

October 2012

## Brain Tumor Classification Using Neural Network Based Methods

Kailash D. Kharat

Govt. Engineering College, Aurangabad, Maharashtra, India., kailashdkharat@yahoo.co.in

Pradyumna P. Kulkarni

Govt. Engineering College, Aurangabad, Maharashtra, India., kul.gokul@gmail.com

Follow this and additional works at: <https://www.interscience.in/ijcsi>



Part of the [Computer Engineering Commons](#), [Information Security Commons](#), and the [Systems and Communications Commons](#)

---

### Recommended Citation

Kharat, Kailash D. and Kulkarni, Pradyumna P. (2012) "Brain Tumor Classification Using Neural Network Based Methods," *International Journal of Computer Science and Informatics*: Vol. 2 : Iss. 2 , Article 9. Available at: <https://www.interscience.in/ijcsi/vol2/iss2/9>

This Article is brought to you for free and open access by Interscience Research Network. It has been accepted for inclusion in International Journal of Computer Science and Informatics by an authorized editor of Interscience Research Network. For more information, please contact [sritampatnaik@gmail.com](mailto:sritampatnaik@gmail.com).

# Brain Tumor Classification Using Neural Network Based Methods

Kailash D.Kharat<sup>1</sup> & Pradyumna P.Kulkarni<sup>2</sup> & M.B.Nagori<sup>3</sup>

<sup>1&2</sup> Govt. Engineering College, Aurangabad, Maharashtra, India.

<sup>3</sup> Dept. of Computer Science and Engg., Aurangabad, Maharashtra, India

E-mail- kailashdkharat@yahoo.co.in , kul.gokul@gmail.com

---

**Abstract** - MRI (Magnetic resonance Imaging) brain tumor images Classification is a difficult task due to the variance and complexity of tumors. This paper presents two Neural Network techniques for the classification of the magnetic resonance human brain images. The proposed Neural Network technique consists of three stages, namely, feature extraction, dimensionality reduction, and classification. In the first stage, we have obtained the features related with MRI images using discrete wavelet transformation (DWT). In the second stage, the features of magnetic resonance images (MRI) have been reduced using principle component analysis (PCA) to the more essential features. In the classification stage, two classifiers based on supervised machine learning have been developed. The first classifier based on feed forward artificial neural network (FF-ANN) and the second classifier based on Back-Propagation Neural Network. The classifiers have been used to classify subjects as normal or abnormal MRI brain images. Artificial Neural Networks (ANNs) have been developed for a wide range of applications such as function approximation, feature extraction, optimization, and classification. In particular, they have been developed for image enhancement, segmentation, registration, feature extraction, and object recognition and classification. Among these, object recognition and image classification is more important as it is a critical step for high-level processing such as brain tumor classification. Multi-Layer Perceptron (MLP), Radial Basis Function (RBF), Hopfield, Cellular, and Pulse-Coupled neural networks have been used for image segmentation. These networks can be categorized into feed-forward (associative) and feedback (auto-associative) networks..

**Keywords**-MRI; Feature Extraction; Feature Selection; Tumor Classification; Feed forward Neural Network; Back-Propagation Neural Network.

---

## I. INTRODUCTION

Early detection and classification of brain tumors is very important in clinical practice. Many researchers have proposed different techniques for the classification of brain tumors based on different sources of information. In this paper we propose a process for brain tumor classification, focusing on the analysis of Magnetic Resonance (MR) images and Magnetic Resonance Spectroscopy (MRS) data collected for patients with benign and malignant tumors. Our aim is to achieve a high accuracy in discriminating the two types of tumors through a combination of several techniques for image segmentation, feature extraction and classification. The proposed technique has the potential of assisting clinical diagnosis.

Necessary preprocessing steps prior to characterization and analysis of regions of interest (ROIs) are segmentation and registration. Image registration is used to determine whether two subjects have ROIs in the same location. However, in this work we do not take into account the location of the tumor in

the classification model so we do not employ registration. Image segmentation is required to delineate the boundaries of the ROIs ensuring, in our case, that tumors are outlined and labeled consistently across subjects. Segmentation can be performed manually, automatically, or semi-automatically. The manual method is time consuming and its accuracy highly depends on the domain knowledge of the operator. Specifically, various approaches have been proposed to deal with the task of segmenting brain tumors in MR images. The performance of these approaches usually depends on the accuracy of the spatial probabilistic information collected by domain experts. In previous work, we proposed an automatic segmentation algorithm that is based on the fuzzy connectedness concept. The main idea is to assign to every pair of voxels,  $x, y$ , in the image, a real number between 0 and 1 indicating their connectedness. Starting with several seed points, all the voxels are automatically assigned to the structure to which they have the highest connectedness value. Utilizing the statistical information cumulated during the segmentation process, this method can provide

satisfying results even in cases where the boundaries of the ROIs cannot be easily identified.

