

# Naïve Bayesian Classification Based Glioma Brain Tumor Segmentation Using Grey Level Co-occurrence Matrix Method

R. Anitha<sup>1</sup>, K. Sundaramoorthy<sup>2</sup>, D. Suseela<sup>3</sup>, T. Suganya Thevi<sup>4</sup>, S. Selvi<sup>5</sup>, Mohammad Aljanabi<sup>6</sup>

<sup>1</sup>Professor, Department of Biomedical Engineering, Jerusalem College of Engineering (Autonomous)

Pallikaranai, Chennai, Tamil Nadu-600100, India

E-mail: anithagodavarthy@gmail.com

<sup>2</sup>Professor, Department of information technology, Jerusalem College of Engineering (Autonomous)

Pallikaranai, Chennai, Tamil Nadu-600100, India

E-mail: ksundaramoorthyphd77@gmail.com

<sup>3</sup>Assistant professor

Department of Artificial Intelligence and Machine Learning

Bannari Amman Institute of Technology (Autonomous)

Sathyamangalam, Tamil Nadu 638401, India. Email: suseelad@bitsathy.ac.in

<sup>4</sup>Assistant Professor, Department of Electronics and Communication Engineering,

PSNA College of Engineering and Technology

Kothandaraman Nagar, Tamil Nadu 624622, India

E-mail: suganyaathevi@gmail.com

<sup>5</sup>Professor, Department of Electrical and electronics Engineering

Panimalar Engineering College, Chennai - 600 123, India, Email: selviselvaraj@gmail.com

<sup>6</sup>Department of computer, College of Education, Alirqia university, 9985+758, Baghdad, Iraq,

Email: mohammad.ca88@gmail.com

**Abstract**— Brain tumors vary widely in size and form, making detection and diagnosis difficult. This study's main aim is to identify abnormal brain images., classify them from normal brain images, and then segment the tumor areas from the categorised brain images. In this study, we offer a technique based on the Nave Bayesian classification approach that can efficiently identify and segment brain tumors. Noises are identified and filtered out during the preprocessing phase of tumor identification. After preprocessing the brain image, GLCM and probabilistic properties are extracted. Naive Bayesian classifier is then used to train and label the retrieved features. When the tumors in a brain picture have been categorised, the watershed segmentation approach is used to isolate the tumors. This paper's brain pictures are from the BRATS 2015 data collection. The suggested approach has a classification rate of 99.2% for MR pictures of normal brain tissue and a rate of 97.3% for MR images of aberrant Glioma brain tissue. In this study, we provide a strategy for detecting and segmenting tumors that has a 97.54% Probability of Detection (POD), a 92.18% Probability of False Detection (POFD), a 98.17% Critical Success Index (CSI), and a 98.55% Percentage of Corrects (PC). The recommended Glioma brain tumour detection technique outperforms existing state-of-the-art approaches in POD, POFD, CSI, and PC because it can identify tumour locations in abnormal brain images.

**Keywords**- Cells, Brain Tumor, Feature extraction, Classification, Segmentation.

## I. INTRODUCTION

Brain tumors are the irregular spots or shapes in MR scans of the head. Cells in brain tumors develop so rapidly that they rupture adjacent cells. The mortality rates associated with brain tumors are very variable and dependent on the area of the brain where the tumor is located. Its seriousness pertains specifically to the dimensions and consistency of the tumor. These days, a cross-sectional scan of the brain is screened using scanning technology. Magnetic Resonance Imaging

