

# FEATURE ANALYSIS FOR BRAIN TUMOR DETECTION USING PROBABILISTIC NEURAL NETWORK CLASSIFIER

<sup>1</sup>Anna P Asgar, Anita Joseph, Anjana Anmmariya Chandy, Anna Xavier,<sup>2</sup>Shyama Sreekumar

<sup>1</sup>Department of Electronics and Communication Engineering, Rajagiri School of Engineering and Technology, Ernakulam, India

<sup>2</sup>Assistant Professor, Department of Electronics and Communication Engineering, Rajagiri School of Engineering and Technology, Ernakulam, India

**Abstract:** The conventional method of brain tumor detection and classification is by human inspection of magnetic resonant images of brain. But it is difficult when large amount of data is to be diagnosed and to be reproducible and also large data needs to be stored in the memory of analyst for immediate remedy suggestion or for continuous monitoring of the tumor at the required last stages. Robust computerized segmentation algorithms can help physicians to detect and classify brain tumor in lesser time with greater accuracy. In this paper, the images are initially pre-processed using median filter. We developed an algorithm for stripping the skull before segmentation process. K means segmentation is done for effective segmentation followed by feature extraction by Gray Level Co-occurrence matrix. Kurtosis, Skewness, Contrast, Entropy, Variance, Standard deviation, Homogeneity, Root Mean Square, Inverse difference moment and Correlation are the various features extracted from Gray Level Co-occurrence matrix. The probabilistic neural network classifier is trained for classification of images into normal, benign or malignant. Finally the area of tumor growth is calculated from which the grade of tumor is identified. The analysis of different features obtained from Gray Level Co-occurrence matrix is carried out to find out the features which can classify the images with more accuracy and helps in removing unnecessary features. The performance of the proposed technique is evaluated using confusion matrix from which 100 % Sensitivity, 88.88% Accuracy, 75% Specificity, 83.33% Positive Predictive Value and 100% Negative Predictive Value is obtained.

**Index terms:** Classification, K-means Segmentation, Gray Level Co-occurrence matrix, Feature Analysis, Probabilistic Neural Network

## I. INTRODUCTION

Brain tumor is an abnormal condition in which cells grow up and multiply uncontrollably in the brain. Brain tumors may be either Benign or Malignant. A Benign tumor is a non-cancerous tumor which does not invade its neighboring tissues or spread around the body. They respond well to treatment and prognosis is usually favorable. A Malignant tumor is a cancerous tumor that may invade its surrounding tissues or spread around the body. It grows worst with the passage of time and ultimately results in the death of a person. Benign tumors usually grow very slowly while malignant tumors grow more quickly in size. Magnetic Resonant Imaging (MRI) of the brain is a painless and safe test that produces detailed images of the brain using a magnetic field and radio waves. When compared to Computed tomography (CT), MRI can be used to visualize finer details in the inner structure of the brain thus making it a special one for brain tumor detection. Manual segmentation of brain tumor from MR images is a very tiring and time consuming process that shows significant differences when performed by different experts. Automatic brain tumor segmentation algorithms for detecting and classifying brain tumor consumes lesser time and is more accurate when compared to manual segmentation algorithms. The main aim of the system is to effectively detect, extract and classify the brain tumor from MR images obtained from patients database and thus to obtain the size and stage of brain tumor. Various statistical textural features of the brain structure is taken into consideration to perform feature analysis and to find out the features which classifies images more accurately followed by performance evaluation. The rest of the paper is organized as follows: Section 2 presents literature survey, Section 3 presents the proposed algorithm, Section 4 presents the results and discussion, Section 5 presents the conclusion and future work.

## II. LITERATURE SURVEY

In recent years, various methods have been proposed for brain tumor segmentation and classification. Damodharan and Raghavan [1] have proposed a methodology for brain tumor detection and classification based on neural network.

