An efficient system based on transforms technique with SVM classifier for diagnosis of tumor in MRI Brain images

K.Kavinkumar[1,] and T.Meeradevi[2]
[1]Assistant Professor, Department of ECE, Kongu Engineering College
Perundurai, Erode -638 060, TamilNadu, India
kavinece07@gmail.com
[2]Professor, Department of ECE, Kongu Engineering College
Perundurai, Erode -638 060, TamilNadu, India

Abstract

Brain Tumour diagnosis is a critical application where lot of emphasis is needed. Conventional image processing constitutes pre-processing, feature extraction and classification mostly in medical imaging image segmentation is popular. In feature extraction, Discrete Wavelet Transform (DWT) and Discrete Fourier Transform(DFT) are used to reduce the coefficient of the image. For Discrete Wavelet Trasform, Haar wavelet is used as a mother wavelet. Three level decomposition are used to represent the feature vector for better classification of MRI brain image. In classification, two different classifiers K Nearest Neighbors (KNN) and Support Vector Machine (SVM) used to classify the tumours in image. sensitivity, specificity, and accuracy are the evaluation metrics to measure the performance of the two classifiers. Comparing the result it is shown that combination of DWT with SVM got the 95.4% accuracy.

Keywords: *Feature extraction – Discrete Wavelet Transform, Discrete Fourier Transform, Feature reduction -PCA, Classification - SVM, KNN*

1. INTRODUCTION

Various imaging modalities available like CT, PET, MRI, X-RAY, etc.MRI is preferred the most preferred because of no radiation risk, high resolution and good SNR. In MRI image formed by measuring signal received due to interaction of magnetic field and radio waves on protons. Noise affects the data coming from acquisition system. Identifying and reducing this is important - to improve validity and accuracy. Methodologies for reducing noise (denoising) are to be considered. We propose a filtering technique based on the Euclidean distance.

In image processing feature extraction is an important step, which is used for reducing the dimensionality. It is the process of transforming given input data into a group of features. This is usually done to gain insight into what is happening with the images and how they can be used to extract desired information. Many extraction methods are used and we propose texture feature extraction.

2. EXISTING METHOD

In existing method [7], feature ex/traction has been done using Multi-Texton Microstructure Descriptor in which four features corresponding to original image, orientation image, multi-texton image and texton structure image are extracted and concatenated to form feature vector of MR brain image. The extracted feature is given as input to support vector machine classifier to classify brain image as normal and tumor.

3. PROPOSED APPROACH

Our proposed system consists of four phases namely, MRI brain image database collection, feature extraction, feature selection and classification. Here two types of methodology is performed as shown in the figure 1 and figure 2 . In proposed method, the feature is extracted from training dataset of 70 normal and tumor brain MR images and these feature vectors are used to train KNN and linear kernel

SVM classifier. To test the classifier 50 brain MR images (normal and tumor) are used and performance is evaluated.

