Brain Tumor Classification using Principal Component Analysis and Probabilistic Neural Network

Sonali B. Gaikwad
ME- CSE
Jawaharlal Nehru Engineering College of Aurangabad, Maharashtra, India

Madhuri S. Joshi, Ph.D
Professor of CSE Department,
Jawaharlal Nehru Engineering College of Aurangabad, Maharashtra, India

ABSTRACT
Abnormal growth of the cell in the brain is the brain tumor. Brain tumor is common and serious disease. The proposed method for tumor classification in magnetic resonance brain image is the human inspection. Magnetic Resonance Imaging (MRI) plays an intrinsic role in the brain tumor disease diagnostic application. Various types of tumor that leads decision complicated. So that correct classification of brain tumor is important to detect the types of tumor. In this paper, Probabilistic Neural network (PNN) is used for brain tumor classification. Decision making was performed in two steps: 1) Feature extraction using Principal Component Analysis (PCA). And 2) Classification is done by Probabilistic neural network (PNN). Brain tumor is classified into three classes: Normal, Benign and Malignant. Again malignant tumor is classified as Glioma and Meningioma. PNN is faster and provide good classification accuracy.

Keywords
Brain Tumor Classification, Principle Component Analysis (PCA), Probabilistic Neural Network (PNN), MRI.

1. INTRODUCTION
Abnormal growth of cells within the brain or the central spinal canal or an intracranial solid neoplasm is called a brain tumor [4]. The main purpose of brain tumor classification is correctly classify the MR image in order to detect which type of tumor that suffered by the patients. Due to tumor there is problem of blood circulation in the brain. Therefore blood tubercles are forms. There are many kind of test are used to detect brain tumor such as MRI, Computed Tomography (CT) scan, Biopsy and many more. Among the entire test MRI has great potential for classification of tumor. In comparison with other diagnostic imaging modalities, such as computerized tomography, MRI provides superior contrast and resolution for different brain tissues [1]. Additionally, more valuable information regarding numerous tissue parameters (proton density, spin-lattice (T1) and spin–spin (T2) relaxation times, flow velocity, and chemical shift), which lead to more accurate brain tissue characterization will be added by MRI.[3].

The correct classification approach leads to the right decision and provide respective treatment to the patients. The key of brain tumor cure is to detect the tumor in its early stage. So that good classification is required. There are various methods to detect brain tumor. In this paper, we studied Probabilistic Neural Network (PNN) algorithm for classification of brain tumor.

In proposed method, the first stage is to extract features from brain MR images using PCA and then train the PNN for classification. The principal component analysis (PCA) method transforms the existing attributes into new ones considered to be crucial in classification [2]. Multiclass classification models for classifying brain tumors are proposed by [1], [3], [7].

The paper is organized in several sections: In the first section, a brief review about brain tumor. In the second section, we discussed Principal Component Analysis for feature extraction. PNN are discussed and implemented in the third section. Methodology of the proposed system is discussed in fourth section. Dataset are discussed in fifth section. In the final section we discussed experimentation and results.

2. PRINCIPAL COMPONENT ANALYSIS
In this paper, feature extraction is done using PCA. The most successful techniques that have been used in image recognition and compression is the Principal Component Analysis (PCA) and it is used to reduce the large dimensionality of the data [1].

Feature extractions are done for training as well as testing phase. The main purpose of MR image recognition system is to identify maximum similarities between training MR images and test MR image.

In the training phase, feature vectors are extracted from each training MR images. Let Ω_1 be a training image of image 1 which has a pixel resolution of M x N (M rows, N columns). In order to extract PCA features of Ω_1, first convert the image into a pixel vector Φ_1 by concatenating each of the M rows into a single vector. The length (or, dimensionality) of the vector $\Phi1$ will be M x N. In this paper, the PCA algorithm is used as a dimensionality reduction technique which transforms the vector Φ_1 to a vector ω_1 which has a dimensionality d where $d << M$ x N. For each training image Ω_i, these feature vectors ω_i are calculated and stored.

In the testing phase, the feature vector ω_j of the test image Ω_j is computed using PCA. In order to identify the test image Ω_j, the similarities between ω_j and all of the feature vectors ω_i's in the training set are computed. The similarity between feature vectors is computed using Euclidean distance. The identity of the most similar ω_i is the output of the image recognizer. If $i = j$, it means that the MR image j has correctly identified, otherwise if $i \neq j$, it means that the MR image j has misclassified [1]. The MR image recognition system is shown in Fig. 1.

