Brain Tumor Detection Based on Hybrid Deep Neural Network in MRI by Wind Driven Water Wave Optimization

[1]A. Niranjan, [2]Dr. R. Vasanth Kumar Mehta,
[1]Assistant professor, Department of ECE,
SCSVMV Deemed University, kancheepur 2Associate Professsor, Department of CSE,
*SCSVMV Deemed University, Kancheepuram, *Email:niranjanaphd1@gmail.com*

Abstract:
Image processing techniques are helpful in several applications mainly in the medical field for the examination of computerized disease. The medical imaging method named Magnetic Resonance Imaging (MRI) is broadly used to record the information about brain tumor for the future clinical study and research. Automatic evaluation approaches are required for the assessment of brain MRI due to its various modality and complexity. This paper presents a novel approach to support the examination of brain MRI. Also, it extracts the edema sector and tumor core from the image using Wind Driven Water Wave Optimization (WDWWO), entropy value and contour based segmentation. Deep Neural Network based Elephant Herding Optimization (DNN-EHO) is proposed for the classification. The proposed work is implemented in MATLAB 2018 platform and the severity of tumor is identified. Further, the WDWWO approach can be validated on Medical Image Computing and Computer Assisted Intervention (MICCAI), 2012 dataset challenge on BRATS (Brain Tumor Segmentation) and attained better outcomes in terms of accuracy, recall, specificity, precision, and F-measure. The testing process for this proposed method is carried out with the T1C, Flair, and T2 modalities. Further, with the BRATS 2012 challenge dataset, WDWWO assisted tumor detection framework is validated and attained better values for F-measure, precision, specificity, sensitivity, and accuracy.

Key terms: *Filtering, Feature extraction, Classification, Segmentation, Tumor detection*

1. INTRODUCTION

In recent medical scanning methods, the MRI scanning is superior to the CT scanning due to the high contrast and sensitivity with respect to various intensity issues in brain image [1]. Because of the low intensity variation between the surrounding cells and tumor cells, the detection of tumor is a difficult task in brain MRI images. Glioblastoma and Glioma are the two categories of brain tumors [2]. The Glioblastoma tumors are detected by employing the conventional methods with high accuracy level and these are the low pixel intensity cells. In adults, the common brain tumor is Glioma and can be classified as four grades. They have irregular boundary regions and are high pixel intensity cells. From the glial cells, they are aroused [3]. Due to the irregular boundary regions, the detection of these tumors from the brain MRI images is a difficult task. The tumor threat relies on the combination of various

factors such as the way of spreading, size and location of the tumor [4].

Malignant and Benign are the two types of brain tumors according to the National Cancer Institute [5]. The cells of the malignant tumor have no big difference when compared to the healthy cells. These tumor cells grow quickly and spread to the surrounding cells. Almost they are referred to as cancer [6]. Whereas the Benign tumors do not spread and they contain abnormal cells that grow slowly with specified limits. Ependymoma, oligodendroglioma, glioblastoma, and astrocytomas are some of the brain tumors. Using the MRI brain images, the detection of brain tumors is complex and more tedious. In certain issues, the abnormal accumulation of fluid within the body is called Edema [7]. Two major aspects here are the grade and the type of tumor which select the features for the segmentation process. This is due to the differences in the appearance such as location, contrast uptake, regularity, shape, etc. [8].

The tumor has been categorized into four grades based on the severity level. The low-grade glioma uses grade 1 and 2 and the high-grade glioma uses the grade 3 or 4. The least dangerous tumor in grade 1 is usually linked to the prolonged survival. When seen through a microscope they are gradually grown and in practically they have a usual appearance. The surgery based treatment is found effective for this type of tumor grade [9]. The cases having grade 1 type of brain tumor are named as gangliocytoma, ganglioglioma, and Pilocytic astrocytoma. Under a microscopic instrument, Grade 2 tumor looks like anomalous and slowly growing. Some disperse into nearby tissues and repeat occasionally as grade high tumor

[10]. Grade 3 tumor is known as malignant, and there is no high contrast among grade 2 and 3 tumors. This tumor has an affinity to frequently replicate as grade 4 and also it is considered as the supreme malignant tumor [11]. While viewing through the microscopic instrument, it quickly duplicates the strange appearance and effectively spread into the nearby tissues. Dead cell zones are presented in the middle of the tumor cells [12].

Different systems for the MRI data classification have been developed to identify the abnormalities in the brain. To separate the normal images from the abnormal images, normally all the developed systems have the pre-processing, extracting features and classification. Several approaches have been introduced to utilize the machine learning approaches. The previous research works developed the hybrid solutions on the basis of principal component analysis, support vector machine, wavelet transform and independent component analysis. To classify the abnormal and normal brain MRI, the KNN and the feed forward back propagation ANN based classifiers are employed [13]. For showing the importance of the processing of brain MRI images on the basis of computer the previous works are important. Although there are many approaches introduced for the brain tumor detection, still, now it is a challenging process. Still, a better method is needed for the detection of brain tumors using the MRI images.

The main aim of this process is to develop an effective segmentation architecture using fuzzy based deep machine learning techniques for predicting

the brain tumor from the MRI images. It also aims to show that the developed model can be used along with the MRI brain medical databases to identify the tumor severity with reasonable accuracy.

The objectives that provide huge support to achieve the prior objectives are listed below:

- To identify the well suitable segmentation model which can assist the clinical experts in predicting the brain tumor using various psychological and physical attributes.
- To determine the future outcomes on the basis of past experiences and current situations.
- To recognize the risky patients with an aim to minimize the cost of treatment and to improve the quality of treatment.

2. RELATED WORKS

Few recent research works performed for brain tumor detection are given as follows:

Manorama Sharma et al. [14] have proposed an approach to retrieve information from brain MRI images using hybrid K-means and Artificial Neural Network (ANN) technique for tumor detection. This work presented an efficient method on the basis of K-means clustering and ANN. GLCM is utilized to extract the features. Fuzzy Inference System was developed by utilizing the extracted features that was followed by thresholding, morphological operator and Watershed segmentation. This method identified the tumor size and the affected part of the MRI brain image.

