# Classification of Brain Tumors using Image Pre-Processing and Artificial Neural Networks

[1]Ajay Kushwaha, [2]Mahesh Kumar Pawar, [3]Anjana Pandey
[1]M.Tech Scholar, [2]Assistant Professor, [3]Assistant Professor
[1]Department of Data Science,
[1]SOIT, RGPV, Bhopal, India

***Abstract :*** The classification of brain tumors has remained complex due to extremely subtle differences in pixel values of cancerous and non-cancerous tumors. Although, normal MRI images and images with tumors can be classified by inspection to some extent, but predicting whether a particular case has cancerous (malignant) type intent of a benign (non-cancerous) intent is almost infeasible. This paper proposes a technique that uses a vigorous image pre-processing technique in conjugation with a probabilistic class of artificial neural network for brain tumor classification. It has been shown that the present approach attains a classification accuracy of 97% (approx). The high value of accuracy can be attributed to the image pre-processing stage and the design of the probabilistic neural design of the Artificial Neural Network.

***Index Terms* – Artificial Neural Network, Image Pre-processing, Malignant Tumor, Benign Tumor, Accuracy.**

## I. INTRODUCTION

Brain tumor classification is a non-trivial task if it is automated. The major challenge that is faced is the fact that brain tumors are:
It should be noted that brain tumors unlike some medical ailments are extremely difficult to diagnose with accuracy. Moreover, even the slightest of diagnostic error can prove to be life-threatening. Recently brain tumor classification has be approached with a new approach i.e. Artificial Intelligence or AI. [3]AI based systems are being explored for automated or assisted classification for brain tumors. The gravity of the fact lies in the following facts. Brain tumors are:
1) Extremely Lethal
2) Very similar in visual representation and pixel values for both malignant and benign tumors.
3) Are severely affected by slight disturbances and noise effects.
Although, various techniques for automated brain tumor classification are being developed, yet the major hindrance remains obtaining high percentage of classification accuracy. The reason for it is the fact that the pixel difference especially between cancerous and non-cancerous tumors is miniscule in nature. The pixel correlation is extremely high leading to failure for most commonly used classifiers. The challenging aspect remains the decision regarding the parameters or features which would guide any system in separating cancerous sets from the non-cancerous ones.[7] Even if such parameters are found, finding a condition that would render the final classification is infeasible. This paper proposes a technique wherein a Probabilistic Neural Network (PNN) is used in conjugation with image pre-processing techniques for achieving classification.

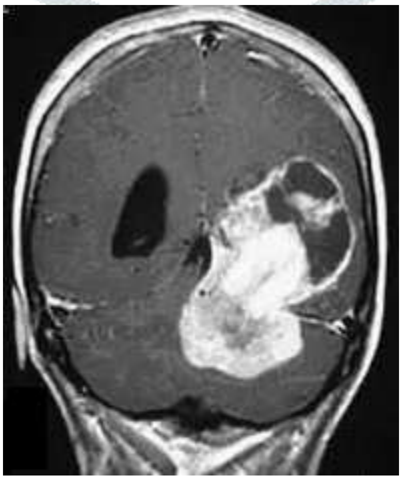


Figure.1 An MRI Image of the Human Brain with a Tumor

Figure.1 shows a typical MRI image of the human brain. It can be seen that a region is affected by tumor (apparently) but predicting whether it is malignant or benign is infeasible. Moreover, it may difficult to detect tumors in the first case.

## II. Data Preparation and Pre-Processing

### 2.1 Data Preparation

The data in this case is in the form of images. The source of the data has been taken as the UCI machine learning repository and www.mathworks.com. The raw data is categorized primarily into two sets:

1) Training data
2) Testing data

The training data set is the set of MRI images which would be fed to the designed artificial neural network at a later stage for training. The testing dataset is the remaining set of MRI images which have not been used for the purpose of training. It should be however noted that the training and testing datasets are disjoint sets and have no image in common.

