© 2016 IJEDR | Volume 4, Issue 1 | ISSN: 2321-9939

Brain Tumor Classification into Normal and Abnormal Using PCA and PNN Classifier

[1]Deven D.Ketkar, [2]Ashish B.Vartak

[1] Lecturer, [2]Assistant Professor

[1]Information Technology Department, 2 Electronics Engineering Department

[1]Finolex academy of management and Technology, Ratnagiri, India

[2]Finolex academy of management and Technology, Ratnagiri, India

---

***Abstract* - There are several number of diseases which are entering in human life due to modern lifestyle, unhealthy food. Also there is no control on fast food eatables. If medical field can take help of several technological aspects, then it will give fine results. Use of automated techniques can give efficient diagnosis. Considering one of the dangerous disease like brain tumor, accurate prediction about tumor malignancy is necessary. We are going to use several Magnetic Resonance Images (MRI's) for training and testing purpose. Classification of tumor is done using two methods Principal Component Analysis (PCA) and Probabilistic Neural Network (PNN) classifier. PCA can be used for Feature extraction of MRI images. Features which we are going to calculate are Mean, Deviation and Eigen Vectors. We are going to take several number of MR images which includes Normal and abnormal images. Features mentioned above corresponding to these images are calculated and with the help of these features we are going to train PNN classifier. When test image comes at the input side, features corresponding to that image are calculated and compared with training dataset and based on similarity between the features, tumor classification is done.**

***Index Terms* - Magnetic Resonance Imaging (MRI), Principal Component Analysis (PCA), Probabilistic Neural Network (PNN), Mean, Deviation, Eigen Vectors, Euclidean Distance.**

I. INTRODUCTION

Today's world is a world of fast life. Also there is no control on proper food, unhealthy food products, fast food eatables etc. The prior aim of medical field is that it should give fast and accurate results of diseases like cancer, alzimer, multiple sclerosis, etc. Hence medical field is taking help of technological aspects so that they can give best results. Most common technique used in medical field is Digital Image Processing which can process the image using different methods and in turn can improve the performance of diagnosis with increase in speed.

II.

One of the most dangerous disease in today's world is brain tumour. It occurs due to abnormal growth of cells inside the brain. These brain tumors have variety of shapes and sizes. Brain tumors can be Normal or Abnormal. The main source to obtain medical images is X-rays and presently MRI. MRI is best known for its characteristics like superior soft tissue differentiation, high spatial resolution and contrast. It does not use harmful ionizing radiation to patients. Medical field can be more superior by using automated and efficient diagnostic technologies in a short period of time. MRIs are examined by radiologists based on visual interpretation and according to that they make decision whether tumor is normal or abnormal. But problem occurs when there is shortage of radiologists and the large number of MRIs are to be analyzed which makes such readings labor intensive, cost expensive and often inaccurate.

The sensitivity of the human eye in analyzing large numbers of MRIs decreases with increasing number of cases, particularly when only a small number of slices are affected. Hence there is requirement for automated systems for analysis and classification of such medical images.

Classification is a process of partitioning an image space into non-overlapping meaningful homogeneous regions. The working of the classifier is only to apply some analytical method on a given set of training samples/data and to get proper separation between required areas. The MRI may contain both normal slices and defective slices. The defective or abnormal slices are identified and separated from the normal slices and then these defective slices are further investigated for the detection of tumor tissues. Such a separation of abnormal and normal slices requires knowledge of a concept called as 'classifier'.