(MRI), Computer Tomography (CT), and Positron Emission Tomography (PET) are the three main types of these scanning methods. The aberrant patterns in brain MR pictures may be seen more clearly with MRI scanning, making it the method of choice in this article. Grey Matter (GM), White Matter (WM), and Cerebro Spinal Fluid (CSF) are the three main types of brain tissue. The tumors can be formed in any type of brain tissues based on the immunity of the patient. There are two stages in tumor analysis as tumor detection stage and tumor diagnosis stage. In tumor detection stage, tumors are

segmented using various image processing techniques. In tumor diagnosis stage, the tumor locality is analyzed with respect to various segmented brain tissues. GM and WM tend to occur in conjunction with less severe malignancies. CSF is associated with tumor grades 3 and 4. With the right treatment at regular intervals, lower-grade cancers are curable and may be managed. Due to the great firmness of higher-grade tumors in relation to the location of CSF, they are incurable. Glioma and meningioma are two examples of high-grade tumors. The Glioma brain tumor is shown in Figure 1(a), whereas the Meningioma is depicted in Figure 1(b).

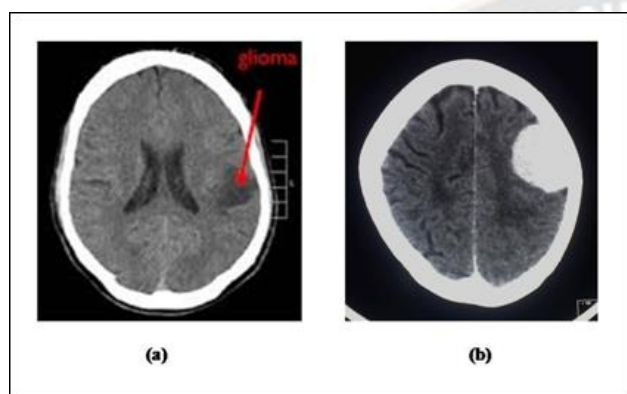


Figure 1 MRI image of tumor affected brain (a) Glioma (b) Meningioma

## II. LITERATURE SURVEY

Lavanyadevi et al. (2017) applied Principal Component Analysis (PCA) technique on brain MR images in order to compress the image patterns. After compressing brain MR data, the scientists used Probabilistic Neural Network (PNN) classification approach to determine if the original picture was normal or not. The authors used GLCM features for these efficient brain MR image classifications. After the source picture has been classified as normal or abnormal, k-means clustering is used on the aberrant brain MR image to locate and isolate the tumor. Shil et al. (2017) employed a thresholding approach like the Otsu binarization method to segregate aberrant patterns from magnetic resonance images of the brain. Then, Discrete Wavelet Transform (DWT) feature was extracted from this threshold image. The PCA approach was used to filter out the retrieved feature set from the MR image of the brain. These aberrant patterns were then classified as tumor-related or non-tumor-related using Support Vector Machine (SVM).

Convolutional Neural Networks were utilised by Pereira et al. (2016) to categorise MR pictures of the brain as normal or pathological. The authors put their suggested approach to the test using BRATS dataset brain pictures. The authors' suggested technique was given a Dice Similarity Coefficient Metric (DSCM) score of 0.88. In 2015, Eman Abdel-Maksoud and colleagues suggested a technique for segmenting brain

tumors from MR images of the brain using a hybrid clustering approach. Using the categorised brain pictures, textural patterns of non-uniform objects were recovered. The authors used the Brain Web data set to test their suggested strategy for detecting and segmenting brain tumors. The authors attained a perfect rate of accuracy and a recall rate of 85.7%. Using the k-means classification approach, Islam and Ahmed (2013) developed a method for efficiently detecting and segmenting tumor areas in brain MR images. In order to evaluate the efficacy of brain MR images for clinical diagnosis, the authors utilised a variety of categorization approaches. Spatial Fuzzy C-means (PET-SFCM) is a method of clustering algorithm suggested by Meena and Raja (2013) for use with Positron Emission Tomography (PET) scan image datasets. The goal function for each cluster is being updated as the algorithm combines spatial neighbourhood data with traditional FCM. Image segmentation is aided by the spatial connection of surrounding pixels. In 2012, Yerpude et al. presented K-means clustering, Expectation Maximization (EM), and Normalized Cuts for picture segmentation (NC).