Abhilash Panda et al [15] introduced MRI segmentation based on discriminative

clustering for the automated detection of brain tumor. In this work, the clustering process represented the tissues in MRI brain image with better variance. The simple technique used to classify the image as normal or else abnormal which was termed as AdaBoost. A simple linear iterative clustering method also used for generating 2D super pixels. This method achieved better accuracy on Brain Web MRI dataset.

Sivagami Periannan and Sudeshna Bhakat [16] proposed Cuckoo Search Algorithm (CSA) and histogram thresholding for the detection of brain tumor. Exact tumor part extraction was done with the proposed techniques and also represented with the careful division of tumor limit along with right visual area of tumor with boundary support. According to the size of tumor, the researchers analyzed that the tumor was actually present or not.

Shiv Naresh Shivhare et al. [17] have introduced an effective recognition and segmentation in MR brain images using the approach named parameter-free clustering. The authors proposed an automatic system for detecting the brain tumor based on the following techniques like parameter free clustering, hole-filling, and morphological dilation operations.

Srinivas and Sasibhushana [18] had proposed segmentation and classification techniques in terms of Fuzzy C Means (FCM) and Support Vector Machine (SVM) to identify the brain tumor. FCM algorithm was used to detect and extract the tumor affected part. The parameters such as MSE (Mean Squared Error) and PSNR (Peak Signal to Noise Ratio) were evaluated.

3. PROPOSED BRAIN TUMOR DETECTION METHODOLOGY

Initially, MRI brain image input is preprocessed and then guided image filtering method is used to filter the input image. At first, the image is classified by means of a hybrid deep learning strategy.
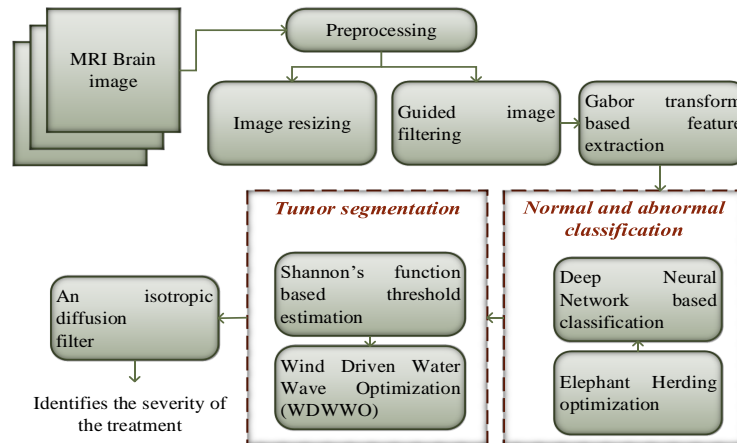
The weight in the DNN (Deep Neural Network) is optimized with Elephant Herding optimization (DNN-EHO). Here, the feature extraction process is done with Gabor transform and then these extracted features are given to the classifier.

Figure 1: Schematic representation of the proposed tumor detection strategy

Schematic outline of the proposed tumor detection strategy is illustrated in Figure 1. Initially, the classifiers are used to categorize the input MRI brain images as normal and abnormal images. Then, these classified images are segmented using wind driven water wave optimization (WDWWO). This optimization method has powerful global search capability and rigorous, great computation exactness. In order to remove the undesirable tissues present in the classified images, erosion is done prior to the segmentation procedure. At last, an anisotropic diffusion filter (ADF) smoothen the image so as to diminish the over-fitting issues.

3.1 Pre-processing
3.1.1 Image resizing:
At first, the images are collected from the dataset for classification. After the transformation, the image needs to resize

for pre-processing. The size of the image is changed to [128×128] in our work. Then the images are filtered using guided image filtering method.

3.1.2 Guided Image filtering:
The filtering result Q and guidance local linear pattern I are the two major postulates of this guided filter [19]. Then, assume a linear adjustment for I of a pixel K positioned at window w_K is described as Q:

$$Q_i = A_K I_i + B_K, \forall_i \in w_K \qquad (1)$$

Here are some linear coefficients which are consistent in window w_K is represented as (A_K, B_K). In this method, a square window having radius 'r' is used. As $\nabla Q = A \nabla I$ therefore both I and Q has an edge and local linear pattern is considered for assurance. The constraints obtained from the filter input P is used for

determination of the linear coefficients (A_K, B_K). Some undesirable factors like noise N or textures are subtracted from input P to obtain the output Q:

$$Q_i = P_i - N_i \quad (2)$$

Identify a solution for minimizing the variation among Q and P, meanwhile the linear pattern needs to be preserved. Apart from that, the associated cost function of window w_K also get reduced:

$$e(A_K, B_K) = \sum_{i \in w_K} \left((A_K I_i + B_K - P_i)^2 + \in A_K^2 \right) \quad (3)$$

Where, $\in$ indicates the regularization parameter, due to this the expansive A_K gets penalized. A model for linear ridge regression and its corresponding solutions are given in equation (4),

$$A_K = \frac{\frac{1}{|w|} \sum_{i \in w_K} I_i P_i - \mu_K \overline{P_K}}{\sigma_K^2 + \in} \quad (4)$$

$$B_K = \overline{P_K} - A_K \mu_K \quad (5)$$

The variance and mean for I in window w_K is represented as σ_K^2 and μ_K respectively. The amount of pixels in window w_K is described as $|w|$ and the mean value for P in w_K is

$$\overline{P_K} = \frac{1}{|w|} \sum_{i \in w_K} P_i$$

represented as . With equation (1) and determined linear coefficients, the filtering result Q_i gets evaluated. For any instance, the pixel i is

associated with the entire overlying windows w_K that covers i. Hence, the estimated Q_i is made available in (6) and then evaluate it for several windows, but it fails to provide the identical results. Due to this, all the possible solutions are average in a simple manner. After calculating (A_K, B_K) for all w_K, the filtered image is computed and it is given in equation (6),