II.LITERATURE SURVEY

The conventional method for computerized tomography and magnetic resonance brain images classification and tumor detection is by human inspection only. Operator assisted classification methods are impractical for large amounts of data. Computerized Tomography (CT) scans and MRIs contain a noise caused by operator performance which can lead to serious inaccuracies in classification. The use of Neural Network techniques shows great potential in the field of medical diagnosis. The Probabilistic Neural Network (PNN) with Discrete Cosine Transform (DCT) was applied for Brain Tumor Classification [1]. The most important point in case of Principal Component Analysis (PCA) is that it is used for feature extraction. PCA can be used with classifiers such as Artificial Neural Network (ANN), PNN, K-Nearest Neighbor classifiers in order to perform various classifications. The work done so far is basically to assist radiologists in marking tumor boundaries and in decision making process for multiclass classification of brain tumors [2]. Extraction of tumor boundaries is done using Gradient Vector Flow (GVF). These segmented Regions of Interest (ROIs) are classified by using (PCA-ANN). Support Vector Machines (SVMs) and Decision Tree (DT)classification is used as a methodology for thecharacterization of the degree of malignancy of braintumor Astrocytoma's (ASTs).The firstlevel is concerned with the detection of low versus high-gradetumors and the second level deals with the detection of lessaggressive as opposed tohighly aggressive tumors [3].

PCA can also be used for face recognition system and great recognition rate can be achieved. Study is made in such a way that identifying the characteristics of face recognition rate when number of training and testing data is varied, the content of noise in the training and test data is varied. Also varying the level of blurriness in the training and test data and the image size. Different databases are used with aligned images [4]. There are two PNN based maximum likelihood classifiers. These classifiers are based on Gram-Charlier series expansion with or without Parzen's windowing technique. The performances of the proposed classifiers can be evaluated in terms of probability of target detection for a number of Gaussian and Non Gaussian noises, and compared with those of the existing neural network classifiers such as Bayesian and Back propagation classifiers [5]. PNN can also be used as a classifier to the automatic classification of underwater objects. Here, a process called multifield feature extraction is used to develop a feature vector. This process involves time domain analysis, time-frequency distribution, spectra and bi-spectra analysis. Underwater target classification can be thought as a problem of small sample recognition, because samples acquired under different conditions often exhibit different clustering characteristics. PNN is chosen to distinguish underwater objects because of its simplicity, robustness to noise, and nonlinear decision boundaries. The PNN classifier is contrasted with a Gaussian classifier and SVM using lake or sea trial data. Experimental results shows that PNN classifier is appropriate to this problem [6]. An algorithm for flaw classification in ultrasonic guided waves signal is presented, in which Wavelet Transform is used in the process of noise suppression and envelop extraction, and the PNN is used for flaw classification of the ultrasonic testing signal. In the process of feature extraction, the necessity of attenuation correction and feature selection of ultrasonic signal is taken into consideration. The comparison of the performances of PNN and Back Propagation (BP) classifier is made which demonstrates that the performance of flaw classification is significantly improved by the use of synthesized algorithm [7]. A self-adaptive method of iris boundary detection is proposed and the method can segment the iris area accurately regardless of the shapes of iris boundaries. A new feature extraction technique based on combination using special Gabor filters and wavelet maxima components is presented. The radial basis function neural network (RBFNN) with a particle swarm optimization (PSO) a novel iris recognition technique with intelligent classifier is proposed for high performance iris recognition.

Combination of RBFNN and particle swarm optimization PSO for an optimized PNN classifier model is presented. The results shows that the proposed algorithm provides superior performance in iris recognition. [8]. Diabetic Retinopathy is a disease caused by complication of diabetes. It is a major cause of Blindness in both middle and advanced age group. Earlier detection of diabetic retinopathy protects patient from vision loss. The early symptom of this blindness is the exudates. Exudates are the liquefied fluid comprising solutes, proteins, cells, or cellular debris leaked from the damaged blood vessels into nearby tissues or on tissue surfaces in the retina. The leakage of such proteins or lipids will result in vision loss to the patients. Identification of the exudates in advance can protect the diabetic patients from blindness. Dilation method is also used by the Ophthalmologists to detect the exudates. But it causes the irritation to the patients' eyes. The method is invented which detects the diabetic retinopathy through detecting exudates by Morphological process in color fundus retinal images and then segregates the severity of the lesions. The severity level of the disease is achieved by PNN classifier [9].

