ABSTRACT:
Image processing has grown into a crucial aspect in fields of biomedical research and clinical practice. The decisive objective of brain tumor imaging study is to abstract the patient related important clinical statistics and their analytical features. The identified quantitative measure of exact tumor area helps the physician/radiologist to treat the patients. In medical image processing the foremost step is image acquisition that is done using MRI scan. MRI scan gives very detailed pictures and is the best technique when it comes to finding tumors in the brain, including if or how much it may be spread into nearby brain tissue. Magnetic resonance imaging classification is a stimulating function mainly due to the variance and complexity in tumor detection. Hence, feature extraction by possibilistic fuzzy C Means clustering and classification using Multilayer Extreme Learning Machine (ML-ELM) is proposed in this work. Also feature sub selection is done with Random Projection. Extreme learning machines perform better than deep neural network classifiers in classifying the dataset as tumorous or not. The comparison between the algorithms proves that the proposed method has higher accuracy, sensitivity and specificity than the other approaches. The proposed approaches are applied for real time datasets and benchmark datasets taken from dataset repositories.


INTRODUCTION
Brain tumor is a cluster of abnormal cells that grows inside of the brain or around the brain. There are two types of brain tumor. One is non-cancerous or benign and other is cancerous or malignant. The National Brain Tumor Foundation (NBTF) for research in the United States estimates that, in children, brain tumors is the cause of one-quarter of all cancer deaths. Also, NBTF states that percentage of people developing brain tumor has increased by 300% over past three decades. Early detection of the brain tumor is very important and it helps the physicians/radiologists in treatment. The tumor will be completely curable if it is identified in the initial stage. Detection of the brain tumor in the initial phases helps in the effective and critical treatment of this ailment. Brain tumor identification essentially comprises Magnetic Resonance Imaging (MRI), Nerve test, Biopsy etc. Medical image processing is the skill of creating visual illustrations of the interior of a body for analysis and medicinal interventions. Medical imaging seeks to reveal internal structures hidden by skin and bones as well as to diagnose and treat disease. In the brain MRI, the tumor may appear clearly but for further treatment, the physician also needs the quantification of the tumor area. Hence, the next step is preprocessing of an acquired magnetic resonance images using a median filter. Preprocessing is done to remove the noise and enhance the image quality for further processing. The median filter is the nonlinear filtering techniques used to preserve the edges of the magnetic resonance brain images and improve the image resolution contrast. After enhancing the quality of MRI brain image it should be segmented. Image segmentation is done using possibilistic fuzzy-CMeans clustering algorithm which assigns a single value to each pixel of an image in order to differentiate multiple regions of an image. The major objective of image segmentation is to divide the image into separate and meaningful regions or clusters that are homogeneous with respect to specific property such as color, texture, gray level, depth, reflectivity etc. After segmentation, dimensionality reduction is done using random projection. Then the features are trained using extreme learning machine to be identified as tumorous or not.