In this method, segmentation of white matter, grey matter, Cerebrospinal fluid, and tumor region is done separately and the outcome had an accuracy of 83% using neural network based classifier. Zanaty [2] presented a hybrid type approach technique which combined FCM, seed region growing, and Jaccard similarity coefficient algorithm for segmented gray matter and white matter tissues measurements from MR images. Chaddad [3] has introduced a technique based on Gaussian mixture model (GMM) using MR images with automatic feature extraction for brain tumor detection. We have used principal component analysis (PCA) and wavelet based features in this method and the GMM feature extraction is performed. It resulted in an accuracy of 87.05% for the T1 and T2-weighted and 84.11% for FLAIR-weighted MR images. Samriti et al. [4] have developed a computer-based technique for identifying the tumor region in the brain using MRI images accurately. The classification was done for identifying whether the tumor is a benign or malignant one. The steps involved in the proposed technique were preprocessing, image segmentation, feature extraction and image classification using NN techniques and the tumor area location has been found. Their technique has been implemented using a user friendly Matlab GUI program. Vinay et al. [5] introduced an algorithm for segmented morphological approach. The abnormalities in brain images are detected using some of the classical image processing tools. Some image enhancement techniques were used to enhance the contrast and normalize the pixel values in

the image after which fast fourier transform and some morphological operations were applied to get the desired results.

### III. PROPOSED ALGORITHM

In this paper, a novel technique has been used for feature analysis which helps in identifying the features that are important and in removing unnecessary features [6]. First pre-processing of input image is done using median filter to remove unwanted signals or noise from it. The image is then subjected to skull stripping before the segmentation process. K means segmentation is done for segmentation followed by feature extraction by Gray Level Co-occurrence matrix (GLCM). Effective classification to normal, benign and malignant is done using Probabilistic Neural Network (PNN) followed by feature analysis and performance analysis. The block diagram of the proposed algorithm is shown in Fig.1

#### A. Pre-processing

Pre-processing on raw data refers to transformation on it, to feed them into artificial neural networks. The standard of Magnetic Resonant(MR) brain tumor images is improved by pre-processing step, thus making these images suited for future processing. It helps in the removal of irrelevant noise, smoothing regions of inner part and maintaining relevant edges. The pre-processing step includes median filtering and skull stripping.

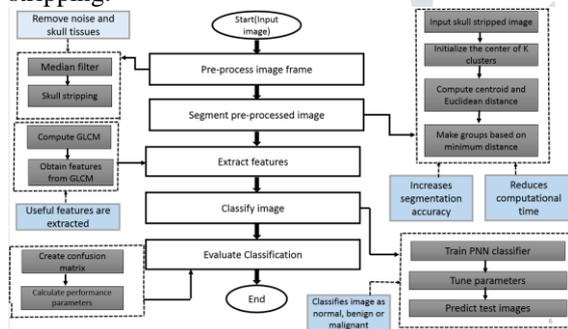


Fig. 1. Flow diagram of the proposed algorithm

1) *Median Filtering*: Median filter is a nonlinear digital filtering technique. It looks for each pixel in an image and its nearby neighbors to decide whether it is a representative of its surroundings or not. A 3\*3 square window is used here. The median value is calculated by sorting all the pixel intensity values from the surrounding neighborhood into ascending order and then replacing the pixel under consideration with the median pixel intensity value. The median filter does not create new unrealistic pixel values as the median value must be a value of one of the pixels in the neighborhood. As a result the median filter is efficient in preserving sharp edges.

2) *Skull Stripping*: MR brain images contain some non-cerebral tissues such as skin, fat, muscle, neck, and eye balls whose presence is a major obstacle for automatic brain image segmentation and analysis techniques [7]. First step in skull stripping is Otsu thresholding which converts a grayscale image to a binary image. It involves iterating through all the possible threshold values to find out a threshold value where the sum of foreground and background spreads is at its minimum.

All pixel intensity values with a value less than the Otsu threshold are considered as background pixels and is given an intensity value zero and the remaining pixels are considered as foreground pixels given a value one. Otsu thresholding is followed by morphological opening. Morphological opening is the dilation of the erosion of an image by a structuring element.

Structuring element used here is diamond. Dilation chooses the highest value by comparing all the pixel values in neighborhood of input image described by structuring element while erosion chooses the lowest value by comparing all the pixel values in the neighborhood of the input image. After morphological opening operation, the object with largest perimeter is chosen as the brain region. Then 1 is assigned to inside and 0 to outside of this object to obtain a mask. This mask is multiplied with the original image to get the skull stripped image.

#### B. Segmentation

Image segmentation is the process of partitioning a digital image into multiple segments. The result of image segmentation is a set of segments that collectively cover the entire image. Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity or texture. Adjacent regions are significantly different with respect to the same characteristics. The segmentation technique used here is K-means segmentation [8].

