A Multi-Scale Convolutional Neural Network Based on Squeeze-net along with ELM for Brain Tumor Classification and Segmentation

R. CAROL PRAVEEN[1], G MOHAN BABU[2] AND M. ANURADHA[3]

[1]*Assistant Professor, Department of ECE, SSM Institute of Engineering and Technology, Dindigul*
[2]*Professor, Department of ECE, SSM Institute of Engineering and Technology, Dindigul*
[3]*Professor, Department of CSE, S.A Engineering College, Chennai*
Corresponding Author: R. CAROL PRAVEEN

Abstract:

The human brain is composed of a myriad number of cells. When this cell grows abnormal and uncontrollable then this will result in the formation of a tumor in the brain as an extra cell mass. Since the brain is protected by the skull, these tumors may cause a malfunction of brain activities. The diagnosis of this type of tumor is a challenging issue in the medical sector. Thus, for this purpose medical experts rely largely on MRI scans for brain tumor identification. The back-propagation approach that has been utilized to train CNN may require many tuning hyper-parameters. These can be tuned with a new methodology ensemble with Extreme Learning Machine (ELM) which may result in a better classification with a more accurate rate. The proposed methodology uses an ensemble-based ELM-based classification model for CNN to obtain feature representations. Thus, an automated system for enhancement, segmentation, and classification of brain tumors is proposed using SqueezeNet-based Multiscale Convolutional Neural Network combined with ELM (SN-MCNN-ELM). The parameterization of the above steps could aid in the early detection of brain tumors for a higher rate of treatment success.

Keywords: *Features Extraction; Classification; Multi-Scale Convolutional Neural Network; Extreme Learning Machine; Ensemble.*

INTRODUCTION:

It is important to diagnose the brain tumor, for the prediction of tumor growth and its treatment based on automatic segmentation and classification. An early diagnosis of the brain tumor involves faster medication reactions that tend to increase the survival rate of patients. From the large database, brain tumors can be located and classified by handling periodic clinical work. The automated recognition, positioning, and classification process is desired and advantageous. To gain information on the tumors (type, shape, scale, position, etc.) many medical imaging techniques have been employed. Main technologies include Magnetic Resonance Imaging, Positron Emission Tomography, Computed Tomography, and Magnetic Resonance Image Spectroscopy. These approaches may be combined to provide more accurate tumor knowledge. Owing to its favorable features, MRI is the most used technique. The scan in

 https://doi.org/10.37896/pd91.4/91487

ISSN: 0369-8963

MRI produces a high level of contrast in soft tissue with hundreds of 2D picture slices without ionizing radiation [2].

Brain Tumor Imaging has exponentially increased in recent decades. In particular, the number of works on the quantification of brain tumors based on images of MRI has considerably increased [2]. The Segmentation of brain tumors has been performed to differentiate the tumor affected tissue and healthy tissue. The division of the brain tumor image into pixels is solved in many BTS applications and the segmentation issue becomes a classification [3].

The paper aims to build and test a profound learning method for the classification and segmentation of brain tumors employing a multi-scale CNN model along with extreme learning methods. The pathological form of these brain tumors is separated and predicted in this model. Two major approaches to tumor segmentation can be seen in the field of Brain tumor segmentation: generative and discriminatory. Generative techniques are based on explicit anatomical modeling, while discriminatory strategies use gold standard segmentations to learn about image features and their relationships [4]. Studies published since the discriminatory method have progressed from classical machinery learning [5] to more modern strategies for the development of deep training [6].

A semi-automatic machine that provides the contour of the tumor contour has been proposed by Sachdeva et al.[7], for example, followed by the calculation of 71 functions using the intense, co-occurrence, and Gabor functions. Furthermore, skull removal is a typical preprocessing step in classical discriminatory approaches, but it poses disadvantages such as parameter selection or the need for prior image knowledge and a high period of computation. As input for classification, the extracted features are used. Two classifications, Support Vector Machine (SVM) and Artificial Neural Network (ANN) have been proposed in the work [8,9]. The precision achieved for SVM was 79.3% to 91.7% and for ANN it was from 75.6% to 94.9%. Analogous to this, Kaur and Chhabra[10] suggested an automated classification scheme of brain tumors using ten features and as a classifying system a Back Propagation Network with a precision of 95.3%. The SOM (Self-Organizing Map) features are investigated using fractal wavelets to achieve an average precision of 90 percent by Iftekharuddin et al.[12].Additional examples of tumor classification pipelines are based on instance learning: A semi-automatic system was developed by Havaei et al.[13] based on a KNN classification scheme. The well-known data collection from BRATS 2013 was used and the second highest score was obtained in full and core testing and Dice comparisons of 0.85 and 0.75 respectively were recorded in total and core tumor areas.

In the last few years, computers can be able to determine features that display data in an optimized manner. This principle is the basis of deep-neural networks [14] that turn problems into data-based problems to solve them from functionality. CNN [15] and Completely

 https://doi.org/10.37896/pd91.4/91487

ISSN: 0369-8963

Convolutional Networks have been used for a variety of different applications within deep neural networks. It is commonly used today in particular for general image transformation [16] and especially for medical image analysis.

PROPOSED METHOD

The proposed work is based on automated brain tumor detection using a deep learning method with the combination of multi-scale CNN. Figure.1, describes the proposed SN-MCNN-ELM architecture. The MRI is a diagnostic method utilized for human anatomical studies. An improvement in the tumor vascularity may cause the advantage of viewing the normal tissue from the tumor cell.

A.Dataset:

The MRI dataset used in the proposed method was obtained through the Kaggle MRI Medical database. It is a publically available dataset. Totally 253 MRI images of the brain are considered and it is divided into two parts, out of which 70% and 30% of datasets are utilized for training and testing sets. Here, we have taken 250 above brain MRI Images and divided our input data into 2 parts.

