IMPACT: International Journal of Research in Engineering & Technology (IMPACT: IJRET)
ISSN(E): 2321-8843; ISSN(P): 2347-4599
Vol. 2, Issue 7, Jul 2014, 93-104

A NOVEL APPROACH FOR BRAIN TUMOR DETECTION USING NEURAL NETWORK

SAID CHARFI[1], REDOUAN LAHMYED[2] & LALITHA RANGARAJAN[3]

[1,2]Student, Department of Computer Science, Manasagangotri, University of Mysore, Mysore, Karnataka, India

[3]Associate Professor, Department of Computer Science, Manasagangotri, University of Mysore, Mysore, Karnataka, India

ABSTRACT

Computer-aided detection/diagnosis (CAD) systems can enhance the diagnostic capabilities of physicians and reduce the time required for accurate diagnosis. The objective of this paper is to review the recent published segmentation and classification techniques and their state-of-the-art for the human brain magnetic resonance images (MRI). The review reveals the CAD systems of human brain MRI images are still an open problem. In the light of this review we proposed a hybrid intelligent machine learning technique for computer-aided detection system for automatic detection of brain tumor through magnetic resonance images. The proposed technique is based on the following computational methods; the histogram dependent thresholding for image segmentation, the discrete wavelet transform for features extraction, the principal component analysis for reducing the dimensionality of the wavelet coefficients, and the feed forward back-propagation neural network to classify inputs into normal or abnormal. The experiments were carried out on 80 images consisting of 37 normal and 43 abnormal (malignant and benign tumors) from a real human brain MRI dataset. The classification accuracy on both training and test images is 90% which was significantly good. The results revealed that the proposed hybrid approach is accurate and fast and robust. Finally, possible future directions are suggested.

KEYWORDS: Neural Networks, MRI Images, HDT Thresholding, Discrete Wavelet Transform, Principal Component Analysis

INTRODUCTION

Early detection of the brain tumor is very important and the motivation for further studies. In the brain magnetic resonance imaging (MRI), the tumor may appear clearly but for further treatment, the physician also needs the quantification of the tumor area. The computer and image processing techniques can provide great help in analyzing the tumor area.

On the other side, computer-aided detection (CAD) has been developing fast in the last two decades. The main idea of CAD is to assist radiologists in interpreting medical images by using dedicated computer systems to provide ‘second opinions’.

Studies on CAD systems and technology show that CAD can help to improve diagnostic accuracy of radiologists, lighten the burden of increasing workload, reduce cancer missed due to fatigue, overlooked or data overloaded and improve inter- and intra-reader variability. The final medical decision is made by the radiologists. Consequently, radiologists expect that CAD systems can improve their diagnostic abilities based on synergistic effects between the radiologist and the computer with medical image analysis and machine learning techniques.

Therefore, the CAD systems should have abilities similar to the radiologists in terms of learning and recognition of brain diseases. For this reason, pattern recognition techniques including machine learning play important roles in the development of CAD systems.

To create a CAD system, the integration of various image processing operations (techniques) such as image segmentation, feature extraction and selection, and classification are essential. Recently, various types of brain computer-aided detection methods have been developed by a number of researchers, including our group using brain MR images based on several types of machine learning classifiers. The challenge remains to provide CAD systems that work in all cases regardless of the quality and the size of the database. CAD systems of human brain MR images are primarily motivated by the necessity of achieving maximum possible accuracy. Our motivation of this study is to improve the performance of a CAD system for human brain tumor detection.

PROPOSED METHOD

- **Framework of the Proposed Work**

Using the reduced incidence matrix we have to make use in order to develop a CAD system with a low computational cost, we have proposed a hybrid intelligent system. The architecture for the proposed CAD brain MRI system is shown in Figure 1. It comprises four main processes for (i) image acquisition and preprocessing, (ii) segmentation of ROI, (iii) feature extraction and selection, and (iv) classification of the selected ROI and performance evaluation.

