A Survey on Brain Tumor Classification Using Artificial Neural Network

M .Queen, T.M. Babi Mol M.E

Abstract_ **Magnetic Resonance imaging (MRI) has become a widely used method of high quality medical imaging. Brain tumor classification is one of the major problems in diagnosing the tumor at early stage. Thus various methods are surveyed in order to obtain better classification accuracy and to reduce the computational time. Since misclassification occurs due to high diversity in tumor appearance and tumor boundaries. The various image processing techniques preprocessing, feature selection and extraction are used to detect exact tumor location. This study classifies brain tumor MRI images automatically as benign, grade1, grade2 and malignant tumor. Classification of tumor is done through Artificial Neural Network. The result of performance ensures that the GLCM with Neural Network Classifier provides about 95% accuracy.**

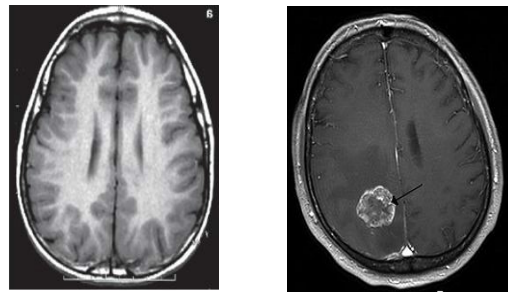
Index Terms: **Brain Tumor, Classification, GLCM, MRI, Neural Network**.

I INTRODUCTION

A brain tumor is any intracranial tumor generated by abnormal and uncontrolled cell division normally found anywhere in brain. Benign brain tumors contain cells that look healthy just like normal cells. They have growth slowly are likely to spread although these tumor may cause damage if they initiate to interfere with normal brain function. But malignant brain tumors have irregular borders and made up of abnormally shaped cells. Figure shows the MRI image of normal brain and tumor brain.

One of the important goals of Artificial Neural Networks is the processing of information similar to humor interaction actually neural network is used when there is a need for brain capabilities and machine idealistic. The advantages of neural network information processing arise from its ability to recognize and model nonlinear relationships between data.

M.Queen,*Communication System , PET Engineering College,India.*
T.M Babimol,*Electronics and Communication Department, PET Engineering College , India*



a)Normal Brain b)Tumor Brain

Fig 1: MRI image of Normal and Tumor Brain

Conventional statistical methods can be used to model nonlinear relationships, but they require complex and extensive mathematical modeling Neural Network provide comparatively easier way to do the same type of analysis. Well design and training of Neural Network make it qualified for decision making operations when it faced with new data outside training data; this will provide ANNs with high reliability exactly like expert person. There are two problems that face ANN's designers in any application comprising of

i) Network structure.
ii) Network generalization.

In designing ANNs, a suitable architecture for the specific application must be well chosen this involves:

a. Choosing a suitable network type .
b. Number of layers,
c. Number of nodes in hidden layers,
d. Activation functions between layers.

Network Generalization means how much the neural network is able to work with different data. Designers of ANN are always faced by the great extend of network generalization, i.e. despite a well designing and training of ANN that decreases the performance error to the least value. Image preprocessing is the first step of image retrieval to ensure accuracy of subsequent steps. The images acquired through different modalities cause many artifacts such as low resolution, noise and extra cranial tissue etc. which reduces the accuracy of acquired result .In order to overcome the above problem preprocessing of an

: 2278 – 1323

image is required..An analysis on filtering techniques such as Gabor & QMF filters for noise is performed by [1].These primitive methods along with reducing the noise blur the important and detailed structure necessary for subsequent steps. In order to increase the processing speed and to reduce the error probability of mammogram images Morphological top hat filtering algorithm is utilized in [2].But these types of filtering are applicable for mammogram images only. To eliminate the noise in images, Gaussian filter is suggested in[3].The advantage of this filter is to reduce the noise as well to increase the contrast and intensity of an image. To improve quality histogram equalization, edge detection, noise filtering and thresholding is proposed in [4]. Diffusion filtering combined with simple non-adaptive intensity thresholding is used to enhance the region of interest in [5]. Apart from noise removal contrast enhancement is considered in Adaptive Histogram Equalization .Thus AHE is proposed for further processing.