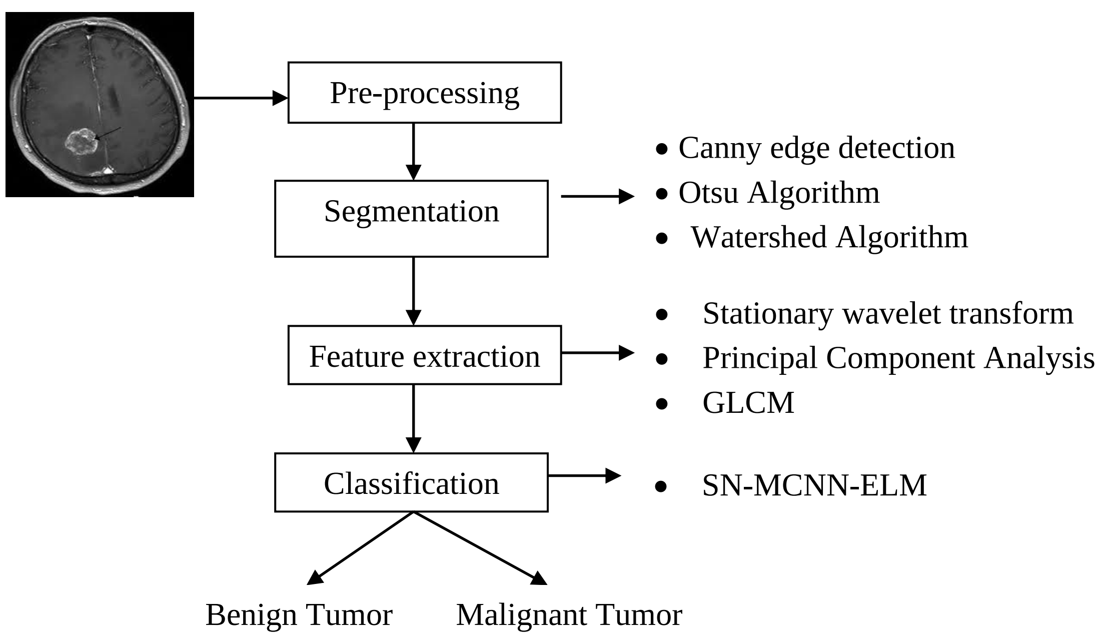

Figure 1. Flow chart of proposed SN-MCNN-ELM based brain tumor detection and classification.

B.Pre-Processing:

The preliminarily step for processing the MRI brain is to improve the image quality of the brain MRI images to make them suitable for further processing. The preprocessing steps based on the adaptive contrast improvement method for modifying the sigmoid feature are performed to enhance the Signal to Noise Ratio using the adaptive filter method. The image smoothing is also utilized for preserving the edge of the image.

 https://doi.org/10.37896/pd91.4/91487

ISSN: 0369-8963

C.Image Segmentation:

The MRI image segmentation is the process of partitioning the image into parts. In this proposed work, segmentation based on the Canny algorithm, Otsu algorithm, and watershed algorithms is implemented.

(i) Canny edge detection:

Canny edge detection[17]is utilized to obtain the information in perspective to structural information. The Canny edge detection method involves the following steps:

- A Gaussian filter is used to smooth the image.
- The intensity gradients are obtained from the image.
- Non-maximum suppression is used to eliminate the edge.
- Double threshold method for obtaining potential edges.
- Lastly, edge tacking hysteresis is used for eliminating weaker and not connected edges.

(ii)Otsu Algorithm:

Otsu's method is introduced by Nobuyuki Otsu [18] and is considered an automatic threshold method used to perform clustering as well as for grayscale image conversion. This method treats the image to have foreground and background pixels classes'.The bi-modal histogram is computed to get the optimum threshold value. The procedure for the Otsu method is:

- The histogram of very pixels is calculated based on the level of intensity.
- Set up initial class probabilities and mean.
- Maximum intensity threshold levels are set from 1 to N.
- Compute the mean and probabilities.
- Update the variance.
- The threshold level which provides maximum inter-class variance is chosen.

(iii) Watershed Algorithm

The pictorial representation of a grayscale image looks like a topographic area with peaks and valleys, wherein image processing the peaks and valleys are representatives of high and low-intensity values respectively. By the Watershed method colored labels (colored water) are used to fill the separated low-intensity value (valley). If the level of water rises depending on the heights (Gradients), the waters with different colors will flow together. Thus, a barrier can be built to stop the water merging process. The process of water flowing and barrier construction will continue until all the peaks are submerged. Finally, the outcomes of this barrier construction imply a segmentation process. The resulted segmentation process may be over-segmented because of the presence of noise or any irregularities. Therefore, a watershed algorithm based on the marker method is implemented in which the regions which imply valley points will merge.

 https://doi.org/10.37896/pd91.4/91487

The purpose of the marker is to make the region that is labeled as foreground or object have a value as'1' with intensity and non-object to have the value '0' with no color. The watershed algorithm is applied after updating the marker labels and boundaries are updated with the '-1'value.

The procedure for the Marker- based watershed segmentation is as follows:

- The segmentation function is calculated.
- The markers with foreground and background pixels are calculated.
- The segmentation functions are updated with only minimal for the marker locations.

D.Feature Extraction

(i)Stationary Wavelet Transform

A wavelet is considered a state function over a finite time interval. This wavelet method is utilized for extracting the features and frequency components. In an image, changes may occur abruptly, it can be analyzed with the help of wavelets. Thus, Stationary Wavelet Transform (SWT) [19] algorithm is utilized instead of the discrete wavelet transform (DWT) to solve the translation-invariance issues. The translation-invariance is computed by eliminating the down and up samples of DWT and upsampling is achieved by filter coefficients with a factor of …… in the jth algorithm level. The proposed work is based on 1-D SWT with j=1.The SWT implementation is depicted in Figure 2. With the increased sample value.

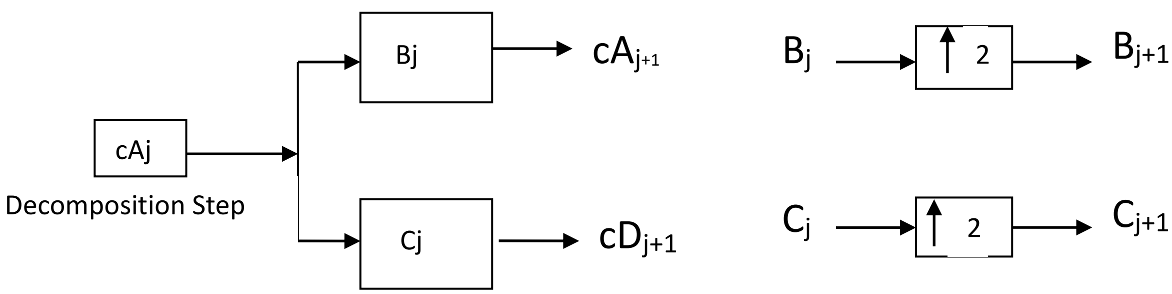

Figure 2. SWT implementation

(ii)Principal Component Analysis (PCA) with Gray-Level Co-occurrence Matrix Algorithm (GLCM)

