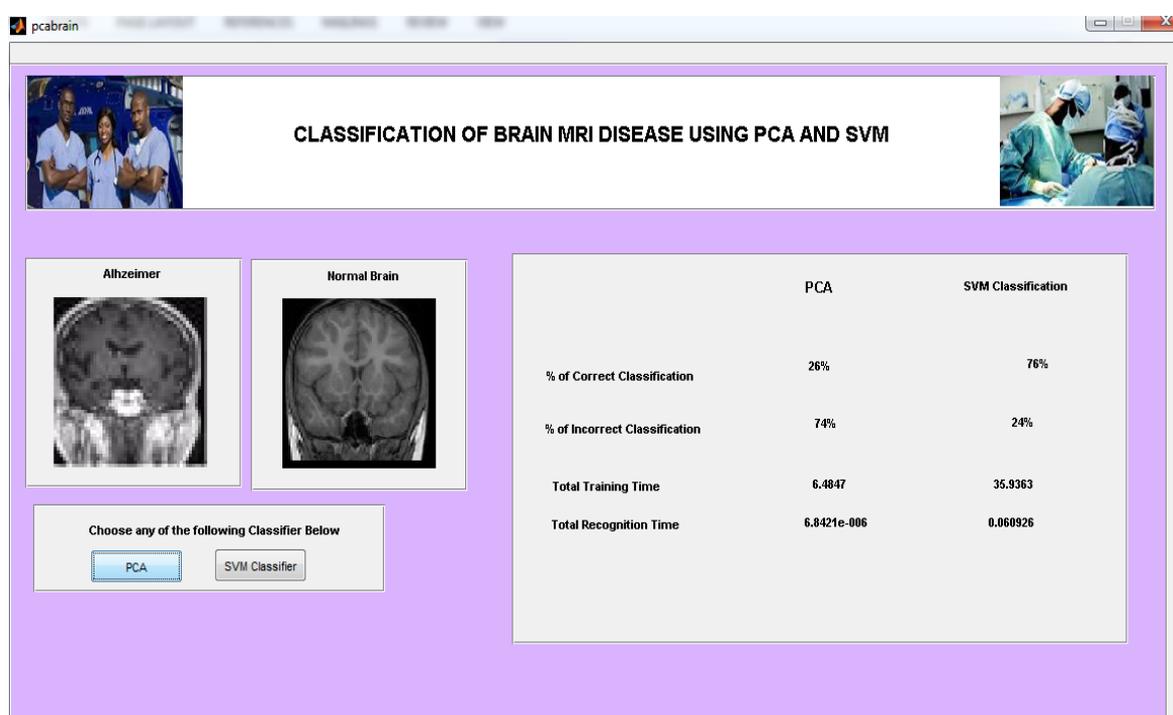


Figure 4.1: User Interface for Analysis

4.3 Training Data

In this work, a total number of 50 images containing both Normal MRI Brain and Alzheimerdisease were used. The features of the images were extracted, and passed into training using principal component analysis, while another set of 50 images were used for testing. The same set of images was passed into the Support Vector Machine, and another set of these images were tested and evaluated. Figures3.2 and 3.3 show typical images of the Alzheimer Disease and Normal MRI Brain, from which features were extracted.

SVM Results

SVM maps input vectors to a higher dimensional vector space where an optimal hyper plane is constructed. The data with linear severability may be analyzed with a hyper plane, and the linearly non separable data are analyzed with kernel functions such as GaussianRBF. The output of an SVM is a linear combination of the training examples projected onto a high dimensional feature space through the use of kernel function. For this work SVM with kernel function linear and RBF (Radial Basis Function) is used for classification of images into two classes namely “Alzheimer” and “Normal MRI Brain”. The labels for these classes are using „1“ and „2” for “Normal” and “Abnormal” respectively.

A total of 50 images were passed into the database for training and testing. The resulting classification shows the SVM result generating a confusion matrix as shown in table 4.1 below. Confusion matrix is a table matrix which shows correct classification, and misclassification. Correct Classification occurs when Normal Brain is classified as Normal and Diseased Brain is classified as Diseased Brain.

Table 4.1: Confusion Matrix for SVM Classification

	Normal Brain	Diseased Brain (Alzheimer)
Normal Brain	19	6
Diseased Brain (Alzheimer)	6	19

PCA Results

The principal components of the images were extracted, and trained while another set of test images were used for testing, the results gotten was poor this is because principal components of an image are not enough for classification. The result is shown in table 4.2 below:

Table 4.2: Results for PCA and SVM

	Correct Classification	Incorrect Classification
PCA	26%	74%
SVM	76%	24%

PCA picks components based on Eigen Values and Eigen Vectors. These components are used for classifications; however, they don't give accurate result during testing, due to redundancy in image. SVM results seem better as compared to the PCA results.

4.4 Training time

The training time is the time required by the system to train and understand the behavior of the images in question, during training patterns are recognized, studied and evaluated until after which the carefully studied pattern are extracted, and used for testing. Training time is of significant importance, as it is used for evaluation of speed and performance. The training and recognition time for both classifiers is shown in table 4.3 below.