Having segmented the ROI and in order to build a classification model, one needs to extract a set of discriminative features from the ROI. Most characterization techniques are based on extracted global visual features that refer to the entire image rather than to regions that are of interest. However, in medical images, feature extraction has to focus on specific regions and capture not only shape but also structural and internal volume properties that can be useful for building a classification model. Megalookonomou et al. proposed a method that efficiently extracts a  $k$ -dimensional feature vector using concentric spheres in 3D (or circles in 2D) radiating out of the ROI's center of mass. The method has been applied successfully to classification and similarity searches of spatial ROIs. In this paper, we propose an approach (see Figure 1) for building a classification from the MR images, and a group of features is extracted. Instead of employing all of the features to build the model, a preprocessing step of feature selection is model performed aiming to remove the redundant features. Based on the statistical information, only the most for informative features extracted from the MR images are utilized in the model building process. In addition, in this

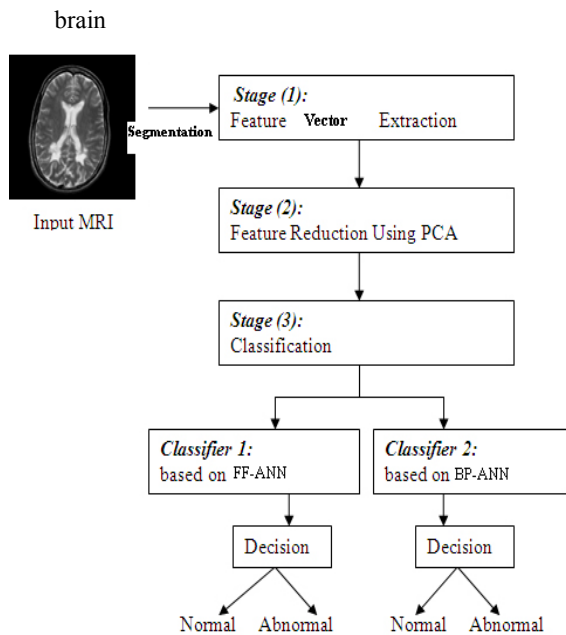


Figure 1. Proposed Work Model

paper, we consider features from other sources (e.g., MRS data) in the classifier training process. This leads to improved classification accuracy.

## II. METHODOLOGY

There are four major steps in the proposed approach for brain tumor classification: (a) ROI segmentation: delineating the boundary of the tumor (ROI) in an MR image; (b) feature extraction: getting meaningful features of the ROI identified in the previous step; (c) feature selection: removing the redundant features; (d) classification: learning a classification model using the features.

### A. Segmentation

Within the segmentation process, each image region confined by a rectangular window is represented by a feature vector of length  $\mathbf{R}$ . These vectors computed for  $\mathbf{Q}$  selected regions are organized in the pattern matrix  $\mathbf{P}_{R,Q}$  and form clusters in the  $R$ -dimensional space. The  $\mathbf{Q}$  pattern vectors in  $\mathbf{P}$  are fed into the input NN layer, while the number  $C$  of the output layer elements represents the desired number of segmentation classes. In each epoch of the network training process, the network weights  $WC,R$  are recalculated by minimizing the distances between each input pattern vector and the corresponding weights of the winning neuron characterized by its coefficients closest to the current pattern. In case that the process is successfully completed, the network weights belonging to separate output elements represent typical class individuals. In this paper, the region segmentation process comprises of training the NN on all image regions extracted by a rectangular sliding window with half overlap, and subsequent exploitation of the trained network for region classification. The algorithm comprises of the following successive steps:

1. Feature vectors computation to create the feature matrix  $P$  using the sliding window
2. Initialization of the learning process coefficients and the network weights matrix  $W$
3. Iterative application of the competitive process and the Kohonen learning rule [10] for all feature vectors during the learning stage
4. NN simulation to assign class numbers to individual feature vectors
5. Evaluation of the regions classification results