III. PROBLEM STATEMENT

The main drawback of manual examination in analyzing MRI's is that it is labor intensive, cost expensive, time consuming, etc. Limitations on sensitivity of human eyes are also a very big problem. The most important thing is that if a radiologist has to examine bulk of MRI's in a short period of time, then it will be difficult for him to give correct result. MRI is new invention in analyzing medical images. If large number of MRI's are to be analyzed in short period of time, then radiologist may give wrong decision or prediction about that` MRI. This can be harmful for life of patients and may cause severe problems. Hence, automated and efficient diagnosis is required so that radiologist can crosscheck his result and can find exact result. Here, we are going to take several number of normal and abnormal MRIs of brain which are training images. We are going to use a PNN classifier. Features corresponding to training images can be calculated using PCA method. Features which we are going to get are Mean, Deviation and Eigen Vectors.

Using these features we are going to train classifier. We can use more number of images to train the classifier so that our classification is good. When input image comes, features corresponding to that image are calculated and compared with training dataset. According to the similarity between testing and training image features, decision is made such that whether given MRI is normal or abnormal.

IV. PROPOSED METHODOLOGY

A. Principal Component Analysis (PCA)

Principal component analysis (PCA) is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set ofvalues of linearly uncorrelated variables called principal components. The number of principal components is less than or equal to the number of original variables defined in such a way that the first principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible) and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to (i.e., uncorrelated with) the preceding components.

Depending on the field of application, it is also named the discrete Karhunen–Loève transform (KLT) in signal processing. PCA can be done by eigenvalue decomposition of a data covariance (or correlation) matrix or singular value decomposition of a data matrix, usually after mean centering (and normalizing or using Z-scores) the data matrix for each attribute.

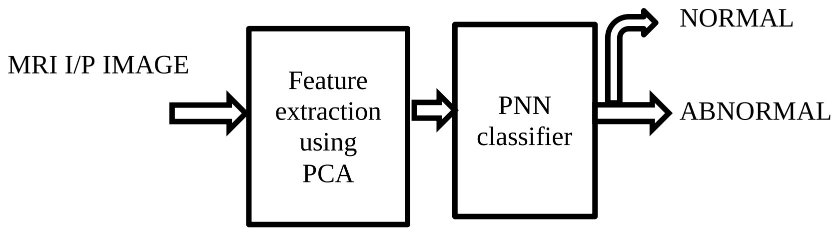

Figure.1 Block diagram of proposed system

PCA is one of the most successful techniques that have been used in image recognition and compression. The purpose of PCA is to reduce the large dimensionality of the data. The task of the MR image recognizer is to find the most similar feature vector among the training set to the feature vector of a given test image. Let T1 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 T1, first convert the image into a pixel vector Φ1 by concatenating each of the *M* rows into a single vector. The length (or, dimensionality) of the vector Φ1 will be *M* x *N*. Here, the PCA algorithm is used as a dimensionality reduction technique which transforms the vector Φ1 to a vector ω1 which has a dimensionality *d* where *d* <<*M*x*N*. For each training image Ti, these feature vectors ωi are calculated and stored. In the testing phase, the feature vector ωj of the test image Tj is computed using PCA. In order to identify the test image Tj, the similarities between ωj and all of the feature vectors ωi's in the training set are computed. The similarity between feature vectors is computed using Euclidean distance. The identity of the most similar ωi is the output of the image recognizer. If i = j, it means that the MR image j has correctly identified, otherwise if i ≠ j, it means that the MR image j has misclassified. Using Principal Component Analysis (PCA), we can perform feature extraction corresponding to collected data set. Features to be calculated-:

1. Mean
2. Deviation
3. Eigen Vectors

The results of a PCA are usually discussed in terms of component scores, sometimes called factor scores and loadings (the weight by which each standardized original variable should be multiplied to get the component score). PCA is the simplest of the true eigenvector-based multivariate analyses. Its operation can be thought of as revealing the internal structure of the data in a way that best explains the variance in the data. If a multivariate dataset is visualized as a set of coordinates in a high-dimensional data space (1 axis per variable), PCA can supply the user with a lower-dimensional picture. This is done by using only the first few principal components so that the dimensionality of the transformed data is reduced. It is also related to canonical correlation analysis (CCA). CCA defines coordinate systems that optimally describe the cross-covariance between two datasets while PCA defines a new orthogonal coordinate system that optimally describes variance in a single dataset. PCA can be thought of as fitting an *n*-dimensional ellipsoid to the data, where each axis of the ellipsoid represents a principal component. If some axis of the ellipse is small, then the variance along that axis is also small and by omitting that axis and its corresponding principal component from representation of then dataset, we loose only a small amount of information.

