

Classification and Segmentation of Brain Tumor using Texture Analysis

QURAT-UL-AIN, GHAZANFAR LATIF, SIDRA BATOOL KAZMI, M. ARFAN JAFFAR,
ANWAR M. MIRZA

Department of Computer Science
FAST National University of Computer and Emerging Sciences
A.K Brohi Road, H-11/4, Islamabad
PAKISTAN

quratul.ain@nu.edu.pk , ghazanfar_384@hotmail.com , sidrabortoolkazmi@nu.edu.pk ,
arfjan.jaffar@nu.edu.pk, anwar.m.mirza@nu.edu.pk

Abstract: - Brain tumor diagnosis is a very crucial task. This system provides an efficient and fast way for diagnosis of the brain tumor. Proposed system consists of multiple phases. First phase consists of texture feature extraction from brain MR images. Second phase classify brain images on the bases of these texture feature using ensemble base classifier. After classification tumor region is extracted from those images which are classified as malignant using two-stage segmentation process. Segmentation consists of skull removal and tumor extraction phases. Quantitative results show that our proposed system performed very efficiently and accurately. We achieved accuracy of classification beyond 99%. Segmentation results also show that brain tumor region is extracted quite accurately.

Key-Words: - Segmentation, Classification, Texture feature, Magnetic resonance imaging (MRI), Support vector machine (SVM), Ensemble base classifier.

1 Introduction

Brain is the kernel part of the body. Brain has a very complex structure. Brain is hidden from direct view by the protective skull. This skull gives brain protection from injuries as well as it hinders the study of its function in both health and disease. But brain can be affected by a problem which cause change in its normal structure and its normal behavior .This problem is known as brain tumor. Brain tumor causes the abnormal growth of the cells in the brain. The cells which supplies the brain in the arteries are tightly bound together thereby routine laboratory test are inadequate to analyze the chemistry of brain. Computed tomography and magnetic resonance imaging are two imaging modalities that allow the doctors and researchers to study the brain by looking at the brain non-invasively [1].

Magnetic Resonance Imaging (MRI) is a medical imaging technique. Radiologist used it for the visualization of the internal structure of the body. MRI provides rich information about human soft tissues anatomy. MRI helps for diagnosis of the brain tumor. Images obtained by the MRI are used for analyzing and studying the behavior of the brain.

Image intensity in MRI depends upon four parameters. One is proton density (PD) which is determined by the relative concentration of water molecules. Other three parameters are T1, T2, and T2* relaxation, which reflect different features of the local environment of individual protons.

Classification is the technique for classifying the

objects into corresponding classes. Ones the brain images acquired they are classified as normal and abnormal. For classification of the images different features of the image are extracted. These features are used for classifying the brain MR image as normal and abnormal. After this classification segmentation is performed on the images for tumor region extraction.

Image segmentation is a process of partitioning an image into different homogeneous regions, so that meaningful information about the image can be obtained and different analysis can be performed on that segmented image. Extraction of brain tumor region requires the segmentation of brain MR images into two segments. One segment contains the normal brain cells and the second segment contains the tumorous cells of the brain. This segmentation is binary in nature. Correct segmentation of MR images is very important because most of the time MR images are not highly contrast thereby these segments can be easily overlapped with each other.

Proposed system combines feature extraction techniques with classification and segmentation techniques for the diagnosis of the brain as normal or tumorous and for extraction of tumor part from the tumorous brain images. Texture feature extraction technique [2][3] are used for the purpose of feature extraction from brain MR images. First order histogram based features and concurrence matrix based features [2] are extracted. These are used independently and combined for classification. Proposed system used

ensemble base classifier [4] for the classification of brain images as normal or tumorous. Last phase of the proposed system is segmentation. Segmentation is performed in two steps. First step removes the skull part from the brain MR image using [5]. After brain portion extraction, tumor region is extracted using FCM [6].

Our proposed technique is fully automatic and robust. No prior knowledge of the image is required about its feature, contents, type and model. Proposed system is very accurate system for diagnosing the brain tumor.

The paper is organized as follows: Related work is represented in section 2. Details of the proposed method are described in section 3. Section 4 contains details of experimentation and results. Conclusion and future work is presented in section 5.

2 Related Work

For classification of brain images as normal and abnormal El-Syed et al [7] proposed a hybrid technique. In the technique[7] first features of the brain MR image are extracted using Discrete Wavelet Transform (DWT) which are then reduced using PCA. Last step is classification based on these features. Two types of classifiers are used for classification one is feed forward back propagation neural network and second is K-nearest neighbors. Maximum 98.6% accuracy is achieved by [7].