4.5 Recognition Time

Recognition of images occurs when the classifier is able to correctly identify and acquaint itself with what it has been trained with. Recognition time plays an important role in classification of medical images because higher recognition time could lead to memory consumption which could affect corresponding results. In the case of a lower recognition time, it yields low memory which is useful and better, and does not affect corresponding results.

Table 4.3: Training and Recognition Time as for both classifiers

	Number of Images	Training Time (in seconds)	Recognition Time (in Seconds)
PCA	50	6.4847	0.006
SVM	50	37.406	0.0589

In table 4.3 above, the training and recognition time is seen and observed for the PCA and SVM. The training time and recognition time of PCA seem smaller and shorter than that of the SVM, hence PCA performs better in terms of training and recognition.

4.6. Performance Evaluation Metrics

The performance evaluation metrics used are: Accuracy, Specificity, Sensitivity, True Positive, True Negative, False Positive and False Negative.

Accuracy- is defined as the accurate value of classification which equals the total sum of correct classification over the total sum of correct and incorrect classification multiplied by 100.

Specificity- is defined as the total division of true positive against the total of true positive and false negative, while sensitivity is defined as true negative against true negative and false positive.

True positive (Tp)- occurs when the correct classification for the right image is done i.e. when a normal brain is classified as normal, while false negative (Fn) occurs when a normal brain is incorrectly classified as diseased.

True negative (Tn)- occurs when the correct classification for the right image is done i.e. when a diseased image is classified as being diseased, while false positive (Fp) occurs when a diseased image is incorrectly classified to be normal.

$$\text{Specificity} = \frac{T_p}{T_p + F_n} * 100; \text{Sensitivity} = \frac{T_n}{T_n + F_p} * 100;$$
$$\text{Accuracy} = \frac{T_n + T_p}{T_p + T_n + F_n + F_p} * 100$$

where, Tp is true positive (i.e. Normal Brain classified correctly as Normal Brain), Fn is false negative (i.e. Normal Brain is incorrectly classified as Diseased), Tn is true negative (i.e. Diseased Brain correctly classified as Diseased), Fp is false positive (i.e. Diseased Brain incorrectly classified as Normal).

From the confusion matrix above, it is seen that our Tp is 19, while Fn is 6, and Tn is 19, while Fp is 6. Hence it is discovered to be

$$\text{Specificity} = \frac{19}{19+6} * 100 = 76\%$$

$$\text{Sensitivity} = \frac{19}{19+6} * 100 = 76\%$$

$$\text{Accuracy} = \frac{19+19}{19+19+6+6} * 100 = 76\%$$

V. Conclusion and Recommendation

5.1 Conclusion

In this work, a classification scheme between PCA and SVM were carried out on diseased image of the brain (Alzheimer) and normal (MRI) brain. The results show that PCA outperforms the SVM.

5.2 Recommendation

The following recommendations are suggested:

- A mapping feature of principal components and support vector machine should be carried out, and results compared.
- The database containing the images should be increased to as much as 400 image samples, and subsequent results evaluated.
- Other classification schemes such as Bayes, KNN, could also be evaluated with the corresponding image disease.
- Other diseases of the brain could also be analyzed, and compared with normal brain.
- Students of Higher Institution should be exposed to MatLab, as a core course of study.

References

- [1] Pham D., Xu C., and Prince J. (2000): "A Survey of Current Methods in Medical Image Segmentation," Annual Review of Biomedical Engineering, 2(3): 315-337, 2000.
- [2] Gladis Pushparathi V. P. and Palani S (2012): "A novel approach for feature extraction and Selection on MRI images for brain Tumor classification " ,Proceeding CCSEA, SEA,CLOUD, DKMP, CS & IT-CSCP 2012, New Delhi: 225-234.
- [3] Neha J. and Karaulia D.S. (2014): "A Comparative Analysis of Filters on Brain MRI Images", International Journal of Advanced Research in Computer Science and Software Engineering, 4(11):893-897.
- [4] Papageorgiou E.I., and Spyridonos P.P. (2008): "Brain Tumor characterization using the soft computing technique of fuzzy cognitive maps" Applied soft computing 8: 820-828.
- [5] Shafaf Ibrahim, Noor Elaiza Abdul Khalid (2011): "Image Mosaicing for evaluation of MRI Brain Tissue abnormalities segmentation study", Int.J.Biology and Biomedical Engineering, 4(5): 181-189
- [6] Karpagam S., and Gowri S. (2011): "Detection of tumor growth by advanced diameter technique using MRI data", Proc. The world congress of Engineering, 1WEC 2011, London, U.K
- [7] Matthew C.Clark, Lawrence O.Hall (1998): "Automatic tumor segmentation using Knowledge based techniques", IEEE Transactions on Medical Imaging, 17(2).
- [8] Carlos A Parra, Khan Iftexharuddin (2003): "Automated brain data segmentation and Pattern recognition using ANN" Proc. CIRAS, 2003