### B. Feature Extraction

The proposed system uses the Discrete Wavelet Transform (DWT) coefficients as feature vector. The wavelet is a powerful mathematical tool for feature extraction, and has been used to extract the wavelet coefficient from MR images. Wavelets are localized basis functions, which are scaled and shifted versions of some fixed mother wavelets. The main advantage of wavelets is that they provide localized frequency information about a function of a signal, which is

particularly beneficial for classification. A review of basic fundamental of Wavelet Decomposition is introduced as follows:

The continuous wavelet transform of a signal  $x(t)$ , square-integrable function, relative to a real-valued wavelet,  $\psi(t)$  is defined as:

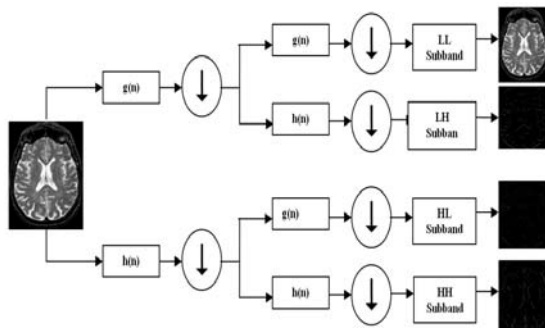
$$(1) W\psi(a,b) = \int_{-\infty}^{\infty} f(x) * \psi_{ab}(t) dx$$

Where  $\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right)$

and the wavelet  $\Psi_{a,b}$  is computed from the mother  $\Psi$  wavelet by translation and dilation, wavelet,  $a$  the dilation factor and  $b$  the translation parameter (both being real positive numbers). Under some mild assumptions, the mother wavelet  $\Psi$  satisfies the constraint of having zero mean.

The eq. (1) can be discretized by restraining  $a$  and  $b$  to a discrete lattice ( $a = 2^b$ ;  $a \in \mathbb{R}_+$ ;  $b \in \mathbb{Z}$ ) to give the discrete wavelet transform (DWT). The discrete wavelet transform (DWT) is a linear transformation that operates on a data vector whose length is an integer power of two, transforming it into a numerically different vector of the same length. It is a tool that separates data into different frequency components, and then studies each component with resolution matched to its scale. DWT can be expressed as.

$$(2) DWT_{x(n)} = \begin{cases} d_{j,k} = \sum (x(n)h^* j(n-2^j k)) \\ d_{j,k} = \sum (x(n)g^* j(n-2^j k)) \end{cases}$$



The coefficients  $d_{j,k}$ , refer to Figure2: DWT Schematically

the detail components in signal  $x(n)$  and correspond to the wavelet function, whereas  $a_{j,k}$  refer to the approximation components in the signal. The functions  $h(n)$  and  $g(n)$  in the equation represent the coefficients of the high-pass and low-pass filters, respectively, whilst parameters  $j$  and  $k$  refer to wavelet scale and translation

factors. The main feature of DWT is multiscale representation of function. By using the wavelets, given function can be analyzed at various levels of resolution. Fig. 2 illustrates DWT schematically. The original image is processed along the  $x$  and  $y$  direction by  $h(n)$  and  $g(n)$  filters which, is the row representation of the original image. As a result of this transform there are 4 subband (LL, LH, HH, HL) images at each scale. (Fig.2). Subband image LL is used only for DWT calculation at the next scale. To compute the wavelet features in the  $_rst$  stage, the wavelet coefficients are calculated for the LL subband using Harr wavelet function.

### C. Feature Selection and Reduction

One of the most common forms of dimensionality reduction is principal components analysis. Given a set of data, PCA finds the linear lower-dimensional representation of the data such that the variance of the reconstructed data is preserved. Using a system of feature reduction based on a combined principle component analysis on the feature vectors that calculated from the wavelets limiting the feature vectors to the component selected by the PCA should lead to a efficient classification algorithm utilizing supervised approach. So, the main idea behind using PCA in our approach is to reduce the dimensionality of the wavelet coefficients. This leads to more efficient and accurate classifier.

The feature extraction process was carried out through two steps: firstly the wavelet coefficients were extracted by the DWT and then the essential coefficients have been selected by the PCA.