For MRI brain image segmentation Sripama [8] proposed fuzzy symmetry based genetic clustering technique. In [8] numbers of clusters are evolved by variable length genetic fuzzy clustering technique. For measuring quality of cluster fuzzy point symmetry based cluster validity index is proposed in this paper. Experiments are performed on different T1, T2 and PD brain images. This technique performs better than FCM and Expectation Maximization algorithm. But this technique does not consider spatial information and sometime does not segment brain image correctly. This technique also does not work properly for the data sets which have same point as a center for different clusters. Zhang proposed a Hidden Markov Random Field Model and the Expectation-Maximization algorithm for segmentation of brain MR images [9]. This is a fully automatic technique for brain MR images segmentation. This method is based on estimation of threshold that is heuristic in nature. Thus most of the time, this method does not produce accurate results. Method in [9] is also computationally very expensive.

Fuzzy c-means (FCM) clustering [6] is an unsupervised system that has been effectively applied in fields such as astronomy, geology, medical imaging, target recognition, and image segmentation for clustering, classifier designs and feature analysis.

Cluster Validity is the procedure of evaluating quantitatively the results of a clustering algorithm. Bezdek proposed two cluster validity indexes [10], the partition coefficient (*PC*) and partition entropy (*PE*), defined as

$$PC = \frac{1}{n} \sum_{j=1}^n \sum_{i=1}^c u_{i,j}^2 \quad (1)$$

$$PE = -\frac{1}{n} \sum_{j=1}^n \sum_{i=1}^c u_{i,j} \log_a(u_{i,j}) \quad (2)$$

PC and *PE* is used to measure the fuzziness of the fuzzy partition matrix, the lower the fuzziness of a partition is, the larger the *PC* value (or the smaller the *PE* value). The two indexes use only the membership values of the fuzzy partition; therefore, they are devoid of connection to the structure of the data set.

3 Proposed Method

The proposed system consists of multiple phases. Fig. 1 shows the details of the proposed system. First MR image for diagnosis is provided to the system as an input. Second step of the proposed system is to extract features from this input image. Texture feature extraction [2] method is used for extracting features from the MR image. After feature extraction, these features independently are used for classification as malignant and benign MR image. Ensemble base classifier[4] is used for classification. It classifies the brain image on the bases of multiple classifiers. No more processing is required once the MR image is determined as benign.

But when the MR image is determined as malignant by the classifier it is further processed for extracting tumor portion from it. For this purpose segmentation is performed on this MR image in two steps. First skull gets removed from this MR images using[5] and brain portion is extracted. After brain portion extraction tumor region is extracted from this using [6]. Following are the details of the proposed system:

3.1 Feature Extraction

The transformation of an image into its set of features is known as feature extraction. Useful features of the image are extracted from the image for classification purpose. It is a challenging task to extract good feature set for classification. There are many techniques for feature extraction e.g. texture Features[2][3], gabor features[11], feature based on wavelet transform[12], principal component analysis, minimum noise fraction transform, discriminant analysis, decision boundary feature extraction, non-parametric weighted feature