#### C. Feature Extraction

Feature extraction is the transformation of original data to a dataset with a reduced number of variables which contains the most discriminatory information. Gray level co-occurrence matrix (GLCM) is used for extracting second order statistical textural features. In the first step, the GLCM is computed and in the other step, the texture features based on the GLCM such as Kurtosis, Skewness, Contrast, Entropy, Variance, Standard deviation, Homogeneity, Root Mean Square, Inverse difference moment and Correlation are calculated [9].

1) Kurtosis: The shape of a random variable's probability distribution is described by the parameter called Kurtosis.

$$\text{KURTOSIS} = \left( \frac{1}{m \times n} \right) \frac{\sum (f(x,y) - M)^4}{SD^4} \quad (3)$$

2) Standard Deviation (SD): The standard deviation is the second central moment describing probability distribution of an observed population and can serve as a measure of inhomogeneity. A higher value indicates better intensity level and high contrast of edges of an image.

$$SD = \sqrt{\left( \frac{1}{m \times n} \right) \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} (f(x,y) - M)^2} \quad (4)$$

3) Entropy: Entropy is calculated to characterize the randomness of the textural image.

$$\text{ENTROPY} = - \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x,y) \log_2 f(x,y) \quad (5)$$

4) Skewness: Skewness is a measure of symmetry or the lack of symmetry

$$\text{SKEWNESS} = \left( \frac{1}{m \times n} \right) \frac{\sum (f(x,y) - M)^3}{SD^3} \quad (6)$$

5) Contrast: Contrast is a measure of intensity of a pixel and its neighbor over the image.

$$\text{CONTRAST} = \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} (x - y)^2 f(x,y) \quad (7)$$

6) Inverse Difference Moment (IDM): Inverse Difference Moment is a measure of the local homogeneity of an image. IDM may have a single or a range of values so as to determine whether the image is textured or non-textured.

$$IDM = \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} \frac{1}{1+(x-y)^2} f(x,y) \quad (8)$$

7) Correlation: Correlation feature describes the spatial dependencies between the pixels.

$$CORRELATION = \frac{\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} (x,y)f(x,y) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (9)$$

8) Homogeneity: Homogeneity describes the closeness of the distribution of elements in the GLCM.

$$HOMOGENEITY = \frac{\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x,y)}{1+(x-y)^2} \quad (10)$$

9) Variance: The dispersion of the values around the mean is represented by variance.

$$VARIANCE = \left( \frac{1}{m \times n} \right) \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} (f(x,y) - M)^2 \quad (11)$$

10) Root Mean Square (RMS): It computes the RMS value of each row or column of the entire GLCM matrix.

$$RMS = \sqrt{\left( \frac{1}{m \times n} \right) \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} (f(x,y))^2} \quad (12)$$

Here  $f(x,y)$  is the second order probability values and

$$MEAN (M) = \frac{1}{m \times n} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x,y) \quad (13)$$

D. Feature Analysis and Selection

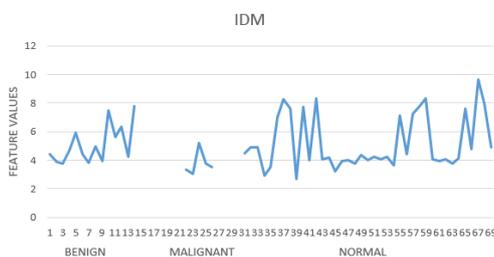


Fig. 2. IDM

Feature selection is the process of determining a subset of the initial features and is done based on feature analysis. The selected features are expected to contain the relevant information from the input data, so that the desired task can be performed by using this reduced representation instead of the complete initial data [10]. Analysis with a large number of variables generally requires a large amount of memory and computation power, also it may result in over fitting of classification algorithm and generalize poorly to new samples.