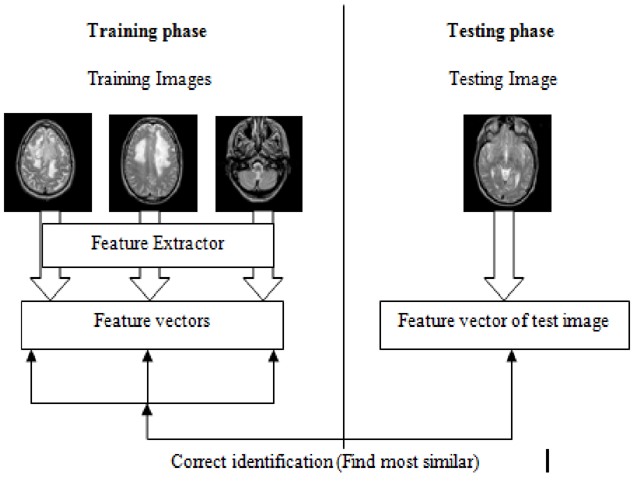

Figure 1: MR image recognition system.

3. PROBABILISTIC NEURAL NETWORK FOR CLASSIFICATION

The probabilistic neural network (PNN) was introduced by Donald Specht. This network is based on the theory of Bayesian classification and the estimation of probability density function. It is necessary to classify the input vectors into one of the two classes in Bayesian optimal manner. This Theory allows for cost function to represent the fact that it may be worse to misclassify a vector that is actually a member of class A than it's classified a vector that belongs to class B is described in book [6]. The Bayes rule such that the input vector belonging to class A is classified as

$$P_A C_A f_A(x) > P_B C_B f_B(x) \qquad (1)$$

Where P_A –Priori probability of occurrence of pattern in class A.

C_A - Cost associated with classifying vectors.

$F_A(x)$ - Probability density functions of class A.

PNN is widely used for classification problem. There are several advantages using PNN instead of Back Propagation (BP) multilayer perceptron. PNNs are much faster than BP multilayer networks. It provides better accuracy. This network is relatively insensitive to outliers and generates accurate predicted target probability scores.

The most important advantage of PNN is that training is easy and instantaneous [1]. 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 [1].

Figure 2 represent the architecture of PNN. The architecture of PNN is made up of 4 types of units.

- Input units
- Pattern units
- Summation units
- Output units

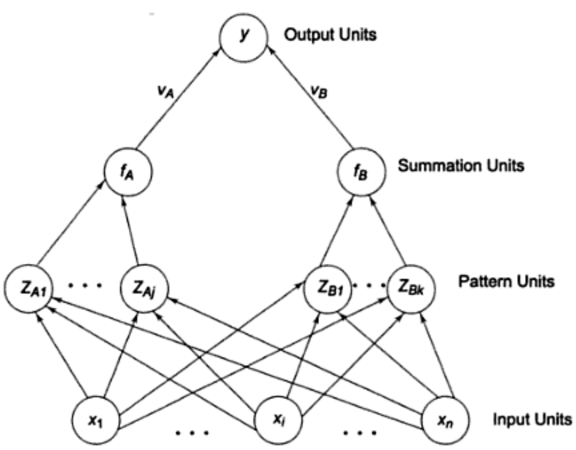

Figure 2: Architecture of PNN

1) *Input units:* Hear input unit x(p), p=1,2,…..P are connected to all pattern units.

2) *Pattern units:* Create pattern unit Z_P .Weight vector for unit Z_P is computed as W_P=x(p). Unit Z_P is either Z_A or Z_B unit.

3) *Summation units:* Connect the pattern unit to summation unit. If x(p) belongs to class A, connect pattern unit Z_P to summation unit S_A. Else connect pattern unit Z_P to summation unit S_B. The weight used by the summation unit for class B is

$$V_B = -\frac{P_B C_B m_A}{P_A C_A m_B} \qquad (2)$$

4) *Output units:* It sums the signals from f_A and f_B. Input vector is classified as class A if the total input to decision unit is positive. y is the final output.

