An Intelligent Classification System for Diagnosing MRI Brain Images using Modified FCM and SVM

R.Somas Kandan
Assistant Professor
K.L.N. College of Engineering
Pottapalayam-630612, India
somaskandan.16@gmail.com

Dr.A.John Sanjeev Kumar
Associate Professor
Thiagarajar College of Engineering,
Madurai-625015, India.
ajscse@tce.edu

Abstract

Automated medical image analysis for identifying tumor at the earlier stage is essential to save human life. Medical Image segmentation and classification plays an important role in the characterization of tumor. The accuracy in recognizing it is highly essential since the treatment planning is based on its identification. A new hybrid technique based on modified Fuzzy C-Means segmentation algorithm and classification based on support vector machine is proposed. The proposed technique consists of four stages mainly segmentation, feature extraction, feature selection and classification. Modified Fuzzy C-means (FCM) algorithm which has an improved computation rate, modified cluster center and updating membership value criterion is used. Several statistical features are extracted to yield a better performance for classification techniques. Feature extraction stage extracts a set of 14 features using GLCM (Gray Level Co-occurrence Matrix). To select the discriminative features among them, Sequential Forward Selection (SFS) algorithm is used. Support Vector Machine (SVM) classifier is used to classify the MR brain images into benign or malignant. Receiver Operating Characteristic (ROC) curve analysis is done for calculating the misclassification rate. The experiment result of proposed system achieves high classification accuracy whose effectiveness is measured in terms of sensitivity and specificity.

Keywords: Modified fuzzy c-mean (FCM), Gray Level Co-occurrence Matrix (GLCM), Sequential Forward Selection (SFS), Support Vector Machine (SVM), and Receiver Operating Characteristic (ROC).

I. INTRODUCTION

Magnetic resonance imaging is used as a valuable tool in the clinical and surgical environment because of its characteristics like superior soft tissue differentiation, high spatial resolution and contrast [1].It does not use harmful ionizing radiation to patients.MRI is efficient as compared with all other imaging techniques in the application of brain tumor detection and identification.[2]

There exists two types of tumors namely Benign and Malignant brain tumors. Benign brain tumors have a homogeneous structure which did not contain cancer cells and they may be either radio logically monitored or completely removed surgically and they do not persist again. The structure of malignant brain tumors is heterogeneous and it contains cancer cells which can be treated with radiotherapy, chemotherapy or a combination thereof, and they are life threatening[3]. Therefore automated and early diagnosis of tumor with MR images is very crucial for saving the valuable life of human. The automated diagnosis involves two major steps: (a) Image segmentation (b) Image classification.

In automatic analysis for brain MR images, segmentation algorithms using computer vision and pattern recognition plays an important role. Accurate segmentation of brain magnetic resonance (MR) images results in three main tissues: grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF) that are fundamental in brain disease diagnosis.

Fuzzy C-means method proposed by Dunn [4] is one of the most widely used clustering methods for image segmentation. It allows the clustering procedures that maintain more information from image than hard clustering methods such as K-means [5] and obtain more accurate results

But FCM method is sensitive to noise. Hence a spatial penalty term is added based on idea that membership values are supposed to be homogenous in a small neighborhood [6]. Inspired by neighborhood expectation maximum (EM) algorithm, Yang [7] introduced another penalty term with respect to fuzzy membership values in neighborhoods. The similar attempt was made by Krinidis and Chatzis [8] in which the penalty term contains not only neighborhood's memberships but also the grey levels.

In [9], spatial information is introduced to compensate for the effect of noise. In [10,11] an objective function is formed to compensate for the grey in homogeneity. In this kind of models, the fuzzy membership value of a pixel is allowed to be influenced by its neighborhood. This approach is named as FCM with spatial constraints (FCM-S).Hence it is noted that there should be a balance between clear cluster edge and robustness to noise. This tradeoff is controlled by some parameters that effects segmentation results.

A generalized FCM clustering algorithm with local information (GFCMLI)is proposed that not only mitigates the disadvantages of standard FCM, but also highly improves the overall clustering performance[12].

In order to obtain more robust and accurate results, a hierarchical strategy is used to construct a more flexibility function, which considers the improved distance function itself as a sub-FCM [13].

Texture features, wavelet transform, and SVM algorithm is used for efficient classification in [14,15]. The proposed technique in [15] is better for tumor detection using first order statistical features .

Texture analysis is proposed to overcome the problem by finding the underlying characteristics of textures. Feature extraction methods [16] are categorized as follows: (1) structural methods, (2) statistical methods, (3) model-based methods; and (4) transform-based methods. Arguably, statistical methods may be the most suitable for characterizing tissues that have random, non-homogeneous structures, such as brain tissues, whose MR images show no apparent regularities. In addition, statistical textural features achieve better discrimination with same classifiers by far less number of relevant but distinguishable features in comparison to other methods of structural approach or wavelet transformation [17]. The most widely used statistical textural features are based on the Gray-Level Co-occurrence Matrix (GLCM) introduced by Haralick et al. [18].He extracted 14 features based on the GLCM. This avoids incorporating highly correlated data into the classification scheme.

Next feature selection is performed to reduce the redundancy. Feature selection is the technique of selecting a subset of relevant features for building robust learning models by removing most irrelevant and redundant features from the data. Feature selection enhances the execution of learning models by easing the impact of the scourge of dimensionality, enhancing generalization capability, speeding up learning process, improving model interpretability [19]. The goal of feature selection is two-fold: to improve the computational efficiency and to reduce the generalization error of the model by removing irrelevant features or noise. Here sequential forward selection (SFS) method is used for selecting discriminative features.