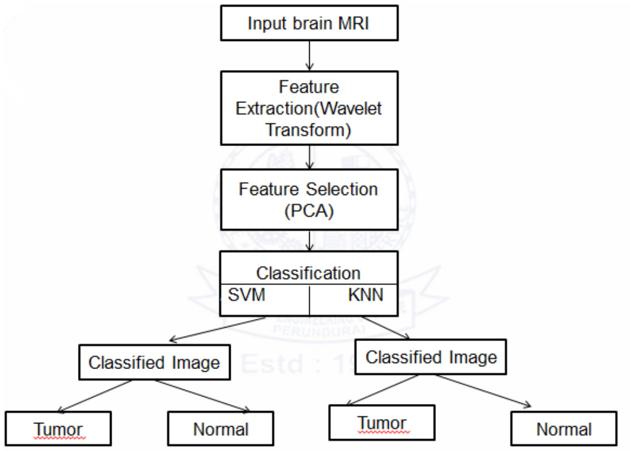

Fig. 1 Block diagram1 of proposed approach

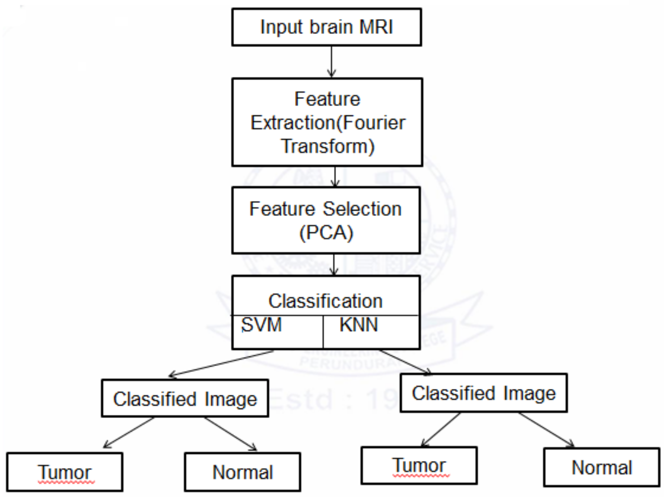

Fig. 2. Block diagram2 of proposed approach

A. Feature Extraction

1. Discrete wavelet transform

Most of the data are not relevant to separate between cancerous and normal tissues, and also, they introduce noise to the system. Moreover, Finding small data sets that are sufficient enough to distinguish between cells of different types is a requirement of diagnosis in practice. Furthermore, it is very important to a pathologist to isolate genes which are potentially intimately related to the tumor makeup and patho-mechanism. From classification point of view, reduction in the dimension of the feature space can help to overcome the risk of overfitting. Overfitting problem arises frequently in tumor classification problem where the dimension of the feature vectors is typically several orders of magnitude larger than the number of training patterns. In such a situation, classification performance on a test set is much more poor than on training set. In our DWT based feature extraction scheme, images are firstly transformed to time-scale domain by multilevel wavelet decomposition. Both of the coefficients extracted and images reconstructed are used as raw features to be selected. One of the image and its three level DWT results are shown in Figure 3, In this paper, Haar wavelet is used as a mother wavelet. Here three level decomposition is used.

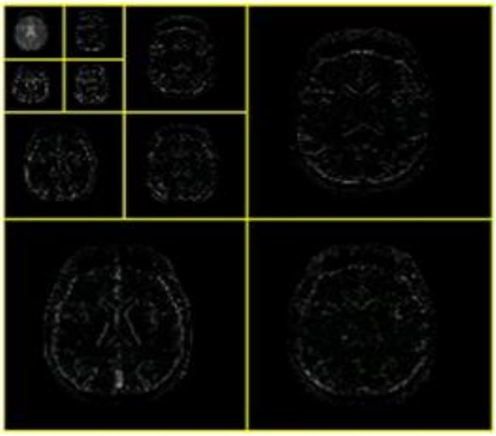

Fig. 3. Discrete Wavelet Transfrom2 level Decomposition

2. Discrete Fourier transform

Since we are dealing with images, and in fact digital images, so for digital images we will be working on discrete fouriertransform. In this transform, the transformed image will include spatial frequency, magnitude, and phase. The spatial frequency directly relates with the brightness of the image. The magnitude of the sinusoid directly relates with the contrast. Contrast is the difference between maximum and minimum pixel intensity. Phase contains the color information.

The formula for 2 dimensional discrete Fourier transform is given below.

$$F(k,l) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i,j)\, e^{-\iota 2\pi (\frac{ki}{N} + \frac{lj}{N})} \quad (1)$$

The discrete Fourier transform is actually the sampled Fourier transform, so it contains some samples that denotes an image. In the equation 1,formula f(I,j) denotes the image, and F(k,l) denotes the discrete Fourier transform. The equation is the 2 dimensional inverse discrete Fourier transform

$$f(a,b) = \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} F(k,l)\, e^{\iota 2\pi (\frac{ka}{N} + \frac{lb}{N})} \quad (2)$$

B. Feature selection

The main linear technique for dimensionality reduction is principal component analysis, performs a linear mapping of the data to a lower-dimensional space in such a way that the variance of the data in the low-dimensional representation is maximized. In practice, the covariance matrix of the data is constructed and the eigenvectors on this matrix are computed. The eigenvectors that correspond to the largest eigenvalues (the principal components) can now be used to reconstruct a large fraction of the variance of the original data. Moreover, the first few eigenvectors can often be interpreted in terms of the large-scale physical behaviour of the system. The original space has been reduced to the space spanned by a few eigenvectors.

C. Classification

1. Support Vector Machine

SVM is a binary classifier which is based on supervised learning. The key concept of SVM is to use the hyper plane to define decision boundaries for separating between data points of different classes. The basic idea of SVM is to classify data between two classes by constructing a hyper plane in high dimensional feature space.

The equation 3 represent the training set data

Consider a training set,

$$\{(x_i, y_i), i=1,2,\dots,n;\ x_i ?\ R^d;\ y_i ?\ \{+1,-1\};\} \quad (3)$$

where $x_i \in R^d$ is input vectors and $y_i \in \{+1,-1\}$ is class labels of MRI brain image.