RELATED WORK:
This review presents a various magnetic resonance image segmentation techniques such as FCM (Fuzzy C-Mean clustering), SOM (Self-Organizing Map), ANN (Artificial Neural Network) .Edge based image segmentation methods utilize the gradient value of the image that are usually sensitive to noise and weak edges [15]. One of the most popular edge based segmentation method is Active contour model (snake), where the external force pushes the snake towards the object. Active contour models (ACM) [15] are based on the deformation of a contour and it will reach object boundaries according to the internal and external energy. The internal energy maintains the contour smoothness while the external energy attracts the contour to the desired boundaries [15]. FVF (Fluid Vector Flow) [13] is an enhancement of Gradient Vector Flow [13]. FVF method is insensitive to the initialization of contour when the contour is far from the object. The directional force dominates and attracts the contour towards the object. When the contour is near the object, the gradient force fits the contour to the object. FVF takes 1 to 5 seconds to process a 256x256 image and the computational time need to be optimized [13]. Continuous-domain convex active contour model [7] uses both the boundary and regional information to find a global optimal solution. ThiNhatAnh Nguyen presented a constrained active contour model [7] which was the extension of convex active contour model and it was used for interactive image segmentation. In general, constrained active contour model has the ability to handle topology changes of the boundary contour [7]. Segmentation takes less than three seconds with a resolution of an image of size 640x480, where optimization costs less than half a second. This method cannot handle transparent or semitransparent boundaries of the image. PDE based active contour model [3] produced 98% accuracy for magnetic resonance images in most of the existing works. Region based image segmentation aims to
identify each region of an object by region descriptor such as color, texture, intensity, motion. Region based segmentation uses statistical measures to construct the constraints which have better performance for noisy and weak object boundaries when compared with edge based image segmentation. Level set segmentation method, initially derives the local intensity clustering property [11] of the image intensities and defines a local clustering criteria function for the image intensities in a neighborhood of each point. This local clustering criterion function was then integrated with respect to the neighborhood center to give a global criterion of image segmentation. In a level set formulation, this criterion defines energy in terms of the level set functions that has represented as a partition of the image domain and a bias field [11]. Yugang Liu presented a multipass level set method [8] which was insensitive to initial user interactions resulted in better performance than normal level of function. In this method hint has been taken by the user through drawing one or at most two rectangular boxes over a target object in every image and iteratively, in order to obtain more accurate segmentation results. The location and size of the rectangular region did not need to be accurate. User interactions typically took less than 5 seconds [8]. Even though multipass level set method produced better results, it required user interaction that may differ from one observer to other. A nonlinear adaptive level set based image segmentation automatically determines the curve to shrink or expand by using Bayesian rule [1]. This method avoids the leakage at weak boundaries and reduces the number of false boundaries. Nonlinear adaptive segmention achieved a smooth front using adaptive direction and probability weighted stopping force. Nonlinear adaptive velocity was designed under two-phase assumption that is not suitable for multiphase segmentation [1]. The velocity variation avoids boundary leakage. Clustering organizes the group of objects into clusters having the same feature, attribute and characteristics that are similar to one another within the cluster and dissimilar to other objects in the clusters. Clustering is also called as segmentation since clustering partitions large objects into groups according to their similarity. K-Means clustering is one of the unsupervised clustering in image segmentation. In k-means clustering method, initially the number of clusters k need to be identified. The center of the k-cluster is chosen arbitrarily which is used to calculate the distance with each pixel [5]. If the calculated distance is closer to the center then, the pixels are moved to that cluster. At last, the center is recalculated and continued to the process until it converges. The execution time for the k-means clustering algorithm was comparatively slow and the computational complexity was reduced. K-means clustering found to be accurate method to extract the ROI (Region of Interest) [5]. Even then the output depends on the size of the partition. The k-means clustering algorithm [8] was tested with a database of 100 MRI brain images and it produced 95% accurate results. Fuzzy C-means clustering is the data clustering method in which data points belongs to single cluster. In this clustering method, a membership function assigns the membership value as zero or one to each pattern data using the membership function. Finally, it divides a collection of n vectors into fuzzy groups and calculates the cluster center in each group. Hence, the cost function of the dissimilarity measure was minimized [2]. X=(x1, x2, …, xn) denotes the image with n pixels to be partitioned into 1<c<n clusters [2]. Zhao Zaixin presented a NWFCM (Neighborhood weighted Fuzzy C-means clustering) method [2] which was used to enhance the efficiency of the image segmentation. NWFCM was more robust to noise when compared with other algorithms [3]. However, FCM produced unsatisfactory results when the noise level increases and it takes huge computational time. Computational time of NWFCM is three to four times faster than the traditional FCM [2] when the size of an image is 308x242. It took 2.26 seconds for computation. In image segmentation, Artificial Neural Network can be defined as a system consisting of non-linear artificial neurons running in parallel that can be one-layered or multi-layered [10]. Generally, ANN has three layers such as input, output and hidden [10]. The system parameters which will change during the operation is called a training phase. When the training phase is completed parameters are fixed and deployed to solve the problem. PNN (Probabilistic Neural Network) combines the best attributes of pattern recognition and feed forward artificial neural network [14]. Probabilistic segmentation is performed by PNN that produces the output with Bayesian posterior probabilities. In ANN, accuracy would not be consistent across slices since each slice need to be checked manually [14]. This method did not consider brain regions neighboring information in third dimension that will leads to poor performance. SOM is the type of artificial neural network for unsupervised learning segmentation [12]. Here, the training process builds map using vector quantization process. SOM consists of two layers. The first layer contains input nodes and second layer contains output nodes that are arranged in a two-dimensional grid format [12]. Every input node is connected with every output node with adjustable weights. The weight of each node will depend on the pattern assigned to that node. SOM [4] maps a higher dimensional input space to lower dimensional input space. Mapping classifies the new input vector automatically [4]. The clear tumor depiction will be provided by using entropy, mean, median, variance, correlation, maximum and minimum intensity values. In SOM, based image segmentation training phase of the neuron is easier and salient features of pattern vectors in data set is used to classify new patterns [4]. It may be an obstacle for the new users which is the drawback of this segmentation algorithm.