The important process in the automated system is brain image classification. In image classification based on the discriminative feature selection MRI brain images are classification into benign or malignant tumor images. Support Vector Machine is used for classifying benign and malignant brain tumors. The basic idea of applying SVMs for solving classification problems can be stated briefly as follows: a) Transform the input space to higher dimension feature space through a non-linear mapping function and b) Construct the separating hyper plane with maximum distance from the closest points of the training set.

In [20], Adaptive thresholding technique is proposed to segment the tumor's region. The rule based classifier was used to classify four types of lesions. The accuracy of classification obtained from this method was 93%, 73%, 84% and 60% for acute stroke, solid tumor, chronic stroke, and for necrosis respectively.

In [21], Principle component analysis is used a feature extraction method and tumor was classified to Benign and Malignant by using Supervisor classifier based Fuzzy Support Vector Machine. The accuracy of Classification was 95.80%.

II. PROPOSED METHOD

In the proposed method, segmentation is performed using modified FCM algorithm, feature extraction using GLCM ,feature selection using Sequential forward selection and classification using Support Vector Machine and ROC Curve analysis is done for calculating misclassification rate. The overall system design of the proposed method is illustrated in Fig.1.

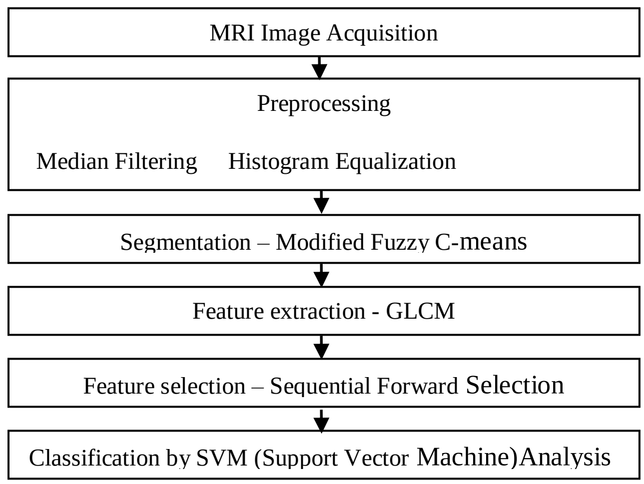
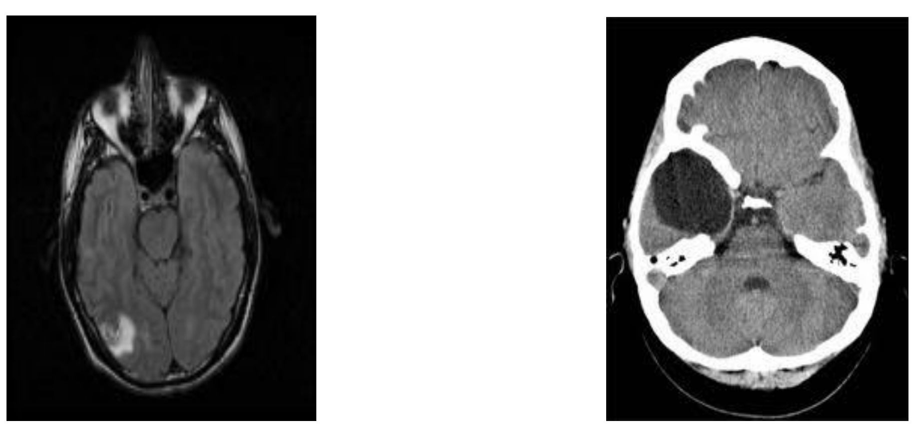


Fig.1 Flow diagram of Proposed Methodology

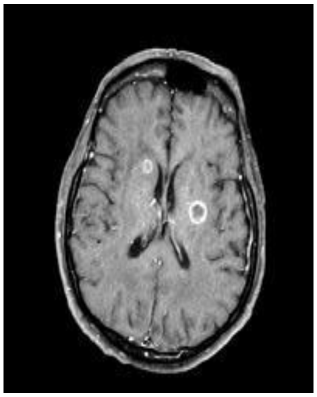
A. Image Acquisition

Image acquisition method has been implemented on real data for human MR images data set. The MR images are collected from the radiologist (Fig.2 and Fig.3).