### 2.2 Data Pre-Processing

The pre-processing of raw data is extremely important so that the neural network later on learns accurately based on attributes of the training data. The various stages of pre-processing used in this approach are given below:

#### 2.2.1 RGB to Gray Scale Conversion

Let the image I be a colored image and be a function of Red (R), Green (G) and Blue (B) values which typically signify the colour or frequency component of the image. Such an image can be represented as:

$$I = f(P_R, P_G, P_B) \qquad (2.1)$$

Here,
I represents the RGB image
f represents a transformational function that makes up the composite image
$P_R, P_G, P_B$ represents the R,G,B values of a Pixel 'P'.

RGB to Gray Scale conversion takes place by assigning weights to the R,G, and B pixel values at a point and summing them up. It can be mathematically represented as:

$$I_{gray} = w_1.I_R + w_2 I_G + w_3 I_B \qquad (2.2)$$

Here,

$I_{gray}$ represents the gray scale image
$w_1, w_2$ *and* $w_3$ represent the weights for the RGB to Gray Conversion

Generally, the values of $w_1, w_2$ *and* $w_3$ are taken to be 0.33, 0.5 and 0.166 respectively.

#### 2.2.2 The Discrete Wavelet Transform (DWT)

Images are prone to degradations by noise effects. The discrete wavelet transform acts like a spectral filter that removes or filter outs the noise in the less significant parts of the image. [1] Mathematically, the generic or continuous version of the discrete wavelet transform, i.e. the continuous wavelet transform is defined as:

$$F(S, P) = \int_{-\infty}^{\infty} \boldsymbol{f}(\mathbf{t}) \, ((\boldsymbol{S}, \boldsymbol{P}, \boldsymbol{t})) \, \boldsymbol{dt} \qquad (2.3)$$

Here,
$F(S, P)$ is the functional dependence of the transform values.
S represents the scaling operation
P represents the positional shifting operation.

Since the continuous wavelet transform generates a staggering amount of co-efficient values, therefore a down-sampled version of the continuous wavelet transform is used which is also termed as the discrete wavelet transform (DWT). The wavelet is helpful for analysis of non-stationary signals which do not follow the Dirichlet's criteria for Fourier methods to work. The wavelet transform has many families and the wavelet function is defined as:

$$W\psi\,(j, k) = \frac{1}{\sqrt{M}} \sum_n S(n).\psi(n)_{j,k} \qquad (2.4)$$

Here,

$\psi(n)_{j,k}$ transforms the continuous time axis into the discrete time axis 'n'

$\frac{1}{\sqrt{M}}$ is the normalizing factor

$S(n)$ is the discrete time signal

**2.2.3 Segmentation**

The complete image doesn't show any difference in images of different classes.[1] The part containing the tumor is to be separated or segmented out from the entire image. This is done in this case based on threshold based segmentation wherein pixel values showing sudden surge above a threshold are separated out. In general, the part of the image containing the tumor would show such a change and hence can be detected.[12]

**2.2.4 Principal Component Analysis (PCA)**

The principal component analysis is a dimensional reduction tool which extracts out the form a large dataset of values, the ones with least correlation in order to avoid redundancy. It can be also thought of as a techniques which extracts out the values which significantly reduce the dimensionality of the data set and only renders values which affect the results the most.[2]

## III. The Probabilistic Neural Network (PNN)

The probabilistic neural network (PNN) is a category of neural network that works on the principle of Baye's theorem of conditional probability. The internal structure of the PNN is similar as that of a general ANN.[2]

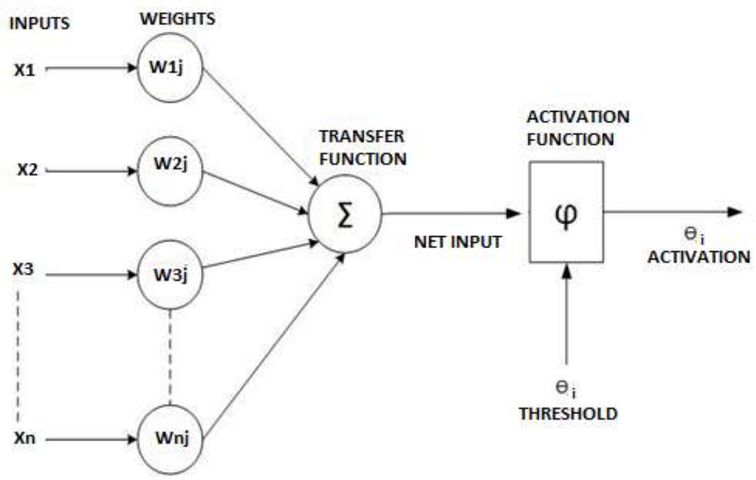



Figure.2 Mathematical equivalent of Probabilistic Neural Network

The output of such a neural network can be given by:

$$y= \sum_{i=1}^{n} X_i \,.\, W_i \;+\; \theta_i \qquad (3.1)$$

Here,
$X_i$ represents the signals arriving through various paths,
$W_i$ represents the weight corresponding to the various paths and
$\theta$ is the bias.