$$Q_i = \frac{1}{|w|} \sum_{K|i \in w_K} (A_K I_i + B_K) \quad (6)$$

Due to symmetry property of the square window (i.e. $\sum_{K|i \in w_K} A_K = \sum_{K \in w_i} A_K$), equation (6) is modelled as,

$$Q_i = \overline{A_i} I_i + \overline{B_i} \quad (7)$$

Where, the mean coefficients for the total windows overlapping i are represented as

$$\overline{A_i} = \frac{1}{|w|} \sum_{K \in W_i} A_K$$

$$\overline{B_i} = \frac{1}{|w|} \sum_{K \in W_i} B_K$$

and . As a spatial shift is produced by the linear coefficients $(\overline{A_i}, \overline{B_i})$ and also due to the changes in (7), ∇I never get scaled again. The gradients of mean filter at any instance are predictable even though they are found much smaller than I which is found closer to that of edges and same as the output, $(\overline{A_i}, \overline{B_i})$. During such circumstances, maintain $\nabla Q \approx \overline{A} \nabla I$ which represents that the abrupt variations in intensity I are conserved regularly in Q.

3.2 Feature extraction based on Gabor transform:

Gabor method can be used to eliminate a large portion of instability in image because of the changes in contrast and lighting, but being robust for slight movements and distortion. Gabor wavelets appear to be a decent estimation for the neurons sensitivity profiles present in visual cortex of higher vertebrates. Generally, these cells will arise in sets with even and odd symmetry as the real and imaginary portion of Gabor-basis wavelets.

A 2D Gabor filter $Q_i(X,Y)$ is expressed as:

$$Q_i(X,Y) = \frac{\tau \upsilon}{\pi} e^{-\left(\tau^2 X^2 + \upsilon^2 Y^2\right)} e^{j2\pi F_0 X^2}$$

(8)

$$X' = X\sin\theta + Y\cos\theta$$

(9)

$$Y' = -X\sin\theta + Y\cos\theta$$

(10)

Where, F_0 indicates the central frequency in the sinusoidal plane wave, and the two parameters τ and υ are used to scale the two axis in the envelope of elliptic Gaussian. The orientations of both the sinusoidal function and the Gaussian envelope are found identical to each other. Normally, the Gabor function is illustrated as Gaussian shaped function. The features are extracted from these filter and is provided as input to DNN-EHO classifier.

3.3 Hybrid deep neural network with Elephant Herding Optimization (DNN-EHO)

The EHO is very effective optimization procedure to train the DNNs when compared to the existing derivative-free optimization procedures, since EHO has a less tuning parameters, convergence speed aids, and well execution over numerous former derivative-free optimization procedures. The arrangement of DNN-EHO is proposed in this paper so as to conquer the previously mentioned issues. It optimizes the DNN framework comprising of at least one auto encoders with a softmax classification layer. This optimization improves the capability of discovering local minima of search space. The ability of EHO's exploring fresh feasible solution applicants situated in extreme from present point at which the algorithm is surrounded. To the best of our insight, our strategy is the first to think about the enhancement of the DNN based structures with the greatest precision.

3.3.1 Deep Neural Network (DNN)

The idea of the DNN depends on a multilayer perceptron (MLP). The completely joined weights of the MLP are initiated by means of unsupervised or supervised pre-training strategy. However, high computational complexity is one of the greatest drawbacks in this model.
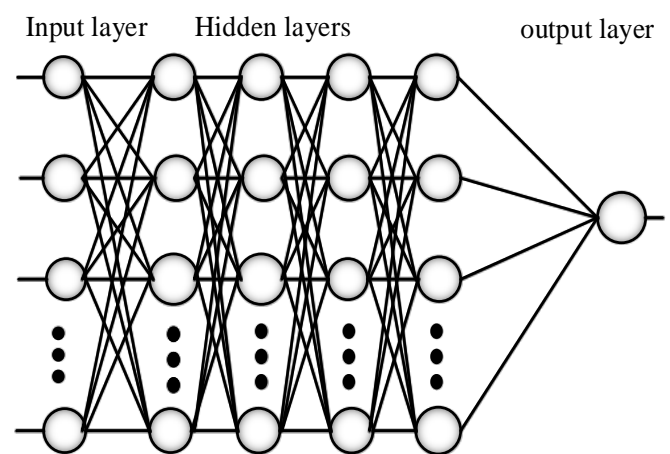

Figure 2: Schematic diagram for DNN

Assume, the input dataset I with $m*n$ dimension and is framed as $Q = \{Q(1), Q(2), ... Q(i)..Q(N)\} \in R_m$. In this demonstration, m represents the dataset length and n denotes the amount of samples taken for estimation. This information matrix is given as input to the versatile auto-encoder. Then, the neural network is framed as the structure illustrated in Fig. 2. The assortment of an activation function is a crucial stage in this procedure which disturbs the executions [37].

For enhancing the performance, we choose the sigmoid activation function in our work because of its impressive nature. Our next goal of this work is to do feature learning and to grasp the feature on input data Q and these features are expressed as given in (11)

$$h(Q(i), W, b) = \sigma(WQ(i) + b), i = 1, 2, ...n)$$

(11)

Also, the output layer function is provided as

$$outputlayer = \sigma\left(W^I h(Q(i) // b), i = 1, 2, ..., n\right)$$

(12)

Moreover, an objective function is introduced in our work to ease the recurrence of input features. The dynamic neurons at the present cycle of hidden layer is controlled by this objective function. Assume, $a_j(i)$ as the activation function of jth hidden layer. We have deliberated back propagation method for the demonstration of programming imperfection forecast. Back propagation is a procedure that propagates the learned weights from external nodes to internal nodes in order to diminish the learning error. It can be attained through computation of the gradient of the network weights. This procedure is processed on input feature set X, where hidden layer can be conveyed as provided in Eq. (13).

$$a = sigmoid(WQ + b)$$

(13)

