Available online at www.ijmrhs.com

ISSN No: 2319-5886

International Journal of Medical Research & Health Sciences, 2017, 6(9): 15-20

k-NN Based Classification of Brain MRI Images using DWT and PCA to Detect Different Types of Brain Tumour

Ayush Arora[1]*, Priyanshu Roy[1], Shwetha MD[2], S Venktesan[2] and Ramesh Babu[2]

[1]*Student, Department of Computer Science and Engineering, Dayananda Sagar College of Engineering, Bangalore, Karnataka, India*

[2]*Professor, Department of Computer Science and Engineering, Dayananda Sagar College of Engineering, Bangalore, Karnataka, India*

**Corresponding e-mail: aayusharora6896@gmail.com*

ABSTRACT

The treatment of any tumour is completely depended what type it is. This paper proposes a set of algorithms which work for the better detection and classification of brain tumour. The MRI image based brain tumour analysis would efficiently deal with classification process for brain tumour analysis. The proposal suggests using three stages namely feature extraction, feature reduction and classification. The discrete wavelet transformation methods (DWT) and principal component analysis (PCA) are used for feature extraction and feature reduction respectively. The k-Nearest Neighbours (k-NN) classifier is used for the classification of the tumour and to tell what type the tumour is Benign or Malignant. By using different methods for feature extraction, feature reduction and classification techniques to extract and calculate the average values of the features of the MRI image and to match the values to the values extracted from the set of input image and classify them to be benign or malignant.

Keywords: Brain tumour, Principal component analysis (PCA), Discrete wavelet transform (DWT), k-Nearest neighbours (k-NN), Classification, Magnetic resonance imaging (MRI)

INTRODUCTION

Magnetic resonance imaging or nuclear magnetic resonance is a technique for creating detailed images of human body. This technique uses a very powerful magnet to align the nuclei of atoms inside the body, and a variable magnetic field that causes the atoms to resonate, a phenomenon called nuclear magnetic resonance. The nuclei produce their own rotating magnetic fields that a scanner detects and uses to create an image. The early techniques used for diagnosis of brain tumours were invasive and sometimes dangerous and have been abandoned for better, non-invasive and high-resolution techniques such as MRI and computed tomography (CT) scans. Neoplasms will often show as differently coloured masses in CT and MRI results [1-3].

Benign brain tumours often show up as hypodense (darker than brain tissue) mass lesions on CT scans. On MRI, they appear either hypodense or isointense (same intensity as brain tissue) on T1-weighted scans, or hyper intense (brighter than brain tissue) on T2-weighted MRI, although the appearance is variable.

Contrast agent uptake, sometimes in characteristic patterns, can be demonstrated on either CT or MRI scans in most malignant primary and metastatic brain tumours.

Pressure areas where the brain tissue has been compressed by a tumour also appear hyper intense on T2-weighted scans and might indicate the presence a diffuse neoplasm due to an unclear outline. Swelling around the tumour known as peri-tumoral oedema can also show a similar result [2-4].

This is because these tumours disrupt the normal functioning of the blood brain barrier (BBB) and lead to an increase in its permeability. However, it is not possible to diagnose high-grade versus low-grade gliomas based on enhancement pattern alone. MRI is always the preferred medical imaging technique when the portrayal deals with soft tissues such as brain tissues, tendons, and ligaments. The biggest advantage of MR Imaging is that it is non-invasive technique. The use of computer technology in medical technology in medical decision support is now widespread and pervasive across a wide range of medical area. A fully automatic normal and automatic human brain classification is of great importance for medical and research purposes [5,6].

PROPOSED ALGORITHM

The features in the MR Images are extracted using the DWT method and further Reduction of the MR image is done by PCA. Once the features are readily extracted then the classification methods (SVM) is run so as to classify which type of tumour it is benign or malignant (Figure 1).

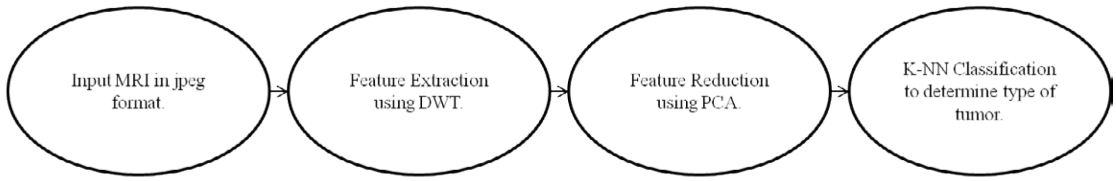

Figure 1 Stage diagram of the classification process

Stage 1: Feature extraction scheme using DWT (Discrete Wavelet Transform)

The feature vector is extracted by DWT in the proposed system. The wavelet is a powerful mathematical tool for feature extraction, and will be used to extract the wavelet coefficient from MR images. Wavelets are functions that are concentrated in time as well as in frequency around certain point. The basic advantage of using the Wavelet Transform method is that it works in such a way that it gives good frequency resolution for low frequency components and high temporal resolutions for high frequency components [7].