**METHODOLOGY**

Tumor segmentation and classification using possibilistic fuzzy C Means clustering and deep
learning classifier is proposed in this study. The proposed technique consists of three modules, namely segmentation module, feature extraction module and classification module. The input MRI (Magnetic resonance Imaging) images are pre-processed and segmented to remove noise using possibilistic fuzzy C means in segmentation module. Median filtering is used in segmentation module for noise removal. Then features are extracted for every segment based on the shape, texture and intensity from segmented regions. After features extraction, important features will be selected using Random Projection (RP) to classification purpose. In training phase, extreme learning machine based classifier is trained with the features of training data and in testing, the features from the segmented image are fed into the trained extreme learning classifier to detect whether the region has brain tumor or not. The block diagram of the proposed technique is given in Fig. 1.

### Segmentation Module:

#### Median filtering:
Median filtering is a preprocessing phase which helps in reducing noise of the input data. It is a digital filtering method that improves the quality of data for future processing. Median filter preserves edges while removing noise. The mechanism of median filter is to travel through the pixels, exchanging each pixel value with the median value of adjacent pixels for each pixel being considered. The pattern of neighbors is termed as the window, which slides, pixel by pixel over the entire image. The median is calculated by sorting the pixel values obtained from the window and exchanging the pixel being considered with the median pixel value. Suppose, let the pixels in the window be represented as \{1_1, 1_2, 1_3, 1_4, 1_5\}. For finding the median, sorting is carried out and then the median value is selected. For example, if the sorted list given by:

Sorted List = \{1_3, 1_1, 1_5, 1_4, 1_2\} given \(1\)

Then Median = 1_5

As we can observe from the example the median is 1_5 as it is in the center of the sorted pixels. The median helps in de-noising the input image for future processing.

### Fuzzy C – Means:
Fuzzy C-means (FCM) is a clustering algorithm with good accuracy and precision. Assume input is denoted as x and the degree of membership of x in the cluster j be denoted by \(Q_{ij}\). Let number of input data is denoted by n and the weighting co-efficient be represented by \(w\). Let the center of the cluster be represented as c and the number of clusters is denoted by mc. The minimization objective function \((F_{min})\) of Fuzzy C- Means (FCM) clustering is defined as:

\[
F_{min} = \sum_{i=1}^{n} \sum_{j=1}^{mc} Q_{ij}^w ||x_i - c_j||^2
\]

Centroids are selected at first as random and then, membership values of the data points are calculated using the formula:

\[
Q_{ij}= \frac{1}{\sum_{m=1}^{mc} \sum_{i=1}^{n} ||x_i - c_m||^{2}}
\]

The new centroid values are calculated using the membership values and are given by:

\[
C_j=\sum_{i=1}^{n} Q_{ij}^w x_i / \sum_{i=1}^{n} Q_{ij}^w
\]

Using the new centroid values, membership values are again calculated and this procedure is continued until it satisfies the equation:

\[
\max \{|Q_{ij}|^w - Q_{ij}^w| < \beta\}
\]

Where, \(\beta\) has the value between 0 and 1.

### Possibilistic Fuzzy C Means Clustering (PFCM):
The Fuzzy C Means clustering algorithm (FCM) has problems dealing with noise and outliers and Possibilistic C Means (PCM) cannot deal with equivalent clusters. Fuzzy Possibilistic C Means (FPCM) cannot deal with large datasets. FPCM lays constraint on the representative values of all the data points in a cluster as 1. FPCM is a hybridization of PCM and FCM and helps in eradicating the problems faced in other fuzzy clustering algorithms [17]. PFCM lead to optimize the following function:

\[
\min (J_{m,n}(U,T,V,x))=\sum_{k=1}^{n} \sum_{i=1}^{c} (au_{ik}^m + bt_{ik}^n)^* ||X_i - V_j||^2 + \sum_{i=1}^{c} g_i \sum_{k=1}^{n} (1-t_{ik})
\]