Where, b represents biases and W denotes the weight of the input feature set. Utilizing this supposition, the mean weight of activation function can be expressed as:

$$P_j = \frac{1}{n}\sum_{i=1}^{n}[a_j(i)]$$

(14)

The mean activation function is near to zero at the beginning of DNN training, due to indolence of the neurons. This can be accomplished by applying a penalty to the mean activation function when it deviates from the crucial value of the average activation function. This can be given as in (15),

$$P_{penalty} = \sum_{j=1}^{s_2} KL(\rho // \rho_j)$$

(15)

Here, s_2 represents the total count of neurons presented in the hidden layer and $KL(.)$ denotes the Kullback–Leibler divergence (KL divergence) and it can be stated as:

$$KL(\rho // \rho_j) = \rho \log\frac{\rho}{\rho_j} + (1-\rho)\log\frac{1-\rho}{1-\rho_j}$$

(16)

If $\rho_j = \rho$ then $KL(\rho\|\rho_j)=0$ or it endures an increment that produces divergence, called an adaptive constant. A cost function is used to determine this adaptive constant and it can be expressed as,

$$C_{adaptive}(w,b)= C(w,b)+ \beta\sum_{j=1}^{s2} KL(\rho // \rho_j)$$
(17)

Where, applied weight penalty is denoted as β which utilizing Kullback–Leibler divergence (KL divergence) classic. It is a very challenging task to identify the significant parameters W and b due to the direct proportionality of the cost function and those two parameters, which can influence the efficiency of the system. To tackle this issue, we frame an optimization problem and it aims to minimize the $C_{adaptive}(w,b)$. The parameters w and b are iteratively updated in the back-propagation method in order to solve the optimization issue. This can be written as:

$$w_{ij}(l)= w_{ij}(l)- \varepsilon \frac{\partial}{\partial w_{ij}(l)}C_{adaptive}(w,b)\Rightarrow X_{new,ci,j}$$
(18)

$$b_i(l)= b_i(l)- \varepsilon \frac{\partial}{\partial b_i(l)}C_{adaptive}(w,b)$$
(19)

Here ε, indicates the learning rate of the DNN model.

3.3.2 Elephant herding optimization (EHO)

We consider the herding activities of the elephants as two operators. A general-purpose global optimization method is formed by a successive idealization of these two operators. The entire global optimization issues are solved so as to create the herding activities of elephants. The idealized laws given below are favored to simplify these issues. Hence, this EHO [20] is utilized here to update the weight of DNN.

1) The population of elephant is a collection of few clans, and every clan consumes a fixed number of elephants.
2) A particular amount of male elephants will depart from their family and stay away from their family groups during each generation.
3) A matriarch is a leader for the elephants who are living together in every clan.

a) Clan updating operator

All the elephants are living together under the control of a matriarch in every clan as stated earlier. Hence, the next location of every elephant live in clan ci is prejudiced by matriarch ci. The position of the elephant j in clan ci can be reorganized as,

$$X_{new,Ci,j} = X_{Ci,j} + \alpha \times (X_{best,\,Ci} - X_{Ci,j})\times r$$
(20)

Where, the old and recently updated location for DNN weight (elephant) j in clan C_i is represented as, $X_{new,Ci,j}$ and $X_{Ci,j}$, respectively. A scale factor to detect the control of matriarch C_i on $X_{Ci,j}$ is [0, 1]. $X_{best,Ci}$ indicates the fittest weight individual in clan ci, r ę [0, 1]. Here, uniform distribution is utilized. The fittest weight is not able to update in every clan using Eq.

(21), i.e., $X_{Ci,j} = X_{best,Ci}$. Then update the position for the fittest one as

$$X_{new,Ci,j} = \beta \times X_{center,Ci} \quad (21)$$

Where, the factor [0, 1] which defines the control of the $X_{center,Ci}$ on $X_{new,Ci,j}$. The information acquired from the entire weights presented in clan Ci is used for the creation of new individual $X_{new,Ci,j}$ in Eq. (21). $X_{center,Ci}$ represents the clan Ci center, and the equation to evaluate it for the d-th dimension is given as

$$X_{center,Ci,d} = \frac{1}{N_{ci}} \times \sum_{i=1}^{n_{ci}} X_{Ci,j,d}$$

$$(22)$$

Where $1 \leq d \leq D$ specifies the dth dimension and D represents the total dimension. N_{Ci} represents the number of weight in clan Ci. $X_{Ci,j,d}$ is the dth of the weight individual $X_{Ci,j}$. The center of clan Ci, $X_{center,Ci}$ can be evaluated via D calculations using Eq. (22).

b) Separating operator

If the male weights presented in the weight collection become teens, then they will vacate their family group and live in some other locations alone. While resolving the optimization issues, these isolation procedures can be molded into isolating operator. Consider, the implementation of isolating operator using weight entities with worst fitness value at every generation so as to increase the searching capacity of EHO technique auxiliary as expressed in Eq. (23)

$$X_{worst,Ci} = X_{min} + (X_{max} - X_{min} + 1) \times rand$$

$$(23)$$

Where, X_{min} and X_{max} are minimum and maximum bound of the location of single elephant respectively. $X_{worst,Ci}$, represents the worst individual (weight) in clan Ci. $rand()$ is a sort of uniform and stochastic appropriation in the range [0, 1]. The EHO technique is implemented using the depiction of clan refreshing and isolating operator and its workflow is outlined as appeared in table (1).

The weight of DNN is more effectively optimized by EHO. The best fitness function parameter is contributed through the essential weight vector of DNN. The weight optimization procedure has the ability to converge optimum or close optimum solution significantly. Hence, the convergence rate of EHO is very fast and its likelihood of scapping is local minima.