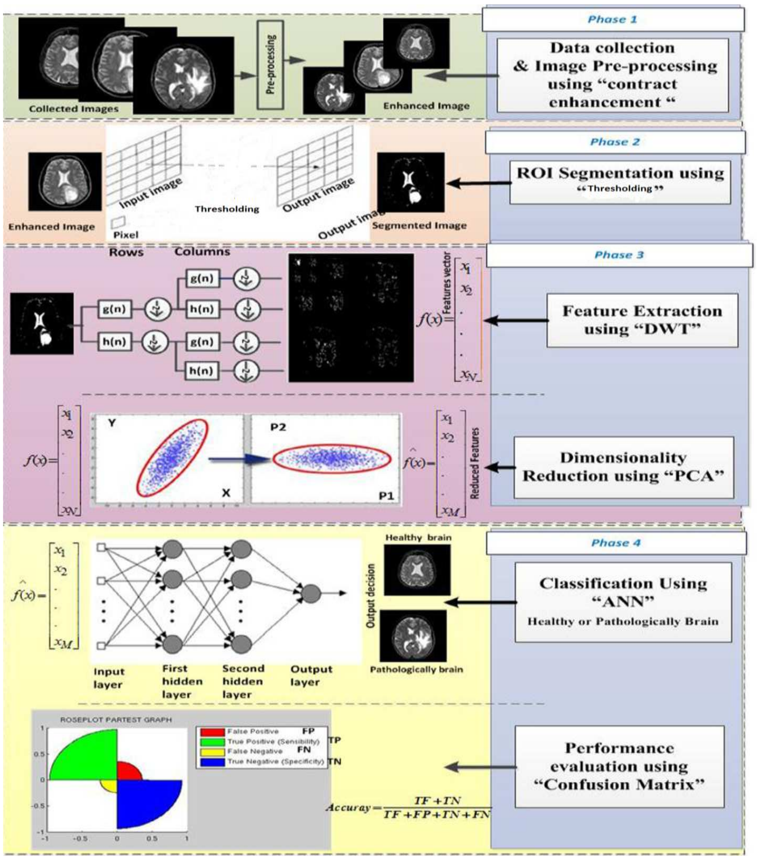

Figure 1: The Proposed Methodology of CAD Brain MRI System

- **Image Acquisition and Pre-Processing**

Image acquisition techniques like MRI, X-Ray, ultrasound, mammography, CT-scan are highly dependent on computer technology to generate digital images. After obtaining digital images, image pre-processing techniques can be further used for analysis of region of interest. A pre-processing is performed in order to remove noise and clean-up the image background. In this stage, preprocessing based on median filter is presented. The preprocessing stage used to improve the quality of the images and make the rest stages more reliable. The median filtering technique is applied to remove the high frequency components in MR images. The advantage of using the median filter is that it removes the noise without disturbing the edges.

- **Segmentation of Region of Interest Using Thresholding**

Image segmentation and defining the region of interest is an important approach and the most time-consuming part of image analysis and processing, which can divide the images into different parts with certain distinctions.

- **Automatic Thresholds**

Automatically selected threshold value for each image by the system without human intervention is called an automatic threshold scheme. This is requirement the knowledge about the intensity characteristics of the objects, sizes of the objects, fractions of the image occupied by the objects and the number of different types of objects appearing in the image.

- **Histogram Dependent Technique (HDT)**

The histogram based techniques is dependent on the success of the estimating the threshold value that separates the two homogonous region of the object and background of an image. This required that, the image formation be of two homogonous and will-separated regions and there exists a threshold value that separated these regions. The (HDT) is suitable for image with large homogonous and will separate regions where all area of the objects and background are homogonous and except the area between the objects and background.

This technique can be expressed as:

$$C(T) = P_1(T)\sigma_1^2(T)+P_2(T)\sigma_2^2(T)$$

Where:

C (T) is the within-group variance.

$P_1(T)$ is the probability for group with values less than T.

$P_2(T)$ is the probability for group with values greater than T.

$\sigma_1(T)$ is the variance of group of pixels with values less than or equal T.

- **Features Extraction Based on Wavelet Transform**

In this study, the feature extraction of MRI images is obtained using the discrete wavelet transform. The wavelet is a powerful mathematical tool for feature extraction (Daube, 1991; Hiremath Shivashankar, & Pujari, 2006). The use of wavelet transform is particularly appropriate since it gives information about the signal both in frequency and time domains. DWT is a frequently used image processing technique which performs the function of transforming images from the spatial domain into the frequency domain. By applying DWT, we are able to decompose an image into the corresponding sub-bands with their relative DWT coefficients.