subject to the constraints \(\sum u_{ik}^m = 1\) where \(i=1\) to \(c\) and \(0<=u_{ik} , t_{ik}<=1\). Here a>0 , b>0 , m>1 and j is the objective function. U is the partition matrix. V is a vector of cluster centers, X is a set of all data points, ||X_i - V_j|| is any norm used to calculate the distance between i^th cluster center and K^th data set n termed as D_{kA}. Here we are using Euclidean Distance. The constants a and b define the importance of fuzzy membership and characteristic values in the objective function and c is center of clusters and x represents a data point. If we increase the importance of membership then that leads us to reduction in significance of characteristic by the same amount. Also, we will see later that the optimal typicality values depend on the magnitude of b. PFCM uses the objective functions of PCM and FCM given in Eq.7 and Eq.8 respectively.

\[
\min (J_{m,n}(U,V,x))=\sum_{k=1}^{n} \sum_{i=1}^{c} u_{ik}^m ||X_i - V_j||^2
\]

\[
\min (P_{m}(T,V,x,G))=\sum_{i=1}^{n} \sum_{k=1}^{c} t_{ik}^m ||X_i - V_j||^2 + \sum_{i=1}^{c} g_i \sum_{k=1}^{n} (1-t_{ik})^m
\]

\[
D_{kA}=\sum_{i=1}^{c} ||X_i - V_j||^{1/2}
\]
If \( b=0 \), and \( g_i=0 \) for all \( i \), then Eq.6 reduces to the FCM optimization problem in Eq.7; while converts it to the usual PCM model in Eq.8. Later, we will see that when \( b=0 \) even if we do not set \( g_i=0 \) for all \( i \) Equation 1 implicitly becomes equivalent to the FCM model. Like FPCM, we get the first-order conditions for extrema of \( J_{an} \) \( U_{ik}=\left(\sum(D_{ik}/D_{jk})^{2m-1}\right)^{-1} \)
\[ 1<=i<=n;1<=k<=n; \]

\[
t_{ik}=\frac{1}{1+((b/g_i) D_{ik})^{2}/p-1}
\]
\[ 1<=i<=c;1<=k<=n; \]

\[
V_{ik}=\sum (au_{ik}^{m_i}+bt_{ik}^{n_i})||X_k
\]
\[ 1<=i<=c; m->\infty; n->\infty \] (12)

PFCM [17] behaves like FCM when the exponents grow without bound. That is, all \( c \) centroids approach the overall mean as \( m->\infty; n->\infty \). Equation 12 shows that if we have higher values for \( b \) than \( a \) then the covariance between two real-valued random variables \( A \) and \( B \) is given by:
\[
CV(A,B) = E[(A-E(A))(B-E(B))]
\]
where,
\[
E(A) = \text{The expected value of} \ A
\]
\[
E(B) = \text{The expected value of} \ B
\]
Entropy \( (E_t) \) can be defined as a degree of irregularity and is given by:
\[
E_t = \sum P_i \log_{2} P_i \] (16)

Where, \( P_j \) is the probability that the difference between 2 adjacent pixels is equal to \( i \).

Energy \( (E_e) \) is defined as a degree of data present in the segmented region or as average of squared intensity values of the pixels. Let the intensity of the pixels be represented by \( R_{ni} \), where \( 0<i<N \) and \( N \) is the number of pixels. The energy can be given by:
\[
E_e = \sum_{i=1}^{N} R_n(i)^2 /N \] (17)

Homogeneity is defined as the state of being homogeneous. It can be given by,
\[
M=\sum_{i=1}^{N} S(i,j)^2 \] (18)

where, \( N \) is the number of grey levels and \( S(i, j) \) is the pixel intensity.

Mean, Variance, Standard Deviation, Median Intensity, Skewness and Kurtosis are intensity based features. After features extraction, important features will be selected using Random Projection to classification purpose [18]. Mean \( (M) \) is simply the average of the entities and obtained by adding all the pixel values of the region divided by the number of pixels in the region. Suppose there are \( N \) number of pixels in the \( i^{th} \) region each having a pixel value \( P_n \), then mean of the \( i^{th} \) region is given by:
\[
M=\sum P_i/N \] (19)

Variance \( (V) \) measures the range of pixels of the image
\[
V= E(X^2) - (E(X))^2 \] (20)

Where, \( E(.) \) is the expectation operation and \( E(X) = M \) (mean). Standard deviation \( (S) \) is the square root of the variance. It also measures the amount of dissimilarity from the average. A low standard deviation indicates that the data points are closer to the mean. Median is the numerical value splitting the pixel values and then picking the middle one. Skewness is a quantification of the irregularity of the probability distribution of a real-valued random
variable about its mean. Kurtosis is defined as the peak value of the probability distribution of a real-valued random variable.