Table 1: Algorithm for EHO

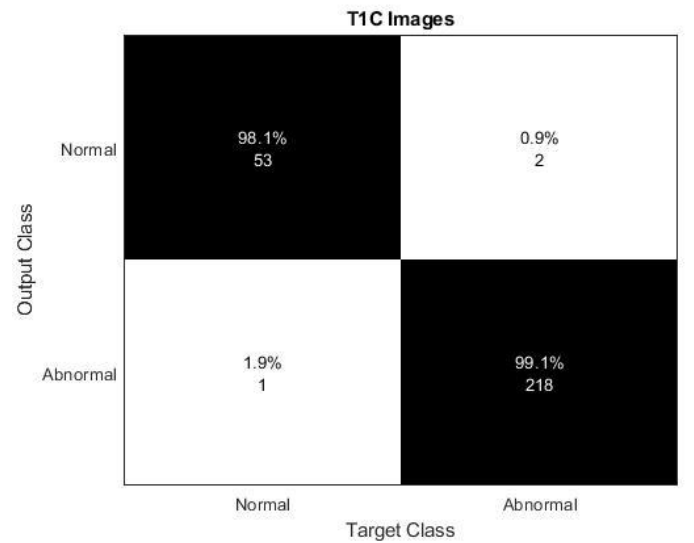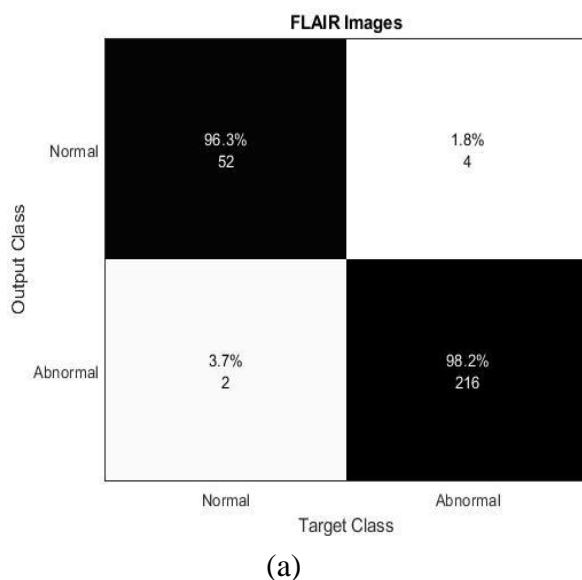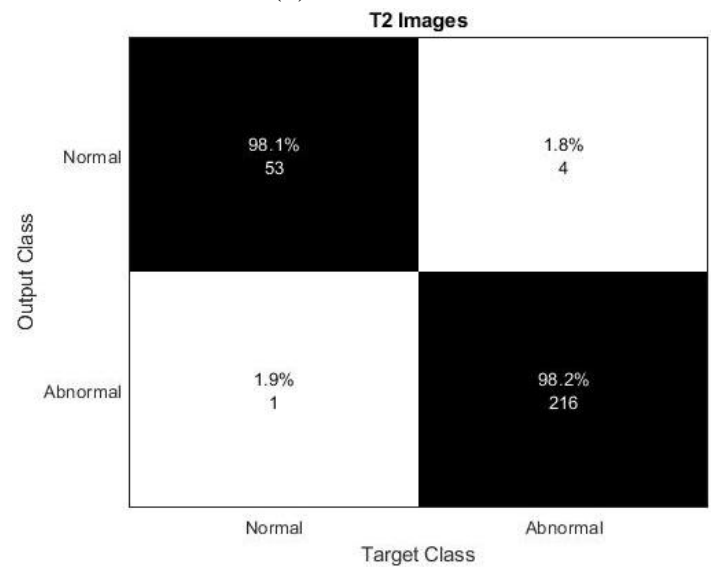| |
|---|
| ***Step 1:*** Initialization. Set generation counter t=1; initialize the population; the maximum generation Max_{gen}. |
| ***Step 2:*** While $t < Max_{gen}$ do Based on fitness, sort all the elephants |
| ***Step 3:*** Clan updating operator for $Ci = 1$ to N_{clan} ($\forall$ clans in whole population) do for $j = i$ to N_{Ci} ($\forall$ elephants in Ci) do |
| ***Step 4:*** Update $X_{Ci,j}$ and use Eq. (21) to generate $X_{new,Ci,j}$. If, $X_{Ci,j} = X_{best,Ci}$ then |

Step 5: Update $X_{Ci,j}$ and Eq. (22) generate $X_{new,Ci,j}$

       end if

       end for j

       end for Ci

Step 6: Separating operator

       For $Ci = 1$ to N_{clan} ($\forall$ clans in whole elephant population) do

Step 7: Using Eq.(23), Replace the worst one from Ci.

       end for Ci

Step 8: Estimate the population with the newly updated positions.

       $t = t + 1$.

Step 9: end while

Generally, the training procedure of DNNs choose the associated weights in-between the neurons so as to reduce the error. The MRI brain images are categorized as normal and abnormal images at the time of DNN-EHO classification. Then, WDWWO algorithm used to segment the image after the classification process.

(a)

(b)

(c)

Figure 3: Confusion matrix for (a) Flair image (b) T1 image and (c) T2 image

The performance of the classifier on a sample data collection with known true values are frequently described by a table called confusion matrix (figure 3). 96.3%, 98.1% and 98.2% of images are identified as true from 53 normal images. Then 98.2%, 99.1% and 98.2% images are identified from 216 abnormal images. Thus, the all-out precision is 98.18%. The segmentation and image smoothing is performed at the post processing stage. We

use WDWWO basis segmentation and an isotropic based diffusion. The value for Flair image of True Positive (TP) is 52, False Negative (FN) is 216, True Negative (TN) is 4 and False Positive (FP) is 2. The value for T1 image of TP is 53, FN is 218, TN is 2 and FP is 1. The value for T2 image of TP is 53, FN is 216, TN is 4 and FP is 1.

3.4 Brain tumor segmentation
3.4.1 Shannon's function:

The uncertainty of a variable is calculated by typically adopting the entropy. The complete information about Shannon entropy is attained from the work of Paul and Bandyopadhyay [24].