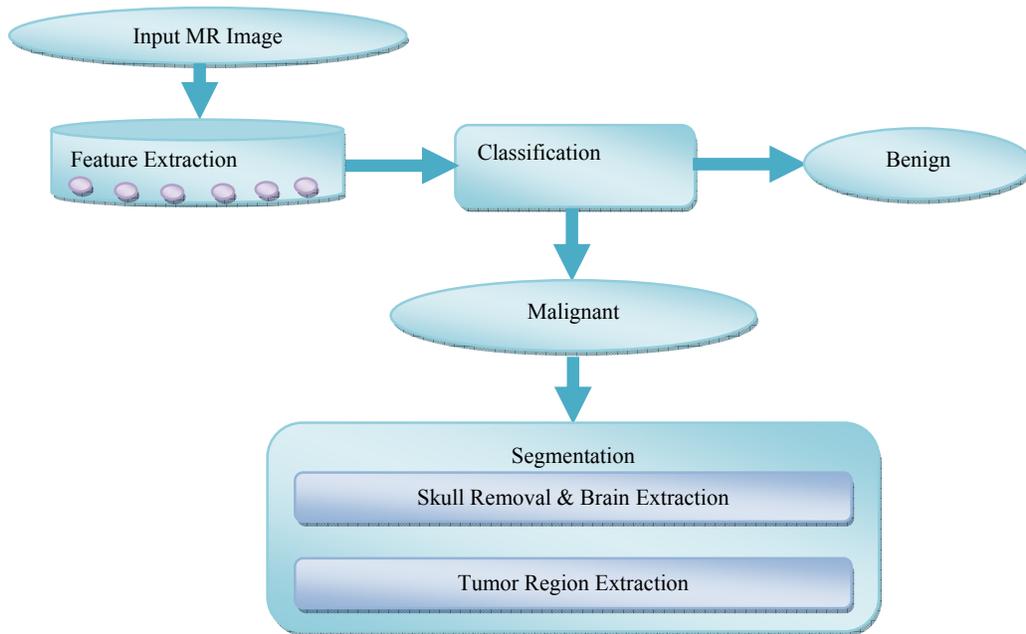


Fig. 1 Proposed System

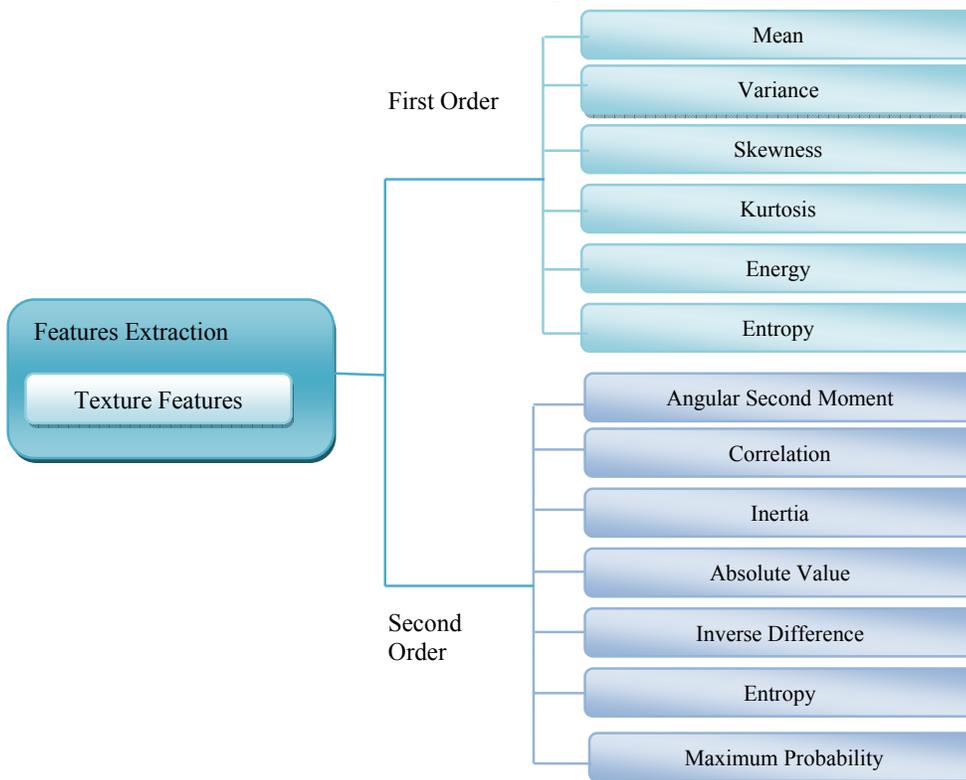


Fig. 2 Feature Extraction

extraction and spectral mixture analysis [13]. We are using texture feature for our proposed system.

3.1.1 Texture Features

Proposed system used two methods for extracting the texture features. First method is based on first order

histogram which is local in nature and second method is based on co-occurrence matrix which is called as second order texture feature. Fig.2 shows the details of the feature extraction process.

3.1.1.1 First-order histogram based features

Histogram of the image gives summary of the statistical information about the image. So first order statistical information of the image can be obtained using histogram of the image. Probability density of occurrence of the intensity levels can be obtained by dividing the value of intensity level histogram with total number of pixels in the image.

$$P(i) = h(i)/NM, \quad i=0,1,..G-1 \quad (3)$$

Where N is number of the resolution cells in the horizontal spatial domain and M is the number of resolution cells in the vertical spatial domain. G is the total gray level of an image.

For quantitatively describing the first order statistical features of the image, useful features of the image can be obtained from the histogram. Mean is the average value of the intensity of the image. Variance tells the intensity variation around the mean. Skewness is the measure which tells the symmetry of the histogram around the mean. Kurtosis is the flatness of the histogram. Uniformity of the histogram is represented by the entropy. Following is the list of features obtained using histogram of the image [2].

$$\text{Mean:} \quad \mu = \sum_{i=0}^{G-1} ip(i) \quad (4)$$

$$\text{Variance:} \quad \sigma^2 = \sum_{i=0}^{G-1} (i - \mu)^2 p(i) \quad (5)$$

$$\text{Skewness:} \quad \mu_3 = \sigma^{-3} \sum_{i=0}^{G-1} (i - \mu)^3 p(i) \quad (6)$$

$$\text{Kurtosis:} \quad \mu_4 = \sigma^{-4} \sum_{i=0}^{G-1} (i - \mu)^4 p(i) - 3 \quad (7)$$

$$\text{Energy:} \quad E = \sum_{i=0}^{G-1} [p(i)]^2 \quad (8)$$

$$\text{Entropy:} \quad H = - \sum_{i=0}^{G-1} p(i) \log_2 [p(i)] \quad (9)$$

3.1.1.2 Co-occurrence matrix based features

Histogram based features are local in nature. These features does not consider spatial information into consideration. So for this purpose gray-level spatial co-occurrence matrix $h_d(i,j)$ based features are defined which are known as second order histogram based features. These features are based on the joint probability distribution of pairs of pixels. Distance d and angle θ within a given neighborhood are used for calculation of joint probability distribution between pixels [2]. Normally $d=1,2$ and $\theta=0^\circ, 45^\circ, 90^\circ, 135^\circ$ are used for calculation. Co-occurrence matrix calculation is illustrated in Fig. 3 for $d=1$. Texture features can be described using this co-occurrence matrix. Following equations define these features [2].