The DWT is implemented using cascaded filter banks in which the low pass and high pass filters satisfy certain specific constraints. The basic scheme of DWT decomposition and its application to MR images is shown in Figure 2. Where the functions h(n) and g(n) represent the coefficients of the high-pass and low-pass filters, respectively.

As a result, there are four sub-band (LL, LH, HH, HL) images at each scale. The LL sub-band can be regarded as the approximation component of the image, while the LH, HL, HH sub-bands can be regarded as the detailed components of the image. For feature extraction, only the sub-band LL is used for DWT decomposition at next scale. Also, the LL sub-band at last level is used as output feature vector. In our algorithm, a two and three level decomposition via Haar wavelet was utilized to extract features. Figure 2 shows a schematic diagram of 3rd level wavelet transform decomposition Haar Wavelet transform along with its conceptual expression diagram. A layout of DWT sub-bands with three-scale dyadic decomposition of Lena image is shown in Figure 2.

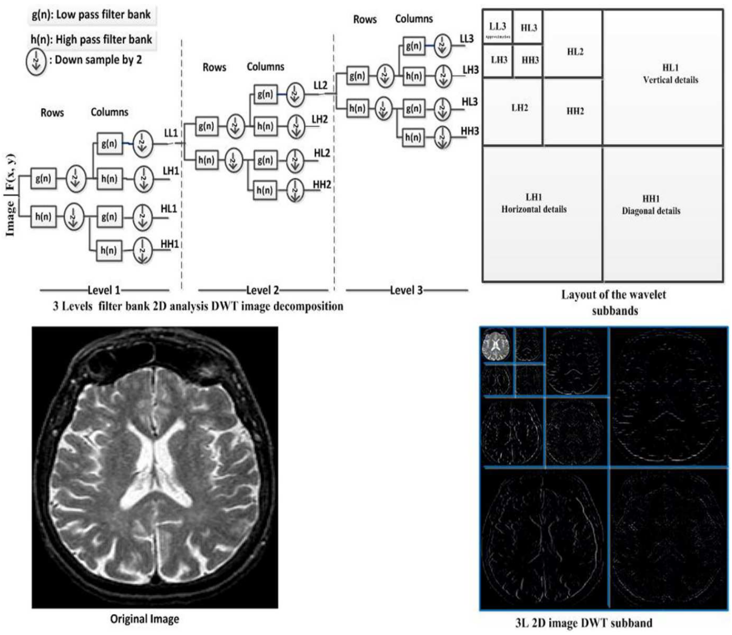

Figure 2: A Schematic Diagram of 3rd Level Wavelet Transform Decomposition

- **Features Reduction Based on Principal Component Analysis (PCA)**

Excessive features increase the computation time and memory storage which sometimes causes some complications in the classification process (the curse of dimensionality), and so it is required to reduce the number of features. The principal component analysis is the most well-known used subspace projection technique as it provides suboptimal solution with a low computational cost and computational complexity. PCA is an efficent strategy for transforming the existing input features of a dataset consisting of a large number of interrelated variables into a new lower-dimension feature space while retaining most of the variations. The input feature space is transformed into a lower dimensional feature space using the largest eigenvectors of the correlation matrix and forms a new set of ordered variables according to their variances or their importance (Jain, Duin, & Mao, 2000).

The implementation of PCA is shown in Figure 3. Additional information and implemented studies about PCA can be reached from (Jolliffe, 1986; Polat & Gunes, 2008; Smith, 2002; Wang & Paliwal, 2003).

PCA is a statistical method used to decrease the data dimensions while retaining as much as possible of the variation present in the data set to process the data faster and effective (Jolliffe, 2002).

This technique has three effects: it orthogonalizes the components of the input vectors so that uncorrelated with each other, it orders the resulting orthogonal components so that those with the largest variation come first, and eliminates those components contributing the least to the variation in the data set. Using a system of feature reduction based on PCA limits the feature vectors to the component selected by the PCA which leads to an efficient classification algorithm. So, the main idea behind using PCA in our approach is to reduce the dimensionality of the wavelet coefficient which results in a more efficient and accurate classifier (Zöllner, Emblem, & Schad, 2012).

