



An Ensemble Classification-Based Methodology Applied to MRI Brain Images for Tumor Detection and Segmentation

Abuzar Qureshi¹(✉), Khan Bahadar Khan¹, Hamza Ali Haider¹,
Rameez Khawaja¹, and Muhammad Yousuf²

¹ Department of Telecommunication Engineering, Faculty of Engineering,
The Islamia University of Bahawalpur, Bahawalpur, Pakistan
{abuxar.coresh, kb.khattak, hamza.ali51214,
rameezkhawaja}@ieee.org

² Department of Electronic Engineering, Faculty of Engineering,
The Islamia University of Bahawalpur, Bahawalpur, Pakistan
myousuf@yahoo.com

Abstract. Automated brain tumor detection is an important application in the medical field. There is a lot of methods developed for this task. In this paper, we have implemented an algorithm which detects the type of brain tumor from MRI image using supervised classification techniques. The major part of the work includes feature extraction using DWT and then reduction of features by using PCA. These reduced features are submitted to different classifiers like SVM, k-NN, Naïve Bayes and LDA. The results from each classifier are then submitted to a voting algorithm that chooses the most frequent result. The dataset for training contains 160 MRI images. The algorithm is processed on 200 * 200 images to reduce processing time. This method is tested and found to be much beneficial and rapid. It could be utilized in the field of MRI classification and can assist doctors to detect the tumor type and diagnose about patient abnormality level.

Keywords: Image segmentation · MRI classification · MATLAB · Image processing · Tumor detection · Feature extraction · Feature reduction

1 Introduction

Nowadays, image processing is a key tool for biomedical engineering. We can use it for examination of different medical images. The most important one is brain tumor detection. The brain tumor is of two types. One which doesn't spread over time called Benign tumor and other which spread with the passage of time is called Malignant tumor.

Image processing techniques are applied for enhancement of the image or to acquire some valuable information from it. Image processing is also used for image segmentation which deals with the extraction of different features of the image. The image segmentation has a vast scope in biomedical engineering. One of the most crucial parts of the biomedical image processing is MRI classification.

The MRI stands for Magnetic Resonance Imaging. It is the imaging technique that produces the high-definition images of anatomical structures of any part of human body, especially in brain. It provides unparalleled view inside human body for clinical decision and biomedical field research. In this algorithm, we use discrete wavelet transform (DWT) [1] to extract features of MRI images. But due to the large storage requirement of wavelet transform, the principal component analysis (PCA) [2] is used. It reduces the dimensionality of the data and reduce computational costs.

Now, the PCA is used to reduce the dimension to get the important features of the MRI image, but we need to classify the data extracted. A lot of techniques are proposed by researchers for the classification of MRI images. These can be described by two classes; one is supervised classification while other is unsupervised. The supervised classification has some superiority over the unsupervised in terms of accuracy. Each approach had achieved good results, but the supervised classification performs better than all in terms of classification accuracy.

In this algorithm, we use two supervised classifiers and two unsupervised classifiers to inspect the accuracy and other parameters between them. In supervised classification, we used the support vector machine (SVM) [3] with linear and Gaussian Radial Based Function (RBF) parameter. The other supervised classifier used is K-nearest neighbor (K-NN) [4]. These both are the popular classification methods based on machine learning basics. The second approach is unsupervised classification in which we use Naïve Bayes (NB) [5] algorithm and Linear Discriminant Analysis (LDA) [6]. These both are the most used unsupervised classification algorithms for biomedical image classification.

Moreover, we use kernel SVMs instead of conventional SVMs. These SVMs are different from conventional SVMs only in terms of the dot product form. All of these classifiers are recommended individually by the researchers and have a very high percentage of accuracy. To make the algorithm more accurate, we use voting method on the results from all the classifiers and chooses the option with more repetition.

We apply segmentation techniques like Otsu Binarization to acquire the segmented image, which is then preprocessed to extract the features. The classification results are subjected to a voting algorithm to select the option with the maximum number of occurrences. We also compared the results of all classifiers to inspect the accuracy of each classifier.

2 Literature Review

Zhang et al. [3], proposed a method for MR brain image classification in which DWT and PCA techniques are employed for feature extraction and reduction. These features are classified using kernel SVM. The purposed method addresses common brain diseases. The input data set is 160 MR brain images with 20 normal and 140 abnormal. The accuracy achieved by RBF kernel comes out to be 99.38% while the linear kernel gives 95% accuracy. So, it is observed that the accuracy rate is high with RBF kernel as compared to linear kernels.