Angular second moment (energy):

$$\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} [p(i,j)]^2 \quad (10)$$

$$\text{Correlation:} \quad \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{ij p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (11)$$

$$\text{Inertia:} \quad \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i-j)^2 p(i,j) \quad (12)$$

$$\text{Absolute Value:} \quad \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} |i-j| p(i,j) \quad (13)$$

$$\text{Inverse Difference:} \quad \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{p(i,j)}{1+(i-j)^2} \quad (14)$$

$$\text{Entropy:} \quad H = - \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} p(i,j) \log_2 [p(i,j)] \quad (15)$$

$$\text{Maximum Probability:} \quad \max_{i,j} p(i,j) \quad (16)$$

3.2 Classification

Classification is the procedure for classifying the input patterns into analogous classes. Selection of a suitable classifier requires consideration of many factors:

- Classification accuracy
- Algorithm performance
- Computational resources

There are basically two types of classification. One is known as unsupervised classification and other is known as supervised classification [13]. Unsupervised classification is the identification of natural groups, or structures, within multi-spectral data. The following characteristics apply to an unsupervised classification

- No extensive prior knowledge of the region is required
- Many of the detailed decisions required for supervised classification are not required for unsupervised classification creating less opportunity for the operator to make errors.
- Unsupervised classification allows unique classes to be recognized as distinct units.

Whereas supervised classification is the process of using samples of known identity to classify samples of unknown identity. The following characteristics apply to a supervised classification:

- Requires detailed knowledge of the area.
- Input patterns are provided with the labels
- Able to detect serious errors by examining training data to determine whether they have been correctly classified.

Proposed system used Ensemble base classifier [4] for the classification purpose. This Ensemble base classifier uses support vector machine for classification. Next section describes the details of the classification

procedure followed by the proposed system.

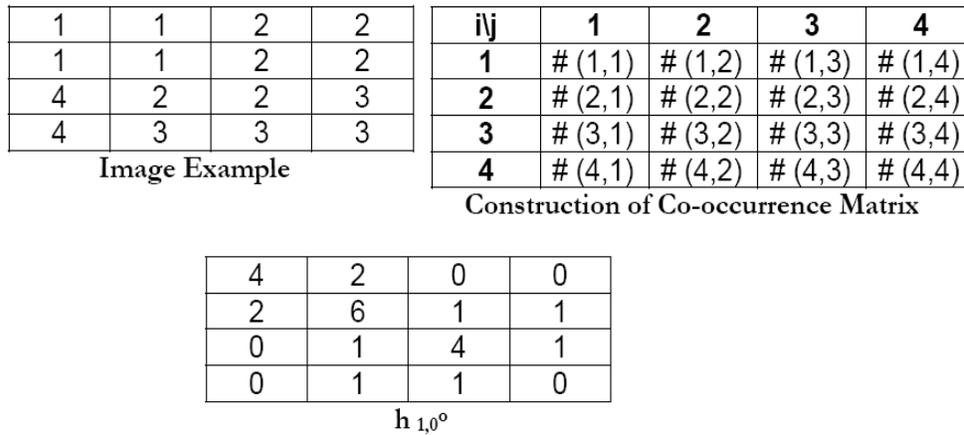


Fig. 3 The spatial co-occurrence matrix calculation

3.2.1 Support Vector Machine

Support vector machine (SVM) is one of the techniques used for the classification purpose. SVM has also been applied on different real world problems such as face recognition [14], text categorization [15], cancer diagnosis[16], glaucoma diagnosis, microarray gene expression data analysis. Proposed system used SVM for binary classification of brain MR image as normal or tumor affected. SVM basically tries to divide the given data into decision surface. Decision surface is a hyperplane which divides the data into two classes. Training points are the supporting vector which defines the hyperplane. The basic theme of SVM is to maximize the margins between two classes of the hyperplane.

3.2.2 Ensemble Base Classification

Ensemble base classifier is used for classification by the proposed system. This is a novel idea proposed by M. Arfan[4]. In this technique a number of classifiers are generated. Results of all these classifiers are merged together using algebraic combination rules. This methodology substantially improves the performance of the overall classification. SVM is used as the base classifier by the ensemble classifier used by the proposed system. Each classifier gives binary results . Cross validation method is used for dividing the dataset into training and testing data. On testing data Holdout testing is performed. For K-fold training and testing first data is arranged into different K-folds. And each fold is further divided into K blocks. Random overlap datasets are generated for each block. Training is performed for k-1 blocks of each fold. Remaining block is used for testing purpose. Fig. 4 show the ensemble base classification.