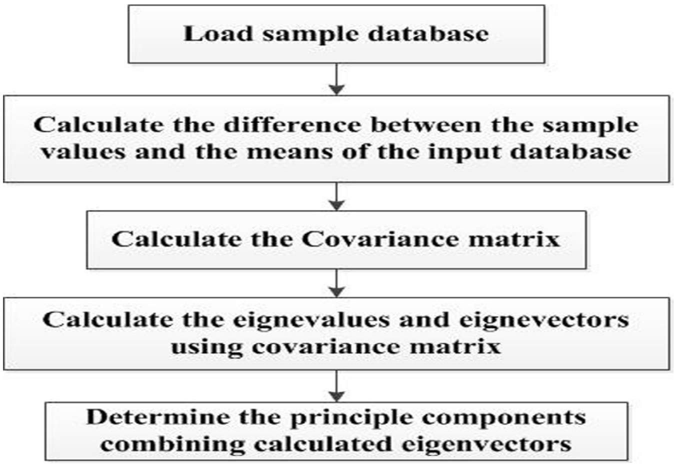

Figure 3: Implementation of PCA

- **Component Analysis (PCA) MRI Images Classification Based on Artificial Neural Networks (ANN)**

Neural networks are widely used in pattern classification since they do not need any information about the probability distribution and the a priori probabilities of different classes. ANN classification system mimics the human reasoning and in some cases, it gives the decision for more than one class to show the possibilities of other diseases. For brain MR image classification, as normal or abnormal, we used a Back-propagation neural network (BPNN) to classify inputs into the set of target categories (normal or abnormal) based on feature selection parameters. BPNN is a supervised learning method which is a non-linear generalization of the squared error gradient descent learning rule for updating the weights of the artificial neurons in a single-layer perceptron, generalized to feed-forward networks (Haykin, 2008).

ANN is a branch of artificial intelligence (AI). It can imitate the way in which a human brain works in processes such as studying, memorizing, reasoning and capable of performing massively parallel computations for data processing and knowledge representation. One advantage of the neural network approach is that most of the intense computation takes place during the training process. Once the ANN is trained for a particular task, operation is relatively fast and unknown samples can be identified.

Generally, an ANN can be defined as a system or mathematical model that consists of many nonlinear artificial neurons running in parallel and may be generated as one-layered or multilayered.

Most ANNs have three layers: input, output, and hidden. The function of the hidden layer is to intervene between the external input and the network output in some useful manner. Detailed theoretical information about ANNs can be found in Haykin (2008). An ANN structure is shown in Figure 4.

Feed forward multilayer neural network (FFNN) is dominantly used. The back propagation algorithm has been used in the training of the FFNN.

Two kinds of signals are identified in this network: The function signals (also called input signals) that come in at the input of the network, propagate forward (neuron by neuron) through the network and reach the output end of the network as output signals; The error signals that originate at the output neuron of the network and propagate backward (layer by layer) through the network. The output of the neural network is described by the following equation:

$$Y = F_0(\sum_{j=0}^{M} W0j\,(F_h\,(\sum_{i=0}^{N} Wji\,Xi)))$$

where W_{0J} represents the synaptic weights from neuron Y in the hidden layer to the single output neuron, X_j represents the ith element of the input vector, F_h and F_0 are the activation function of the neurons from the hidden layer and output layer, respectively, W_{ji} are connection weights between the neurons of the hidden layer and the inputs. The learning phase of the network proceeds by adaptively adjusting the free parameters of the system based on the mean square error E, described by the following equation, between predicted and measured path loss for a set of appropriately selected training examples:

$$E = \frac{1}{2}\sum_{i=1}^{m}(Yi - di)^2$$

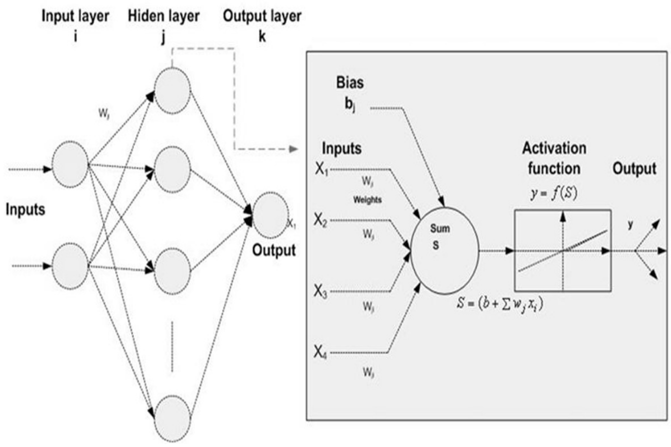

Figure 4: Artificial Neural Network (ANN) Architecture

Where y_i is the output value calculated by the network and d_i represents the expected output. When the error between network output and the desired output is minimized, the learning process is terminated and the network can be used in a testing phase with test vectors. At this stage, the neural network is described by the optimal weight configuration, which means that theoretically ensures the output error minimization.

PERFORMANCE EVALUATION

Quantitative evaluation of the proposed system and it is performance were analyzed using different statistical measures.

Confusion matrix was used to calculate the performance of the CAD system. Confusion matrix contains information about actual and predicted classifications. For evaluating the proposed algorithm based on the confusion matrix, we used the metrics of sensitivity (measures the proportion of actual positives which are correctly identified), specificity (measures the proportion of negatives which are correctly identified), and accuracy (as given in the equations below).

Sensitivity (true positive rate) = TP / (TP + FN)

Specificity (false positive rate) = TN / (TN + FP)

Accuracy (percent of all samples correctly classified)

= (TP + TN) / (TP + TN + FP + FN)

Where: TP: (true positives) is the correctly classified positive cases, TN: (true negative) is the correctly classified negative cases, FP: (false positives) is the incorrectly classified negative cases and FN: (false negative) is the incorrectly classified positive cases.

The receiver operating characteristic (ROC) curve is the plot that displays the full picture of trade-off between the sensitivity and (1-specificity) across a series of cut-off points. Area under the ROC curve (AUC) is considered as an effective measure of inherent validity of a diagnostic test. Partest plot and Rose plot test method is one of the methods in ROC curve method.

Here come some performance results of our Neural Network.

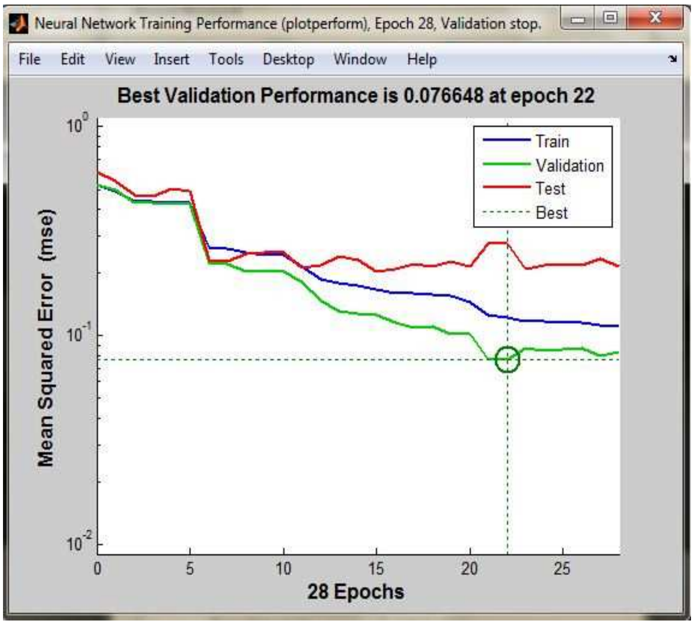

Figure 5: Neural Network Training Performance Plot

The default performance function for feed-forward networks is mean square error *mse*-the average squared error between the network outputs *a* and the target outputs *t*. It is defined as follows:

$$F = mse = \frac{1}{N}\sum_{i=1}^{N}(e_i)^2 = \frac{1}{N}\sum_{i=1}^{N}(t_i - a_i)^2$$

In the graph above we can see the development of the *mse* during the 28 iterations for the training, testing and validation steps. In this graph the *mse* is going down towards the best results expected. At the epoch 22 we get the minimum *mse* which is 0.076648.