4. METHODOLOGY

The main purpose of proposed method to detect tumor automatically from the MR images. The PNN classifier presented good accuracy, very small training time, robustness to weight changes, and negligible retraining time [1].

The proposed method is made up of six steps which are starting from input data to output. In the first step, use MR image as input image. Convert input image into graylevel and resize it into 256 x 256 size. This process is performed in second step. In the third step, perform PCA for feature extraction. Features are extracted from training as well as testing MR images. In the forth step, PNN is train for classification of MR images. Finally performance based on the result will be analyzed at the end of the development phase. The proposed brain tumor classification method is shown in Fig. 3. The final output of the proposed method is either normal, benign or malignant. Further malignant are classified either as Glioma or Meningioma.

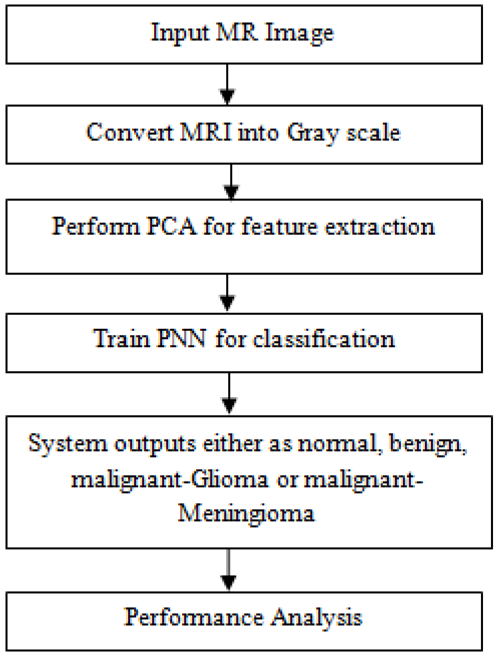

Figure 3: The proposed method of PCA-PNN.

5. DATASET

In this study, it has obtained some brain MR images from Government hospital of Aurangabad and most of the images from Sahyadri Hospital of Pune, Maharashtra, India. Dataset contains 70 training samples (Normal=25, Benign=25, and Malignant=20) and 35 testing samples for classify tumor as normal, benign, or malignant. It also contains 24 training MR images and 20 testing samples for glioma and meningioma type classification.

6. RESULTS AND DISCUSIONS

In this proposed method, various individualize experiments were performed and the sizes of the training and testing sets were determined the classification accuracy.

The dataset was divided into four datasets- Two datasets for training and two datasets for testing. The first training dataset was used to train MR images for normal, benign and malignant. This dataset contain 70 MR images. The second training dataset which contain 24 brain tumor MR images of Glioma and Meningioma. The testing dataset was used to verify the accuracy of trained PNN network for brain tumor classification.

The MATLAB R2009a is used to implement PNN. In this proposed method smoothing factor was used to examine the accuracy of the classifier. The spread value (SV) was used as smoothing factor and examined the classifier accuracy when different values of spread value were used. If S.V. is near zero, the network will act as a nearest neighbour classifier, and the network will take into account several nearby design vectors if its value becomes larger [1].

In this study, PNN were examined using testing datasets. Clinical and performance measurements with respect to normal, benign and malignant are shown in Table 1 and Table 2 respectively. In both the table the accuracy of the classifier was examined for different values of the S.V.

True positives (TP) refer to the positive tuples that were correctly labeled by the classifier, while true negatives (TN) are the negative tuples that were correctly labeled by the classifier. False positives (FP) are the negative tuples that were incorrectly labelled. Similarly false negatives (FN) are the positive tuples that were incorrectly labeled [5].

Table 1. Clinical measurements of PNN with respect to normal, benign and malignant.

SV	Training Sample	Testing Sample	TP	TN	FP	FN
10^2	70	35	9	11	4	11
10^3	70	35	10	11	3	11
10^4	70	35	10	12	3	10
10^5	70	35	10	13	3	9
10^6	70	35	10	13	3	9
10^7	70	35	12	22	1	0
10^8	70	35	12	19	1	3

The sensitivity and specificity measures can be used for correct classification. Sensitivity is also referred to as the *true positive (recognition) rate* (that is, the proportion of positive tuples that are correctly identified), while specificity is the *true negative rate* (that is, the proportion of negative tuples that are correctly identified) [5]. These measures are defined as

$$\text{Sensitivity} = \frac{t\_pos}{pos} \qquad (3)$$

$$\text{Specificity} = \frac{t\_neg}{neg} \qquad (4)$$

The accuracy is a function of sensitivity and specificity:

$$\text{Accuracy} = \text{Sensitivity}\ \frac{pos}{(pos+neg)} + \text{Specificity}\ \frac{neg}{(pos+neg)}$$

Table 2. Performance measurements of PNN with respect to normal, benign and malignant.