The probabilistic neural network classifies based on the Baye's theorem of conditional probability given by:

$$P(M/N) = \frac{P(N/M).P(M)}{P(N)} \qquad (3.2)$$

Here,
P(M) represents the individual probability of an event M
P(N) represents the individual probability of an event N
P(M/N) represents the probability of M given N is true
P(N/M) represents the probability of N given M is true

The technique basically computes the probability of a given data sample to belong to different defined classes and makes the final decision based on the maximum probability of belongingness to a class.

The designed probabilistic neural network is fed with the feature values of the pre-processed data which trains the neural network in accordance to the defined classes. The conceptual structure of the designed PNN is given below:

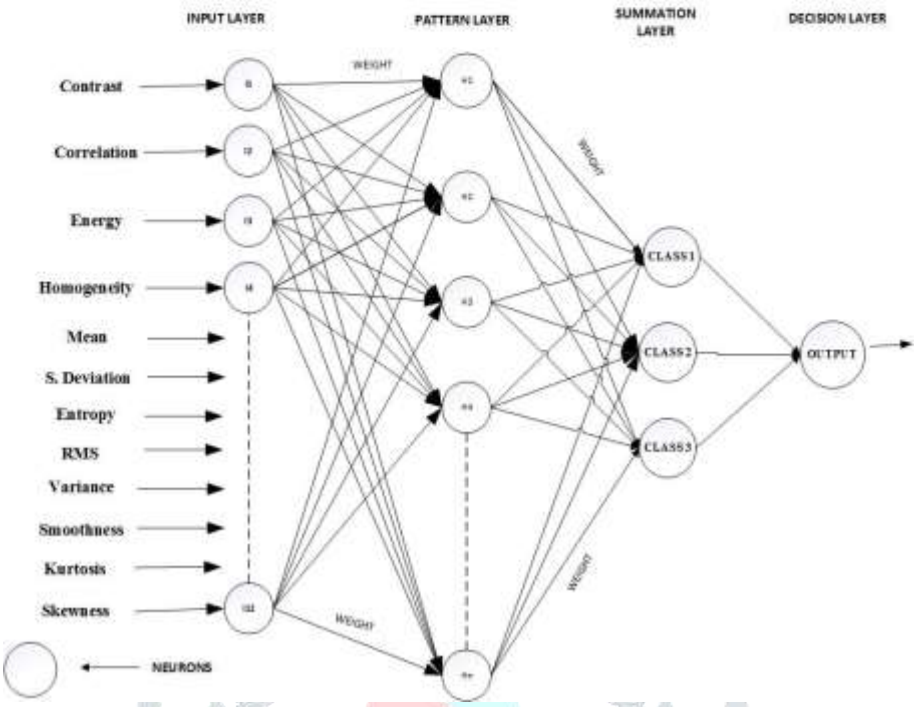



Figure.3 Conceptual Internal Structure of PNN

It can be seen that the proposed structure is fed with twelve features which are:
1) Contrast
2) Correlation
3) Energy
4) Homogeneity
5) Mean
6) Standard Deviation
7) Entropy
8) RMS value
9) Variance
10) Smoothness
11) Kurtosis
12) Skewness

## IV. PROPOSED TECHNIQUE

The proposed technique can be expressed as a sequential implementation of the following step:

Step.1: Divide data into training and testing data sets.
Step.2: Apply RGB to Gray Scale Conversion.
Step.3: Apply Segmentation
Step.4: Apply DWT
Step.5: Apply PCA
Step.6: Compute twelve attributes or features for the entire training dataset
Step.7: Design a probabilistic neural network (PNN)
Step.8: Train PNN with training dataset
Step.9: Test PNN using the testing dataset
Step.10: Compute Accuracy

The aforesaid steps are presented in the form of a flowchart in the following section.