3.3 Segmentation

After the classification phase segmentation of the malignant images is performed into two steps.

3.3.1 Skull removal and brain region extraction

M.Stella and Blair proposed [5] a method which performs automatic segmentation of the Brain from MR images and this method performs segmentation even in the presence of RF inhomogenities. This method follows an integrated approach for segmentation which is based on prior knowledge, anisotropic filters and snake contouring techniques.

3.3.2 Brain Tumor Extraction

After brain region extraction brain tumor is extracted using FCM. For accurately finding the number of clusters, FCM algorithm is iterated for a range of hypothesized numbers of clusters. And best option for cluster is chosen based on cluster validity measure. Some cluster validity measures, the partition coefficient (PC) and the partition entropy (PE), Bezdek [6],[10] yielded amazingly good results for some of the test images. This might be partially due to the special nature of the data, which is not common in clustering problems: the data is 1-D and at least one entry exists at each possible point. On the other hand, accepting the fact that different values might fit the given data (e.g., for segmentation at different detail), a threshold on the validity measure should be chosen below/above which is accepted.

We have used a technique based upon fuzzy entropy that locates optimal and dynamic threshold according to

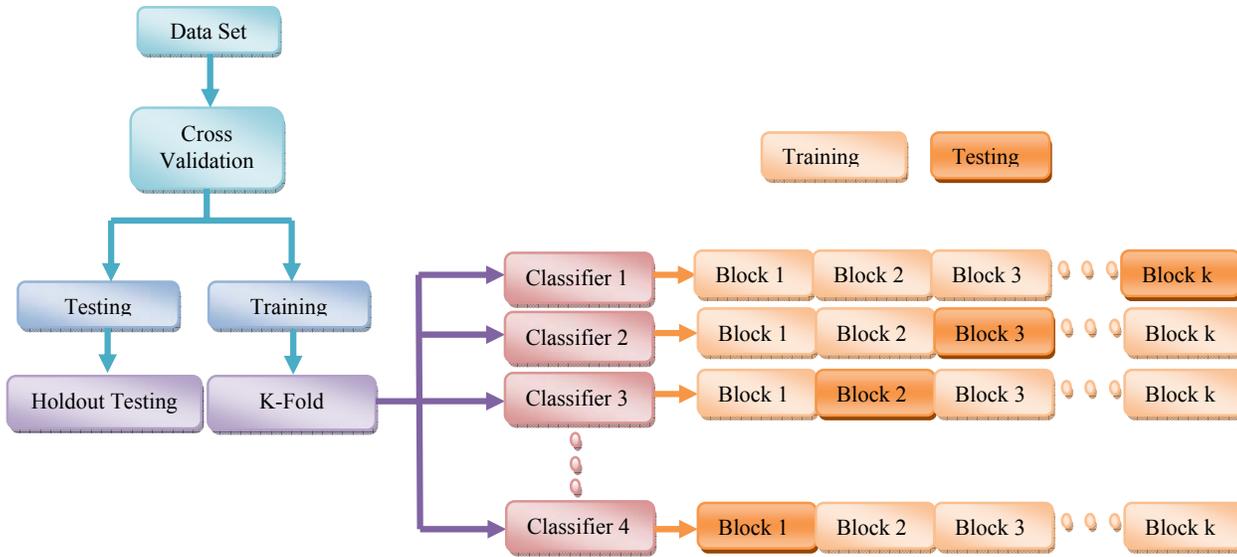


Fig. 4 Architecture of Ensemble Base Classifier

Technique	AANLIB Data Accuracy (%)	Real MRI Data Accuracy (%)
Texture Combined+ Ensemble Classifier	99.2	99.63
First Order + Ensemble Classifier	96.87	96.62
Second Order + Ensemble Classifier	98	97.13
Texture Combined+ SVM	97.82	97.19
First Order + SVM	96.2	96.53
Second Order + SVM	97	96.89
Texture Combined+ ANN	95.4	97.78
First Order + ANN	92.5	93.41
Second Order + ANN	91.42	92.22
DWT+ PCA+ Ensemble Classifier	99.54	98.64
DWT+PCA+ANN[7]	95.4	94.3
DWT+PCA+KNN[7]	98.2	97.5
DWT+SOM[7]	95.13	94.72
DWT+SVM[7]	97.25	96.64

Table1. Comparison of classification performance for the proposed technique and recently other work

the clusters find out by using eminent method of FCM (Fuzzy C-Mean). Input of this phase is fuzzy membership matrix from prior phase. It also gives an objective function that is used to find out adaptive, optimal and dynamic threshold which is used in the next phase to segment the image. We have used logarithmic fuzzy entropy based error functions [10] equation (4) that is depicted in Fig 3. The thresholds values are obtained from the objective function, as the gray levels with the maximal levels of fuzziness respectively.