Random projection (RP) is a significant technique for reducing dimensionality that utilizes indiscriminate subspaces. It uses small number of hidden nodes but the accuracy of the classification is increased. It creates a random matrix for data representation in lower dimensional space. RP method is based on Johnson-Lindenstrauss (JL) lemma [16]. The lemma states that any N point set lying in d-dimensional Euclidean space can be embedded into a r-dimensional space, with r \( \geq O(\varepsilon^{-2}\ln(N)) \), without distorting the distances between any pair of points by more than a factor \( 1 \pm \varepsilon \), where \( \varepsilon \in (0, 1) \). Notation in matrix format is given as,

\[
\mathbf{X} = \mathbf{Y}
\]

Where \( \mathbf{Y} \) is the original data set of \( N \), \( \mathbf{X} \) is the projection of the data into a lower, r-dimensional subspace, and \( \mathbf{M} \) is the random matrix that satisfies the JL lemma. Orthogonality is a property that can be used to create similarity between vectors. Hence in equation 21 making \( \mathbf{M} \) orthogonal helps in creating vectors closer to each other. However it may increase computation time and can be ignored when dealing with very high-dimensional space.

**Classification module:**
After feature extraction module, the image is classified into tumorous or non-tumorous with the use of Extreme Learning Machine (ELM). Guang-Bin Huang [19] [21] and colleagues introduced Extreme Learning Machines (ELM) for a Single Layer Feed Forward Network with a fast learning speed and good generalization capability with minimal human intervention. Hidden nodes are chosen randomly in ELM. They outperform Single Layer Feed forward networks and Support Vector Machines as analyzed by G.-B. Huang et al [19] [21][22]. ELM allocates weights of output nodes systematically and they have faster training speed. They have good generalization performance but the ability to handle big data is not optimal. Hence, multilayer ELM (ML-ELM) was proposed by Kasun et al [20] in 2013. ML-ELM uses ELM-AE (ELM-Auto Encoder) as building blocks in each layer for learning. ELM-AE reproduces the input signal and activation functions can be either linear or nonlinear. The input data is mapped to L-dimensional ELM feature space, and the network output is

\[
f_i = \sum b_i h(x) = h(x) b
\]  (22)

where \( b = [b_1, \ldots, b_L]^T \) is the output weight matrix between the hidden nodes and the output nodes, \( h(x) = [g_1(x), \ldots, g_L(x)] \) are the outputs of hidden node for input \( x \), and \( g_i(x) \) is the output of the \( i \)-th hidden node. Given \( N \) training samples \( \{(x_i, t_i)\}_{i=1}^{N} \), the ELM can re-solve the following learning problem:

\[
\mathbf{H}b = T
\]  (23)

where \( T = [t_1, \ldots, t_N]^T \) are target labels, and \( \mathbf{H} = [h^T(x_1), \ldots, h^T(x_N)]^T \). We can calculate the output weights \( b \) from

\[
b = \mathbf{H}^T T
\]  (24)

where \( \mathbf{H}^T \) is the Moore-Penrose generalized inverse of matrix \( \mathbf{H} \). We can add a regularization to improve the performance of the ELM as,

\[
\mathbf{B} = ((\mathbf{I/C}) + \mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T T
\]  (25)

To perform unsupervised learning using ELM input data is used as output data. Orthogonalization is a property that helps in binding parameter closer to each other. Hence choosing weights and biases orthogonally helps in increasing performance of ELM-AE [20]. In ELM-AE, input data can be projected to a different or equal dimension space, as shown by the Johnson- Lindenstrauss lemma and calculated as

\[
h = g(a . x + m) \quad a^T a = L, \quad m^T m = 1
\]  (26)

where \( a = [a_1, \ldots, a_L] \) are the orthogonal random weights, and \( m = [m_1, \ldots, m_L] \) are the orthogonal random biases between the input and hidden nodes. ELM-AE’s output weight is responsible for learning the transformation from the feature space to input data. For sparse and compressed ELM-AE representations, we calculate output weights \( b \) as follows:

\[
b = ((\mathbf{I/C}) + \mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{X}
\]  (27)

where \( \mathbf{H} = [h_1, \ldots, h_N] \) are ELM-AE’s hidden layer outputs, and \( \mathbf{X} = [x_1, \ldots, x_N] \) are its input and output data.