SVM maps the d-dimensional input vectors from input space to high dimensional feature space using non linear function $\Phi(.)$. The separating hyper plane is defined as $w^T \Phi(x) + b = 0$, with w as weight vector having dimension equal to $\Phi(x)$ and b is bias [11].

The separating hyper plane can be defined in many ways when the data are linearly separable. However, SVM is based on the maximum margin principle in which the aim is to construct a hyper plane by satisfying maximal distance between the two classes. The equation 4 and 5 represents the class values. Same way equation 6 represents the multi-class mathematical model
The SVM start with following formulations

$$w^T \Phi(x_i) + b \geq +1 \text{ for } y_i = +1 \quad (4)$$
$$w^T \Phi(x_i) + b \leq -1 \text{ for } y_i = -1 \quad (5)$$

$$y_i (w^T \Phi(x_i) + b) \geq 1, \quad i=1,2,\dots,n \quad (6)$$

The classifier is defined as

$$fn(x) = \text{sign}(w^T \Phi(x) + b) \quad (7)$$

For each training data x_i, the function yields $fn(x_i) \geq 0$ for $y_i = +1$and yields $fn(x_i) \leq 0$ for $y_i = -1$ [6]. An example of SVM classification with an optimal hyper plane that minimizes the separating margin between the two classes are illustrated by points noted by 'o' s and 'Δ's is shown in Fig. 4. Support vectors are elements of the training dataset that lie on the boundary of hyper planes of the two different classes.

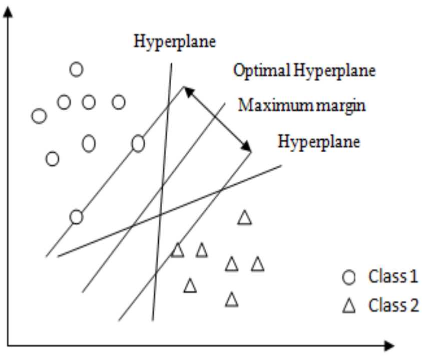

Fig. 4. Illustration of SVM

2. K Nearest Neighbors

K Nearest Neighbor method is mostly used method for classification of image. An object is classified based on the distance from its neighbors. If k = 1, then the object is classified simply as the class of its nearest neighbor. The K Nearest Neighbor classification method is based on nearest distance of neighbor classes. Based upon the distance, the K Nearest Neighbors selects only k-nearest neighbor classes. Finally the majority vote is then taken to predict the class for an object [3]. The most common method used in k-nearest-neighbor for measuring distance is Euclidean. The Euclidean distance is used to measure the distance between the test data and train data. Then based on the distance measure the object is classified to one of the predefined classes. This same procedure is followed for classification of MR brain image as normal and tumor.

Training and Testing Process: To train the classifier, we need some data features to identify the class of brain tumor. These data features train the classifier and the classifier will identify the type of tumor. The data features which is chosen for training the classifier is three level decompositions. Both the KNN and SVM classifier are trained with set of images using these features and tested with another set of testing images from which the classifier identifies MRI brain image class as normal and tumor.

4. EXPERIMENTAL RESULTS

A. Materials

The image data set for experiment contains 60 normal MRI brain image and 60 MRI brain tumor image which is collected from Brainweb database. In the proposed method the brain image dataset is divided into two sets such as 1) training dataset and 2) testing dataset. The training dataset is used for learning the classifier and to evaluate the performance of proposed system testing dataset is used.

B. Results

In this section the experimental results of the proposed method using MRI brain images with and without tumor are described. The proposed system is implemented in MATLAB 2014a.

5. PERFORMANCE ASSESSMENT OF PROPOSED TECHNIQUE

The classifiers are trained with training dataset (35 normal and 35 tumor) images and the classification accuracy is calculated with the testing (25 normal and 25 tumor) images. In the testing phase, the testing dataset image is given to the proposed technique to find the class of brain images. The obtained result of classifier without and with tumor is shown in Figure. 5 and Figure. 6 which classified the MRI image as normal and tumor image respectively.