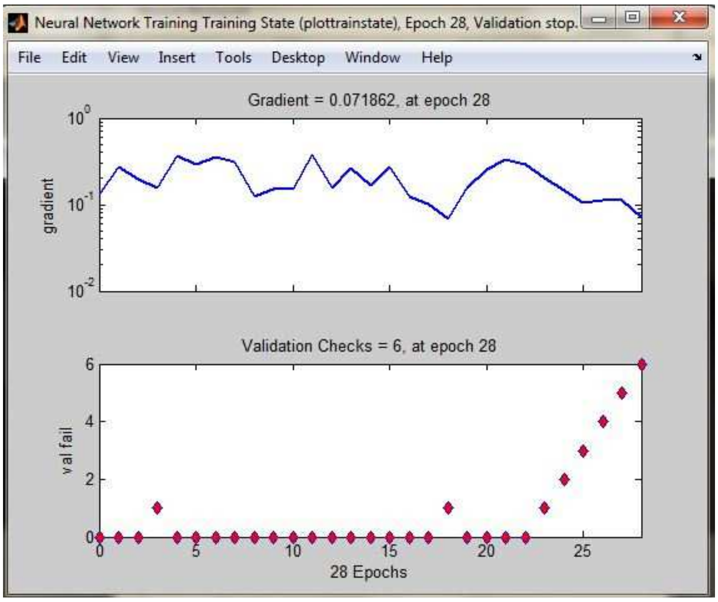

Figure 6: Neural Network Training State Plot

The above figure shows two graphs, the first one up describes the results of the gradient value in every iteration.

The more the gradient value is near to 0 the more the performance of our network is going up.

The second graph represents the number of validations checks in each epoch, we can see that after the epoch 22 the number of validation checks increases rapidly and it reaches 6 at the iteration 28.

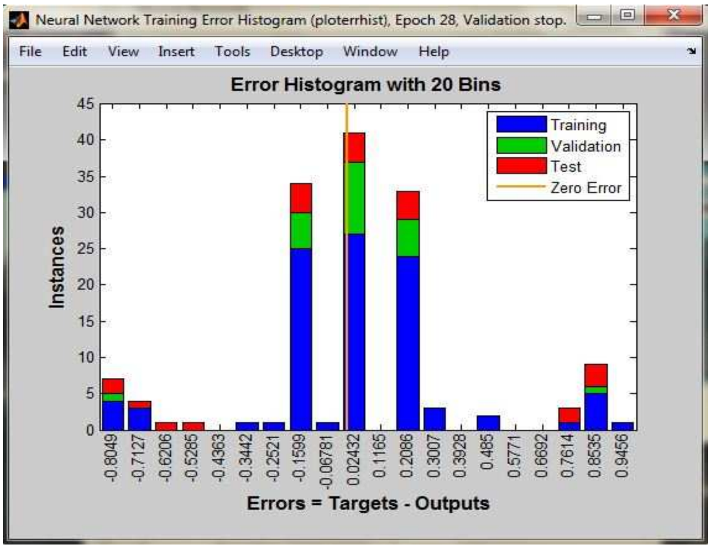

Figure 7: Neural Network Training Error Histogram Plot

This graph represents the Error Histogram which calculated as the difference between the targets of our neural networks and the actual outputs.

Figure 8: Neural Network Training, Testing and Validation States Confusion Matrices

This figure shows the confusion matrix for the training, testing and validation steps, and the all confusion matrix of the neural networks using this matrix we can calculate the specificity, sensitivity and accuracy of our artificial neural network.

A confusion matrix contains information about actual and predicted classifications done by a classification system. Performance of such systems is commonly evaluated using the data in the matrix.

The green indicates the number of inputs assigned correctly to their classes, and the red indicates the misclassification of the inputs.

The black and blue cells indicate the overall results.

The following figure shows the ROC graphs for the training, testing and validation phases of the classifier created. At the last the overall ROC of the system.

A ROC represents the classes by the False Positive Rate in function of the True Positive Rate.