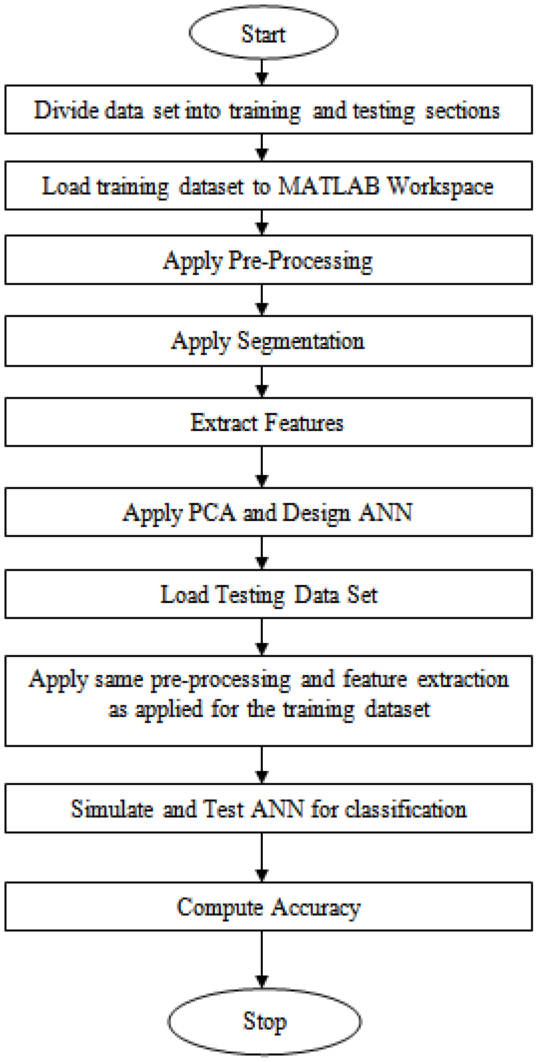



Figure.4 Flowchart of Proposed System

## V. RESULTS AND DISCUSSION

The results obtained are presented in this section. The system is designed on MATLAB 2017a.

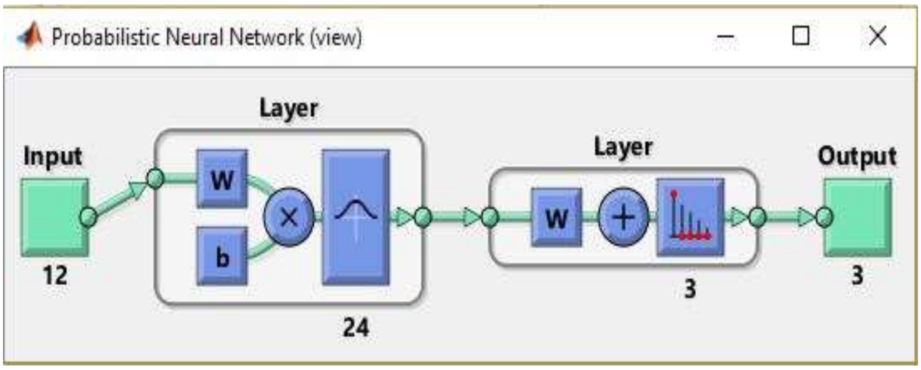
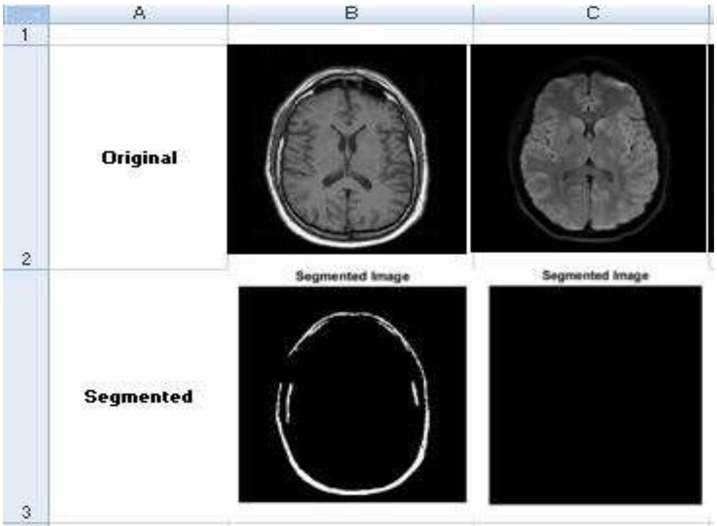



Figure.6 Design of Probabilistic Neural Network

It can be seen that the input layer of the proposed PNN has 12 neurons corresponding to the 12 feature values. The output layer has 3 neurons corresponding to the three classes i.e.