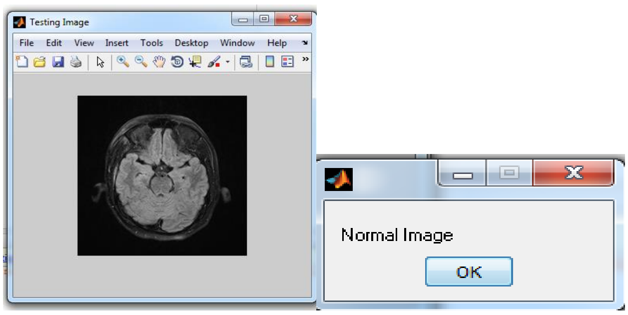

Fig. 5 . MRI brain tumor image and classified output

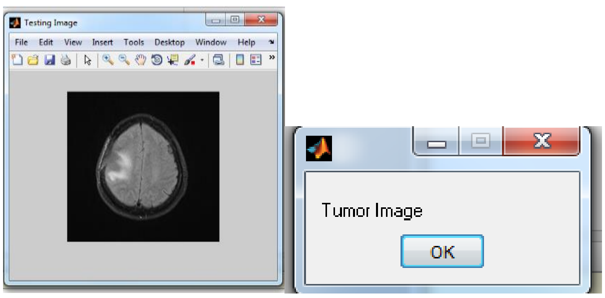

Fig. 6. MRI brain tumor image and classified output

The above figure 5 and 6 are the classified out obtained for the data used.

TABLE 1: PERFORMANCE ANALYSIS

Classifier	Feature Extraction	Sensitivity (%)	Specificity (%)	Accuracy (%)
KNN	FFT	88.8	70.8	69.9
	DWT	90.4	69.6	79.5
SVM	FFT	77.2	80.3	74.3
	DWT	100	91.3	95.4

The performance metrics sensitivity, specificity, and accuracy of SVM classifier KNN classifier with three feature extraction methods is shown in Table 1. From the above table it is understood that SVM with DWT provide better results

The obtained classified output from SVM and KNN classifer is resolved through performance metrics such as sensitivity, specificity, and accuracy [7].

Sensitivity = tp/(tp+fn)
Specificity = tn/(tn+fp)
Accuracy = (tp+tn)/(tp+tn+fp+fn)

True positive (tp): Tumor people correctly identified as having the condition
False positive (fp): Normal (healthy) people incorrectly identified as tumor
True negative (tn): Normal (healthy) people correctly identified as healthy
False negative (fn): Tumor people incorrectly identified as normal (healthy)

6. CONCLUSION

The proposed method for feature extraction is DFT and DWT and two machine learning algorithms is used to classify the MR brain images as normal and abnormal. While using DWT small information in medical images are retained and gives needed information for classification. The accuracy of proposed method DWT with SVM classifier is 95.4%.

References

1. Hemalata V.Bhujle and Basavaraj H.Vadavadagi, “NLM based magnetic resonance image denoising – A review”, Biomedical Signal Processing and Control, vol. 47, pp. 252-261, 2019.
2. Li Liu ; Songyang Lao, “Median Robust Extended Local Binary Pattern for Texture Classification”, IEEE Transactions on Image Processing, vol. 25 , no. 3 , pp.1368-1381, 2016.
3. M. Calonder, V. Lepetit, M. Ozuysal, T. Trzcinski, C. Strecha, P. Fua, “BRIEF: Computing a local binary descriptor very fast”, IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 7, pp. 1281-1298, 2012.
4. J.Mohan, V.Krishnaveni, YanhuiGuo, “A survey on the magnetic resonance image denoising methods”, Biomedical Signal Processing and Control, vol. 9, pp.56-69, 2014.
5. Gravel Pierre, Gilles Beaudoin, Jacques A. De Guise, “A method for modelling noise in medical images’, Medical Imaging IEEE Transactions , vol. 23, no.10, pp. 1221- 1232, 2014.
6. Nooshin Nabizadeh, Miroslav Kubat, "Brain tumors detection and segmentation in MR images : Gabor detection and segmentarion in wavelet vS. Statistical features", Elsevier, computers and Electrical Engineering vol. 45, pp. 286-301, 2015.
7. Ms Jahnavi, sree priya kurup, " A novel approach to detect brain tumor in MRI images using hybrid technique with SVM classifiers", IEEE International conference on Recent Trends in Electronics , Information & Communication Technology, pp. 546-549, 2016.
8. K.Kavin kumar, T.Meeradevi, S.Maheswaran,”An Efficient Method for Brain Tumor Detection Using Texture Features and SVM Classifier in MR Images”, Asian Pacific Journal of Cancer Prevention, vol. 19, no. 10, pp. 2789-2794,2018.
9. Maheswaran Shanmugam, Sunilvignesh Nehru and Sathesh Shanmugam, "A wearable embedded device for chronic low back patients to track lumbar spine position",Biomedical Research 2018; Special No. :Special Section: Computational Life Sciences and Smarter Technological Advancement, vol. 2018, pp. S118-S123,2018.
10. Sudheesh K.V, and Basavaraj, "Texture feature abstraction Based on Assessment of HOG and GLDM features for Diagonising Brain Abnormalities in MRI images", Global Journal of computer science and technology, vol.8, pp. 26-30, 2018.