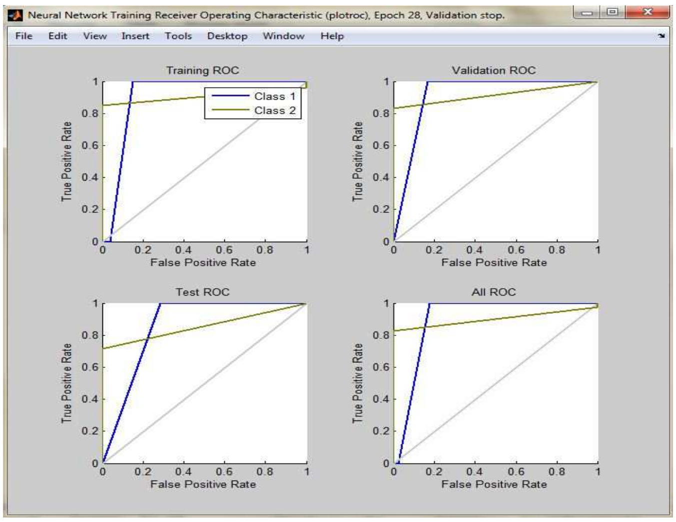

Figure 9: Neural Network Training, Testing and Validation States ROC Plot

A perfect test would show points in the upper-left corner, with 100% sensitivity and 100% specificity.

From the graphs we can see the for the class 1 reaches 100 percent sensitivity and 100 percent specificity very fast whereas for the class 2 it takes some time, but generally the results are quite good.

EXPERIMENTATIONS RESULTS

The images were randomly selected as there are total 80 images consisting of 37 normal and 43 abnormal brain images.

The network configuration is $N_I*N_{HI}*2$(Figure 4 shows the general architecture of the FFNN used for classification), N_I is the input layer and N_{H1} is the hidden layer. The network configuration is $N_I*N_{HI}*2$, such that a Two-layers network with 10 input neurons for the feature vectors selected from the wavelet coefficients by the PCA, 10 neurons in the hidden layer and two neurons in the output layer was used to represent normal and abnormal human brain. Table 2 shows the network parameters used for training.

The experimental results for normal and abnormal classification are listed in Table 1. The effectiveness of our approach has been demonstrated in Table 3, with a small set of data. The classification accuracy of 90%, sensitivity of 82%, specificity of 100%. Also, the accuracy of a model in making predictions is evaluated regularly using an ROC analysis. An ROC curve is generated by combining the true positive fraction (sensitivity) and false positive fraction (1-specificity) by setting different thresholds. A quantitative measure of the accuracy of the classification technique is obtained by finding the area under the ROC curve (AUC) which varies between 0.0 indicating poor classification performance, and 1.0 indicating high classification performance. The area of the ROC curve results implied that our hybrid technique can provide a consistency high accuracy for classification of human brain MR images. In this methodology the AUC = 90.

The results show that our method obtains quite perfect results on both training and test images.

Table 1: Setting of Training and Testing Images

Total No. of Images	No. of Images in Training Set(60)		No. of Images in Testing Set(20)	
	Normal	Abnormal	Normal	Abnormal
80	20	28	17	15

Table 2: The Network Parameters Used for Training

Parameters	ANN
Number of input layer units	10
Number of hidden layer	1
Number of first hidden layer unit	10
Number of output layer units	2
Maximum number of epochs to train	1000
Minimum performance gradient	1e – 10
Learning rate	0.01

Table 3: Classification Rates

TP	TN	FP	FN	Sensitivity (%)	Specificity (%)	Accuracy (%)
33	31	7	0	100	82	90

CONCLUSIONS AND FUTURE WORK

Magnetic Resonance Imaging (MRI) is an important examination and diagnosis method for brain tumors in medical imaging. With a sound mechanism and clear imaging of soft tissues, the doctor on the patient's diagnosis can be scientific and rational, to grasp the exact progression of the disease state, which would set out the appropriate treatment, surgery and following-up to a series of disease control measures.

Computer-aided analysis is to reduce the workload of doctors, to improve the diagnostic accuracy of the para-medical analysis, and meanwhile to improve the automatic degree in practice.

With the advance of computational intelligence and machine learning techniques, computer-aided detection attracts more attention for brain tumor detection. It has become one of the major research subjects in medical imaging and diagnostic radiology. In this study, we reviewed current studies of the different segmentation, feature extraction and classification algorithms. In particular, this paper reviews recent papers which are between 2006 and 2013. In light of this, we proposed a hybrid technique for processing of MRI brain images. The proposed technique first applies automated thresholding as a front-end processor for image segmentation and detecting the region of interest, and then employs the discrete wavelet transform to extract features from MRI images. Moreover the principal component analysis is performed to reduce the dimensionality of the wavelet coefficients which results in a more efficient and accurate classifier. The reduced features are sent to back-propagation neural network to classify inputs into normal or abnormal based on feature selection parameters.