**RESULTS AND DISCUSSION:**
Real Time MRI images are collected from publicly available domains. The proposed brain tumor detection technique is analyzed with the help of experimental results.

**Experimental set up and evaluation metrics:** The proposed technique is implemented using MATLAB on a system having the configuration of 4 GB RAM and 2.8 GHz Intel i-5 processor. The metrics used to evaluate the proposed technique are sensitivity, specificity and accuracy. In order to find these metrics, we first compute some of the terms of True Positive (TP), True Negative (TN), False Negative (FN) and False Positive (FP).

The evaluation metrics of sensitivity, specificity and accuracy can be expressed in terms of TP, FP, FN and TN. Sensitivity is the proportion of true positives that are correctly identified by a diagnostic test. It shows how good the test is at detecting a disease:

\[
\text{Sensitivity} = TP / (TP + FN)
\]  (28)

Specificity is the proportion of the true negatives correctly identified by a diagnostic test. It suggests how good the test is at identifying normal (negative) condition:

\[
\text{Specificity} = TN / (TN + FP)
\]  (29)

Accuracy is the proportion of true results, either true positive or true negative, in a population. It measures the degree of veracity of a diagnostic test on a condition:

\[
\text{Accuracy} = (TN + TP) / (TN + TP + FN + FP)
\]  (30)

**EXPERIMENTAL RESULTS:**

![Figure 3: Input MRI images.](image1)

![Figure 4: MRI images after Median and Weiner Filtering.](image2)
Performance analysis: The performance of the proposed technique is analyzed with the use of evaluation metrics of sensitivity, specificity and accuracy. Table 2 shows the performance of the proposed method. Metrics are evaluated by varying the hidden neurons from 50 to 200 in step of 25. Figure 7 to 9 shows graphs of sensitivity, specificity and accuracy respectively. The graphs are taken by varying hidden neuron in the neural networks.

Table 2: Comparative Performance of the different classifiers in terms of sensitivity, specificity and accuracy

<table>
<thead>
<tr>
<th>S.No</th>
<th>Classifier</th>
<th>Sensitivity%</th>
<th>Specificity%</th>
<th>Accuracy%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DNN</td>
<td>95</td>
<td>93</td>
<td>95</td>
</tr>
<tr>
<td>2</td>
<td>ELM</td>
<td>91</td>
<td>94</td>
<td>95</td>
</tr>
<tr>
<td>3</td>
<td>ML-ELM</td>
<td>94</td>
<td>97</td>
<td>97</td>
</tr>
</tbody>
</table>

Figure 7: Chart showing Performance of the different classifiers in terms of sensitivity, specificity and accuracy

Table 3: Comparative Performance of the different classifiers in terms of training and testing accuracy

<table>
<thead>
<tr>
<th>S.No</th>
<th>Classifier</th>
<th>Training Accuracy%</th>
<th>Testing Accuracy%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DNN</td>
<td>94</td>
<td>95</td>
</tr>
<tr>
<td>2</td>
<td>ELM</td>
<td>96</td>
<td>95</td>
</tr>
<tr>
<td>3</td>
<td>ML-ELM</td>
<td>96</td>
<td>97</td>
</tr>
</tbody>
</table>

Figure 8: Chart showing performance of the different classifiers in terms of training accuracy
CONCLUSION
In this study, tumor classification using possibilistic fuzzy C means and multilayer extreme learning classifier is proposed. The image is segmented using Clustering and classified using extreme learning classifier. The proposed technique is evaluated using sensitivity, specificity and accuracy, training and testing time. Results show that multi-layer extreme learning machines provide better results than deep learning classifiers and extreme learning machines. For future work the methodology can be tested on big data and more real datasets.

Some optimization techniques can be implemented for improving the performance of the proposed system.

REFERENCES

22. S.N.Deepa, “ B.Arunadevi ,”Extreme Learning Machine For Classification Of Brain Tumor In 3D MRI Images” ,Informatol. 46, 2013., 2, 111-121