Feature Extraction stage is an important stage that uses algorithms and techniques to detect and isolate various desired portions or shapes of a given image .Binarization technique is used in[1] Binarization approach has been applied for detection of cancer .Gray level Co occurance Matrix is used in [2]. This 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.[5]Feature extraction technique used is PCA .It is one of the most successful techniques that have been used in image recognition and in compression. The purpose of PCA is to reduce the large dimensionality of the data. Thus GLCM is proposed for Feature Extraction Technique. Classification is to performance classify the type of tumor .Classification accuracy mainly depends on the training of the network. There is different variety of classifiers and their performance is evaluated based on the classifier accuracy.

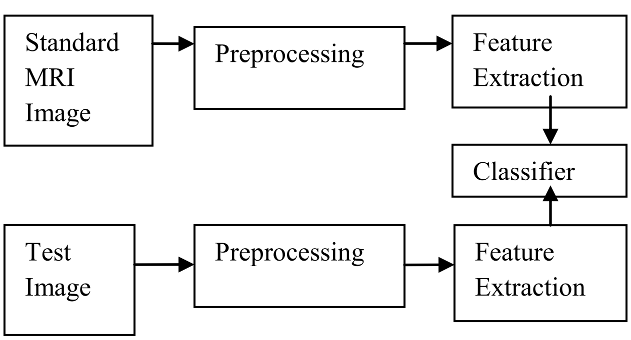


Fig 2:General System Flow Design

II METHODOLOGY

A.PROBABILISTIC NEURAL NETWORK

This network provides a general solution to pattern classification problems by following an approach developed in statistics, called Bayesian classifiers [3,4].PNN is adopted for it has many advantages. Its training speed is many times faster than a BP network. PNN can approach a Bayes optimal result under certain easily met conditions [4]. Additionally, it is robust to noise examples. Advance hybrid PNN such as done by Georgiadas et all [5] aimed to improve brain tumor characterization on MRI by using PNN and the non-linear transformation of textured features. This method employs a two level hierarchical decision tree discriminate the metastatic brain tumor cases from the gliomas and meningiomas (primary brain tumor) cases. At each level, classification was performed using two different LSFTPNN classifier. LSST-PNN then was compared with the support Vector Machines with Radial Basis Kernel (SVM-RBF) and the Artificial Neural Network (ANN) classifiers. The most important advantage of PNN is that training is easy and instantaneous [4]. Weights are not “trained” but assigned. Existing weights will never be alternated but only new vectors are inserted into weight matrices when training. So it can be used in real-time. Since the training and running procedure can be implemented by matrix manipulation, the speed of PNN is very fast. The network classifies input vector into a specific class because that class has the maximum probability to be correct. The PNN has three layers: the Input layer, Radial Basis Layer and CompetitiveLayer.

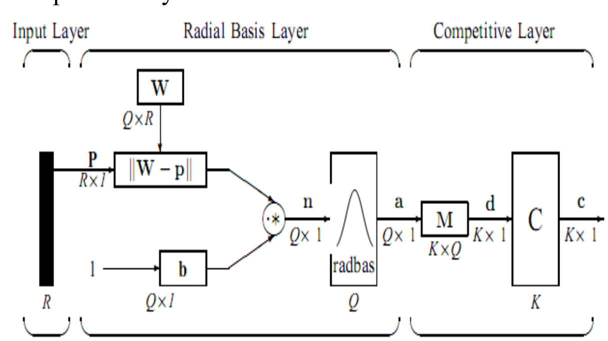


Fig 3:Network structure

International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 3 Issue 11, November 2014

Radial Basis Layer evaluates vector distance between input vector and row weight vectors in weight matrix. These distances are scaled by Radial Basis Function nonlinearly. Then the Competitive Layer finds the shortest distance among them, and thus finds the training pattern closest to the input pattern based on their distance.