A preliminary evaluation on MRI brain images shows encouraging results, which demonstrates the robustness of the proposed technique. We have realized a large number of algorithms that could also be applied to the developed system and compare the results with this one. According to the experimental results, the proposed method is efficient for automated diagnosis of brain diseases. Our proposed method produces classification accuracy of 90% with 100% sensitivity rate and 82% specificity rate. These experiment results show that the proposed classifier method can successfully differentiate between healthy and pathologically cases and can increase the diagnostic performance of human brain abnormality.

The challenge remains to provide generalized CAD systems that work in all cases regardless of database size and quality. So, CAD system remains an open problem.

As present there is not a public test database of high capacity of MRI sequences internationally, and common evaluation criteria are still relatively simple, so in the future work, we will attempt to contact other laboratories carrying out similar research to share MRI data and work together in order to establish a large-capacity, open test database, and continue to test the proposed method with the aid of the new database to validate its effectiveness and further improvement.

There are several future directions which might further improve the CAD systems for human brain MR images:

- The acquisition of large databases from different institutions with various image qualities for clinical evaluation and improvement in the CAD systems.
- Improve the classification accuracy by extracting more efficient features and increasing the training data set.

- There is still much room for additional researcher to utilize other machine learning techniques and integrate them into a hybrid one system.
- Further experiments and evaluation are therefore desirable to establish whether the proposed approaches have generic applications.

Finally we hope that in the entire study, existing computer-aided analysis and evaluation systems on medical images can be improved and new evaluation criteria can be proposed and applied in the future works.

REFERENCES

1. Danciu, M. G. (2013). *A hybrid 3D learning-and-interaction-based segmentation approach applied on CT liver volumes.* Radio Engineering. 22(1), 100-113.
2. Daubechies, I. (1991). Ten lectures on wavelets. *CBMS-NSF Series in Applied Mathematics* (SIAM).
3. Duda, R., Hart, P., & Stork, D. (2001). *Pattern classification*. New York: John Wiley & Sons Inc.
4. El-Dahshan, E. A., Hosny, T., & Salem, A. B. M. (2010). Hybrid intelligent techniques for MRI brain images classification. *Digital Signal Processing*, 20(2), 433–441.
5. Haykin, S. (2008). *Neural networks and learning machines* (3rd ed.). New Jersey: Pearson Prentice Hall.
6. Jain, A. K., Duin, R. P. W., & Mao, J. (2000). Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22, 4–37.
7. Jolliffe, I. T. (2002). *Principal component analysis* (2nd ed.). New York: Springer Verlag.
8. Mathworks. (2012). *Neural Networks toolbox*. Accelerating the pace of engineering and science.
9. M.C. Clark, L.O. Hall, D.B. Goldg of, et al. Automatic Tumor Segmentation Using Knowledge-based Techniques [J]. *IEEE Transactions on Medical Imaging*, 1998, 17: 238-251.
10. M.R. Kaus, S.K. Warfield, P.M. Black, et al. *Automated Segmentation of MR ages of Brain Tumors* [J]. Radiology, 2001, 218: 586-591.
11. M. Sezgin, B. Sankur. Survey over Image Thresholding Techniques and Quantitative Performance Evaluation [J]. *Journal of Electronic Imaging*, 2004, 13 (1): 146–165.
12. N. Otsu. A Threshold Selection Method from Gray-Level Histograms [J]. *IEEE Transactions on System, Man and Cybernetics*, 1979, 9 (1): 62-66.
13. R. C. Gonzalez, R. E. Woods, "*Digital Image Processing*", 2nd Ed.
14. Tanga, H., Wua, E. X., Mab, Q. Y., Gallagherc, D., Pereraa, G. M., & Zhuang, T. (2000). MRI brain image segmentation by multi-resolution edge detection and region selection. *Computerized Medical Imaging and Graphics*, 24(6), 349–357.
15. Zöllner, F. G., Emblem, K. E., & Schad, L. R. (2012). SVM-based glioma grading: Optimization by feature reduction analysis. *Zeitschrift für Medizinische Physik*, 22(3), 205–214.