El-Sayed et al. [7], presents a hybrid technique for brain MRI classification. For feature extraction and reduction, DWT and PCA are used respectively. For classification,

two classifiers are used. The first classifier based on feed-forward back-propagation artificial neural network (FP-ANN) and the second is based on k-nearest neighbor (k-NN). The success rate of FP-ANN and k-NN is 97% and 98% respectively. Andrés et al. [8], proposed the SOM-FCM based method for MRI segmentation. Features are extracted from GLCM and histogram of the 3-D image. SOM training with fuzzy clustering is employed to classify the input data. The input data contain T1 MR Images with ages between 7 to 71 years old. The results of SOM-FCM classification provide high accuracy.

Saritha et al. [9], proposed the classification method for MR brain images using combined wavelet entropy-based spider web plots and probabilistic neural networks. The wavelet entropy-based spider web is used to extract the features and the probabilistic neural network classifies the MR images. This classification accuracy of this algorithm found out to be 100%. Chandra et al. [10], presented the brain MR images classification with SVM classifier and compared it with another classifier AdaBoost. The input dataset contains 86 abnormal and 48 normal images. The dataset is used for MRI training and then classified by SVM as well as AdaBoost. The accuracy of SVM and AdaBoost comes out to be 92.71% and 89.31% respectively.

3 Proposed Model

Classification of brain MR images is mostly used to detect the type of tumor in the brain. Our model to identify the tumor is shown in Fig. 1. The model has two major parts; pre-processing and classification. Pre-processing includes two blocks of feature extraction and feature reduction. The segmented image is subjected to DWT for feature extraction and PCA is employed for feature reduction. In classification, we have employed four classifiers to individually classify the data.

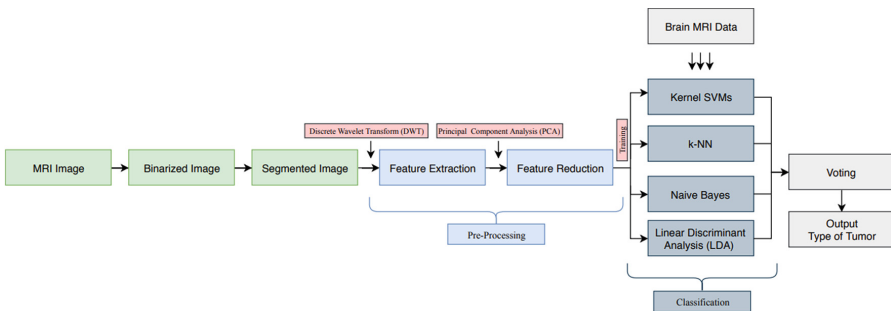


Fig. 1. Block diagram of the proposed model

SVM and k-NN classifier belong to supervised classification while Naïve Bayes and LDA belongs to unsupervised classification. Then the individual results of all classifiers are subjected to voting algorithm which selects the option with maximum numbers of occurrence.

3.1 Feature Extraction (DWT)

For feature extraction, we use a discrete wavelet transform (DWT), which is given by

$$W_{\psi}(a, b) = \int_{-\infty}^{\infty} x(t)\psi_{a,b}(t)dt$$

In the case of 2D images, the DWT individually apply on each dimension causes the result to be 4 sub-bands (LL, LH, HH, HL).

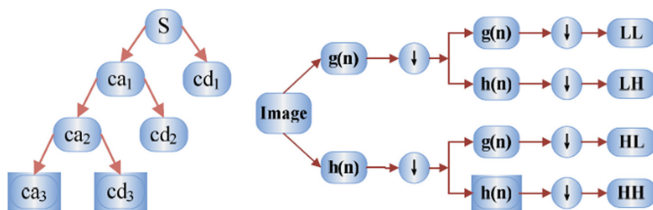


Fig. 2. Two-dimensional wavelet transform tree [3]

Thus, we use a 3-level wavelet decomposition tree. 2D DWT results in 4 sub-bands (LL, LH, HH, HL) and the sub-band LL is used for next DWT. The right portion of Fig. 2 [3] represents the 4 sub-bands obtained as the result. This LL sub-band is regarded as approximation component of the image while the other three sub-bands are regarded as detailed component of the image. So, the approximation component is again subjected to 2D DWT to repeat the process. Thus, the wavelet transform provided us the hierarchal framework to interpret the details of image.

3.2 Features Reduction (PCA)

Excessive feature increases the computational time and requires more storage. The classification and decision become more complicated with large number of features. Therefore, it slows down the executing process. This complication refers to the curse of dimensionality.