B.FEED FORWARD ANN CLASSIFIER

A three layer Neural network is created with 500 nodes in the first (input) layer, 1 to 50 nodes in the hidden layer, and 1 node as the output layer. The number of nodes is varied in the hidden layer in a simulation in order to determine the optimal number of hidden nodes. This was to avoid over fitting or under fitting the data. Due to hardware limitations, ten nodes in the hidden layer were selected to run the final simulation .Figure shows the design of the Feed Forward Neural network .The 500 data points extracted from each subject were then used as inputs of the neural networks. The output node resulted in either a 0 or 1, for control or patient data respectively. Since the nodes in the input layer could take in values from a large range, a transfer function was used to transform data first, before sending it to the hidden layer and then was transformed with another transfer function before sending it to the output layer. In this case, a tan sigmoid transfer function was used between the input and hidden layer, and a log sigmoid function was used between the hidden layer and the output layer.

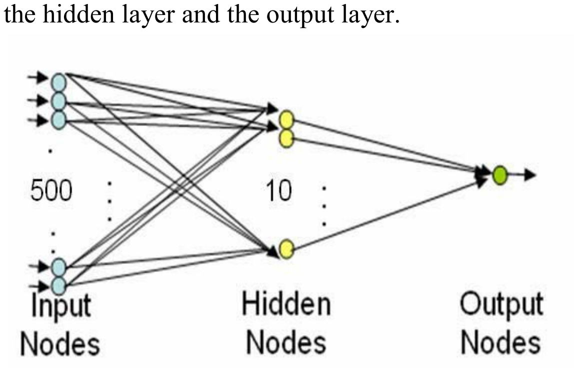


Fig 4:Feed Forward Neural Network

The weights in the hidden node needed to be set using "training" data. Therefore, subjects were divided into training and testing datasets. Out of the 69 subjects, 2 random patients and 2 random controls were selected as test data, while the rest of the dataset was used for training. Training data was used to feed into the neural networks as inputs and then knowing the output, the weights of the hidden nodes were calculated using back propagation algorithm. 120 trials were performed on the same Neural Network, selecting 65 subjects randomly every time for retraining and 4 remaining subjects for testing to find accuracy of Neural network prediction.

BACK PROPAGATION ARTIFICIAL NEURAL NETWORK CLASSIFIER

The most widely used neural-network learning method is the BP algorithm. Learning in a neural network involves modifying the weights and biases of the network in order to minimize a cost function. The cost function always includes an error term a measure of how close the network's predictions are to the class labels for the examples in the training set. Additionally, it may include a complexity term that reacts a prior distribution over the values that the parameters can take. The activation function considered for each node in the network is the binary sigmoid function defined (with s = 1) as output = 1/(1+e-x), where *x* is the sum of the weighted inputs to that particular node. This is a common function used in many BPN. This function limits the output of all nodes in the network to be between 0 and 1. Note all neural networks are basically trained until the error for each training iteration stopped decreasing.

.

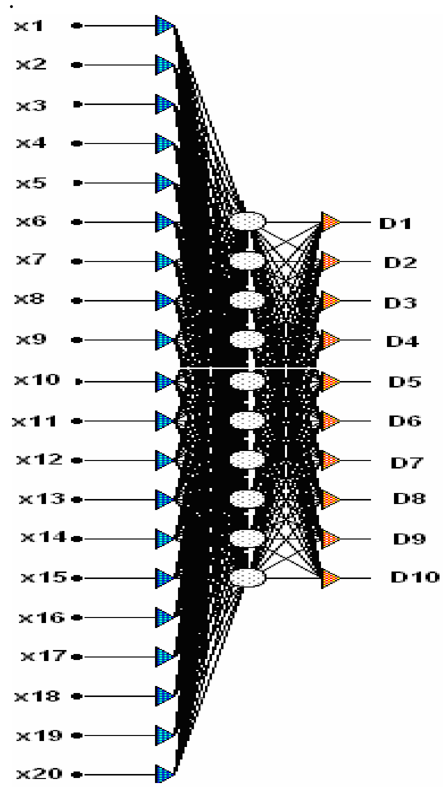


Fig 5:Back Propagation Network

: 2278 – 1323

Figure 4 shows the architecture of the specialized network for the prediction of stroke disease. The complete set of final data (20 inputs) are presented to the generic network, in which the final diagnosis corresponds to output units. The net inputs and outputs of the j hidden layer neurons
can be calculated as follows