A graph-based approach called Normalized Cut was compared to the two prior unsupervised learning techniques. MRI scans of the brain taken from the Brain Online database. E. The complexity of the neurological system makes it difficult to create an automated system for diagnosing brain tumors, as stated by Aarthi et al. (2022). Brain tumor diagnosis across diverse patient populations is a common use of the many data mining approaches outlined by Reddy et al. (2018). In this paper, we use two different data mining classification techniques to predict the type of brain tumor a patient has. The first method is called the Nave Bayesian classification technique. This method has already been utilised to determine the type of tumour, and it also enables the examination of historical data taken from data sets. As a result, it is able to provide neurologists with assistance as they make their predictions. Prediction research shows that the decision tree outperforms the naive Bayes classifier in terms of both speed and accuracy.

## III. MATERIALS AND METHODS

### 3.1 Materials

The magnetic resonance scans of the brain in this article are from the BRATS 2015 database (2015). There are many MR pictures of the brain stored in this collection, collected from patients in medical facilities all around the globe. These brain MR images are categorized into different modules as normal, abnormal and non-diagnosed. In this paper, 125 brain MR images are accessed from normal category and 75 brain MR images are accessed from abnormal category. All these brain tumor affected images are also having manually tumor segmented regions as gold standard images.

### 3.2 Methods

#### 3.2.1 Preprocessing

During acquisition of brain MR images in health centers, the brain MR images are affected by impulse noises. These noises are resembled with tumor pixels and hence they should be removed before tumor detection process starts. In this investigation, noise in the original MR scans of the brain was eliminated using an adaptive median filter.

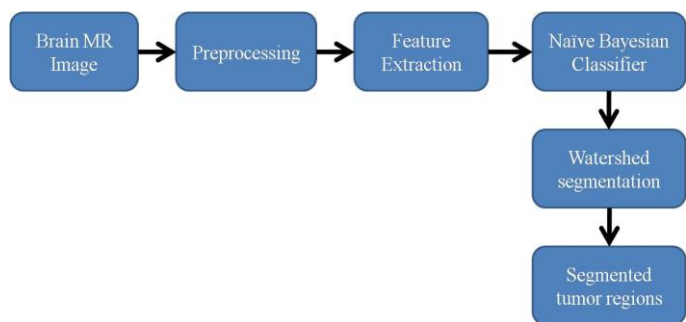


Figure 2 Flow chart for Segmenting brain tumours

#### 3.2.2 Feature Extraction

In this publication, MR images of the brain with the noise removed are analysed in order to obtain GLCM and probabilistic characteristics.

##### (i) GLCM features

Each pixel's connection to its neighbours is codified by the texture characteristics. For each possible orientation, a unique pattern will be generated. In this study, we use the GLCM matrix format to extract 45-degree-angled texture patterns from a preprocessed picture. Normal and pathological brain images may be distinguished from one another based on the contrast, energy, homogeneity, and correlation that are produced from the GLCM matrix. The normal brain texture values and aberrant brain texture values are shown in Table 1 below. It is clear from looking at Table 1 that the recovered GLCM patterns vary from usual to anomalous. The training dataset's normal and pathological brain pictures are used to create these texture properties.

Table 1 Features of the extracted GLCM

GLCM parameters	Normal brain MR image	Glioma brain MR Image
Contrast	2.734	1.384
Energy	3.738	4.234
Homogeneity	-6.573	-1.754
Correlation	-7.864	-1.354

##### (ii) Probabilistic features

The probabilistic characteristics draw attention to the pixel-level differences between the tumor area and the rest of the image. In the training dataset, these characteristics are retrieved from both normal and pathological MRI scans of the

brain. To achieve a high classification rate, the preprocessed brain MR image is analysed to extract probabilistic characteristics such as similarity metric, border metric, entropy, kurtosis, and grey level feature, as shown in the following equations.

$$\text{Similarity Metric (SM)} = \frac{\sum_{i=1}^N |I(i) - I(i-1)|}{M \times N} \quad (1)$$