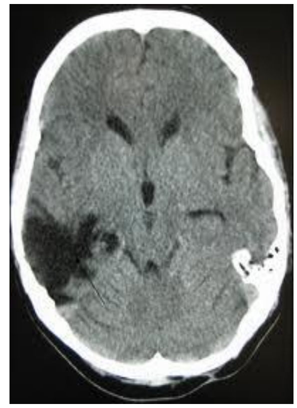


b) Benign c**) Malignant**

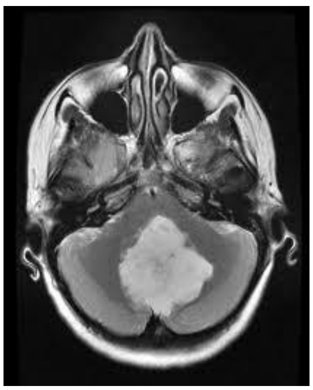
Fig.2 Sample Images



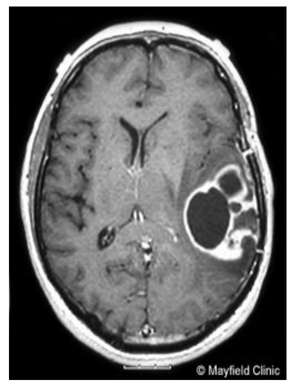
Test image 1



Train image 1



Test image 2



Train image 2

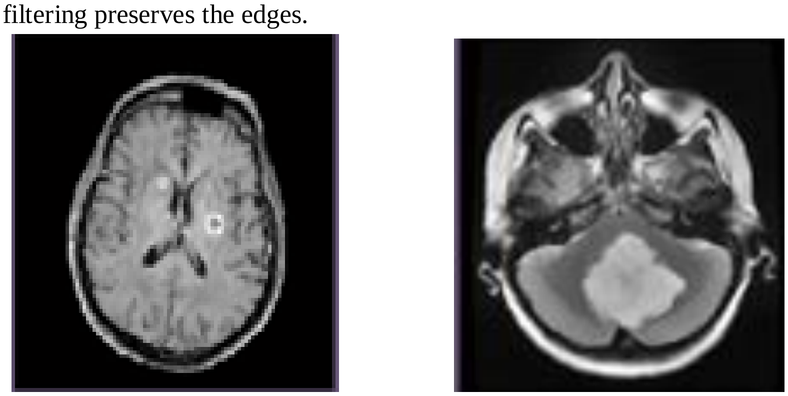
Fig.3 Samples of testing and training data sets

B. Preprocessing

Initially the input image is given to a median filter in order to reduce the noise and obtained output is histogram equalized to enhance the contrast

1) Median filtering

In the median filtering operation, the pixel values in the neighborhood window are ranked according to intensity, and the middle value (the median) becomes the output value for the pixel under evaluation. When Compared to the mean filter, median filter is less sensitive to extreme value. Therefore the extreme values are more effectively removed. Median filtering preserves the edges.



(a) (b)

Fig.4 Median Filtering images (a) Benign image (b) Malignant image

2) Histogram equalization

It enhances the contrast of images by transforming the values in an intensity image, or the values in the color map of an indexed image, so that the histogram of the output image approximately matches a specified histogram.

(a) (b)

Fig.5 Histogram equalization (a) Benign (b) Malignant

III. SEGMENTATION

A. FCM clustering

In FCM clustering , pixels are allocated to each class by using fuzzy memberships. An Image is divided into c clusters where the dimension of the images is N pixels. This image is represented by X=(x1, x2,…,xn) where xi represents multispectral (features) data. Here the optimization is a minimization problem whose objective function defined as

$$J=\sum_{i=1}^{N}\sum_{j=1}^{N} u_{ij}^{m} \parallel x_j - v_i \parallel^2 \quad (1)$$

The fuzzyness of the resulting is controlled by the parameter m and the value of m is assigned as two. When the neighboring pixels have similar features, the probability of getting same cluster is comparatively high. In clustering spatial information plays an important role but the standard FCM algorithm does not fully utilize it.

To exploit the spatial information, a modified membership function is used and it is defined as,

$$u_{ij} = \frac{u_{ij}^{m} s_{ij}^{n}}{\sum_{k=1}^{c} u_{kj}^{m} s_{kj}^{n}} \quad (2)$$

Where $S_{ij} = \sum_{k \in N(x_j)} u_{ik}$

is called spatial function, and N(xj) represents a square window centered on pixel xj in the spatial domain. The Spatial functions in two homogeneous region tends to improve the original membership, and the clustering result remains unchanged. To correct the misclassified pixels from noisy regions , two steps are followed in FCM algorithm.

The first step is to calculate the membership function in the spectral domain and the second step is to map the membership information of each pixel to the spatial domain and then compute the spatial function from that. The iteration proceeds with the new membership function that is incorporated with the spatial function. The iteration is stopped when the maximum difference between two cluster centroids is less than a threshold value (=0.01).

B. The Modified FCM algorithm (MFCM)

Step 1: Set the number of clusters c and the parameter m.

$$J_m(u,v) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{m} d^2\,(x_j, v_i) \qquad (3)$$

Initialize the fuzzy Cluster centroid vector

$\boldsymbol{V} = [\boldsymbol{v_1}, \ldots\ldots\ldots, \boldsymbol{v_c}]$ randomly and set $\varepsilon = 0.01$

Step 2: compute u_{ij}

$$u_{ij} = \left[\sum_{k=1}^{c}\left[\frac{d(x_j, v_i}{d(x_j, v_k)}\right]^{2/(m-1)}\right]^{-1} \qquad (4)$$

Step 3: compute v_i

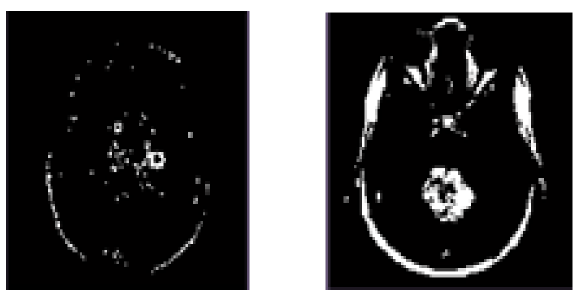
$$v_i = \frac{\sum_{j=1}^{n} u_{ij}^{m} x_j}{\sum_{j=1}^{n} u_{ij}^{m}} \qquad (5)$$

Step 4: update u_{ij}

Step 5: update v_i

Repeat Steps 4 and 5 until the following termination criterion is satisfied:

$|v_{new} - v_{old}| < \epsilon$



(a) (b)

Fig.6 Segmented Image (a) Benign (b) Malignant

C. Feature extraction

The procedure for extracting textural properties of image in the spatial domain was presented by Haralick. The gray level Co-occurrence matrix method considers the spatial relationship between pixels of different gray levels.