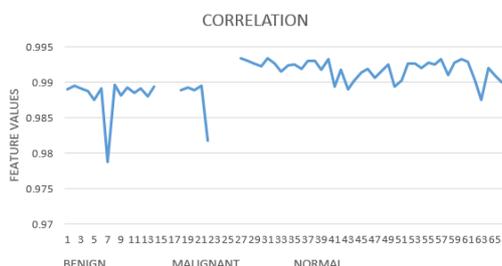


Fig. 3. Correlation

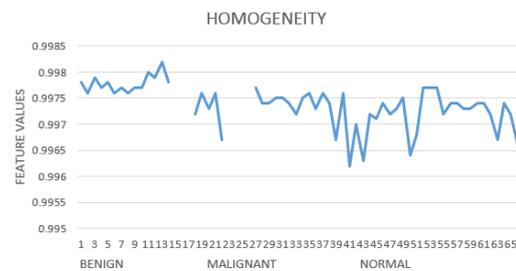


Fig. 4. Homogeneity

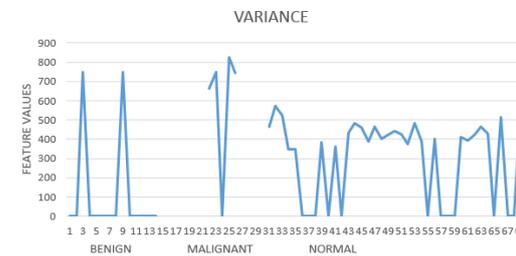


Fig. 5. Variance

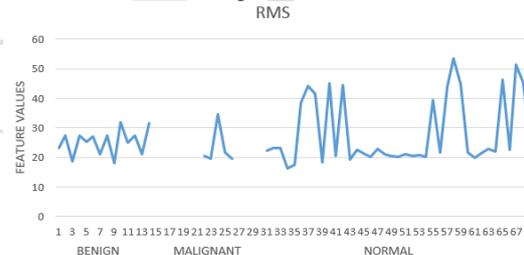


Fig. 6. RMS

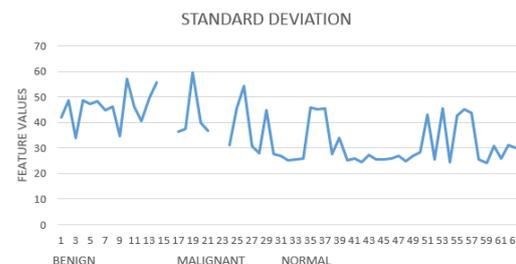


Fig. 7. Standard Deviation

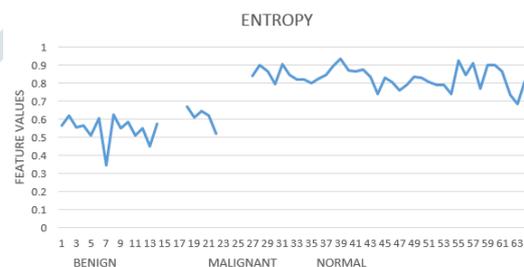


Fig. 8. Entropy

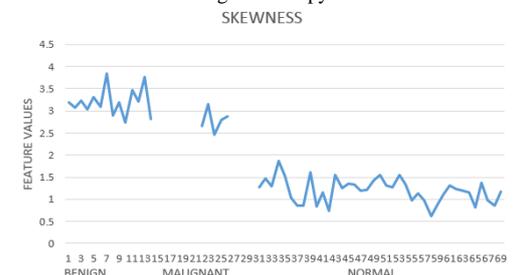


Fig. 9. Skewness

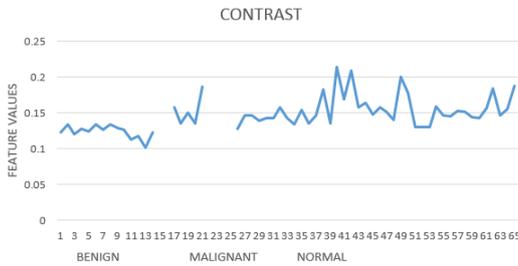


Fig. 10. Contrast

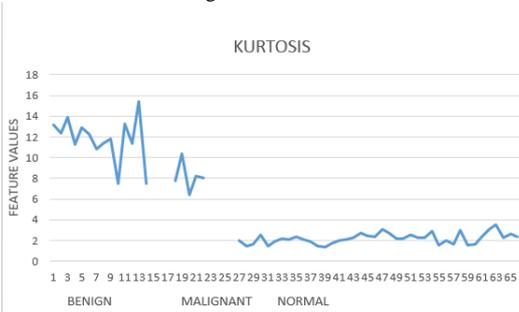


Fig. 11. Kurtosis

Various plots for different features considered are depicted in Figures 2,3,4,5,6,7,8,9,10 and 11. Kurtosis, Contrast, Skewness, Entropy and Variance shows distinct range of feature values for normal, benign and malignant tumor and are selected as the best features for classification. Homogeneity, Correlation, IDM, Standard deviation and RMS shows close range of feature values for normal, benign and malignant tumor and are considered as the worst features for classification.