To find the axes of the ellipse, we must first subtract the mean of each variable from the dataset to center the data around the origin. Then, we compute the covariance matrix of the data, and calculate the eigenvalues and their corresponding eigenvectors of this covariance matrix. Then, we must orthogonalize the set of eigenvectors, and normalize each to become unit vectors. Once this is done, each of the mutually orthogonal, unit eigenvectors can be interpreted as an axis of the ellipsoid fitted to the data.

B. Probabilistic Neural Network (PNN)

A probabilistic neural network (PNN) is a feed forward neural network, which was derived from the Bayesian networkand a statistical algorithm called Kernel Fisher discriminant analysis. It was introduced by D.F. Specht in the early 1990s.PNN is a type of Radial Basis Function (RBF) network, which is suitable for classification of patterns. The architecture has four layers, an input layer, a hidden layer, a pattern layer and an output layer. The pattern layer constitutes a neural implementation of a Bayes classifier, where the class dependent Probability density Functions (PDF) are approximated using a Parzen estimator. Parzen estimator gives the PDF by minimizing or reducing the expected risk in classifying the training set incorrectly. Hence, with the use of Parzen estimator, the classification gets closer to the true underlying class density functions as the number of training samples increases. The pattern layer is made of a processing element corresponding to each input vector in the training set. Each output class must consist of equal number of processing elements otherwise some classes may be inclined falsely which will result in poor classification results. Each processing element in the pattern layer is trained once. An element is trained in such a way that it will return a high output value when an input vector matches the training vector. In order to obtain more generalization or accuracy, a smoothing factor is included while training the network. This smoothing factor is also called as a spread value. The pattern layer classifies the input vectors based on competition, where only the highest match to an input vector wins and generates an output. Hence only one classification category is generated for any given input vector. If there is no relation between input patterns and the patterns programmed into the pattern layer, then no output is generated. If we compare PNN to the feed forward back propagation network, training of PNN is very much simpler. Basically, probabilistic networks classify on the basis of Bayesian theory, hence it is necessary to classify the input vectors into one of the two classes in a Bayesian manner. In a PNN, the operations are organized into a multilayered feed forward network with four layers. PNN is mostly used in classification problems. When an input is present, the first layer computes the distance from the input vector to the training input vectors. This produces a vector where its elements indicate how close the input is to the training input. The

second layer sums the contribution for each class of inputs and produces its net output as a vector of probabilities. Finally, a complete transfer function on the output of the second layer picks the maximum of these probabilities, and produces a 1 (positive identification) for that class and a 0 (negative identification) for non-targeted classes.
If the probability density function (pdf) of each of the populations is known, then an Unknown, X, belongs to class "i" if

$$f_i(x) > f_j(x), \text{ all } j \neq i \quad (1)$$

Input layer

Each neuron in the input layer represents a predictor variable. In categorical variables, *N-1* neurons are used when there are *N* number of categories. It standardizes the range of the values by subtracting the median and dividing by the interquartile range. Then the input neurons feed the values to each of the neurons in the hidden layer.

Pattern layer

This layer contains one neuron for each case in the training data set. It stores the values of the predictor variables for the case along with the target value. A hidden neuron computes the Euclidean distance of the test case from the neuron's center point and then applies the RBF kernel function using the sigma values.

Summation layer

For PNN networks there is one pattern neuron for each category of the target variable. The actual target category of each training case is stored with each hidden neuron; the weighted value coming out of a hidden neuron is fed only to the pattern neuron that corresponds to the hidden neuron's category. The pattern neurons add the values for the class they represent.

Output layer

The output layer compares the weighted votes for each target category accumulated in the pattern layer and uses the largest vote to predict the target category.

C. Advantages of PNN

PNN is a classifier which maps any input pattern to a number of classifications. PNN is famous for its fast training process. PNN converges to an optimal classifier as the size of the training set increases. Training samples can be added or removed without extensive retraining.