$$net_j^h = \sum_{i=1}^{N+1} W_{ji} \, X_i$$
$$Y_j = f\left(net_j^h\right)$$

Calculate the net inputs and outputs of the k output layer neurons are

$$net_k^0 = \sum_{j-1}^{j+1} V_{kj} \, Y_j$$

$$Z_K = f(net_k^0)$$

Update the weights in the output layer (for all k, j pairs)

$$V_{kj} \leftarrow V_{kj} + c \curlywedge (d_k - z_k) z_k (1 - z_k) y_j$$

Update the weights in the hidden layer (for all i, j pairs)

$$W_{ji} \leftarrow W_{ji} + c \curlywedge^2 Y_j (1 - y_j) X_i \sum_{k-1}^{k} (d_k - z_k) z_k (1 - z_k)$$

Update the error term

$$E \leftarrow E + \sum_{k-1}^{k} (d_k - z_k)$$

and repeat from Step 1 until all input patterns have been presented (one epoch). If *E* is below some predefined tolerance level, then stop. Otherwise, reset *E = 0*, and repeat from Step 1 for another epoch.

D.Pulse Coupled Neural Network

A PCNN neuron contains two main compartments: the Feeding and Linking compartments Each of these communicates with neighboring neurons through the synaptic weights M and W respectively. Each retains its previous state but with a decay factor. Only the Feeding compartment receives the input stimulus, S. The values of these two compartments are determined by,

$$F_{ij}[n] = F_{ij}[n-1] + S_{ij} + V_L \sum M_{ijkl} Y_{kl}[n-1]$$
$$L_{ij}[n] = L_{ij}[n-1] + V_L \sum W_{ijkl} Y_{kl}[n-1]$$

Where $F_{ij}[n]$ is the Feeding compartment of the (*i, j*) neuron embedded in a 2D array of neurons, and $L_{ij\,[n]}$ is the corresponding Linking compartment. Y_{kl} 's the outputs of neurons from a previous iteration $[n-1]$. Both compartments have a memory of the previous state, which decays in time by the exponent term. VF and VL are normalizing constants. If the receptive fields of *M* and *W* change then these constants are used to scale the resultant correlation to prevent saturation.

The states of these two compartments are combined in a second order fashion to create the internal state of the neuron, *U*. The combination is controlled by the linking strength, *β*. The internal activity is calculated by,

$$U_{ij}[n] = F_{ij}[n][1 + \beta] L_{ij}$$

The internal state of the neuron is compared to a dynamic threshold θ, to produce the output, Y, by [n]
Y=1 if $U_{ij\,[n]} > \theta_{ij}$ &Y=0 otherwise

The threshold is dynamic in that when the neuron fires (Y>θ) the threshold then significantly increases its value. This value then decays until the neuron fire again.

E. Neural Network Classifier

The Back Propagation technique is akin to supervised learning as the network is trained with the expected replies. Each iteration modifies the connection weights in order to minimize the network error .Adjustment of weights, layer by layer is calculated from the output layer back to the input layer. The learning rate plays a major role in training. When the rate is low then the convergence of the weight to an optimum is very slow. When the rate is too high the network can oscillate .The number of records in the dataset and the number of iterations that determine the training duration. The number of neurons in the hidden layer was empirically determined .First the number of neurons is increased one by one from 1 to 20then five by five to 60.This gives 28 different number of configuration .For each configuration 10 training session were performed with randomly different initial weight factors .Neural network are an alternative method for classification .They work with large number of qualitative variables such as behaviors, provided that they can be coded and they are able to used on linear linked variables .Moreover once they are saved NN can easily be reused to immediately classify additional records as with discriminate analysis.