The method calculates a GLCM by calculating how often a pixel with a certain intensity i occurs in relation with another pixel j at a certain distance d and orientation. Each element (i, j) in the GLCM is the sum of the number of times that the pixels with value i occurred in the specified relationship to a pixel with value j in the raw Image. Co-occurrence matrices are calculated for four directions: 0, 45, 90 and 135.

D. Feature selection

SFS algorithm is a bottom-up search procedure which starts from an empty feature set S and gradually adds features selected by some evaluation function that minimizes the mean square error (MSE). In each iteration, the feature to be included in the feature set, it is selected among the remaining available features of the feature set, which have not been added to the feature set. So, the new extended features set should produce a minimum classification error compared with the addition of any other feature.

Sequential Forward Selection

Starting from the empty set, sequentially add the feature x^+ that results in the highest objective function J (Y_k+x^+) when combined with the features Y_k that have already been selected.

1.Start with the empty set $Y_0 = \{Ø\}$

2.Select the next best feature $x^+ = \arg\max_{x€Y_k} [J(Y_k+x)]$

3.Update $Y_k+1=Y_k+x^+$; and increment the value of k and go to step 2.

SFS is widely used for its simplicity and speed.

E. Classification

Applying appropriate classification technique is the key success for image classification. The most effective classification technique is Support Vector Machines based on kernel methods, which is very efficient in solving intricate categorization issues in many kinds of application areas

SVM classifier

(i)Data setup: Dataset contains three classes (Normal, Benign and Malignant), each N samples. The data is 2D plot original data for visual inspection.

(ii)SVM with linear kernel of range (-t, 0) is taken in order to find the best parameter value C using 2-fold cross validation (50% of data is used for training and remaining set of data is used to test the image).

(iii)After finding the best parameter value for C, the entire data is again trained using this parameter value.

(iv)Plot support vectors which is close to the hyper plane.

(v) Plot decision area i.e. cases with one category of the target variable are on one side of the plane and cases with the other category are on the other side of the plane.

SVM maps input vectors to a higher dimensional vector space where an optimal hyper plane is constructed. Among the many hyper planes available, there is only one hyper plane that maximizes the distance between itself and the nearest data vectors of each category. This hyper plane which maximizes the margin is called the optimal separating hyper plane and the margin is defined as the sum of distances of the hyper plane to the closest training vectors of each category.

In the case of linear separable data, the SVM tries to find among all hyper planes that minimize the training error, the one that separates the training data with maximum distance from their closest points

$w \bullet x + b = 0$

with w and b weight and bias parameters respectively. In order to define the maximal margin hyperplane (MMH) the following constrains must be fulfilled:

Minimize $½ \|w\|^2$ with $y_i(w \bullet x_i + b = 0) \geq 1$

This is a classic nonlinear optimization problem with inequality constraints. It can be solved by the karush-kuhn-Tucker (KKT) theorem by introducing Lagrange multipliers maximize $\sum_{i=1}^{l} a_i - \frac{1}{2}\sum_{i,j=1}^{l} y_i y_j a_i a_j x_i^T x_j$

Subject to $\sum_{i=1}^{l} a_i y_i = 0$ and $a_i \geq 0$

The solution of w is:

$$w = \sum_{i=1}^{l} a_i y_i x_i$$

The only nonzero solutions define those training data (usually a small percentage of the initial data set) that are necessary to form the MMH and are called support vectors. The optimal hyper plane theory is generalized for non-linear overlapping data by the transformation of the input vectors into a higher dimensional feature space through a mapping function

$x_i \in R^n \rightarrow z(x) = [a_1\Phi_1(x), a_2\Phi_2(x), \ldots.. a_n\Phi_n(x)]^T \in R^f$

The KKT conditions transform to

Maximize $\sum_{i=1}^{l} a_i - \frac{1}{2}\sum_{i,j=1}^{l} y_i y_j a_i a_j \; K(x_i, x_j)$

Subject to $\sum_{i=1}^{l} a_i y_i = 0$ and $a_i \geq 0$

The optimization problem is solved using the MATLAB optimization toolbox

IV. RESULTS AND DISCUSSIONS

The implementation results and the performance analysis are discussed in this section, The proposed algorithm is implemented using Matlab R2013a. The MR brain images with tumor and non tumor images are obtained from the various MRI datasets .Totally 80 images, out of which 40 images are considered as tested images and remaining 40 images are considered as training images. Fig.4, Fig.5 show the preprocessed image using median filter and histogram equalization.

The segmentation process is done using Modified FCM algorithm where the dimensionality of the input is highly reduced. The segmented image output using Modified FCM algorithm is illustrated in Fig.6 and Fig.7. The segmented output is then analyzed using feature extraction.

The Feature extraction is used for dimensionality reduction where the reduced dataset is used to extract the relevant information. The statistical features include energy, entropy. GLCM matrix is constructed from which different statistical features are obtained. Here Sequential Forward Selection algorithm is used as an automatic feature selection algorithm.