To reduce features, we use principal component analysis (PCA) which is a useful tool to reduce the dimensions of the dataset having a large number of unrelated variables. It is done by transformation of the dataset to new variables in terms of their significance. The PCA technique has three stages; first, it performs orthogonality method on the components of the input vector to make them uncorrelated with each other. Second, it creates the order of resulting orthogonal component with the first largest variation. Third, it removes the components of vector whose contribution is least to the dataset.

The features after extraction and reduction are:

Mean: The mean value presents the contribution of pixel intensity to the whole image. For benign tumors, the mean is less than that of malignant. As an image is basically a matrix with rows and column vectors. So, the mean can be calculated by

$$\text{Mean of Matrix} = \frac{\text{Sum of all elements of matrix}}{\text{Total number of elements}}$$

Standard Deviation: It is the measure of variation and dispersion in an image. It can be simply calculated as the square root of the variance of the Matrix.

$$\text{Standard Daviation} = \sqrt{\text{Variance}}$$

Entropy: The entropy is the measure of the degree of randomness and disorder in an image. The entropy of an image can be easily interpreted from its histogram.

RMS: RMS is a root-mean-squared value. It is the RMS value of each row for each column.

Variance: The matrix obtained from subtracting its mean from its each element has another mean which is called variance of the matrix. Generally, it is the measure of how far a data set is spread out. In image processing, it is used to find that how every individual pixel varies from the center pixel as well as neighbor pixels.

Smoothness: It is the measure of the average value of an image with noise removed. It is often used to reduce the noise in the image.

$$\text{Smoothness of Image} = 1 - \frac{1}{1 + (\text{sum of elements})}$$

Kurtosis: Kurtosis is the measure of the highest peaks in the image. In other words, it can be described as the measure of heaviness and thickness of the given data. It determines the noise in the image with respect to resolution.

Skewness: Skewness is the measure of symmetry in distribution which tells us about the glossiness and darkening of the image surface. The dataset is symmetric when it looks same from left and right side of the center point.

IDM: Stands for Inverse Difference Movement.

It is a type of image textural feature which deals with the discontinuities in image.

Contrast: Contrast is the difference in luminance or color that makes an object distinguishable. It is the difference in color and brightness of the objects lying in same field of view.

Correlation: Correlation is the process of moving the filter mask on the screen and taking the largest magnitude value.

Energy: Energy can be calculated by the square root of uniformity which is the summation of each element of the image's matrix. It is the mean squared value of image.

Homogeneity: It refers to the surface of images having similar characteristics.

3.3 Ensemble Classification

Classification of the biomedical images can be done by employing either supervised or unsupervised classification techniques [11]. In supervised classification, the user can provide the custom sample pixels for decision, but in unsupervised classification, the results only based on software analysis. In this proposed model we have implemented four different classifiers to compare their results. The two classifiers are from supervised classification class while the other two are from unsupervised class.

Support Vector Machine (SVM): SVM is a support vector machine, whose introduction is the landmark in the field of machine learning. It belongs to supervised classification class. The major benefits of SVM are mathematical tractability, high accuracy, and direct geometry decision. There are a lot of SVMs, but the best among them is kernel SVM. In the proposed model, we use kernel SVM with two different parameters; linear and RBF. Traditional SVM uses hyperplane to classify the data. In kernel SVM, the algorithm is almost same but each dot product between vectors is replaced by non-linear kernel function.

K-Nearest Neighbor (k-NN): k-NN is one of the trusted algorithms for classification and belongs to the supervised classification class. Its work based on the minimum distance of query instance to the training samples. It provides a variety of distance measuring techniques like Euclidian or Hamming distance.

Naïve Bayes Classifier: Naïve Bayes algorithm works on conditional probability. Its working is based on the Bayes theorem on conditional probability, which is given as:

$$P(H|E) = \frac{P(E|H) * P(H)}{P(E)}$$

So, this classifier predicts the data on the basis of the predicted class. It belongs to unsupervised classification class. The major advantages of this classifier are high scalability, less use of training data and easy to implement.

Linear Discriminant Analysis (LDA): Linear Discriminant Analysis (LDA) can be used as a classifier as it makes the assumption about the input data. It makes assumption that whether the data is Gaussian or not. And the variance of each attribute of the acquired set of features. It also falls in the unsupervised classification class.

4 Results and Discussion

In this study, we have developed the algorithm for tumor detection using DWT+PCA+Classifiers. We apply different image processing techniques including grayscale of the image, OTSU binarization of the image and then filtration of the image to get the segmented image. The results are shown in Fig. 3.