Logarithmic Entropy

$$H(A) = -\frac{1}{n \ln 2} \sum_i \left[\mu_A(x_i) \ln \mu_A(x_i) + (1 - \mu_A(x_i)) \ln (1 - \mu_A(x_i)) \right] \quad (17)$$

4 Experimentation and Results

The proposed method has been implemented using the Matlab environment on Core 2 Duo, processor speed 1.6 GHz. The proposed system has been tested on the data set available at web [17]. It has also been tested on dataset of real brain MR images consisting of normal and tumorous brain images.

Table 1 shows the results of all experimentations. We first perform classification of the dataset. Once the classification is done, the malignant brain images are further segmented for extraction of tumor region from these brain MR images.

Performance of each classifier is measure in terms of confusion matrix, sensitivity, specificity and accuracy. Sensitivity is a measure which determines the probability of the results that are true positive such that person has the tumor.

Specificity is a measure which determines the probability of the results that are true negative such that person does not have the tumor.

Accuracy is a measure which determines the probability that how much results are accurately classified.

For classification purpose, first each classifier is tested using first order feature set. Performance of each classifier is measured for this feature set. Fig.5 shows the performance of each classifier when first order histogram base feature set is used for classification. Results show that Ensemble base classifier gives higher accuracy as compared to other classifiers. Second order texture feature set is also tested against each classifier. Fig.6 shows the performance of all classifier using second order feature set. When we combined the first order and second order texture feature and perform

classification on these combined features set then the results improve tremendously. Fig.7 shows the performance of each classifier against combined texture feature set. By comparing out results with the other recently work [7], for each case ensemble base classifier gives the higher accuracy. It also gives very accurate sensitivity and specificity rates as compared to [7].

We also tested the performance of ensemble base classifier for the feature extracted by using the DWT and then reduced using PCA. Ensemble base classifier also gives higher accuracy in this case also as compared to the recently proposed technique [7].

Classification phase provides the images which are malignant to the segmentation phase as input. Segmentation phase of the proposed system quite accurately extract the tumor region from these malignant brain images. Fig. 9 shows the results of segmentation phase. All images show that tumor region which is very high in contrast is quite accurately identified and extracted by the proposed system.

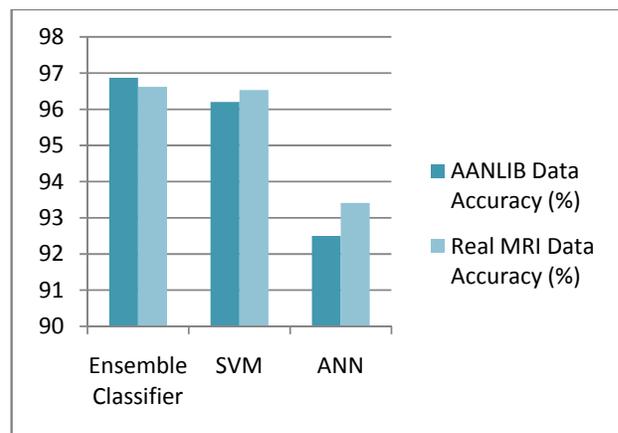


Fig.5 Graph shows each classifier's performance using First Order Texture Features.

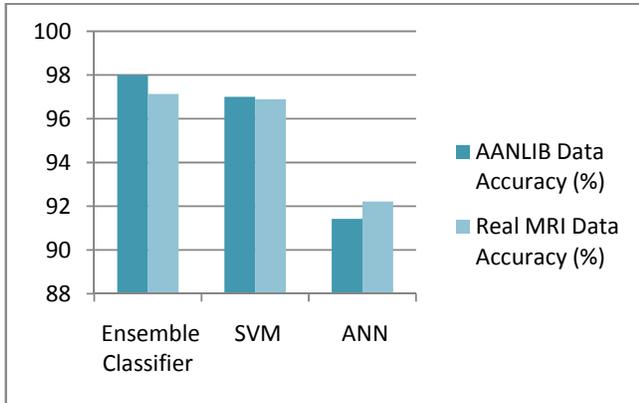


Fig. 6 Graph shows each classifier’s performance using Second Order Texture Features.