PCA [20] is a transformation based on the orthogonal principle for making the correlated variables linearly uncorrelated variables. Initially, predominant variances with large values are computed, and subsequently, successive components are computed orthogonally. This will minimize the feature dimension leading to the computation reduction. This statistical method gives the spatial relationship among the pixels by computing its Gray level covariance matrix. GLCM is a measure of obtaining the relationship between the frequency and spatial components which can be calculated from the image texture. The features like mean, standard deviation, contrast, correlation, energy, Homogeneity,Inverse Difference Movement, variance, smoothness, and entropy are extracted in this proposed method. The measure of GLCM is also based on calculating the distance between two pixels. The grayscale representation of any imageis

 https://doi.org/10.37896/pd91.4/91487

ISSN: 0369-8963

provided as I(x,y) with an intensity level0,1…,N-1**.** GLCM is expressed as q(a,b) with 'y' and 'x' are row and column respectively.'A' gives the total number of rows and 'B' gives total number of columns. The mean and the standard deviation are represented as μ and σ.
Correlation: computed to find the grey level linear dependency of nearby grey levels

$$C_r = \sum_{y,x=1}^{A,B}((b-\mu)(a-\mu)q(b,a))$$

Contrast provides the intensity calculation of the near pixels for the whole image.

$$C_0 = \sum_{y,x=1}^{A,B}|y-x|^2 q(a,b)$$

Energy provides the summation of the squared pixels in an image.

$$Eg = \sqrt{\sum_{y=0}^{b-1}\sum_{x=0}^{a-1} f^2(y,x)}$$

Homogeneity computed the relationship between each pixel and its diagonal pixels from the GLCM matrix.

$$\sum_{x,y=0}^{N-1}(P_{a,b}/1+(a-b)^2)$$

Mean calculates the average value of all pixels in an input image to the total pixels count.

$$M = \frac{1}{(a\,x\,b)}\sum_{y=0}^{b-1}\sum_{x=0}^{a-1} f(y,x)$$

Standard Deviation gives the pixel dispersion of the image as a second-order moment.

$$\rho = \sqrt{\frac{1}{(a\,x\,b)}\sum_{y=0}^{b-1}\sum_{x=0}^{a-1}(f(y,x)-M)^2}$$

Entropy provides the texture details present in an image.

$$E = -\sum_{y=0}^{b-1}\sum_{x=0}^{a-1} f(y,x)\log_2 f(y,x)$$

Root Mean Square provides the square root of the sum of all pixels.

$$\sqrt[2]{\sum_{x,y=1}^{A,B}(|q(x,y)|)^2 / E}$$

Variance gives the expectation of the squared deviation of the mean of the pixel.

$$\frac{1}{A+B} * \sum_{x=1}^{A}\sum_{y=1}^{B}(q(x,y)-M)$$

 https://doi.org/10.37896/pd91.4/91487

ISSN: 0369-8963

Inverse Difference Movement (IDM) provides local homogeneity whereby, its value will be high when the pixels values are uniform.

$$I_{dm} = \sum_{y=0}^{b-1}\sum_{x=0}^{a-1}\frac{1}{1+(y-x)^2}f(y,x)$$

Coarseness is a method of extracting the roughness of the image from its textual information.

$$C = \frac{1}{2^{b+a}}\sum_{y=0}^{b-1}\sum_{x=0}^{a-1}f(y,x)$$

Classification:

An extreme learning machine classifier ensemble is utilized to improve classification performance. The proposed automatic classification method has been described in this Section. A convolution neural network (CNN) is predominantly used in the classification, detection, and recognition of images. Due to its efficient and intelligent structure, CNN can generate high accuracy rates for big data. In this proposed work, the Squeeze-Net architecture-based CNN model is used.

(i)Extreme Learning Machine

Extreme Learning Machines (ELMs) is one of the learning algorithms used to train single-layer feed-forward neural networks [21]. In the ELM approach, the input weights were selected extemporary according to the CDF (Continuous Distribution Function) and the output was calculated using the minimum norm solution. The method for ELM training is depicted in Algorithm 1.

'X_m' and 'w' are the matrices which represent't'training samples of'd' dimension as *'t x d'* and interdependences between the input and hidden layer as *'d x l'*respectively. '*t*x *l*' is a matrix represents the bias vector duplication,andarerandomly chosen based on thenormaldistribution. $H_a$isconsidered as thehiddenmatrix,while$P(\cdot)$isanarbitrary piecewise continuous function that provides the approximation proficiency of ELM.The output matrix is provided with α matrix, H_a^+means the Moore–Penrosepostulated inverseof the $H_a$matrix .The target matrix is provided by T_m which represents the target vectors (labels) 't', and gives the number of classes 'c' T_m=D_x.

Algorithm 1 ELM algorithm.

Initialize DatasetD_x , target T_m , and hidden node *l*.
Calculate the hidden matrix
Compute the hidden matrix$H_A = P\ (X_m W_a + \ b)$.
Chose the input weights and the bias randomly.
Calculate the weights at the output $\propto = H_a^+ T_m$
Compute the ELM parameters W_a,b and α
End

Here there is the minimum norm solution which is used to reduce the classification error and

https://doi.org/10.37896/pd91.4/91487

the weights of the norms.The reason for choosing the extremelearningmachineis due to intrinsicfeatures like the speed of execution,universalapproximation of capability,andgood overallperformance.

(ii)SN-MCNN-ELMfor ensemble classification

Several models have been introduced as ELM-based ensembles to improve the performance of a single model. The multi-model combination of the proposed method is provided by Algorithm 2, which includes discrete ELM models, and the output was calculated by combining ELM and CNN with parameters.

Algorithm 2 ELM-based ensemble training for classification.

InitializeDatasetD_x ,Number of individual modelsM_i, and hidden node *l*.
For m= 1 to M_i do
Chose the input weights as$W_a^{(m)}$ and biases as $b^{(m)}{}_a$
Calculate the hidden matrix$H_a^{(m)} = P\ (X_m W_a^{(m)} + \ b^{(m)})$
Calculate the output weights $\propto^m = H_a^{+(m)} T_m$
Calculate the outputs $O_x^{(m)} = H_a^{(m)} \alpha^{(m)}$
End for
Calculate the overall hidden matrix$H_x\ = [O^{(1)}, O^{(2)}, \dots \dots O^{(M_i)}]$
Calculate the fusion parameters$F_x = {H_x}^+ T_m$
Return the ensemble parameters$F_x, W_a^{(m)}$,$b^{(m)}$and $\propto^m$, for m=1,2…….M_i

(iii) Training the CNN using ELM

The training of a convolutional neural network using the Extreme Learning Machine Algorithm is known as ELM-CNN. The CNN model consists of convolutions, pooling, and fully connected layer components. In this study, fully connected layers are determined by the convolution layer.

The CNN is utilized as a pile of layers for performing functions such as convolution, pooling, and irregularity. Each layer composes of an input and output feature map. The SqueezeNet is introduced to make subgroups of neural networks with few fitting parameters. SqueezeNet is used instead of AlexNet since its parameter levels are very less i.e 0.5 MB. This comparison is made with ImageNet also, providing the outcomes very similar to Alexnet for image classification.