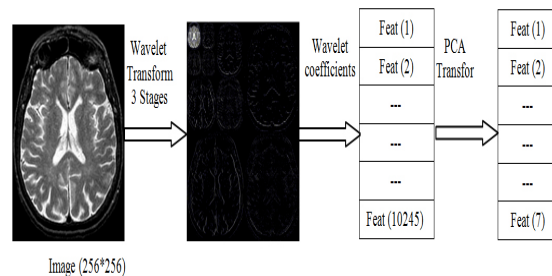


Figure3: Schematic diagram for the used feature extraction and reduction scheme

## III. MODEL LEARNING

### A. Feed Forward Artificial Neural Network (FF-ANN) Based Classifier

A three layer Neural network was created with 500 nodes in the first (input) layer, 1 to 50 nodes in the hidden layer, and 1 node as the output layer. We varied the number of nodes 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 2 shows the design of the Feed Forward Neural networks used in this research.

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.

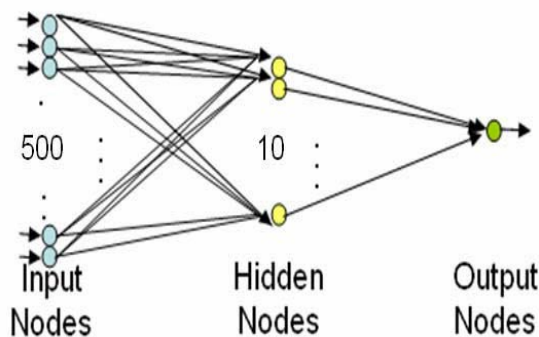


Figure 4: Feed Forward Neural Network

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.

#### B. Back Propagation Artificial Neural Network (BP-ANN) Based 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 sigmoidal function defined (with  $s = 1$ ) as  $\text{output} = 1/(1+e^{-x})$ , 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.

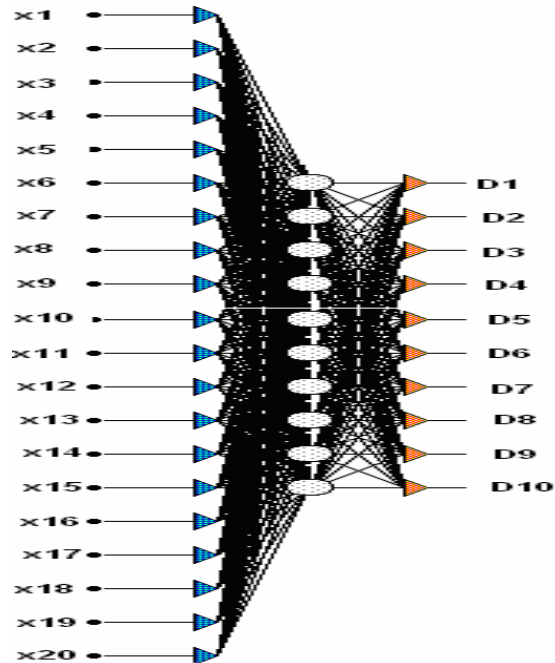


Figure 5: Back Propagation Neural Network

Figure 5 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

$$\text{net}_j^h = \sum_{i=1}^{N+1} W_{ji}x_i$$

$$y_j = f(\text{net}_j^h)$$

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

$$net_k^o = \sum_{j=1}^{J+1} V_{kj} y_j$$

$$Z_k = f(net_k^o)$$

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

$$V_{kj} \leftarrow V_{kj} + c \lambda (d_k - Z_k) Z_k (1 - Z_k) y_j$$

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

$$W_{ji} \leftarrow W_{ji} + c \lambda^2 y_j (1 - y_j) x_i \left( \sum_{k=1}^k (d_k - Z_k) Z_k (1 - Z_k) V_{kj} \right)$$

Update the error term

$$E \leftarrow E + \sum_{k=1}^k (d_k - Z_k)^2$$

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.

#### IV. CONCLUSIONS

In this paper, we propose two approaches for Brain Tumor Detection based on artificial neural networks. The networks were categorized into feed-forward neural networks and Back propagation neural Network.