The method as per the following: Consider, an image frame having size $m \times n$. Then, the pixel's gray value positioned at (X,Y) can be represented as $f(X,Y), X \in \{1,2,..,m\}, Y \in \{1,2,..,n\}$.
Then, the total gray levels present in the whole image l is represented as L and the complete gray level set is described as $\{0,1,2...L-1\}$. The normalized histogram is represented as $E = \{e_0, e_1, ... e_{L-1}\}$. To resolve the issue of multi-thresholding, this equation is rewritten as;

$$E(t) = e_0(t_1) + e_1(t_2),...,e_{L-1}(t_{k-1})$$
(24)

Thus the objective function can be defined as:

$$f(t) = \max(E(t))$$
(25)

In which $f(t)_{T*}$ is the optimal threshold. WDWWO optimization method is used to accomplish this optimal result.

3.4.2 Wind Driven Water Wave Optimization (WDWWO)

The WDWWO [21] is used for the extracted image segmentation. The searching capability is improved by this WDWWO algorithm (i.e. to obtain optimal solution), moreover it also minimizes the iteration number, computational burden, and total number of searching agents. During each iteration, a propagation is performed by each wave. On the basis of original wave, a new wave gets generated by dynamically adjusting the searching ability. This dynamic adjustment is carried out to balance the local and global search.

To overcome the demerit of traditional WWO, the WDO velocity is modified. Naturally, the motion of water wave gets accelerated due to the wind to attain the best optimal solution. It is a feasible and most effective algorithm to eliminate the optimization issues. The wave during each iteration may perform propagation. With original wave x as base, the new wave x' gets generated.

Due to propagation, the waves (abnormal images) having high fitness may get exploited within the small area and those having low fitness may explore the large areas. The convergence is easily performed by WWO, due to this the solutions may fall within the local optimum. The global searching capability of WWO is high therefore it eliminates search stagnation and determines the global optimal solution. Then, set the t value as 1s to attain the location update equation and it is given below:

$$\sigma'(Q) = \sigma(Q) + rand(-1,1).\lambda L(Q).\left(\overline{W_{new}}.\Delta t\right)$$
(26)

Where, the search space length is denoted as $L(Q)$, and the random number that are uniformly distributed is represented as $rand(-1,1)$.

For $f(\sigma') > f(\sigma)$, the wave x gets replaced with the new wave σ' and the height is set as H_{max}. However, the reduction in wave height and reservation of wave x indicates the energy loss.

Where $\sigma'(Q)$ denotes the location of data σ', $\sigma(Q)$ denotes the location of original data σ, and the speed of wind is represented as $\overline{W_{new}}$. The presence of new abnormal images σ' beyond the boundary may effectively reassign the random location in entire search space. Therefore, the wave wavelength is given as

$$\lambda = \lambda.\xi^{-(f(\sigma)-f_{min}+\varepsilon)/(f_{max}-f_{min}+\varepsilon)}$$
(27)

Where, the minimum and maximum fitness values are represented as f_{min} and f_{max} respectively and the wavelength reduction coefficient is represented as ξ. The infinitesimal positive integer that eliminates the division with 0 is represented as ε.

a) Refraction

During wave propagation, the wave energy reduction continues till reaching the maximum wave height. The major goal of refraction is search stagnation avoidance. Therefore, the location update is expressed as,

$$\sigma'(Q) = N\left(\frac{\sigma^*(Q)+\sigma(Q)}{2}, \frac{|\sigma^*(Q)-\sigma(Q)|}{2}\right)$$
(28)

Where, N (μ, σ) generates the Gaussian random number, and the optimal solution is denoted as σ^*. After refraction, the wave height is reset as H_{max} for the new wave σ' and the equation for wavelength update is given as

$$\lambda' = \lambda \frac{f(\sigma)}{f(\sigma')}$$
(29)

b) Breaking

Maximization of wave energy may automatically increase the steeper of wave crest, therefore finally the entire wave gets spit into a sequence of solitary waves. In this particular approach, the k dimensions are randomly selected and the solitary waves are produced from those selected dimensions. Then, the location update is given as,

$$\sigma'(Q') = \sigma(Q) + N(0,1).\beta L(Q)$$
(30)

Where, the breaking coefficient is represented as β. The fitness value attained by the wave $\sigma*$ is found greater than the produced solitary waves, therefore the resultant wave $\sigma*$ is found retained. Or else the location of wave $\sigma*$ is taking over by the generated solitary wave. Finally based on the threshold, the brain tumor is segmented and predicted the size of the tumor.

This entire process is clearly described in the WDWWO algorithm. The WDWWO algorithm is given in table 2:

Table 2: Algorithm for WDWWO

Input: Set of abnormal images
Output: Best segmented brain image
Step 1: Randomly initialize population and some main parameters
Step 2: While, termination condition is not reached do
Step 3: for each $\sigma \in P$ do
Step 4: Introduce WDO Velocity into the propagation operation, propagate σ to new σ' using equation (26)
Step 5: if $f(\sigma') > f(\sigma)$ then
Step 6: if $f(\sigma') > f(\sigma^*)$ then
Step 7: Break σ' using (30)
Step 8: Update σ^* with σ'
Step 9: Replace σ with σ'
Step 10: else
Step 11: Decrease $\sigma.H$ by one
Step 12: if $\sigma.H = 0$ then
Step 13: Refract σ to σ' using (28) and (29)
Step 14: Update the wavelengths with (30)
Step 15: Return σ^*

Next to the segmentation process, the image will be smoothened in order to identify the severity of the treatment.

3.5 An isotropic diffusion filtering

The anisotropic diffusion [22] is an iterative approach, which is largely applied to distinguish edge or noise. It is normally referred to as a nonlinear filtering operation. During each iteration, the image gradient is calculated for edge detection, then a single diffusion coefficient on the basis of gradient value is evaluated for each pixel. The diffusion of low gradient values are authorized with a maximum diffusion coefficient, however, the diffusion of high gradient values are limited with weak coefficient.