Whereas, in a preprocessed MRI of the brain,  $I(i)$  is the centre pixel in a  $3 \times 3$  window, and  $I(i-1)$  is the prior pixel centred at  $I$ .  $M$  and  $N$  stand for the image's width and height, respectively.

$$\text{Boundary Metric (BM)} = \sum_{i=1}^N |I(i) - I(i-1)|^2 \quad (2)$$

$$\text{Edge boundary strength} = \frac{\sum [\mu_x(i) + \mu_y(i)]}{M \times N}$$

$$\text{Entropy (H)} = \sum_{i=0}^{N-1} P(i) \cdot \log_2 P(i)$$

$$\text{Kurtosis} = \sigma^{-4} \sum_{i=0}^{N-1} (i - \mu)^4 \cdot P(i) - 3$$

whereas,  $\mu$  is the mean of the  $3 \times 3$  window over the center pixel.

$$M = \sum_{i=0}^{N-1} I \times P(i)$$

$$\text{Grey level feature} = \frac{1}{N} \sum_{i=0}^{N-1} \frac{I(i,j)}{i^2}$$

The feature vector used by the classifier to distinguish between normal and pathological Glioma brain pictures is constructed using the retrieved GLCM characteristics and probabilistic features.

#### 3.2.3 Naïve Bayesian classification

This article uses a supervised classification approach to identify brain MR images from Glioma. This algorithm has probabilistic and statistical assumptions as its starting point. The Bayesian classifier receives as input the mutually correlated characteristics retrieved from the MR brain picture. The main advantages of this Bayesian classifier are given in the following points.

- When brain MR images include many extracted characteristics, neural networks and support vector machines struggle to classify pixels. High feature dimensions cause these algorithms to fail classification. In this research, Bayesian classification technique classifies brain MR image pixels for both low and high dimension derived feature sets.
- The computational time for training and testing of this classification algorithm is superior to other conventional classification algorithms.
- This classification algorithm reduces the high dimensional extracted feature set by removing the irrelevant features from the feature set.

The Bayesian classification technique uses Bayes theorem, which states,

The feature data is represented as 'X' and we assume that its classification label is not known. Further, we assume that



its hypothesis is 'H'. Then, the Bayes theorem states that the classification label of feature data 'X' may be 'B' and its probability for classification label is given in the following equation as,

$$P(B|X) = \frac{P(X|B) \cdot P(B)}{P(X)}$$

where as,

P (B|X) is represented the expected classification label for the feature data and P (B) is the probability of classification label. P (X|B) is the likelihood of the classification label and P(X) is the probability of the prediction for classification label belonging to feature data X. The Naïve Bayesian classification technique uses Bayes theorem to categorise MR brain image pixels as normal or cancerous based on recovered feature set. These stages illustrate this classification method.

**Step 1:**

Assign classification label for all extracted features in training dataset images, which is in the form of  $X = \{x_1, x_2, \dots, x_n\}$ . The number of images in training dataset is represented by 'n'.

**Step 2:**

Assume the following classification labels for the extracted features in testing brain MR image as,  $B = \{B_1, B_2, \dots, B_m\}$  and  $C = \{C_1, C_2, \dots, C_m\}$  Where m is the sample dataset's pixel count for an MR picture of the brain.

The classification label 'B' represents the pixel with normal features and the classification label 'C' represents the pixel with abnormal features as tumor.

**Step 3:**

The pixel in image is classified using the following equation as,

$$P \rightarrow 'B' \text{ if } P(B_i|X) > P(B_j|X) ; \text{ where, } 1 \leq j \leq m;$$

$$P \rightarrow 'C' \text{ if } P(B_i|X) \leq P(B_j|X) ; \text{ where, } 1 \leq j \leq m;$$

**3.2.4 Watershed Segmentation**

The segmentation approach is used to separate tumor locations in aberrant brain pictures after categorization. This article segments brain MR tumor areas using watershed segmentation. Steps use this watershed segmentation approach.