Next, we compute the DWT of the segmented image to extract the features. The three-level DWT technique makes the feature vector of 1024 values. These extra features are reduced using the PCA technique. The results of feature reduction are shown in Fig. 3 and Table 1.

Next, these features are classified by different classifiers and the results shown in Table 2. After the detection of tumor type, all results are subjected to a voting algorithm to choose the right option. The results of all classifiers are stored in a 1D array. The voting algorithm inspects that array and chooses the option with more occurrences.

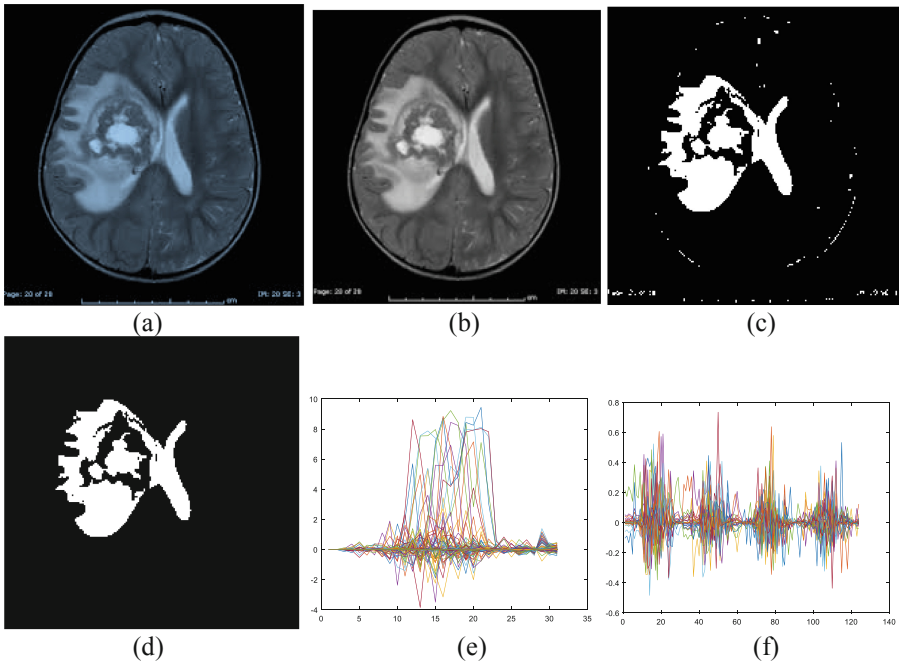


Fig. 3. (a). Brain MRI image, (b). Grayscale image of MRI, (c). OTSU Binarized Image, (d). Segmented Image, (e). Extracted features from DWT, (f). Reduced feature using PCA.

Figure 3 represents the image processing steps on the MR images of the brain. Figure 3 (d) represents the segmented image from which we extract the features. After extraction of features, Fig. 3 (e) shows the plot of the extracted features which are then reduced and shown in Fig. 3 (f). So, the reduction in features is approximately about 92%.

Table 1. Some useful extracted features

MRI image	Type of tumor	Entropy	Kurtosis	Correlation	Smoothness	RMS	Energy
IMG001	Benign	3.17346	7.32819	0.199005	0.920457	0.0898	0.7621
IMG002	Benign	3.26983	7.95668	0.093089	0.897422	0.0898	0.7685
IMG003	Benign	3.07565	7.7971	0.089525	0.904047	0.0898	0.7556
IMG004	Benign	2.74648	10.9703	0.097650	0.718636	0.0898	0.7861
IMG005	Malignant	3.67005	5.62099	0.076926	0.918462	0.0898	0.7577
IMG006	Malignant	3.52392	6.52204	0.073405	0.928384	0.0898	0.7402
IMG007	Malignant	3.62834	5.32384	0.095075	0.913222	0.0898	0.7378
IMG008	Malignant	3.19429	9.73182	0.142678	0.951576	0.0898	0.7604

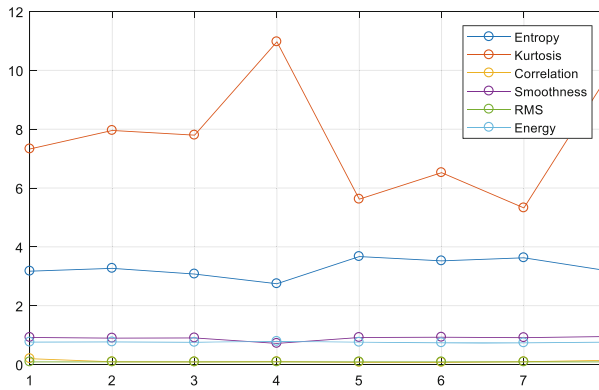


Fig. 4. Graphical representation of Table 1.