The statistical texture features are extracted by GLCM matrix. The Fourteen features are extracted namely Contrast,Correlation,Correlation,ClusterShade,Dissimilarity,Energy, Entropy, Homogeneity, Homogeneity, Maximum probability, Sum of squares, Autocorrelation, Sum of average and Sum of variance. Feature selection uses sequential forward selection which is the simplest greedy search algorithm. This algorithm selects only predominant features for classification.

SVM Classifier is used to classify MR images benign and malignant images. The accuracy of classification is measured in terms of sensitivity and specificity (Fig.8 and Fig.9). The values of true positive(TP),true negative (TN), false positive (FP) , false negative (FN), sensitivity, specificity, accuracy for the proposed method are compared with the existing methods like Classifier Performance Comparison in tab.1. The comparison graph is also shown in Fig.10.

The accuracy results shows the proposed algorithm is the best choice for tumor classification

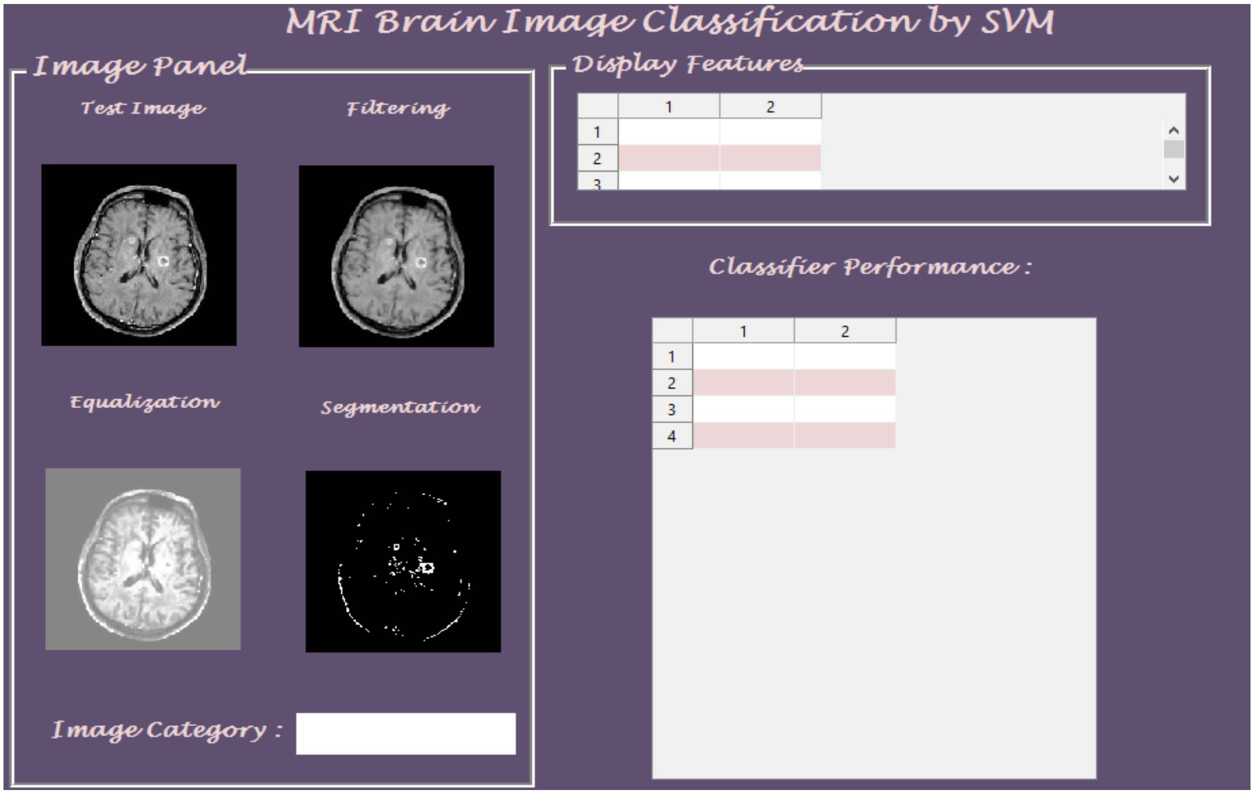


Fig.7 Segmented output for the given test image

The classification follows feature extraction stage. SVM Classifier is used such that with one category of the target variable are on one side of the plane and cases with the other category are on the other side of the plane. The vectors near the hyper plane are the support vectors which are shown in the Fig. 11.

4.1 Classifier Performance analysis: Accuracy and Mis-classification rate

The Accuracy (or Power) is the probability that the test correctly classifies the subjects; the Mis-classification rate is its complement to 1. It considers both the Precision (positive predictivity) and the Sensitivity of the test to compute the score: P is the number of correct results divided by the number of all returned results. S is the number of correct results divided by the number of results that should have been returned. The F1 score can be interpreted as a weighted average of the Precision and Sensitivity, where an F1 score reaches its best value at 1 and worst score at 0. Accuracy is a measure which determines the probability that how much results are accurately classified.