SV	Training Sample	Testing Sample	Sensitivity	Specificity	Accuracy
10^2	70	35	69.23%	50%	57%
10^3	70	35	76.92%	50%	60%
10^4	70	35	76.92%	54.54%	62.85%
10^5	70	35	76.92%	59%	65.71%
10^6	70	35	76.92%	59%	65.71%
10^7	70	35	92.30%	100%	97.14%
10^8	70	35	92.30%	86%	88.57%

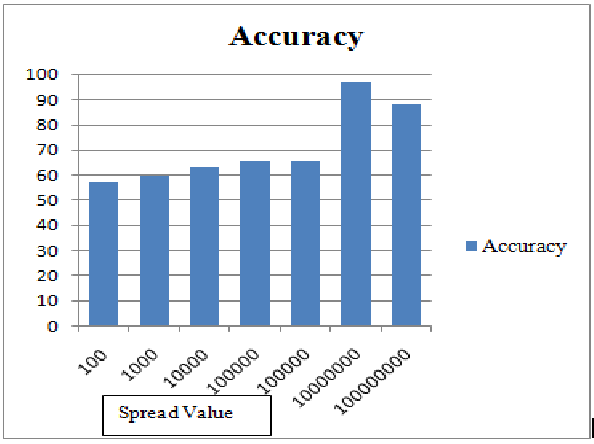

Figure 4: Accuracy of normal, benign and malignant MR images on the basis of performance measurement.

In Table 2, 70 training samples and 35 testing samples have been used. The classifier accuracy varies from 57% to 97.14%. With 90 training samples and 35 testing samples, the results remain unchanged.

Therefore 70 training samples can suffice to get the acceptable accuracy and it also reduces the classifier training time.

Where S.V. $=10^7$, accuracy is 97.14 %.

Table 3 and Table 4 show results for Glioma and Meningioma.

Table 3. Clinical measurements of PNN with respect to Glioma and Meningioma.

SV	Training Sample	Testing Sample	TP	TN	FP	FN
10^2	24	20	12	6	2	0
10^3	24	20	12	6	2	0
10^4	24	20	12	6	2	0
10^5	24	20	12	6	2	0
10^6	24	20	14	6	0	0

Table 4. Performance measurements of PNN with respect to Glioma and meningioma.

SV	Training Sample	Testing Sample	Sensitivity	Specificity	Accuracy
10^2 - 10^5	24	20	85.7%	100%	90%
10^6	24	20	100%	100%	100%

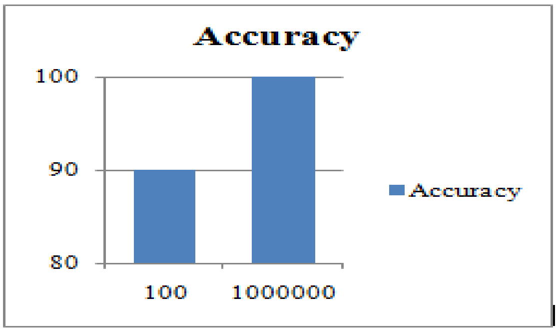

Figure 5: Performance Accuracy of Glioma and Meningioma.

In this work, Table 1 and Table 2 shows clinical and performance measurements related to normal, benign and malignant whereas further classification are shown in Table 3 and Table 4. PNN was developed to examine the accuracy with different spread value ranges from 10^2 to 10^7 and 10^2 to 10^6 as shown in Table 2 and Table 4 respectively. Classifier accuracy reached to 100% when S. V. was 10^6.

Figure 6 show normal MR image which indicate MR brain image without tumor. Figure 7 and Figure 8 represent benign and malignant MR images. Brain MR Images of Glioma and Meningioma are shown in Figure 9 and Figure 10 respectively.