D. Disadvantages of PNN

Even though PNN has many advantages, it has several disadvantages too. PNN has large memory requirements. Also, there occur problems corresponding slow execution of network. PNN requires a representative training set even more than other types of neural networks. One of the most important point in case of PNN is that training set should be thoroughly representative of the actual population for effective classification. If we are going to add or subtract training samples, it is similar to adding or removing of neurons in pattern layer. If we are going to increase training set, PNN asymptotically converges to Bayes classifier.

V. PNN TRAINING

PNN is a useful neural network architecture with slightly different in fundamentals from back propagation. The architecture is feed forward in nature which is similar to back propagation, but differs in the way that learning occurs. PNN is supervised learning algorithm but includes no weights in its hidden layer. Each hidden node represents an example vector, with the example acting as the weights to that hidden node. These are not adjusted at all. PNN consists of an input layer, which represents the input pattern or feature vector. The input layer is fully interconnected with the hidden layer, which consists of the example vectors (the training set for the PNN). The actual example vector serves as the weights as applied to the input layer.

Finally, an output layer represents each of the possible classes for which the input data can be classified. However, the hidden layer is not fully interconnected to the output layer. The example nodes for a given class connect only to that class's output node and none other. One other important element of the PNN is the output layer and the determination of the class for which the input layer fits. This is done through a winner-takes-all approach. The output class node with the largest activation represents the winning class. While the class nodes are connected only to the example hidden nodes for their class, the input feature vector connects to all examples, and hence influences their activations. Hence, it is the sum of the example vector activations which determines the class of the input feature vector.

In PNN algorithm, calculation of the class-node activations is a simple process. For each class node, the example vector activations are summed, which are the sum of the products of the example vector and the input vector. The hidden node activation, shown in the following equation is the product of the two vectors (E is the example vector, and F is the input feature vector).

$$h_i = E_i F \qquad (2)$$

The class output activations are then defined as:

$$C_j = \frac{\sum_{i=1}^{N} e^{h_i - 1/\gamma^2}}{N} \quad (3)$$

where N is the total number of example vectors for this class, h_i is the hidden-node activation, and γ is a smoothing factor. The smoothing factor is chosen by doing experimentation. If the smoothing factor is too large, details can be lost, but if the smoothing factor is too small, the classifier may not generalize well. There is no real training that occurs since the example vectors serve as the weights to the hidden layer of the network. If we have given an unknown input vector, the hidden node activations are computed and then summed at the output layer. The class node with the largest activation determines the class to which the output feature vector belongs. As no training required, classifying an input vector is fast, depending on the number of classes and example vectors that are present. It is also easy to add new examples to the network by simply add the new hidden node, and its output is used by the particular class node. This can be done dynamically as new classified examples are found. The PNN also generalizes very well when noisy data set is present [13].

VI. EXPERIMENTATION AND RESULTS

We took 90 images for training which consists of Normal and Abnormal images. Our classification includes classification of normal and abnormal images. The value of sigma (spread) is varied from 100 to 48×10^5 . Different values of *TP*, *TN*, *FP* and *FN* are found out. For different values of sigma, we have calculated Sensitivity, Specificity and Accuracy. For = 100, we have got lowest accuracy which is 82.54 %. At =48×10^5, we have got the highest accuracy which is 97.11 %. After that, as the value of sigma increases, accuracy goes on decreasing.