During the training phase, the proposed Squeezenet-based CNN architecture, chose the filter randomly through which the data are transmitted to the output block. The convolution operation is computed as the kernel multiplication according to the input as portrayed in Figure 3.The advantages of SqueezeNet architecture are as follows: Efficient distribution of

ISSN: 0369-8963

parameters; the working principle of CNN is directly proportional to the parameter values.

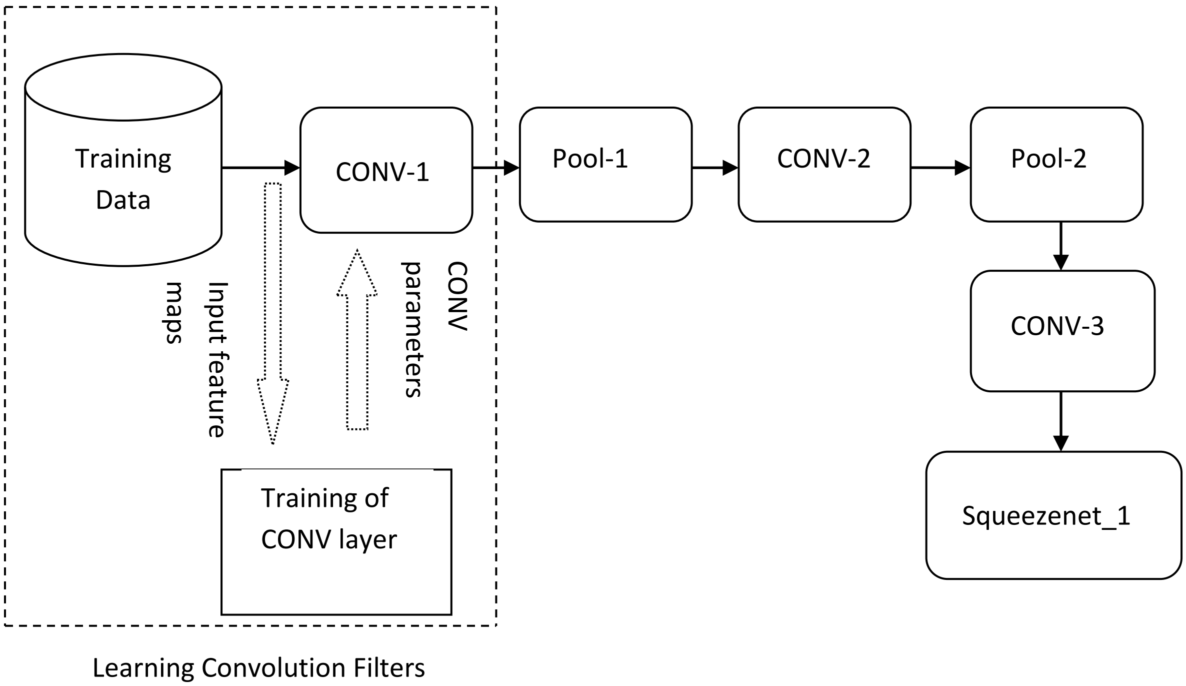

Figure 3.Overall methodology for the proposed SN-MCNN-ELMapproach

.

Algorithm 3:Procedure for training the Convolution layer using ELM.

Initialize the training data for producing feature map
Normalize the training setD_N
Chose the target set$T_m = [D_N|1]$
Generate randomly the input weight W_a and bias b
Compute the Hidden matrix$H_a = P \ (X_m W_a + \ b)$
Compute the output weight$\propto = H_a^+ T_m$
Implement Filters and bias $\left[f_{mat}^{T_m} \middle| b^{T_m}\right] = \propto$
Filters matrix $f_{mat}^{T_m}$= reshape (f_{mat}).
Return convolutional parameters $f_{mat}^{T_m}$ and b^{T_m}.
Result
End

The training procedure of an SN-MCNN-ELM looks like alayer-wise process so that the process can be stopped at any instant to obtain the features accordingly and it is summarized in Algorithm 3. In MCNN-ELM, the classifier can be replaced with any classifier at the final layer.

 https://doi.org/10.37896/pd91.4/91487

ISSN: 0369-8963

Compatibility with any system: FPGAs needonly 10 MB of chip memory for the process. This would be possible with the SqueezeNet architecture when compared with other architectures.The features which are extracted from the above methods are fed into MCNN along with the Extreme learning Machine algorithm for classifying benign and malignant brain tumors. The combination of these two methods is preferred for its high-performance accuracy. ELM can be considered asa single hidden layer composed ofa feed-forward neural network.The input hidden layer weights are selected and the output is obtained accordingly. The hidden layer contains the activation functions likes, Gauss,Sigmoid, and Sinus with along with linear activation functions as output are implemented. The appropriate activation function andneurons counts are sigmoid and 1500 respectively. The weights of these input layers don’t have any effect on the output.An ELM network algorithm can be computed as follows [22]:

$$f^T = \sum_{m=1}^{M} \propto_m P \ (X_m W_a + \ b)$$

where W_a is the weights between the input and the hidden layer, b are the bias values.α_m values are the weights of both the hidden layer and the output and P (.) is the activation function.

EXPERIMENTAL RESULTS

The experimental procedure involves three modules. First, features are extracted, and then the classification is based on the ensemble classification method and compared with state-of-the-art models

The pre-processing step involves the conversion of images into grayscale through the adaptive filter methodfor eliminating the noise and image enhancement can be implemented through the adaptive contrast improvement method.One of the major causes for the evolution of numerous brain tumor detection is no predominant or exact method has been followed for the segmentation process. Thus in this paper, segmentation algorithms namely the Canny algorithm, Otsu algorithm, and watershed algorithms.

The segmentation process based on Otsu's Thresholding algorithm and Canny edge detection algorithm provides similar results with smaller details in the image. Thus, this type of bit-wise segmentation method may reduce the Mis-segmentation in the non-tumor region. However, the segmentation process of the watershed method helps to partition the tumor-affected region into multiple segments. The partitioning of the non-tumor regions has been reduced by this type of segmentation process. Thus, combinations of segmentation methods results provide enhanced results in terms of shape, area, and variation of the intensity values among the edges of the tumor-affected regions. An example of the test image which is used for implementing the proposed work is depicted in Figure4.

The feature is extracted to obtain the brain image sets features that would provide high accuracy in Tumor in the brain image. The features are extracted through the PCA along with GLCM to extract the spatial feature and the texture features like mean, standard deviation,

 https://doi.org/10.37896/pd91.4/91487

contrast, correlation, energy, Homogeneity,Inverse Difference Movement, variance, smoothness, and entropy are extracted in this proposed method.