The anisotropic diffusion is expressed as:

$$\begin{cases} \dfrac{\partial \sigma'(x,y,t)}{\partial t} = div\big(C\|\nabla\sigma'(x,y,t)\|\big)\nabla\sigma'(x,y,t) \\ \sigma'(x,y,0) = \sigma'_0(x,y) \\ \dfrac{\partial\sigma'}{\partial N}\big|_{\partial\sigma'} = 0 \end{cases}$$

(31)

Where, the diffusion function represented as $C(.)$ is a nonnegative, monotonous decreasing function, i.e., $C(0) = 1$. The differences provided by the closest neighbors in four different directions (n, s, e, w) and the diffusion coefficient are evaluated as follows:

$$\nabla_n\sigma'_{i,j} = \sigma'_{i-1,j}, C_{n_{i,j}} = G\big(\|\nabla_n\sigma'_{i,j}\|\big)$$

(32)

$$\nabla_s\sigma'_{i,j} = \sigma'_{i-1,j}, C_{s_{i,j}} = G\big(\|\nabla_s\sigma'_{i,j}\|\big)$$

(33)

$$\nabla_e\sigma'_{i,j} = \sigma'_{i-1,j}, C_{e_{i,j}} = G\big(\|\nabla_e\sigma'_{i,j}\|\big)$$

(34)

$$\nabla_w\sigma'_{i,j} = \sigma'_{i-1,j}, C_{w_{i,j}} = G\big(\|\nabla_w\sigma'_{i,j}\|\big)$$

(35)

Where the function G is defined as:

$$G\big(|\nabla\sigma|\big) = E\left(-\left(\left\|\dfrac{\nabla_\sigma}{K}\right\|\right)\right)^2$$

(36)

The zones having weak gradients are distinguished from strong gradient zones using the threshold K. Here, the value of

K is set as 0.1. Therefore the pixel value is modified using:

$$\sigma'(X,Y) = \sigma'(X,Y) + \lambda[C_n.\nabla_n\sigma' + C_s.\nabla_s\sigma' + C_e.\nabla_e\sigma' + C_w.\nabla_w\sigma']_{x,y}$$

(37)

With $0 \le \lambda \le 1/4$

For such case, the modified model is represented as the regular anisotropic diffusion. After that, the segmented tumor portion area is greater than a hundred means, the corresponding patient needs severe treatment. If the area is less than hundred means the corresponding patient are suggested for the severe treatment. Based on the convex area [23], the tumor volume is to be calculated for identifying the severity of the treatment. So the patient with higher volume of tumor should be suggested to the suitable treatment.

Tumor size calculation:

Diameter-based measurement (DB): The anteroposterior diameter (d1) and craniocaudal diameter (d2) are calculated for the segmented portion. In the axial plane, the lateral diameter (d) is determined.

DB tumor volume (V) is calculated using ellipsoid formula (V = d1 × d × d2 × π / 6).

4. RESULTS AND DISCUSSION

The experimental results obtained with the proposed brain MRI segmentation technique is presented and analyzed in this part. Implementation is carried out in the MATLAB 2018 (a) simulation platform. Initially, this proposed approach is implemented over 128 × 128 sized brain MRI images. Then, this entire process is validated with the BRATS 2012 challenging dataset images (72 images recorded from 6 patients with T2, T1C,

and Flair modalities) having 216 × 160 sized MICCAI brain tumor segmentation images. The brain MRI images are initially preprocessed and filtered. Then, the reliability of this pre-processing phase is analyzed using the image quality metrics. From the available dataset, select an image for testing purpose and then pre-process that selected image to improve the image regions. Then, via the classification results, we can identify abnormal images in the BRATS 2012 data set. Finally, implement an appropriate segmentation technique in post-processing stage for tumor region extraction from pre-processed image. Finally, compared the segmented tumor sections to identify the treatment level (whether it needs severe treatment or not).

4.1 Evaluation metrics

The performances are measured using specificity, F-measure, accuracy, sensitivity, and precision metrics. The formula for the parameters are given below:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

(38)

$$precision = \frac{TP}{TP + FP}$$

(39)

$$Sensitivity = \frac{TP}{TP + FN}$$

(40)

$$Specificity = \frac{TN}{TN + FP}$$

(41)

$$F - measure = \frac{2 \times precision \times sensitivity}{precision + sensitivity}$$

(42)

The below section shows the image representation for the various steps involved in the process.