Sensitivity = TP/(TP+FN)
Specificity = TN/(TN+FP)
Accuracy = (TN+TP)/(TN+TP+FN+FP)
Where
TP : True Positive
TN : True Negative
FN : False Negative
FP : False Positive

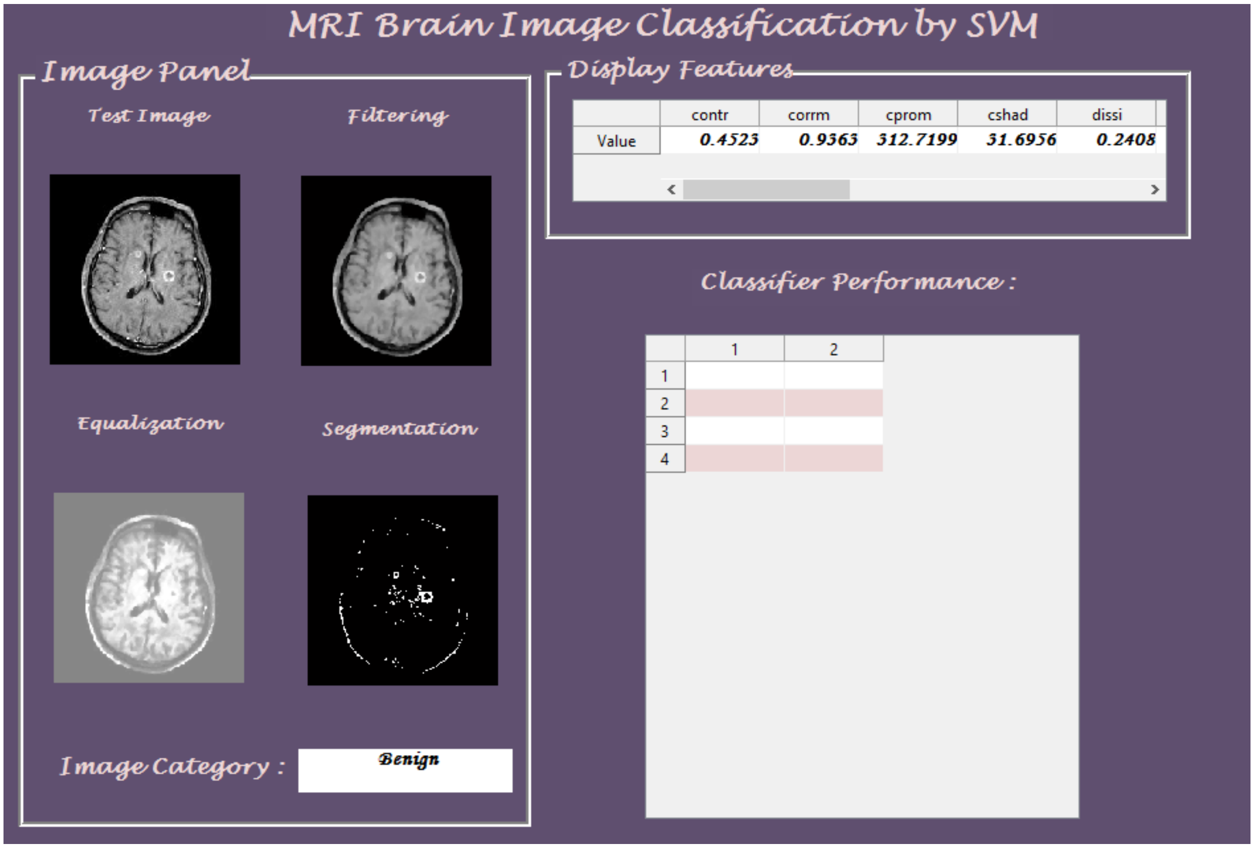


Fig. 8 Classified output for the given test image 1

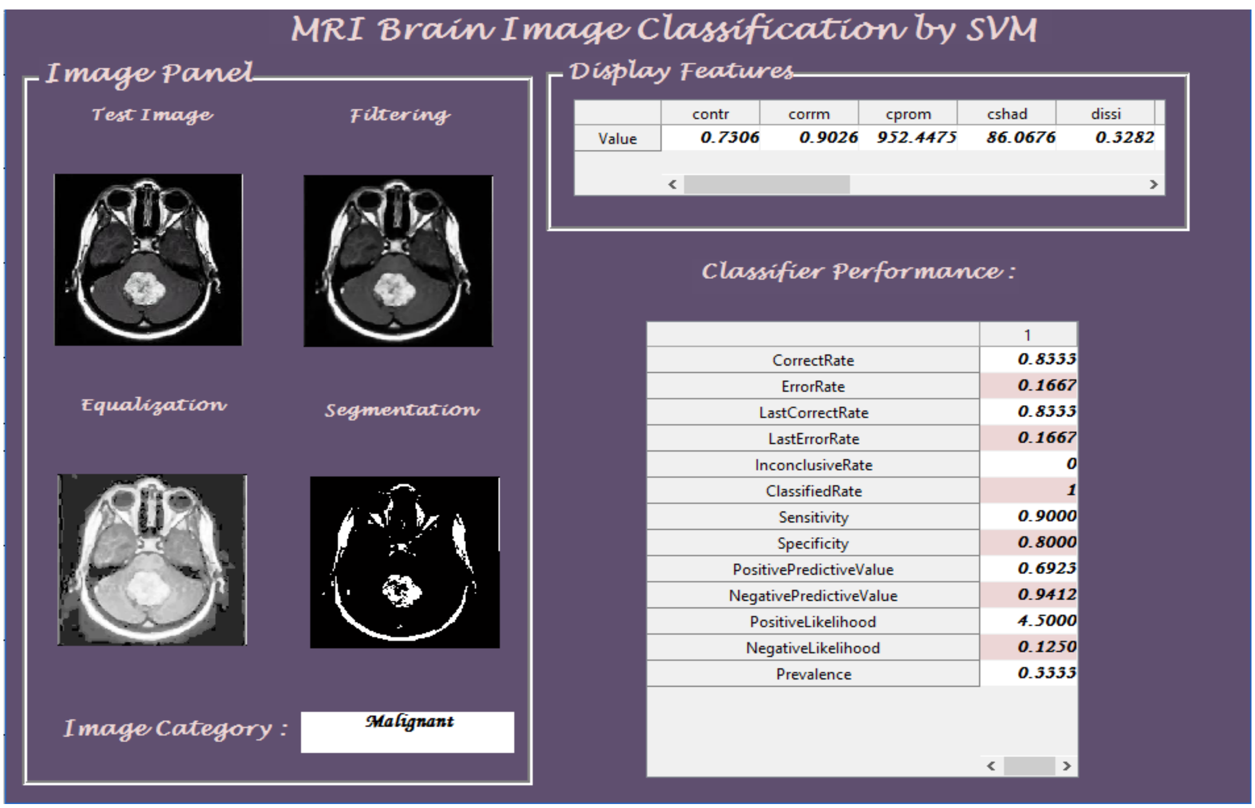


Fig. 9 Classified output for the given test image 2

TABLE 1. CLASSIFIER PERFORMANCE COMPARISON

Parameters	Texture feature method	Texture feature + FFNN	Proposed
True positive (TP)	22	20	25
False Positive (FP)	3	3	1
True Negative (TN)	11	11	13
False Negative (FN)	4	6	1
Sensitivity	0.85	0.77	0.96
Specificity	0.79	0.79	0.92
Accuracy	0.83	0.78	0.96

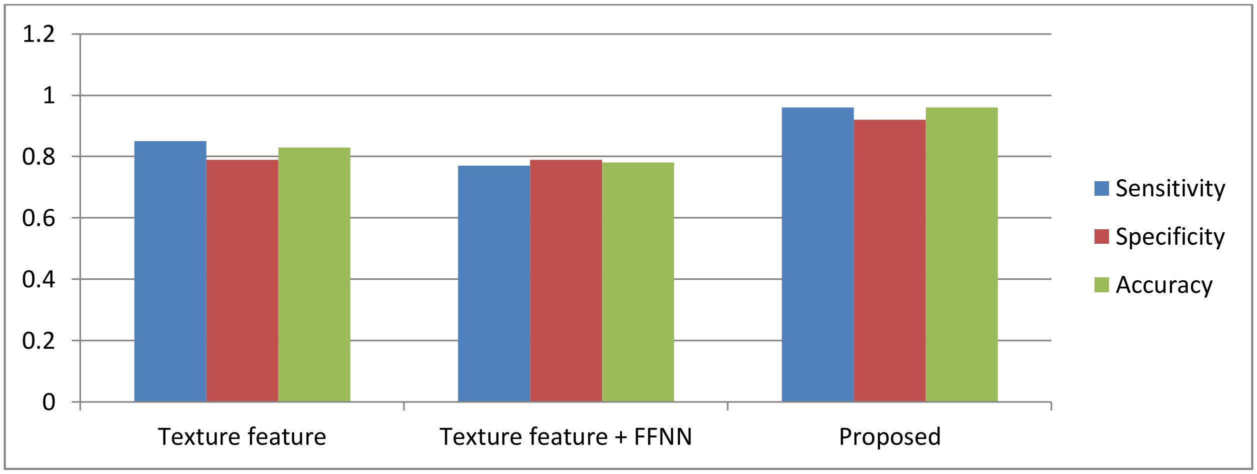
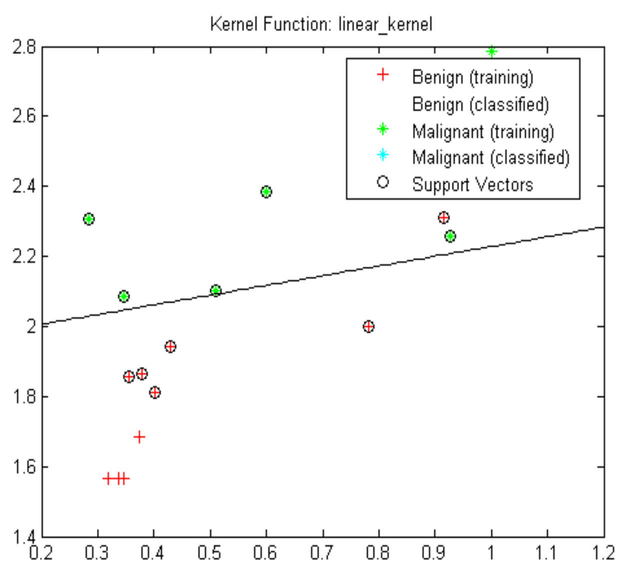


Fig. 10 Performance of Proposed algorithm

V. ROC Curve Analysis

Performance of each test is characterized in terms of its ability to identify true positives, while rejecting false positives. The accuracy of a diagnostic test can be summarized in terms of an ROC curve The Receiver Operating Characteristic (ROC) curve is a popular tool in medical imaging research. The ROC curve for the given test image is shown in the Fig.11.

Sensitivity



1-Specificity

Fig. 11: Receiver Operating Characteristic (ROC) curve

It is observed that the straight line is hyper plane that separates Benign and Malignant brain images. Below the hyper plane are the clusters of benign images and above the hyper plane are clusters of malignant images.