A review of fundamental of wavelet decomposition is introduced as follows:

Initially the process is started by selecting a mother wavelet from 'Haar', 'Daubecheis', 'Morlet' etc. The signal is now translated into shifted and scaled versions of this mother wavelet. This analysis divides the signal into approximate and three detailed signals. The approximate sub signal shows the general trend of pixel values a detailed sub signals on the horizontal, vertical, and diagonal details.

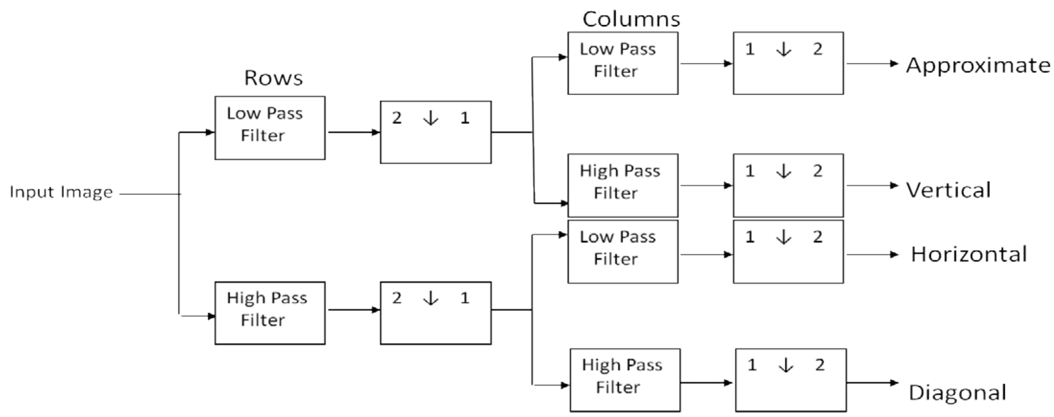

Figure 2 Discrete Wavelet Transformation (DWT) diagram

In the method of DWT, the rows are initially fed with low pass and high pass filter and then down sampled then both are filtered again by both low pass and high pass. The results are down sampled again and we get 4 different images with approximate, vertical, horizontal and the diagonal details respectively (Figure 2).

The MATLAB Function used for the DWT is DWT2 (X,'wname') where X is the matrix and wname is the name of the mother wavelet. We use the Daubecheis mother wavelet. This function gives out 4 components an approximate component matrix and the details on the horizontal, vertical and the diagonal matrices.

Discrete wavelength transform algorithm

1) The DWT2 function is applies to the red, green and the blue components separately.
2) Now combine the approximate, horizontal, vertical and the diagonal components of the derived matrices using the red, blue and the green components.
3) Now the approximate, horizontal, vertical and the diagonal components are viewed. As the approximate component is double it has to be divided by 255.
4) A standardizing function is used to visualize the images better.
5) Now the above steps are repeated once again so as to get the sampled outputs in the columns.

Stage 2: Feature reduction using Principal Component Analysis (PCA)

PCA is one of the most commonly used forms of dimensional reduction. When a set of data is given, PCA works in a way so as to preserve the variance and find the linear lower-dimensional representation of the data. The main purpose of the using PCA in this approach is to reduce the dimensionality of the wavelet coefficients. Using PCA for feature reduction of the derived feature vectors should lead to efficient and accurate classification algorithm. The size of the input matrix is reduced from (1024) to (7). The matrix now consists of these components.

Principal component analysis algorithm

Let X be an input data set (X: matrix of dimensions M X N).

1) Calculate the empirical mean u[m] = (1/N)∑ i=1 to N, X [m,n].
2) Calculate the deviations from the mean and store the data in the matrix B [M N]:, B=X-u.h, where h is a 1 x N row vector of all 1's: h[n] =1 for n=1....N.
3) Find the covariance matrix C: C = (1/N) $B.B^*$.
4) Find the eigenvectors and eigen values of the covariance matrix V−1CV = D, where V: the eigen vectors matrix, D: the diagonal matrix of eigen values of C, D [p,q] = λm for p=q=m is the m^{th} eigen values of the covariance matrix C.
5) Rearrange the eigenvectors and eigen values λ1 ≥ λ2 ≥ λ3 ≥ ...... ≥ λN.
6) Choosing components and forming a feature vector. Save the first L columns of V as the M x L matrix W: W[p,q]=V[p,q] for n=1....M, q=1....L where 1≤ L ≤ M.
7) Deriving the new data set the eigenvectors with the highest eigen values are projected into space, this projection results in a vector represented by fewer dimension (L<M) containing the essential coefficients.