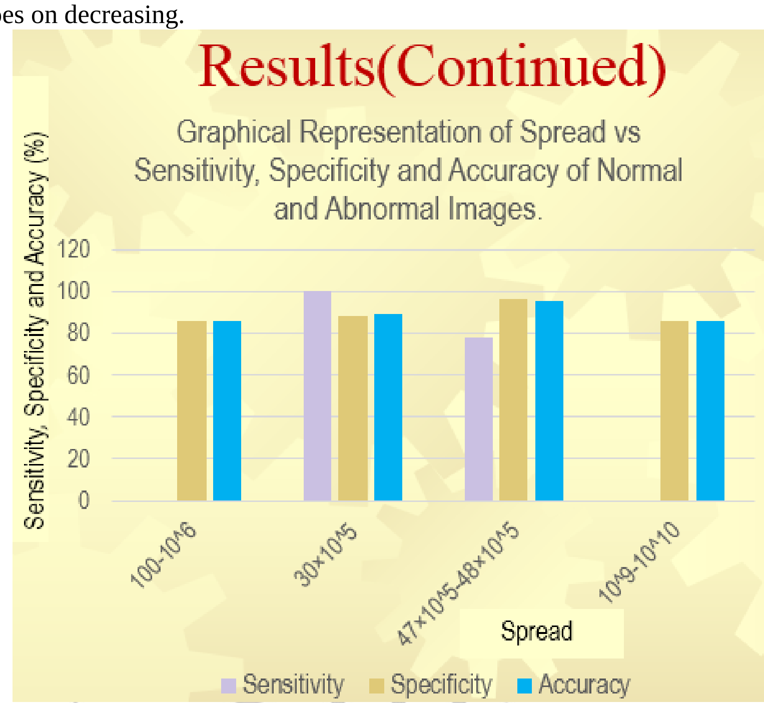

Figure.2 Graphical representation of spread vs sensitivity, specificity and accuracy

VII. REFERENCES

[1] D Sridhar, Muralikrushna, "Brain Tumor Classification Using Discrete Cosine Transform and Probabilistic Neural Network", 2013 International Conference on Signal Processing, Image Processing and Pattern Recognition [ICSIPR], pp.1-5

[2] Vinod Kumar, JainySachdeva, Indra Gupta, NiranjanKhandelwal, Chirag Kamal Ahuja, "Classification of Brain Tumors Using PCA-ANN", 20111 IEEE Conference, PP-1079-1083.

[3] D.Glotsosl, P. Spyridonos, P.Petalas, D.Cavourasz,V. Zolota, P.Dadioti, Lekk, G. Nikiforidis, "A hierarchical decision tree classification scheme for brain tumor astrocytomaGrading using Support Vector Machines", 3rd International Symposium on Image and Signal Processing and Analysis(2003), pp-1034-1038.

[4] Bruce Poon, M Ashraful Amin, Hong Yan, "PCA based Face Recognition and Testing Criteria", Proceedings of the Eighth International Conference on Machine Learning and Cybernetics, Baoding, 12-15 July 2009, pp-2945-2949.

[5] Moon W. Kim, Mohammad Arozullah, "GENERALIZED PROBABILISTIC NEURAL NETWORK BASED CLASSIFIERS",U.S. Government work not protected by U.S. copyright.pp-III 648-III 653.

[6] JieTian, ShanhuaXue, Haining Huang, "Classification of Underwater Objects Based on Probabilistic Neural Network", 2009 Fifth International Conference on Natural Computation,pp.38-42.

[7] Xiaohui Zhang,Guangmin Sun, Haomiao Liu, and Qite Wang, "Flaw Classification in Ultrasonic Guided Waves Signal Using Wavelet Transform and PNN Classifier", IEEE Conference, pp-978-982.

[8] Jin Liu, Xiao Fu, Xingbin Yao, "PSO-RBFNN Based Optimized PNN Classifier Model", 2012 2nd International Conference on Computer Science and Network Technology, pp.456-459.

[9] Mahendran G, Dhanasekaran R and Narmadha Devi K.N, "Identification of Exudates for Diabetic Retinopathy based on

Morphological Process and PNN Classifier", International Conference on Communication and Signal Processing, Apri[3-5, 2014, India, pp-1117-1121.
[10] MohdFauzi Othman and MohdAriffananMohdBasri, "Probabilistic Neural Network. For Brain Tumor Classification", 2011 Second International Conference on Intelligent Systems, Modelling and Simulation, pp.136-138.
[11] Jacec M. Zurada, "*Introduction to Artificial Neural System".* Jaico Publishing House.
[12] Simon Haykin, "*Neural Networks"*, Second Edition, A comprehensive Foundation, Prentice Hall of India Private Limited.
[13] M. Tim Jones, "*Artificial Intelligence A Systems Approach"*, Infinity Science Press, 2008