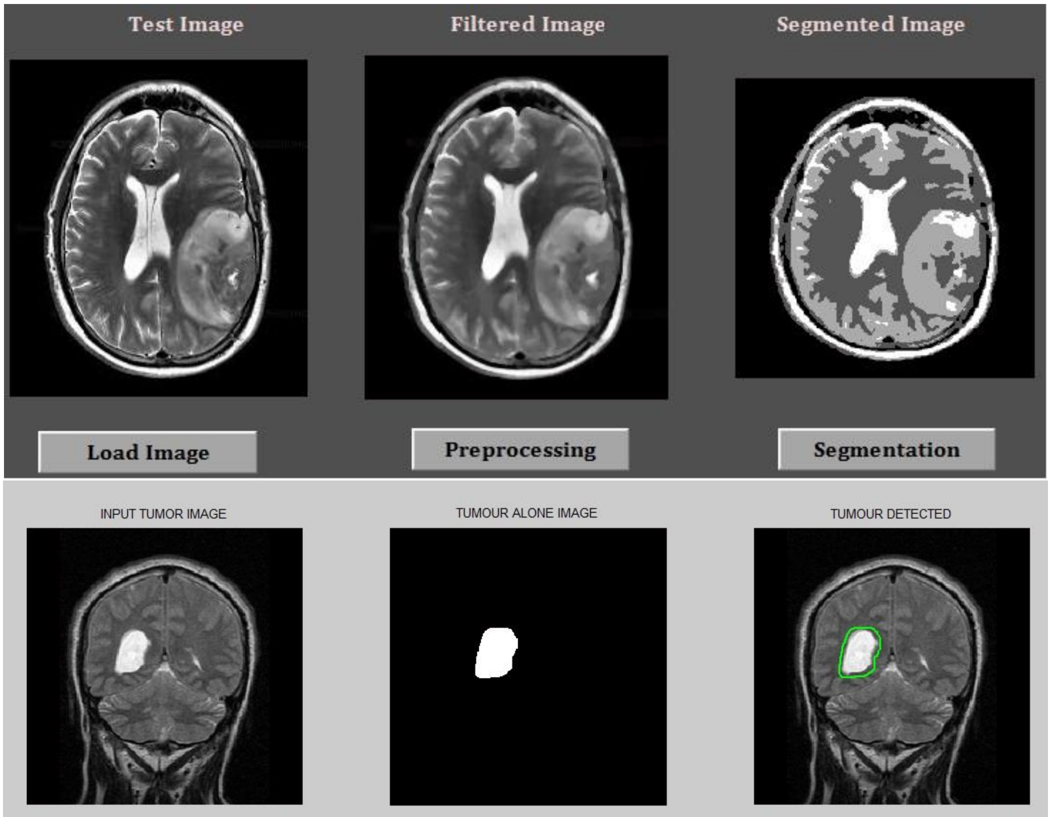

Figure 4. Result of a sample test image.

The performance can be improved in ELM when the number of nodes is increased. Thus in this method, two different CNN models are provided in Table.1 from the smallest network of about 0.4×10^6 range to the largest network about 34.9×10^6 range.The selection of these values is to evaluate the function of the SN-MCNN-ELM for all the ranges to improve the performance and for having a comparison between the SN-MCNN-ELM and CNN.

Table 1. A description of the implemented SN-MCNN-ELM model

Layers	**Training epochs**	**Learning rate**	**CONV-1**	**Pool-1**	**CONV-2**	**Pool_2**	**CONV-3**	**Pool_3**	**SQUEEZE NET _1**	**CONV-4 (FC)**
Small CNN	20	0.001	5 × 5 × 1 × 20	Method :max Size: 2 × 2	5 × 5 × 20 × 50	Method :max Size: 2 × 2	4 × 4 × 50 × 500	Method :max Size: 2 × 2	1x1	1 × 1 × 500 × 6
Large CNN	30	0.001	5 × 5 × 1 × 200	Method :max Size: 2 × 2	5 × 5 × 200 × 500	Method :max Size: 2 × 2	Filters: 4 × 4 × 500 × 4000	Method :max Size: 2 × 2	1x1	1 × 1 × 4000 × 6

https://doi.org/10.37896/pd91.4/91487

MATLAB 2017 is used for executing the proposed algorithm on the computer with a CPU with an Intel i5 core and 64GB of RAM. The MatConvNet library of MATLAB is utilized for performing CNN algorithms.

In this method, the performance of MCNN was evaluated from the features extraction module. To learn about the impact of the extracted features independent classifiers are compared.For the training of the other classifier, Matlab Libraries were used, and a five-fold cross-validation method is also used. For training, the momentum value is assigned as 0.9, and the batch size as 100. It is also observed that changes in epochs numbers don't have any influence on the performance.

The test accuracy rate and its equivalent training period of SN-MCNN-ELM and CNN algorithms are given in Table.2 for both large and small CNN model.

Table 2. Performance comparison between SN-MCNN-ELM and CNN on the dataset.

Layers	**TRAINING TIMING (s)**				**TEST ACCURACY (%)**			
	Large		**Small**		**Large**		**Small**	
	SN-MCNN-ELM	**CNN**	**SN-MCNN-ELM**	**CNN**	**SN-MCNN-ELM**	**CNN**	**SN-MCNN-ELM**	**CNN**
Pool-1	4.80	451.9	1.29	150.1	57.03	55.34	57.34	52.06
CONV-2	16.22	455.7	2.12	151.5	53.71	54.49	54.77	54.77
Pool-2	17.31	491.1	2.25	130.7	53.40	58.62	57.72	54.20
CONV-3	28.29	498.8	2.69	123.5	56.18	57.10	54.80	52.08
SQUEEZE NET _1	30.80	481.5	2.75	122.9	56.43	51.72	56.48	51.48

For extracting relevant features in the CNN algorithm, it is important to train the model for classification. Thus, the training time is very much predominate in the CNN classification mode. But the proposed SN-MCNN-ELM is a layer-wise methodology, the features are calculated in-prior to training phase, thus the layers which are near to the desire feature alone is trained. The time required for training time in SN-MCNN-ELM is reduced comparatively.

From the above table it is understood that the working performance of the SN-MCNN-ELMalgorithm is more when compared with the CNN algorithms. The extreme learning machines perform better by increasing the hidden nodes but it is not predominant in the common CNN.The SN-MCNN-ELM algorithm is 55-120 times faster than
CNN to get better efficiency, even for the limited network.