**Step 1:** Cover the categorised MR brain picture with a 3x3 window.

**Step 2:** Compute gradients of the pixels in sub window.

**Step 3:** Create marker line on the gradient image and then construct edge map using its gradient values.

**Step 4:** Proceed steps 1 to 3 till the end pixels in an classified image.

**Step 5:** Apply threshold on the gradient map image. The image is segmented based on the threshold value.

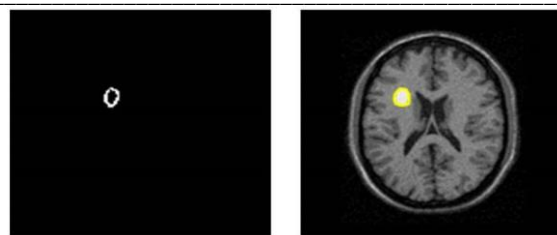


Figure 3 (a) Segmented tumor region (b) Overlay of segmented tumor region over source MR brain image

In Figure 3(a), the tumour area was isolated using the watershed segmentation approach, and in Figure 3(b), the segmented tumour region is superimposed on the original MR brain picture. The raw brain MR images taken from the public dataset (Figure 4(a)), the ground truth pictures (radiologist-marked tumor images) Figure 4(b), and the segmented image of the tumor using the proposed technique (Figure 4(c)) are all shown below.

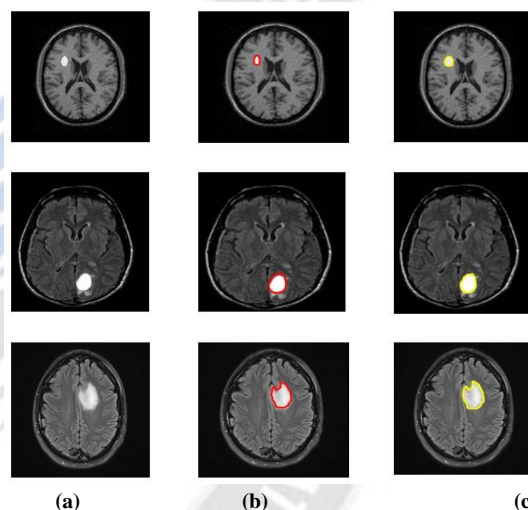


Figure 4 (a) Source brain MR images (b) Ground truth images (c) Tumor segmented image by proposed method.

**IV. RESULTS**

In this research, we utilise MATLAB R2016 to analyse the results of the Glioma tumor identification and segmentation method. The proposed method's classification effectiveness is measured by calculating the classification rate, which is the proportion of correctly labelled brain MR images to the entire amount of magnetic resonance brain pictures in the database. Percentages may have a number between zero and one hundred. In this article, the suggested classification system is used on a dataset of 200 brain MR images received from a publicly available source. There are a total of 200 magnetic resonance imaging (MRIs) of the brain, 125 of which represent "normal" brain tissue and 75 representing Glioma. The approach presented here was successfully applied to 73 Glioma brain MR pictures and 124 normal brain MR images. As a result, the suggested approach has a classification

rate of 99.2% for normal brain MR pictures and a rate of 97.3% for Glioma brain MR images. As can be shown in Table 2, the suggested technique achieves an overall classification rate of 98.25% when applied to a dataset consisting of 197 out of 200 brain MR images.

Table 2 Classification rate comparisons between suggested and existing methods

Methodologies	Number of brain MR images correctly classified	Classification Rate (%)
Proposed method	197	98.25
Neural networks	192	96
SVM (linear)	187	93.5
SVM (RBF)	185	92.5

The suggested method's efficacy is measured using the Critical Success Index (CSI), the Percentage of Corrects (PC), the Probability of Detection (POD), and the Probability of False Detection (POFD) (PC). These performance evaluation parameters are computed using contingency table represented in Table 3.