Figure 4 and Figure 5 have shown the accuracy with respect to S.V for different classes of brain tumor.

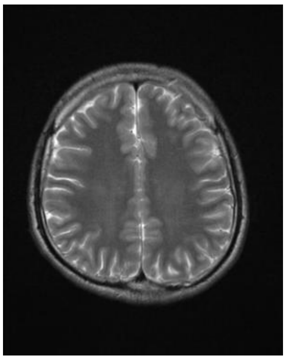

Figure 6: Normal MR image.

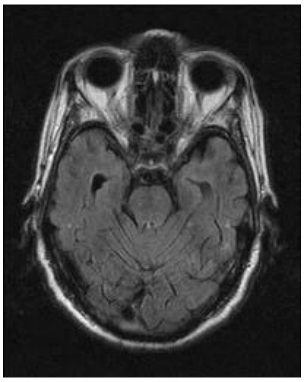

Figure 7: Benign MR image.

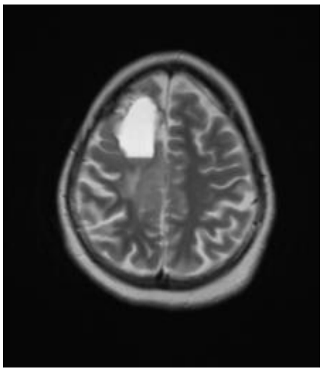

Figure 8: Malignant MR image.

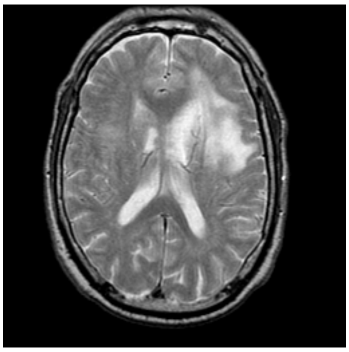

Figure 9: Meningioma MR image.

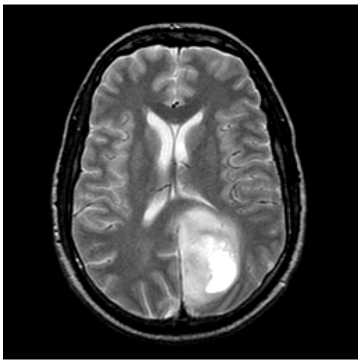

Figure 10:Glioma MR image.

7. CONCLUSION

In this paper brain tumor MR image is automatically classified into normal, benign and malignant category using PNN classifier. The malignant image is further diagnosed into meningioma and glioma type of tumor. Maximum accuracy of 97.14% and 100% is obtained with spread value $=10^7$ and 10^6 as shown in Table 2 and Table 4 respectively. This work will act as supportive tool for radiologists and will help doctor for fast diagnosis based on which the treatment plan can be decided.

Future Scope:

We can apply this method on foreign patient images and also we can try for more features for better accuracy.

8. REFERENCES

[1] M.F. Othman and M. A. M. Basri, "Probabilistic Neural Network for brain tumor Classification", *IEEE Second International Conference on Intelligent Systems, Modelling and Simulation*, 136 – 138, 2011.

[2] N. Kwak, and C. H. Choi, "Input Feature Selection for Classification Problems", *IEEE Transactions on Neural Networks*, 13(1), 143–159, 2002.

[3] P. Georgiadis and G. Kagadis, "Non-linear Least Squares Features Transformation for Improving the Performance of Probabilistic Neural Networks in Classifying Human Brain Tumors on MRI" ICCSA 2007, LNCS 4707, pp. 239 – 247, 2007, Springer-Verlag Berlin Heidelberg 2007.

[4] D. SRIDHAR, "Brain Tumor Classification Using Discrete Cosine Transform and Probabilistic Neural Network", 2013 *IEEE International Conference on Signal Processing, Image Processing and Pattern Recognition*, 978-1-4673-4862-1/13, 2013 IEEE.

[5] J. Han and M. Kamber, "Data Mining: Concepts and Techniques", 2006.

[6] S. N. Sivanandam, S. N Deepa, "Introduction to Neural Network using MATLAB 6.0", 2006.

[7] V. Kumar, J. Sachdeva, "Classification of brain tumors using PCA-ANN" 978-1-4673-0126-8/11 2011 IEEE.