After evaluating each element one after the other, in this section, the whole proposed approach that mixes the capabilities extraction using the SN-MCNN-ELM set of rules are evaluated and the ELM primarily based ensemble for classification. We examined different features on both the small and large networks. The hyper-parameters used in an ELM based ensemble; number of individual ELM models 'M_i', and the number of hidden nodes 'l', which might be decided by using 5-fold cross-validationmethod.The extracted features were fed into Squeeze Net pertained architecture using ELMclassifier.The following evaluation metrics are computed in this proposed work, such as Accuracy, Precision, Sensitivity, specificity, Youden Index and F1-score as portrayed in Table 3.

Table 3.Performance comparison of proposed method with other classifiers.

Classifier		Accuracy %	Precision %	Sensitivity %	F_1Score%	Specificity%	Youden Index%
Proposed Method (SN-MCNN-ELM)		**98.31**	**98.76**	**98.04**	**98.41**	**98.60**	**95.73**
SVM	RBF kernel	94.96	92.43	93.56	95.03	93.23	89.45
	Linear kernel	83.53	86.23	85.73	86.34	87.56	71.45
	Polynomial kernel	87.37	89.10	89.03	91.23	81.23	81.34
KNN		93.65	92.73	93.67	94.54	93.43	93.45
Naïve Bayes		80.67	81.54	82.89	81.54	81.92	63.67
Decision Tree		92.45	93.12	93.78	94.23	94.87	90.65
Neural Network		94	92.45	93.34	93.54	94.59	90.76

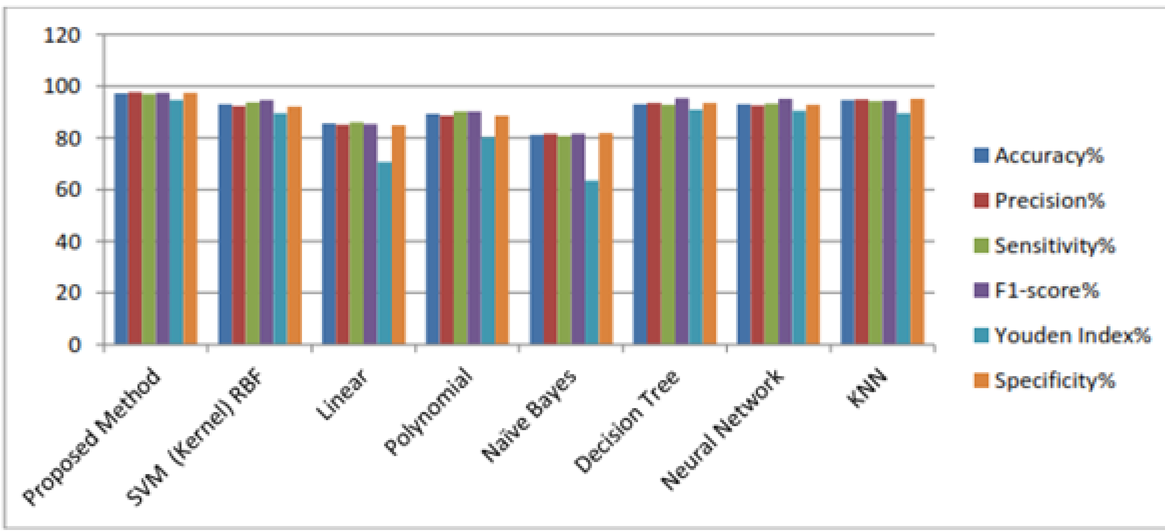

Figure 5.Comparison of the performance metrics of the proposed and existing classification methods.

From the above table it can be concluded that the proposed work outperforms than other classifier with accuracy of 98.31%.A deep learning method with the combination of SN-MCNN-ELM is implemented in this work to obtain best performance metrics to reduce the time complexity.The value of Specificity and Sensitivity parameters seems to be quite similar percentage-wise. The result of sensitivity represents excellent determinative variation of benign tumors, since it is the calculation of total positive divide by the total actual benign tumor. Specificity provides the results of total negative divide by the total actual malignant tumor to determine the determinative value of malignant tumor detection. The differences between the true and false positives are provided by Youden-Index. The precision rate is the calculation of the actual positives among all predicted positives. The F1 score is computed for determining the balance between the sensitivity and precision when distribution of class is uneven. Overall, comparison of the performance metrics which are implemented in this work is shown in Figure 5.

In this method, Graphical User Interface is developed using the MATLAB 2017a as shown in Figure 6.