Table 3 Contingency table for proposed method validation

Observation by proposed method	Observation from expert radiologist	
	Tumor pixel	Non-tumor pixel
Tumor pixel	A	B
Non-tumor pixel	C	D

The number of correctly detected pixel ration in classified MR brain image is defined by POD and its value varies from 0 and 255. The number of wrongly detected pixel ration in classified MR brain image is defined by POD and its value varies from 0 and 255. CSI is the percentage of pixels that were properly detected by both the radiologist and the suggested algorithm. The percentage of correctly classified pixels is defined by PC. These performance evaluation metrics are given in the following equations as,

$$POD = \frac{A}{A+C}$$

$$POFD = \frac{B}{B+D}$$

$$CSI = \frac{AB}{A+B+C}$$

$$PC = \frac{A+D}{N}$$

Table 4 Evaluation of the Proposed Methodology's Performance

Performance evaluation metrics	Experimental results (%)
POD	97.64
POFD	92.18
CSI	98.37
PC	98.75

Table 4 contrasts the suggested approach of tumor identification and segmentation with the standard methods currently used to identify and segment brain tumors. The POD, POFD, CSI, and PC performance metrics are analysed in Table 5. This work proposes a technique for detecting and

segmenting tumors, and its results show that it can get as high as 97.64% POD, 92.18% POFD, 98.37% CSI, and 98.75% PC.

## V. DISCUSSIONS

The conventional neural networks classifies 192 brain MR images correctly over 200 brain MR images and achieves 96% of classification rate. The conventional neural networks classifies 187 brain MR images correctly over 200 brain MR images and achieves 93.5% of classification rate. The classification rate of this method is low due to its low training rate and less number of training images with low pixel resolution. While differentiating Glioma brain pictures from normal brain images, the scientists also utilised an unsupervised classification strategy. Both linear and Radial Basis Function options are available for use with the SVM classifier (RBF). The SVM classifier with linear mode obtained 93.5% of classification rate by correctly classifying 187 images over 200 images. The conventional SVM classifier in RBF mode classifies 185 brain MR images correctly over 200 MRI images and achieves 92.5% of classification rate.

Table 5 Performance comparisons of suggested and established techniques

Methodology	Performance evaluation metrics			
	POD (%)	POFD (%)	CSI (%)	PC (%)
Proposed method	97.64	92.18	98.37	98.75
Nilesh Bhaskarrao Bahadure et al. (2017)	92.10	86.38	96.74	95.28
Sreedhanya et al. (2017)	91.27	85.28	94.57	94.67
Alfonse et al. (2016)	93.28	88.37	96.45	96.16

The conventional method of finding and dividing Glioma brain tumours, The SVM classification technique was employed by Nilesh Bhaskarrao Bahadure et al. (2017) to differentiate tumour pictures from normal brain MRI images. Results showed a POD of 92.1%, POFD of 86.38 %, CSI of 96.74%, and a PC of 95.28% using this approach. Low-resolution brain MRI scans are not a good fit for this technique. Using hybrid classification technology, Sreedhanya et al. (2017) were able to identify brain cancers in MRI scans of the brain with a 91.27% probability of detection (POD), 85.28% probability of false detection (POFD), 94.57% sensitivity (CSI), and 94.67% specificity (PC). Only in aberrant brain MRI scans was the inner area of the tumor region borders found using this approach. Glioma brain tumor pictures were detected and classified from normal brain MRI images using a support vector machine (SVM) classification technique (Alfonse et al., 2016). The POD, POFD, CSI, and PC rates for the traditional approach were 93.28 %, 88.37%, 96.45 %, and 96.16 % respectively. This method's primary

drawback is that it cannot properly categorise photos of the brain's inside.

## VI. CONCLUSIONS

In this work, we use a Naive Bayesian classification strategy to identify and segment Glioma brain tumors. To boost the suggested method's classification rate, the GLCM and probabilistic features are retrieved from the original brain picture. The suggested method's efficacy is measured using the Critical Success Index (CSI), the Percentage of Corrects (PC), the Probability of Detection (POD), and the Probability of False Detection (POFD) (PC). This work proposes a technique for detecting and segmenting tumors, and its results show that it can get as high as 97.64% POD, 92.18% POFD, 98.37% CSI, and 98.75% PC.