E. Classification

Probabilistic Neural network (PNN) is a feed forward neural network which performs classification by learning from data and do not use rule sets [6]. NN can generalize using previous data and learn from past experiences. They have advantages like learning by themselves, tolerating fault and searching for the optimum. They perform well on difficult, multivariate, nonlinear and noisy domains, such as brain tissue segmentation, where it becomes more difficult to use decision trees, or rule-based systems. The four layers of PNN are Input layer, Pattern layer, Summation layer and Output layer. The input layer consist of input neurons to which various features extracted using GLCM is given as input. The pattern layer computes distances from the input vector to the training input vectors and produces a vector whose elements indicate how close the input is to a training input. The summation layer sums these contributions for each class of inputs to produce as its net output a vector of probabilities. The output layer picks the maximum of these probabilities, and produces '1' for malignant class, '2' for benign class and '3' for normal class.

F. Calculation of Area and Grade

Area of the tumor is calculated from the segmented image [11].

$$\text{Area of single pixel (A)} = \frac{1}{VR} \times \frac{1}{HR} \quad (14)$$

Where Vertical resolution (VR) = Horizontal resolution (HR) = 120dpi

$$\text{Area of tumor} = A \times N \quad (15)$$

where N is the number of white pixels in the image after segmentation.

Stage can be found out from the area of the tumor as follows:  
 Stage 1: The tumor is small (less than 2cm) and well localized.  
 Stage 2: The size of tumor ranges between 2cm and 4 cm.  
 Stage 3: The size of the tumor has exceeded 4cm.

IV. RESULTS AND DISCUSSION

Image dataset of 74 images were taken from various hospitals and confirmed by radiologist. Out of the 74 MR images 59 images were used for training and 15 images were used for testing the PNN. All the images are having 512x512 pixels size and 120dpi vertical and horizontal resolution. MATLAB which is developed by MathWorks is the tool used for the implementation of the proposed technique. The sample experimental results obtained from the proposed technique are depicted in Figures 12, 13, 14 and 15.

The performance analysis is done using confusion matrix and the various parameters such as Accuracy, Sensitivity, Specificity, Positive Predictive Value (PPV) and Negative Predictive Value (NPV) are calculated as in (16), (17), (18), (19) and 20.

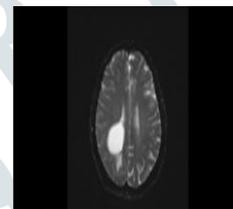


Fig. 12. Input image



Fig. 13. Median filtered image

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \quad (16)$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (17)$$

$$\text{Specificity} = \frac{TN}{FP+TN} \quad (18)$$

$$\text{PPV} = \frac{TP}{TP+FP} \quad (19)$$

$$\text{NPV} = \frac{TN}{TN+FN} \quad (20)$$

where TP is True Positive, FP is False Positive, TN is True Negative and FN is False Negative.

From Fig.16 TP=5, TN=3, FP=1, FN=0

The values of performance parameters when Kurtosis, Skewness, Variance, Contrast and Entropy, that are considered as the best features, are given as input values to PNN is: Accuracy= 88.88%, Sensitivity= 100%, Specificity=75%, PPV= 83.33%, NPV=100%.

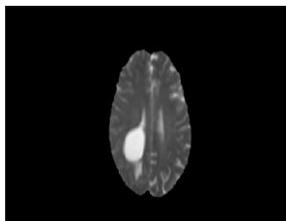


Fig. 14. Skull stripped image



Fig. 15. Segmented tumor image

|           |           | ACTUAL |        |           |
|-----------|-----------|--------|--------|-----------|
|           |           | NORMAL | BENIGN | MALIGNANT |
| PREDICTED | NORMAL    | 6      | 0      | 0         |
|           | BENIGN    | 0      | 5      | 0         |
|           | MALIGNANT | 0      | 1      | 3         |

|           |           | ACTUAL |           |
|-----------|-----------|--------|-----------|
|           |           | BENIGN | MALIGNANT |
| PREDICTED | BENIGN    | 5      | 0         |
|           | MALIGNANT | 1      | 3         |

Fig. 16. Confusion matrix of best features

|           |           | ACTUAL |        |           |
|-----------|-----------|--------|--------|-----------|
|           |           | NORMAL | BENIGN | MALIGNANT |
| PREDICTED | NORMAL    | 6      | 0      | 0         |
|           | BENIGN    | 0      | 4      | 0         |
|           | MALIGNANT | 0      | 2      | 3         |

|           |           | ACTUAL |           |
|-----------|-----------|--------|-----------|
|           |           | BENIGN | MALIGNANT |
| PREDICTED | BENIGN    | 4      | 0         |
|           | MALIGNANT | 2      | 3         |

Fig. 17. Confusion matrix of worst features

From Fig.17 TP=4,TN=3, FP=2,FN=0.