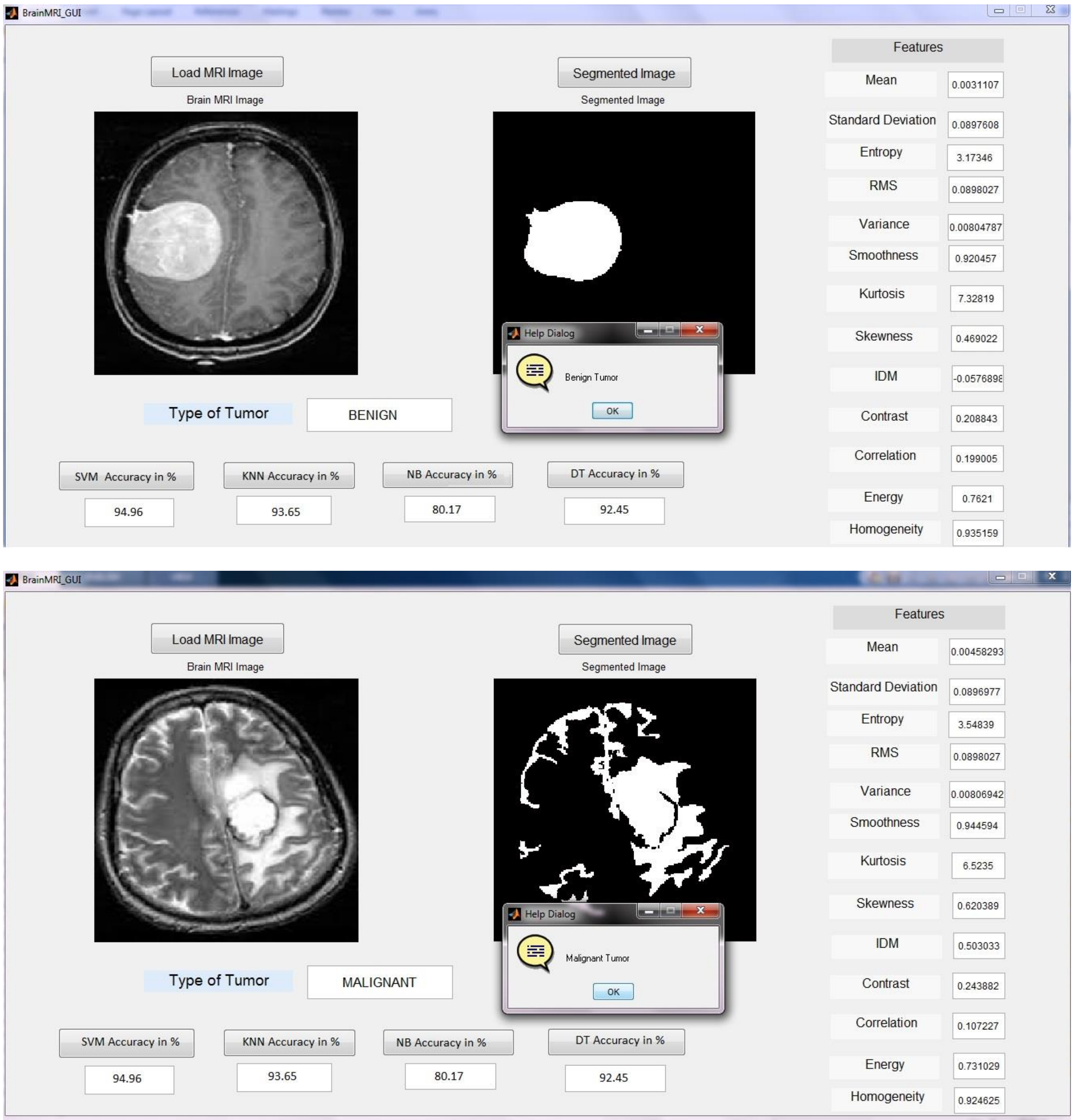

Figure 6. GUI of the proposed method detect the brain tumor.

The proposed ELM method is an automated learning framework with the combination of Squeeze Netbased multi-scale CNN will provide reduction in the classification error and

ISSN: 0369-8963

weight norms .The comparison with other classifiers namely SVM,KNN,Naiyes Bayes and Decision tree has been implemented. The parameter will have influence on the execution and itstimeis provided inFigure 7.It is understood from the figure that, large set could provide better performance. The ensemble based ELM is consider as fast learning algorithm thus grid search based is utilized to obtain the hyper-parameters.

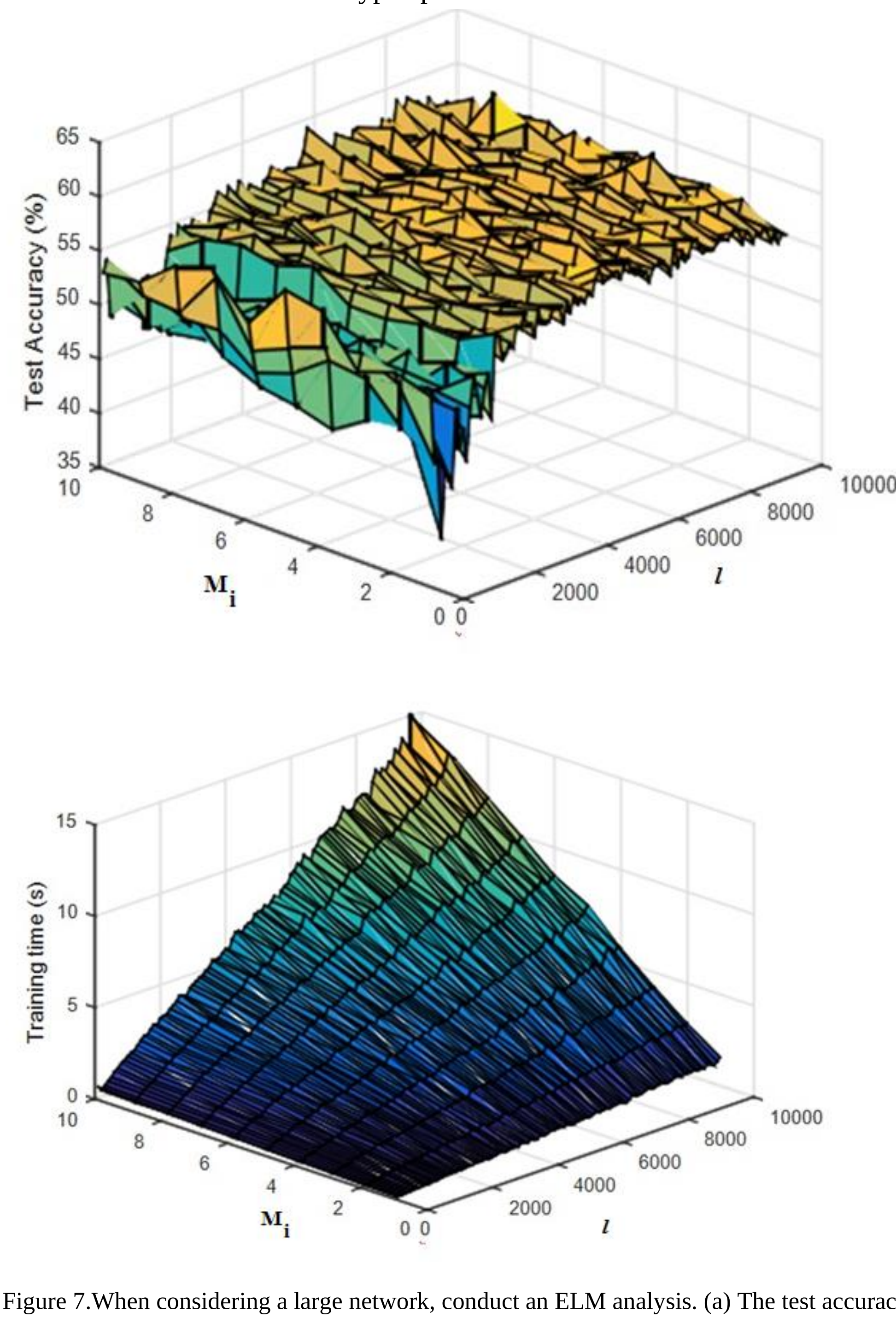

Figure 7.When considering a large network, conduct an ELM analysis. (a) The test accuracy and (b) the training time for Squeeznet_1.

 https://doi.org/10.37896/pd91.4/91487

Figure 8. depicts the proposed methodology's confusion matrix and ROC which is based on CONV-3 features computed using a large network.