The above Algorithm shows the steps involved for extracting principal components of the input classifier to the two classifiers: the main reason for carrying out feature extraction process in two steps. Firstly, the extraction of wavelet coefficients by DWT and then the selection of essential components by PCA.

Stage 3: k-Neural Network (k-NN) classifier

k-NN classifier is one of the simplest classification techniques. K closest training vectors are determined for an input feature vector X in the process according to a suitable distance metric. The vector X is assigned the class to which the majority of this k nearest neighbours belongs to. The k-NN algorithm is based on a distance and a voting function in k nearest functions. Euclidean distance is the metric employed in the classifier. k-NN algorithm consists of two phases one is training phase and other one testing phase. In training phase, the data points are fed in a n-dimensional space where n can be any positive integer greater than 1. These training data points have labels associated with them that denominate their class. In testing phase, unlabelled data called data points are given and the algorithm generates the nearest data points to the unlabelled point. The Nearest Neighbour classifier selects a single neighbour in the data set and proceeds further whereas the k-Nearest Neighbour classifier selects k neighbours and then takes the majority of the data set type as the result for the test case. Therefore, it provides good performance for optimal values of k. Finally, the algorithm returns the class of majority of that list [8-10].

k-nearest neighbours algorithm

1) Determine a suitable distance metric.
2) In the training phase: Stores all the training data set P in pairs (according to the selected features) P = (yi, ci), i=1. . .n, where yi is a training pattern in the training data set, ci is its corresponding class and n is the amount of training patterns.
3) During the test phase: Computes the distances between the new feature vector and all the stored features (training data).
4) The k-nearest neighbours are chosen and asked to vote for the class of the new example. The correct classification given in the test phase is used to assess the correctness of the algorithm. If this is not satisfactory, the k value can be tuned until a reasonable level of correctness is achieved (Figure 3).

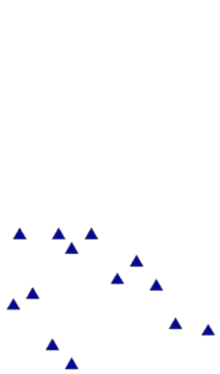

Figure 3 k-Nearest neighbours classification diagram

RESULTS

The images in Figures 4 and 5 show brain MRI of benign type and brain MRI of malignant type respectively.