The values of performance parameters when IDM, Homogeneity, SD, RMS and correlation, that are considered as the worst features, are given as input values to PNN is: Accuracy=77.77%, Sensitivity= 100%, Specificity=60%, PPV=66.66%,NPV=100%.

## V. CONCLUSION AND FUTURE WORK

We proposed a computer aided system for the classification of brain MR images into normal, benign and malignant followed by calculation of area and grade of tumor from the segmented image. As a part of pre-processing we used median filtering which helps in removing noises from the input image. We used a skull stripping algorithm based on Otsu threshold technique to improve the skull stripping performance. The skull stripped image was segmented using K means segmentation and various features were extracted using GLCM. The resulting images of segmentation were further used for calculation of area and grade of tumor. Feature analysis of the extracted features was done to find out the best and worst features for classification. The best and worst features were given as separate inputs to the probabilistic neural network for classification and the best features resulted in an accuracy of 88.88% and the worst features resulted in an accuracy of 77.77% which validates the process of feature analysis. The

computational time taken for the detection of tumor was 8 seconds. From the experiments performed on different images, it is concluded that the proposed brain tumor detection algorithm is fast and accurate when compared with manual segmentation. In the future work, we would like to show that as the area of the tumor increases only limited number of efficient features are required for fast and accurate classification.

## REFERENCES

- [1] Selvaraj Damodharan, Dhanasekaran Raghavan, "Combining Tissue Segmentation and Neural Network for Brain Tumor Detection", *International Arab Journal of Information Technology*, 2015, Vol.12, no. 1, pp. 42–52
- [2] E. A. Zanaty, "Determination of gray matter (GM) and white matter (WM) volume in brain magnetic resonance images (MRI)," *International Journal of Computer Applications*, vol. 45, pp. 16–22, 2012.
- [3] A. Chaddad, "Automated feature extraction in brain tumor by magnetic resonance imaging using gaussian mixture models," *International Journal of Biomedical Imaging*, vol. 2015, Article ID 868031, 11 pages, 2015.
- [4] Samriti, Paramveer Singh, "Brain Tumor Detection Using Image Segmentation", *International Journal of Engineering Development and Research*, Volume 4, ISSN:2321-9939, pp.1923-1925.
- [5] Parameshwarappa, Nandish.s, "A Segmented Morphological Approach to Detect Tumor in Brain Images", *International Journal of Advanced Research in Computer Science and Software Engineering*, 2014, Volume 4, ISSN:2277-128, pp.408-412
- [6] N. Varuna Shree, T. N. R. Kumar. March 2018, "Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network", *Brain Informatics*, Volume 5, Issue 1, pp 23–30.
- [7] Demirhan.A, Toru.M, Guler.I, "Segmentation of Tumor and Edema along with Healthy Tissues of Brain Using Wavelets and Neural Networks", *IEEE Journal of Biomedical and Health Informatics*, 2015, pp.2168-2194.
- [8] Swapnil R. Telrandhe, Amit Pimpalkar, Ankita Kendhe, "Detection of Brain Tumor from MRI images by using Segmentation and SVM", *World Conference on Futuristic Trends in Research and Innovation for Social Welfare*, Volume 10 Issue I, 2016, ISSN: 2537-8453, pp.112-123.
- [9] N. B. Bahadure, A. K. Ray, H. P. Thethi, "Image Analysis for MRI Based Brain Tumor Detection and Feature Extraction Using Biologically Inspired BWT and SVM", *International Journal of Biomedical Imaging*, vol.17, Article ID 9749108, pp.152–165.
- [10] Neha Baraiya, Hardik Modi, Himanshu Patel, "An Efficacious Graphical User Interface Implementation for Automatic Classification of Brain Tumor from Magnetic Resonance Imaging Images Using Image Processing", *International Journal for Research in Applied Science and Engineering Technology*, Volume 6 Issue I, January 2018, ISSN: 2321-9653, pp.1238-1246.
- [11] Hunnur, S. S., Raut, A., Kulkarni, S. (2017), "Implementation of image processing for detection of brain tumors", 2017 *International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 278–283