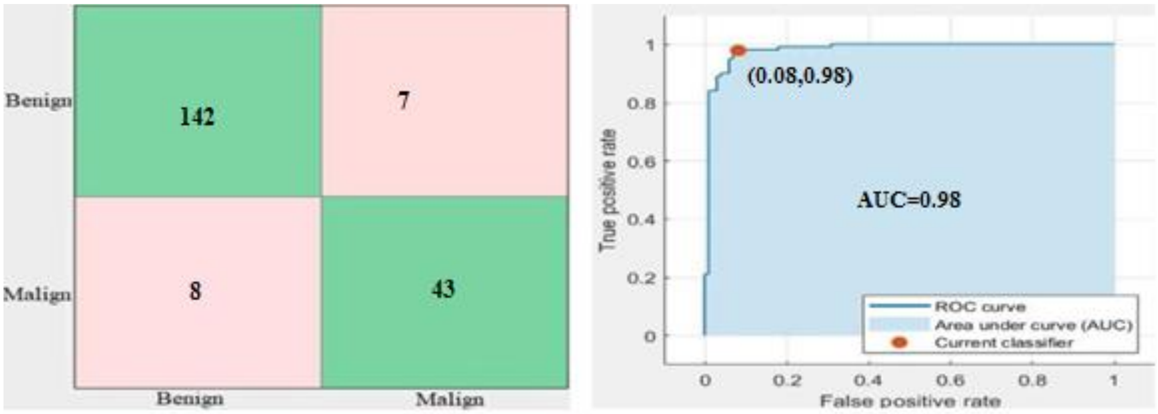

Figure. 8. Confusion matrix and ROC of proposed SN-MCNN-ELM method.

CONCLUSION:

In this paper, an automated extreme machine learning approach for detecting brain tumours in MRI images is proposed. The proposed method combines the work of CNN and ELM to extract and classify features. The segmentation performance can be increased using various segmentation method. In this work one of the most pertained Multi-scale Convolutional neural network method using squeeze net is implemented along with ELM.With the aid of this approach, tumor are detected and classified as benign and malignant tumor from the Kaggle brain MRI dataset. In future, the brain tumor affected region can be calculated in percentage area. In addition, this methodology have to be implemented with other classifiers for other medical process like skin lesion segmentation.

REFERENCE:

[1] Hussain, S., Anwar, S. M., & Majid, M. (2018). Segmentation of glioma tumors in the brain using deep convolutional neural network. *Neurocomputing*, *282*, 248-261.

[2] Sarbanes Data, Dr. Monish Chakra borty, Brain Tumor Detection from Pre-Processed MR Images utilizing Segmentation Techniques, IJCA Special Issue on second National Conference-Computing, Communication and Sensor Network, CCSN, 2011 .

[3] S. Roy and S.K. Bandyopadhyay, “Detection and Quantification of Brain Tumor from MRI of Brain and its Symmetric Analysis”, International Journal of Information and Communication Technology Research, KY, USA, June 2012.

[4] Hazra, A. Dey, S. Gupta and M. Ansari, “Brain tumor detection based on segmentation using MATLAB”, 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), 2017.

 https://doi.org/10.37896/pd91.4/91487

[5] V Ramakrishna Sajja and N Gnaneswara Rao, "Segmentation of Medical Images using Enhanced K-Means, Wavelet Transformations and Morphological Operations", CIE46 Proceedings, 29-31 October 2016, Tianjin / China.

[6] Işın, A., Direkoğlu, C., &Şah, M. (2016). Review of MRI-based brain tumor image segmentation using deep learningmethods. *Procedia Computer Science*, *102*, 317-324.

[7] Sachdeva, J., Kumar, V., Gupta, I., Khandelwal, N. and Ahuja, C.K., 2013. Segmentation, feature extraction, and multiclass brain tumor classification. Journal of digital imaging, 26(6), pp.1141-1150.

[8] V Ramakrishna Sajja, Gnaneswara Rao Nitta, "Experimental Approache for Detection of Brain Tumor Grade Using Svm Classification", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7, Issue-5S4, February 2019, Retrieval Number: E10160275S419/19©BEIESP, PP.79-85.

[9] Jeevanantham, V, MohanBabu, G. Detection and diagnosis of brain tumors-framework using extreme machine learning and CANFIS classification algorithms. Int J Imaging Syst Technol. 2021; 31: 540– 547. https://doi.org/10.1002/ima.22479

[10] Kaur, N. and Sharma, M., 2017, August. Brain tumor detection using self-adaptive K-means clustering. In 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS) (pp. 1861-1865). IEEE.

[11] Rao, C.S. and Karunakara, K., 2021. A comprehensive review on brain tumor segmentation and classification of MRI images. *Multimedia Tools and Applications*, *80*(12), pp.17611-17643.

[12] Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R. and Lanczi, L., 2014. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE transactions on medical imaging*, *34*(10), pp.1993-2024. 1.

[13] Ashokkumar, S. R., Anupallavi, S., Premkumar, M., & Jeevanantham, V. (2021). Implementation of deep neural networks for classifying electroencephalogram signal using fractional S-transform for epileptic seizure detection. International Journal of Imaging Systems and Technology, 31(2), 895-908.

[14] Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.M. and Larochelle, H., 2017. Brain tumor segmentation with deep neural networks. *Medical image analysis*, *35*, pp.18-31. 1.

[15] Ashokkumar, S. R., Anupallavi, S., MohanBabu, G., Premkumar, M., & Jeevanantham, V. (2022). Emotion identification by dynamic entropy and ensemble learning from electroencephalogram signals. International Journal of Imaging Systems and Technology, 32(1), 402-413..

[16] S. Pereira, Adriano Pinto, Victor Alves, and Carlos A. Silva, "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images", IEEE TRANSACTION ON MEDICAL IMAGING, VOL 35, MAY 2016. 1.

ISSN: 0369-8963

[17] Premkumar, M., Ashokkumar, S. R., Jeevanantham, V., Pallavi, S. A., Mohanbabu, G., & Raaj, R. S. (2022). Design of cost-effective real time tremor alerting system for patients of neurodegenerative diseases. Materials Today: Proceedings, 57, 1989-1994.

[18] Premkumar, M., Jeevanantham, V., & Anupallavi, S. (2021, November). Design of an EEG-based transceiver with data decomposition for IoHT applications. In 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) (pp. 258-262). IEEE.

[19] Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, *9*(1), pp.62-66.

[20] Shensa, M, J., "The discrete wavelet transform: wedding the a trous and Mallat algorithms". IEEE Transactions on signal processing, 40(10), 2464-2482, 1992.

[21] Premkumar, M., Sundararajan, T. V. P., & Mohanbabu, G. (2022). Dynamic Defense Mechanism for DoS Attacks in Wireless Environments Using Hybrid Intrusion Detection System and Statistical Approaches. Tehnički vjesnik, 29(3), 965-970. 5.

[22] SR, A. (2022). An efficient framework for glioma tumor classifications and diagnosis using proposed CNN architecture. International Journal of Nonlinear Analysis and Applications.

 https://doi.org/10.37896/pd91.4/91487