Table 1 and Fig. 4 represents the different feature values for some MR images samples. This shows the trend of values difference between Benign and Malignant tumors. For example, we can see that the entropy for Benign tumor is always less as compared to Malignant tumor. As the Entropy is the measure of randomness in image, it is less for Benign tumors.

Table 2 represents the results of all the classifiers. This provides the data of 8 images in which 4 are Benign and 4 are malignant tumors. From the table, we can infer that the Linear SVM and LDA classifier are the most accurate. Moreover, k-NN is also more accurate as compared to Naïve Bayes.

Table 2. Classification results

MRI image	Type of tumor	Results of classifiers				
		SVMs		k-NN	Naïve Bayes	LDA
		Linear	RBF			
IMG001	Benign	Benign	Malignant	Benign	Benign	Benign
IMG002	Benign	Benign	Malignant	Benign	Benign	Benign
IMG003	Benign	Benign	Benign	Benign	Benign	Benign
IMG004	Benign	Benign	Malignant	Benign	Malignant	Benign
IMG005	Malignant	Malignant	Malignant	Malignant	Malignant	Malignant
IMG006	Malignant	Malignant	Malignant	Benign	Benign	Malignant
IMG007	Malignant	Malignant	Benign	Malignant	Benign	Malignant
IMG008	Malignant	Malignant	Malignant	Malignant	Malignant	Malignant

5 Conclusion and Future Work

A hybrid model is proposed to detect the type of tumor from the MR images of the brain. This shows accurate results and reduces human efforts as well. The proposed model employees DWT and PCA for feature extraction and reduction from MRI of brain. On basis of features, the type of tumor is detected using different classifiers. For more accuracy, voting algorithm is used. From the results, we see that the Linear SVM and LDA classifiers detect the accurate type of tumor. k-NN classifier is also accurate in most of the cases. And the RBF SVM has the least accuracy for the given dataset. In future, we can use other methods for feature extraction and feature reduction and compare their results with the existing ones. More advancements can be made to make the model more integrated and accurate.

References

1. Tufekci, Z., Gowdy, J.N.: Feature extraction using discrete wavelet transform for speech recognition. In: Proceedings of the IEEE SoutheastCon (2000)
2. Kambhatla, N., Leen, T.K.: Dimension reduction by local principal component analysis. *Neural Comput.* **9**(7), 1493–1516 (1997)
3. Zhang, Y., Wu, L.: An MR brain images classifier via principal component analysis and kernel support vector machine. School of Information Science and Engineering, Southeast University, Nanjing, China (2012)
4. Cocosco, C.A., Zijdenbos, A.P., Evans, A.C.: A fully automatic and robust brain MRI tissue classification method. *Med. Image Anal.* **8**(1), 93–94 (2004)
5. Zhou, X., et al.: Detection of pathological brain in MRI scanning based on wavelet-entropy and Naïve Bayes classifier. In: International Conference on Bioinformatics and Biomedical Engineering (2015)
6. Gladis Pushpa Rathi, V.P., Palani, S.: Brain tumor MRI image classification with feature selection and extraction using linear discriminant analysis. In: Computer Vision and Pattern Recognition (2012)

7. El-Dahshana, E.-S.A., Hosny, T., Salem, A.-B.M.: Hybrid intelligent techniques for MRI brain images classification. *Dig. Sig. Process.* **20**(2), 433–441(2010)
8. Ortiz, A., Palacio, A.A., Górriz, J.M., Ramírez, J., Salas-González, D.: Segmentation of brain MRI using SOM-FCM-based method and 3D statistical descriptors. *Communications Engineering Department, University of Malaga, 29004 Malaga, Spain* (2013)
9. Sariitha, M., Paul Joseph, K., Abraham, Mathew, T.: Classification of MRI brain images using combined wavelet entropy based spider web plots and probabilistic neural network. *Pattern Recogn. Lett.* **34**(16), 2151–2156 (2013)
10. Chandra, S., Bhat, R., Singh, H., Chauhan, D.S.: Detection of brain tumors from MRI using Gaussian RBF kernel based support vector machine. *Department of CSE & IT, Jaypee University of IT, Solan, HP, India* (2009)
11. Ozer, S., et al.: Supervised and unsupervised methods for prostate cancer segmentation with multispectral MRI. *Int. J. Med. Phys. Res. Pract.* **37**(4), 1873–1883 (2010)