## REFERENCES

- [1] Alfonse, M., & Salem, A. B. M. (2016). An automatic classification of brain tumors through MRI using Support Vector Machine. *Egyptian Computer Science Journal*, 40, 11–21.
- [2] BRATS Database. (2015). <http://braintumorsegmentation.org/>
- [3] Abdel-Maksoud, E., Elmogy, M., & Al-Awadi, R. (2015). Brain tumor segmentation based on a hybrid clustering technique. *Egyptian Informatics Journal*, 16(1), 71–81. <https://doi.org/10.1016/j.eij.2015.01.003>
- [4] Islam, S., & Ahmed, M. (2013). Implementation of image segmentation for natural images using clustering methods. *International Journal of Emerging Technology and Advanced Engineering*, 3, 175–180.
- [5] Lavanyadevi, R., Machakowsalya, M., Nivethitha, J., & Kumar, A. N. (2017). Brain tumor classification and segmentation in MRI images using PNN, IEEE International Conference on Electrical, Instrumentation and Communication Engineering, Karur, Tamil Nadu, India (pp. 1–6). <https://doi.org/10.1109/ICEICE.2017.8191888>
- [6] Meena, A., & Raja, K. (2013). Spatial Fuzzy C-means PET image segmentation of neurodegenerative disorder spatial Fuzzy C-means PET image segmentation of neurodegenerative disorder. *Indian J. Comput. Sci. Eng.*, 4, 50–55.
- [7] Bahadure, N. B., Ray, A. K., & Thethi, H. P. (2017). Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM. *International Journal of Biomedical Imaging*, 2017, 9749108. <https://doi.org/10.1155/2017/9749108>
- [8] Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Transactions on Medical Imaging*, 35(5), 1240–1251. <https://doi.org/10.1109/TMI.2016.2538465>
- [9] Shil, S. K., Polly, F. P., Hossain, M. A., Iftekhar, M. S., Uddin, M. N., & Jang, Y. M. (2017). An improved Brain Tumor Detection and Classification Mechanism, International Conference on Information and Communication Technology Convergence, Jeju Island, South Korea (pp. 54–57). <https://doi.org/10.1109/ICTC.2017.8190941>
- [10] Sreedhanya, S., & Pawar, C. S. (2017). An automatic brain tumor detection and segmentation using hybrid method. *International Journal of Applied Information Systems*, 11(9), 6–11. <https://doi.org/10.5120/ijais2017451641>
- [11] Yerpude, A., & Dubey, S. (2012). Colour image segmentation using K-medoids clustering. *Int. J. Comput. Technol. Appl.*, 3, 152–154.
- [12] Chitra, T., Sundar, C., & Gopalakrishnan, S. (2022). Investigation and Classification of Chronic Wound Tissue images Using Random Forest Algorithm (RF). *International Journal of Nonlinear Analysis and Applications*, 13(1), 643–651. doi: 10.22075/ijnaa.2021.24438.2744
- [13] E. Aarthi, S. Jana, W. Gracy Theresa, M. Krishnamurthy, A. S. Prakaash, C. Senthilkumar, S. Gopalakrishnan (2022), Detection and Classification of MRI Brain Tumors using S3-DRLSTM Based Deep Learning Model. *IJEER* 10(3), 597-603. DOI: 10.37391/IJEER.100331.
- [14] Reddy, Danda & Harshitha, Chinta & Belinda, Carmel. (2018). Brain tumor prediction using naïve Bayes' classifier and decision tree algorithms. *International Journal of Engineering and Technology (UAE)*. 7. 137-141. 10.14419/ijet.v7i1.7.10634.
- [15] Zaw, Hein & Maneerat, Noppadol & Win, Khin. (2019). Brain tumor detection based on Naïve Bayes Classification. 1-4. 10.1109/ICEAST.2019.8802562.