III PERFORMANCE EVALUATION

Thus after evaluating the performance of different classifiers with respect to classifer accuracy.It is found the Neural network classifier has better classifer accuracy.The following graph gives clear view of different the classifiers classification accuracy.Finally this study clearly shows that Neural

: 2278 – 1323

Network Classifier is better for Brain tumor classification.

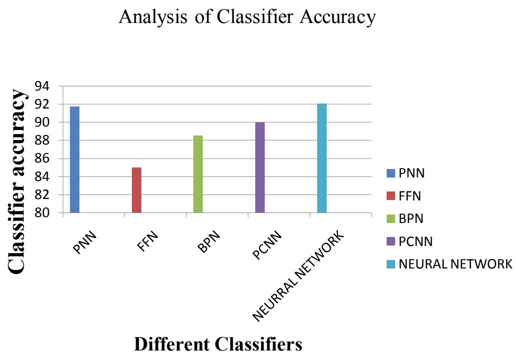


Fig 6: Performance Analysis of Different Classifiers

IV CONCLUSION AND FUTURE WORK

Thus an optimal classifier for the detection and classification of brain tumor was surveyed. The proposed technique gives very promising results comparing with other techniques. In future this technique can be used to detect other type of cancer like breast cancer.

V REFERENCES

[1]AlaaELEYAN and Hasan DEMIREL, "Cooccurence matrix and its statistical features as a new approach for face recognition", Turk J Elec Eng & Comp Sci, Vol. 19, No. 1, 2011

[2] N. Albert Singh, V. Amsaveni," Detection of Brain Tumor using Neural Networks", International Journal ICCCNT , vol 5 No 2 ,May 2013 On Image Processing

[3] D. Bhattacharyya and T. H. Kim, "Brain tumor detection using MRI image analysis," Communications in Computer and Information Science, vol. 151, pp. 307-314, 2011

[4] Chaplo .S, 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, vol. 1, no. 1, pp. 86–92, 2006.

[5] Dipali M. Joshi, Dr.N. K. Rana,"Classification of Brain Cancer Using Artificial Neural Network", IEEE transaction on Image Processing, Vol.7,No.2,February 2013.

[6] Eltaher Mohamed Hussein, Dalia Mahmoud Adam Mahmoud," Brain Tumor Detection Using Artificial Neural Network", Journal on Science and Technology vol,13Dec 2013.

[7] H. Khotanlou, O. Colliot, J. Atif, and I. Bloch, "3D brain tumor segmentation in MRI using fuzzy classification, symmetry analysis and spatially constrained deformable models," Fuzzy Sets and Systems, vol. 160, no. 10, pp. 1457-1473, 2009.

[8] Meiyan Huang, Wei Yang, Yao Wu, Jun Jiang, Wufan Chen, Senior Member, IEEE, and Qianjin Feng" Brain Tumor Segmentation Based on Local Independent Projection-based Classification "2013 IEEE Transactions on Biomedical Engineering

[9] Mohd Fauzi Othman and Mohd Ariffanan Mohd Basri," Probabilistic Neural Network For Brain Tumor Classification" 2011 Second International Conference on Intelligent Systems, Modelling and Simulation.

[10] R.Nayak, L.C.Jain and B.K.H. Ting. Artificial Neural Networks in Biomedical Engineering:A Review. In Proceedings Asia-Pacific Conference on Advance Computation.2001

[11] VinodKumar,JainySachdeva,"Classification of brain tumors using PCA-ANN" 2011 World Congress on Information and Communication Technologies.

[12] D.F. Specht, "Probabilistic Neural Networks" Neural Networks, vol. 3, No.1, pp. 109-118, 1990.

: 2278 – 1323