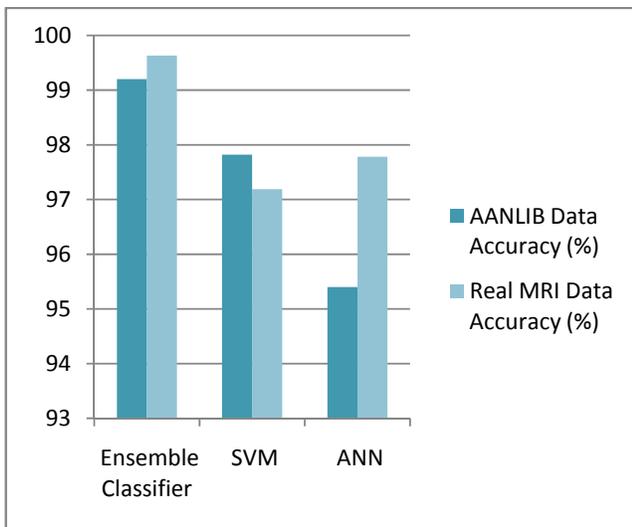


Fig. 7 Graph shows each classifier’s performance using Combined Texture Features

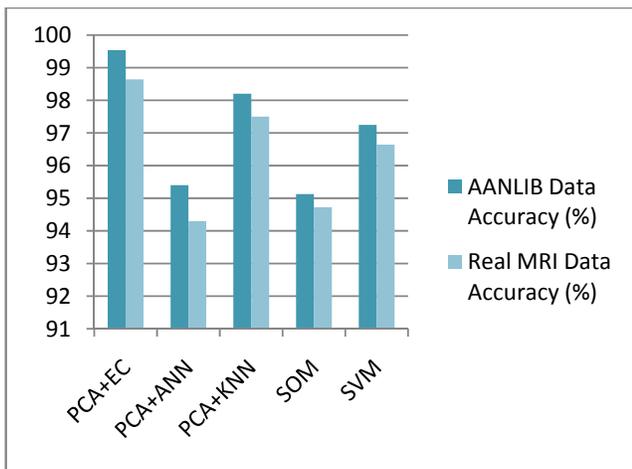


Fig. 8 Graph shows each classifier’s performance using DWT features, EC represents Ensemble base Classifier