CONCLUSION

In this paper, MRI brain image segmentation is performed using Modified fuzzy C-Mean algorithm. In order to get a better classification rate, different statistical feature were extracted by using GLCM. SFS is used to select the discriminative features among them. SVM is used to classify the input, which is MRI Brain image into Benign and Malignant images. Afterwards the Malignant images are further separated as benign and malignant category. This kernel technique will help to get more accurate result. In the proposed work about 96% are classified accurately and 4% are misclassified. This is the limitation of the proposed work. To overcome the problem of SVM training, hybrid classifiers can be used. Some more texture features can also be added to improve the classification accuracy. In future, 3D image Analysis can be done for classification of MRI brain images.

References

[1] S. M. Smith *et al.*, ``Advances in functional and structural MR image analysis and implementation as FSL," *Neuroimage*, vol. 23, pp. S208_S219, Dec. 2004.

[2] S. Debette and H. S. Markus, ``The clinical importance of white matter hyperintensities on brain magnetic resonance imaging: Systematic review and meta-analysis," *Bmj*, vol. 341, p.c3666, Jul. 2010.

[3] Armstrong, T.S., Cohen, M.Z., Weinbrg, J., Gilbert, M.R.: 'Imaging techniques in neuro oncology', Semoncnur, 2004, 20, (4), pp. 231–239

[4] Dunn, J.C.: 'A fuzzy relative of the isodata process and its use in detecting compact well separated clusters', *J. Cybern.*, 1973, **3**, (3), pp. 32–57

[5] Mignotte, M.: 'A de-texturing and spatially constrained K-means approach for image segmentation', *Pattern Recognit. Lett.*, 2011, **32**, pp. 359–367

[6] Pham, D.L.: 'Fuzzy clustering with spatial constraints'. IEEE Proc. of Int. Conf. on Image Processing, 2002, pp. 65–68

[7] Yang, Y.: 'Image segmentation by fuzzy c-means clustering algorithm with a novel penalty term', Comput. Informatics, 2007, 26, pp. 17–31

[8] Krinidis, S., Chatzis, V.: 'A robust fuzzy local information c-means clustering algorithm', IEEE Trans. Image Process., 2010, 19, (5), pp. 1328–1337

[9] Despotovic, I., Vansteenkiste, E., Philips, W.: 'Spatially coherent fuzzy clustering for accurate and noise-robust image segmentation', IEEE Signal Process. Lett.,2013, 20, (4), pp. 295–298

[10] Ahmed, M.N., Yamany, S.M., Farag, A.A., et al.: 'Bias field estimation and adaptive segmentation of MRI data using a modified fuzzy c-means algorithm'. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 1999, vol. 1, pp. 250–255

[11] Ahmed, M.N., Yamany, S.M., Mohamed, N., et al.: 'A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data', IEEE Trans. Med. Imaging, 2002, 21, (3), pp. 193–199

[12] Kashif Hussain Memon1, Dong-Ho Lee1, "Generalised fuzzy c-means clustering algorithm with local information", IET Image Processing,2016

[13] Yunjie Chen1, Jian Li1, Hui Zhang2 , Yuhui Zheng2, Byeungwoo Jeon3, Qingming onathan Wu4,"Non-local-based spatially constrained hierarchical fuzzy C-means method for brain magnetic resonance imaging segmentation", IET Image Processing,2016

[14] T. Torheim, E. Malinen, K. Kvaal et al., "Classification of dynamic contrast enhanced MR images of cervical cancers using texture analysis and support vector machines," IEEE Transactions on Medical Imaging, vol. 33, no. 8, pp. 1648–1656, 2014.

[15] J. Yao, J. Chen, and C. Chow, "Breast tumor analysis in dynamic contrast enhanced MRI using texture features and wavelet transform," IEEE Journal on Selected Topics in Signal Processing, vol. 3, no. 1, pp. 94–100, 2009.

[16] A. Materka and M. Strzelecki, ``Texture analysis methods_A review," Tech. Univ. Lód¹, Lodz, Poland, Tech. Rep. COST B11, Brussels, 1998, pp. 9_11.

[17] N. Aggarwal and R. K. Agrawal, ``First and second order statistics features for classification of magnetic resonance brain images," *J. Signal Inf. Process.*, vol. 3, no. 2, p. 146, 2012.

[18] R. M. Haralick, K. Shanmugam, and I. Dinstein, ``Textural features for image classification," IEEE Trans. Syst., Man, Cybern., Syst., vol. SMC-3, no. 6, pp. 610_621, Nov. 1973.

[19] P. Pudil, J. Novovicov´a, and J. Kittler, “Floating search methods in feature selection,” *Pattern Recognition Letters*, vol. 15, no. 11, pp. 1119–1125, 1994.

[20] N.M. Saad, Syed A.R. Abu-Bakar and A.F. Muda,“Automatic Brain Lesion Detection and Classification Based On Diffusion-Weighted Imaging Using Adaptive Thresholding and A Rule Based Classifier”, International Journal of Engineering and Technology (IJET), Vol 6 No 6 Dec 2014-Jan 2015.

[21] A. Jayachandran and R. Dhanasekaran, “Brain Tumor Detection And Classi-fication of MR Images Using Texture Features And Fuzzy SVM Classifier “, Research Journal of Applied Sciences, Engineering and Technology 6(12), 2264-2269, 2013.