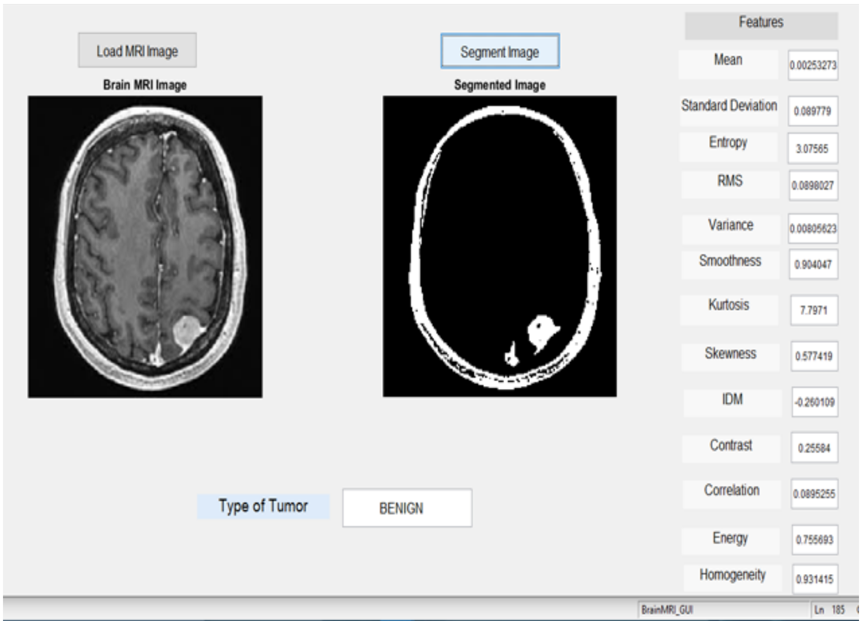

Figure 4 Result of a brain MRI of benign type

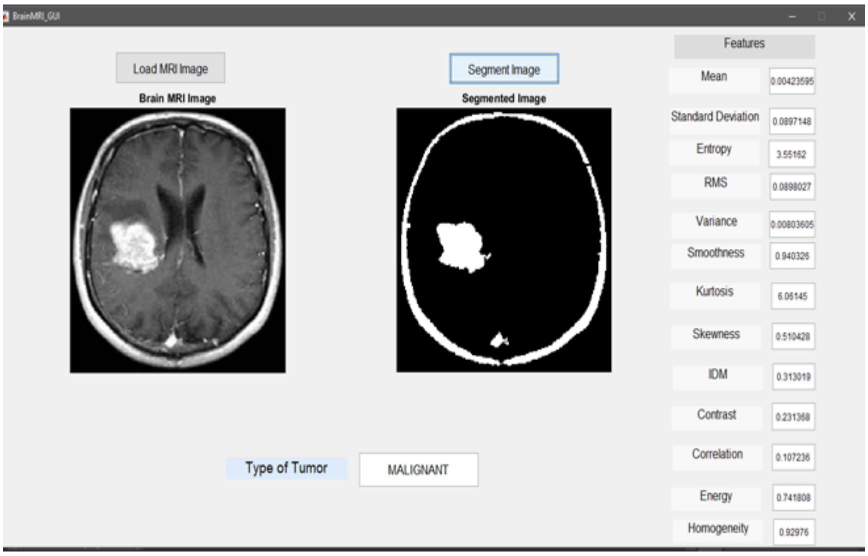

Figure 5 Result of a brain MRI of malignant type

CONCLUSION

An image in the jpeg format of the provided MRI brain image in Dicom format is obtained. The jpeg image is then fed into the code so as to get the type of the tumour as the output. Then the image is segmented using the simple conversion codes. Then feature matrix is extracted from the segmented image using discrete wavelet transform and then the features are reduced by principal component analysis. Then finally k-NN classification technique is used to determine whether the tumour is benign or malignant.

Further enhancements

The further focus is to increase the accuracy of the above algorithms when combined together and also processing Dicom images directly instead of using jpeg images.

REFERENCES

[1] Fletcher-Heath, Lynn M., et al. “Automatic segmentation of non-enhancing brain tumors in magnetic resonance images.” *Artificial Intelligence in Medicine* 21.1 (2001): 43-63.

[2] Chaplot, Sandeep, L. M. Patnaik, and N. R. Jagannathan. “Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network.” *Biomedical Signal Processing and Control* 1.1 (2006): 86-92.

[3] Kara, Sadık, and Fatma Dirgenali. “A system to diagnose atherosclerosis via wavelet transforms principal component analysis and artificial neural networks.” *Expert Systems with Applications* 32.2 (2007): 632-640.

[4] Maitra, Madhubanti, and Amitava Chatterjee. “Hybrid multiresolution Slantlet transform and fuzzy c-means clustering approach for normal-pathological brain MR image segregation.” *Medical Engineering & Physics* 30.5 (2008): 615-623.

[5] Karibasappa, K., and S. Patnaik. “Face recognition by ANN using wavelet transforms coefficients.” *IE (I) Journal-CP* 85 (2004): 17-23.

[6] Hiremath, P.S., S. Shivashankar, and Jagadeesh Pujari. “Wavelet based features for color texture classification with application to CBIR.” *International Journal of Computer Science and Network Security* 6.9A (2006): 124-133.

[7] Sengur, Abdulkadir. “An expert system based on principal component analysis, artificial immune system and fuzzy k-NN for diagnosis of valvular heart diseases.” *Computers in Biology and Medicine* 38.3 (2008): 329-338.

[8] Kociołek, Marcin, et al. “Discrete wavelet transform-derived features for digital image texture analysis.” *International Conference on Signals and Electronic Systems*. 2001.

[9] Latifoğlu, Fatma, et al. "Medical diagnosis of atherosclerosis from Carotid Artery Doppler Signals using principal component analysis (PCA), k-NN based weighting pre-processing and Artificial Immune Recognition System (AIRS)." *Journal of Biomedical Informatics* 41.1 (2008): 15-23.

[10] El-Dahshan, E. A., Abdel-Badeeh M. Salem, and Tamer H. Younis. "A hybrid technique for automatic MRI brain images classification." *Studia Univ. Babes-Bolyai, Informatica* 54.1 (2009): 55-67.