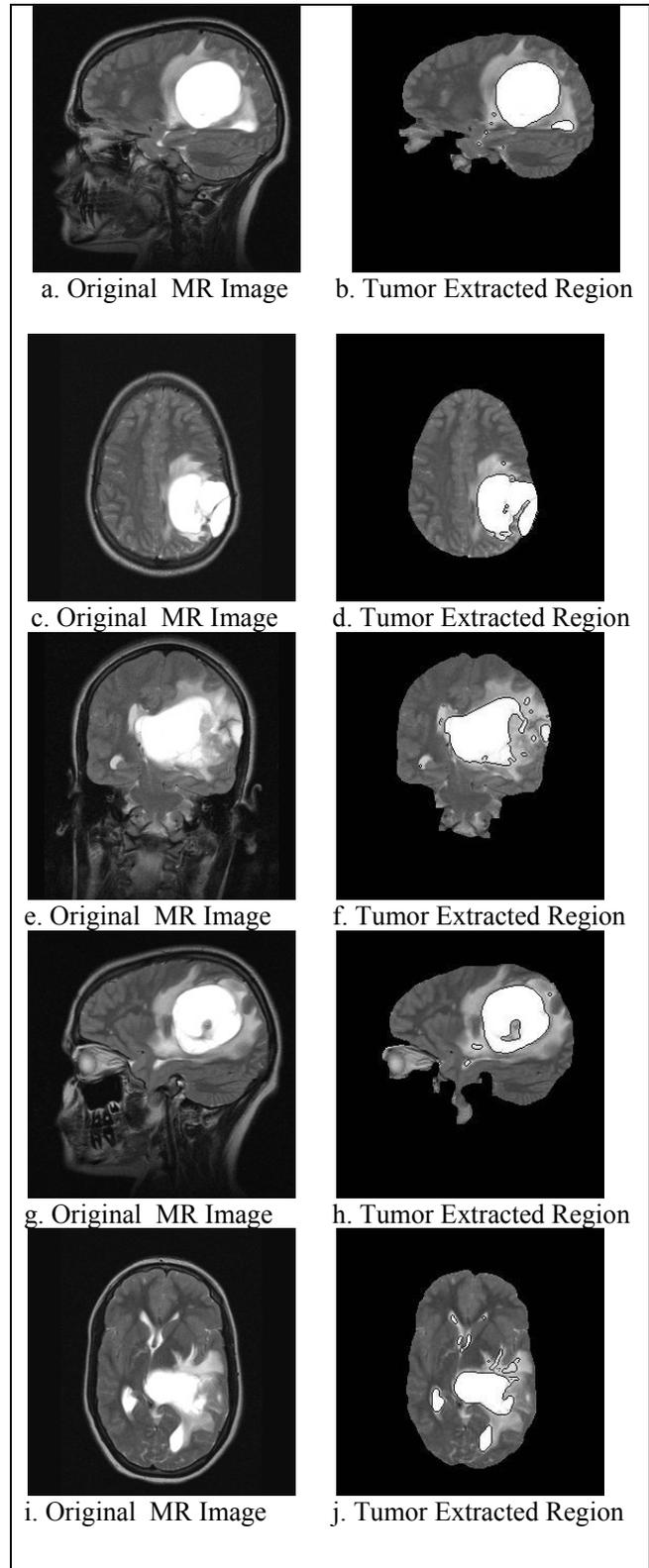


Fig. 9 Tumor region is extracted from the original malignant brain MR image

5 Conclusion

Proposed system is developed for diagnosing the brain tumor from brain MR images. This system performs this diagnosis in multiple phases. First phase of diagnosis is texture feature extraction which consists of first order and second order texture feature extraction. These extracted features are used for classification. In classification phase proposed system used ensemble base classifier for classifying brain images as benign and malignant. Once the images are determined as malignant these are further processed for tumor extraction from them. Tumor extraction is performed in the segmentation phase. Segmentation phase is a multi step phase. It first removes the skull part of the brain and then extracts the tumor region using FCM and gives us the boundary of the tumor.

All experiments show that the proposed system gives exceptionally good results as compared to the recently proposed techniques. We achieved accuracy of classification more than then 99%. We also achieved very accurate results of segmentation which effectively extract the tumor region from brain MR images.

Future work of this research is to measure the thickness and area of the tumor extracted region.

Acknowledgement

The authors would like to thank Dr. Usman for providing the brain MR scans from Abrar MRI center Rawalpindi. The authors would also like to thank Dr. Tauqeer a radiologist of Holy Family Hospital Rawalpindi for helping us in understanding the brain MR images interpretation.

References:

- [1] Milan Sonka, Satish K. Tadikonda, Steve M. Collins, Knowledge-Based Interpretation of MR Brain Images, *IEEE Transaction on Medical Imaging*, Vol. 15, No. 4, 1996, pp. 443-452.
- [2] Andrzej & Michal, Texture Analysis Methods – A Review, Technical University of Lodz, Institute of Electronics, COST B11 report, Brussels 1998
- [3] Haralick, R.M., K. Shanmugan, and I. Dinstein, Textural Features for Image Classification, *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-3, 1973, pp. 610-621
- [4] M. Arfan Jaffar, Zeshan Hayder, Ayyaz Hussain and Anwar M. Mirza, An Intelligent Ensemble Based Systems for Breast Cancer Diagnosis, *Proceedings of 2009 International Conference on Computer Engineering and Applications, Philippine*, 2009, pp. xx-xx
- [5] M. Stella & Blair, Fully Automatic Segmentation of the Brain in MRI, *IEEE transactions on medical imaging*, VOL. 17, NO. 1, 1998, pp. 98-107.
- [6] J. C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms New York: Plenum Press, 1981.
- [7] El-Sayed A, El-Dahshan, Abdel-Badeeh M. Salem and Tamer H. Younis, A hybrid technique for automatic MRI brain images classification, *Studia Univ, Babeş Bolyai, Informatica*, Vol LIV, 2009
- [8] Sriparna Saha and Sanghamitra Bandyopadhyay, MRI Brain Image Segmentation by Fuzzy Symmetry Based Genetic Clustering Technique, *Evolutionary Computation*, 2007, pp. 4417-4424
- [9] Yongyue Zhang, Michael Brady, and Stephen Smith, Segmentation of Brain MR Images Through a Hidden Markov Random Field Model and the Expectation-Maximization Algorithm, *IEEE Transaction on Medical Imaging*, vol. 20, No. 1, 2001, pp. 45-57
- [10] V. Boskovitz, Hugo Guterman, An adaptive neuro fuzzy system for automatic image segmentation and edge detection, *IEEE Transaction on fuzzy systems*, 2002, pp. 247-262.
- [11] Liu & Wechsler, Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition, *IEEE Trans. Image Processing*, Vol. 11, 2002, pp. 467-476.
- [12] M. Kociołek, A. Materka, M. Strzelecki P. Szczypiński, Discrete wavelet transform-derived features for digital image texture analysis, *Proc. of International Conference on Signals and Electronic Systems*, Lodz, Poland, 2001, pp. 163-168.
- [13] D. LU & Q. WENG, A survey of image classification methods and techniques for improving classification performance, *International Journal of Remote Sensing*, Vol. 28, No. 5, 2007, pp 823-870
- [14] Jennifer Huang, Volker Blanz, and Bernd Heisele, Face Recognition Using Component-Based SVM Classification and Morphable Models, *LNCS 2388*, 2002, pp. 334-341.
- [15] Joachims, T. Text categorization with support vector machines. Technical report, LS VIII Number 23, University of Dortmund, 1997.
- [16] T.S. Fury, N. Cristianini, N. Duffy, Support vector machine classification and validation of cancer tissue samples using microarray expression data, *Proc of BioInformatics*, Vol. 16, No. 10, 2000, pp. 906-914
- [17] Harvard Medical School, Web: data available at <http://med.harvard.edu/AANLIB/>.