The purpose is to develop tools for discriminating malignant tumors from benign ones assisting decision making in clinical diagnosis. The proposed approach utilizes a combination of these two neural network techniques and is composed of several steps including segmentation, feature vector extraction and model learning. These two methods can then be used to filter out non-suspecting brain scans as well as to point out suspicious regions that have similar property as the tumor regions.

#### ACKNOWLEDGMENT

We would like to thank **Prof. M. B. Nagori**, for valuable guidance at every step in making of this paper. She motivated us & boosted our confidence & we must admit that the work would not have been accomplished with her guidance & encouragement.

#### REFERENCES

- [1] K.M. Iftexharuddin, On techniques in fractal analysis and their applications in brain MRI, in: T.L. Cornelius (Ed.), Medical imaging systems: technology and applications,

Analysis and Computational Methods, vol. 1, World Scientific Publications, 2005, ISBN 981-256-993-6.

- [2] L.P. Clarke, R.P. Velthuizen, M.A. Camacho, J.J. Heine, M. Vaidyanathan, L.O. Hall, R.W. Thatcher, M.L. Silbiger, MRI segmentation: methods and applications, *Magn. Reson. Imaging* 13 (3) (1995) 343–368.
- [3] J.C. Bezdek, L.O. Hall, L.P. Clarke, Review of MR image segmentation techniques using pattern recognition, *Med. Phys.* 20 (4) (1993) 1033–1048.
- [4] H.S. Zadeh, H.S. Windham, Optimal linear transformation for MRI feature extraction, *IEEE Trans. Med. Imaging* 15 (1996) 749–767.
- [5] H.S. Zadeh, J.P. Windham, A comparative analysis of several transformations for enhancement and segmentation of magnetic resonance image scene sequences, *IEEE Trans. Med. Imaging* 11 (N3) (1992) 302–318.
- [6] D. Wang, D.M. Doddrell, A segmentation-based partial-volume-compensated method for an accurate measurement of lateral ventricular volumes on T1-weighted magnetic resonance images, *Magn. Reson. Imaging* 19 (2001) 267–272.
- [7] X. Zeng, L.H. Staib, R.T. Schultz, J.S. Duncan, Segmentation and measurement of the cortex from 3-D MR images using coupled surfaces propagation, *IEEE Trans. Med. Imaging* 18 (10) (1999) 927–937.
- [8] Khan M. Iftexharuddin et al., Fractal-based brain tumor detection in multimodal MRI, *Appl. Math. Comput.* (2008), doi:10.1016/j.amc.2007.10.063
- [9] N. Pal and S. Pal, "A review on image segmentation techniques," *Pattern Recognition*, vol. 26, pp. 1277-1294, 1993.
- [10] Josin G. M and Liddle P.F. Neural Network Analysis of the Pattern of Functional Connectivity between Cerebral Areas in Schizophrenia. *Biological Cybernetics*, 2001; 84, 117-122.
- [11] M. Sasikala and N. Kumaravel, "Comparison of Feature Selection Techniques for Detection of Malignant Tumor in Brain Images", *IEEE Indicon 20C Conference India*, Dec. 2005.
- [12] Yan Zhu and Hong Yan, "Computerized Tumor Boundary Detection Using a Hopfield Neural Network", *IEEE Transactions on Medical Imaging*, 16, No. 1, February 1997.
- [13] Alan Wee-Chung Liew and Hong Yan, "An Adaptive Spatial Fuzzy Clustering Algorithm for 3-D MR Image Segmentation", *IEEE Transactions on Medical Imaging*, 22, No. 9, September 2003.
- [14] Nicolae Duta and Milan Sonka, "Segmentation and Interpretation of MR Brain Images: An Improved Active Shape Model", *IEEE Transactions on Medical Imaging*, 17, No. 6, December 1998.

- [15] Madiha J. Jafri,Vince D.Calhoun," Functional Classification of Schizophrenia Using Feed Forward Neural Networks" IEEE EMBS Annual International Conference New York City, USA, Aug 30-Sept 3, 2006.
- [16] Jayashri Joshi1, Mrs.A.C.Phadke," Feature Extraction and Texture Classification in MRI", International Conference [ICCT-2010], 3rd-5th December 2010
- [17] el-sayed a. El-dahshan, abdel-badeeh m. Salem, and tamer h. Younis," a hybrid technique for automatic mri brain Images classification" BABES\_{BOLYAI, INFORMATICA, Volume LIV, Number 1, 2009