1) Normal
2) Malignant
3) Benign




| | A | B | C |
|---|---|---|---|
| 4 | | 1 | 2 |
| 5 | MRI Type | Normal | Normal |
| 6 | Contrast | 0.423863636 | 0.417613636 |
| 7 | Correlation | 0.080305851 | 0.193570981 |
| 8 | Energy | 0.702853822 | 0.709520919 |
| 9 | Homogeneity | 0.912159091 | 0.914299242 |
| 10 | Mean | 0.003935541 | 0.007328694 |
| 11 | Standard_Deviation | 0.10655652 | 0.106376924 |
| 12 | Entropy | 2.743369367 | 2.575002567 |
| 13 | RMS | 0.106600358 | 0.106600358 |
| 14 | Variance | 0.011361868 | 0.011315766 |
| 15 | Smoothness | 0.879123104 | 0.931240417 |
| 16 | Kurtosis | 6.961505164 | 8.09675679 |
| 17 | Skewness | 0.711892665 | 0.861298531 |

Figure.7 Original and Segmented MRI images and feature values

Figure.7 depicts the original and segmented forms of 2 MRI image samples. Their corresponding feature values are also shown. It can be observed that the feature values vary by very small amounts in different cases. Hence pattern recognition becomes complex.

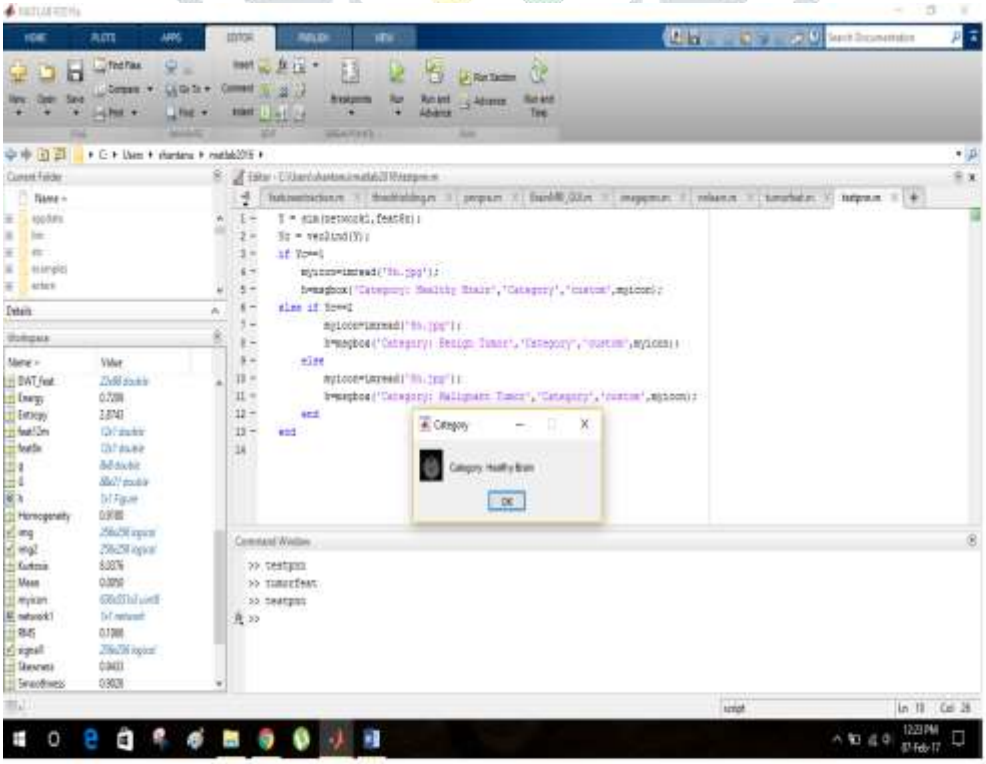


Figure.8 MATLAB Command window on classification pop-up

The accuracy computed for the dataset is computed as:

$$Ac = \frac{TP+TN}{TP+TN+FP+FN} \tag{10}$$

Here
Ac represents accuracy
TP represents true positive
TN represents true negative
FP represents false positive
FN represents false negative
It was found that, for the given dataset
Ac= 17+9/17+9+0+1 = .9629 = 97% (approx) having 9 normal, benign and malignant cases respectively corresponding to the testing class.

## VI. CONCLUSION

It can be concluded that the proposed system attains high accuracy of classification due to the probabilistic nature of the PNN which has emerged as an extremely effective classification tool for non-stationary and abruptly changing signals and datasets. The rigorous image pre-processing is done which helps in computing the feature values more precision which finally train the neural network. The exhaustive feature set ensures that the predictive classification is holistic in nature and maintains its accuracy across various datasets which can be tested.

## REFERENCES

**[1]** Hasan A., Dr.Jumaa D.& Dr.Bashk S.(2016) **"**Brain image classification based on discrete wavelet transform and probabilistic neural network" International Journal of Scientific & Engineering Research, Volume 7, Issue 5

**[2]** Gaikwad S.& Joshi M.(2015) "Brain Tumor Classification using Principal Component Analysis and Probabilistic Neural Network", International Journal of Computer Applications (0975 – 8887) Volume 120 – No.3.

**[3]** Mala K.& Alagappan V(2015), "Neural network based texture analysis of CT images for fatty and cirrhosis liver classification**"** Applied Soft Computing Volume 32, , Pages 80–86.

**[4]** Menzes B.,(2015), "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)", IEEE Transactions on Medical Imaging, VOL. 34, NO. 10.

**[5]** Dr.Kulhalli K. & Kolge V(2014), "Primary Level Classification of Brain Tumor using PCA and PNN", International Journal on Recent and Innovation Trends in Computing and Communication Volume: 2 Issue.

**[6]** Popescu L & Sasu I(2014),"Feature extraction, feature selection and machine learning for image classification: A case study" International Conference on Optimization of Electrical and Electronic Equipment (OPTIM), IEEE.

**[7]** Sridhar D. & Krishna M(2013), "Brain Tumor Classification U sing Discrete Cosine Transform and Probabilistic Neural Network", International Conference on Signal Processing, Image Processing and Pattern Recognition [ICSIPR].

**[8]** Sridhar D. & Krishna M (2013), "Brain Tumor Classification Using Discrete Cosine Transform And Probabilistic Neural Network" International Conference on Signal Processing, Image Processing and Pattern Recognition [ICSIPR] IEEE.

**[9]** Dahab D.,Ghoniemy S. & Selim G.(2012),"Automated Brain Tumor Detection and Identification Using Image Processing and Probabilistic Neural Network Techniques" International Journal of Image Processing and Visual Communication Volume (Online) 1 , Issue 2.

**[10]** Hua X.(2012), "Human–computer interactions for converting color images to gray" National Journal on Neurocomputing 85 , Elsevier, PP 1–5.

**[11]** Kolge V.& Kulhalli K.(2012), "PCA And PNN Assisted Automated Brain Tumor Classification", IOSR Journal of Electronics and Communication Engineering (IOSR-JECE), PP: 19-23.

**[12]** Rao K., Stephen M.,& Phanindra D.(2012), "Classification Based Image Segmentation Approach" International Journal of Computer Science And Technology, Vol. 3, Issue 1.

**[13]** Othman M. & Basri M(2011), "Probabilistic Neural Network for Brain Tumor Classification", Second International Conference on Intelligent Systems, Modeling and Simulation, IEEE Computer Society.

**[14]** Othman M. & Basri M.(2011), "Probabilistic Neural NetworkFor Brain Tumor Classification" , Second International Conference on Intelligent Systems, Modelling and Simulation.

**[15]** Ruikar S., & Doye D.(2010), " Image Denoising using Wavelet Transform",International Conference on Mechanical and Electrical Technology (ICMET),IEEE,pp. 509-515.