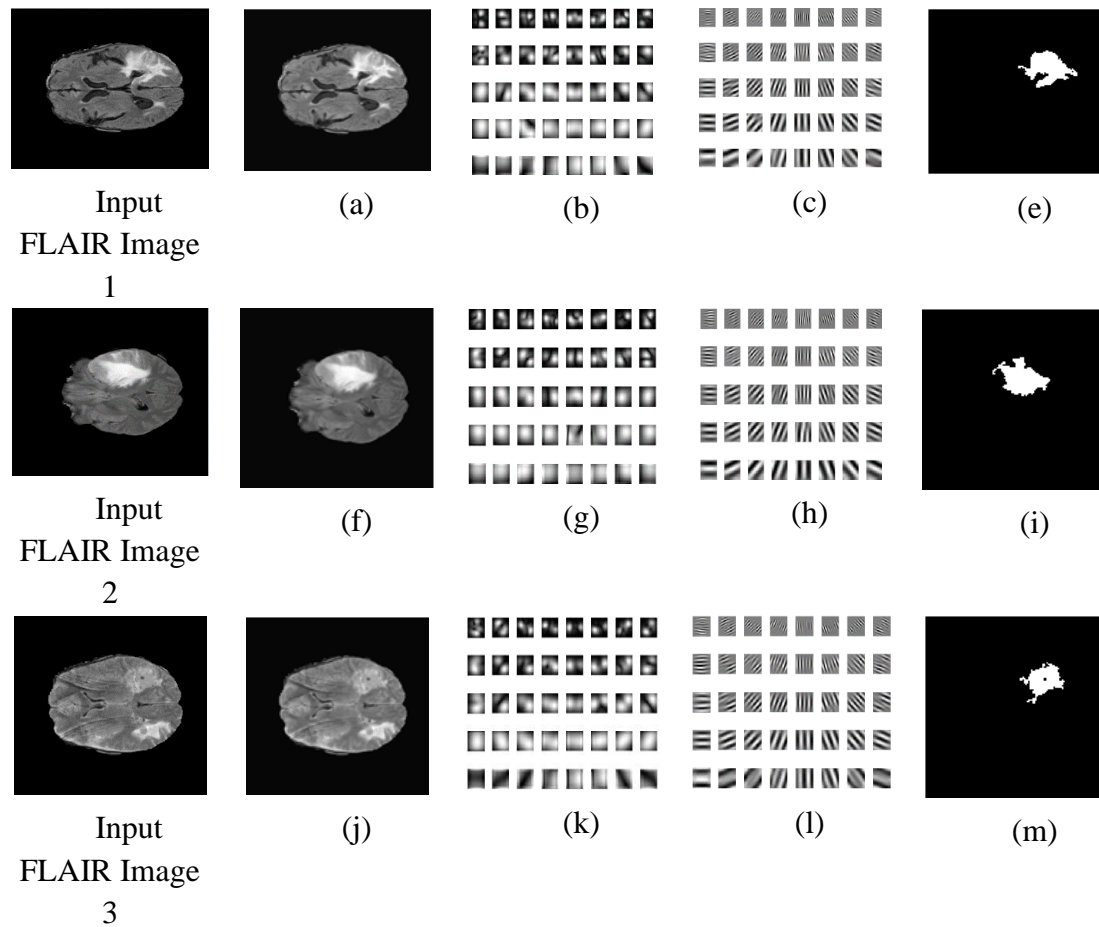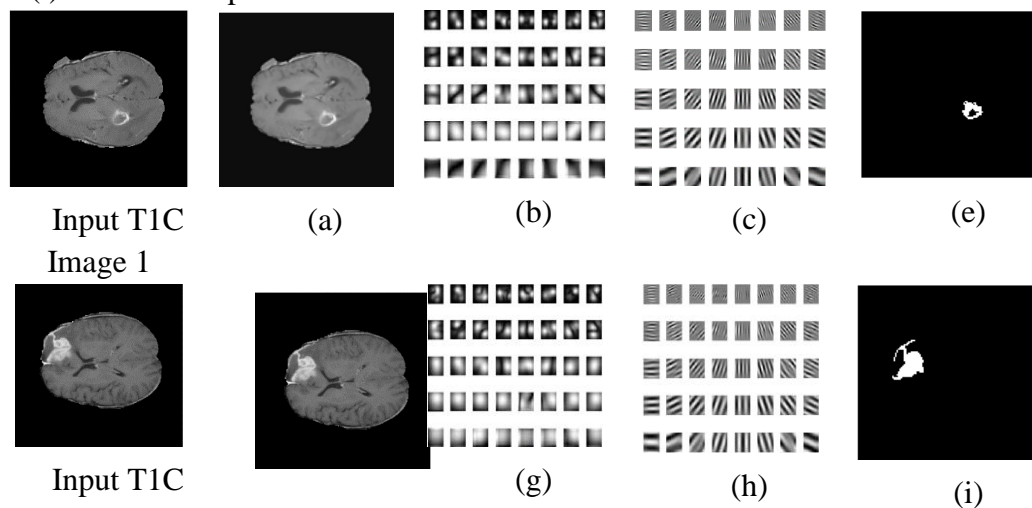| Input FLAIR Image 1 | (a) | (b) | (c) | (e) |
| Input FLAIR Image 2 | (f) | (g) | (h) | (i) |
| Input FLAIR Image 3 | (j) | (k) | (l) | (m) |

Figure 4: BRATS image dataset registered with FLAIR Images

In the above figure (4), FLAIR images from the BRATS data set are taken for the evaluation. Here figures (a), (f) and (j) represented the Guided filtered images. Figure (b), (g) and (l) represented the magnitude of Gabor filter then figures (c), (h) and (l) are the real part of Gabor filter image. After classification, the abnormal images are segmented and the exact tumor portions are detected with the WDWWO algorithm and which are given in figures (e), (i) and (m).
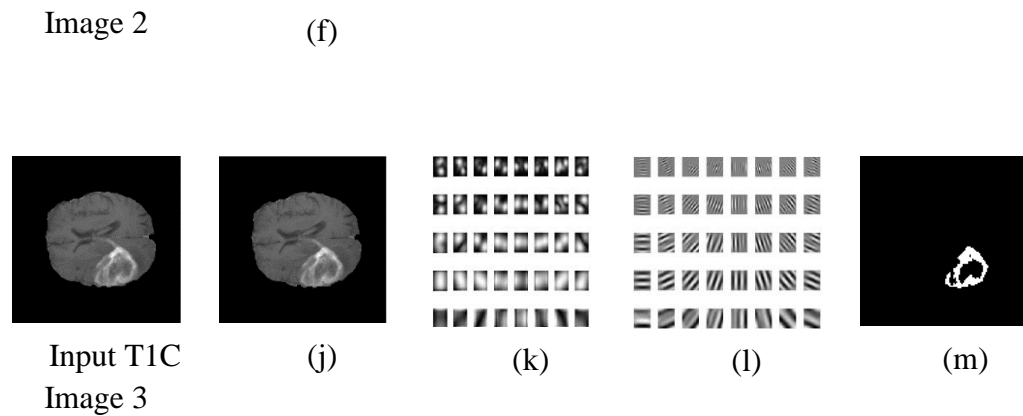

| Input T1C Image 1 | (a) | (b) | (c) | (e) |
| Input T1C | | (g) | (h) | (i) |

Image 2   (f)

Input T1C   (j)   (k)   (l)   (m)
Image 3

Figure 5: BRATS image dataset registered with T1C Images

In this above figure (5), T1C images from the BRATS data set are taken for the evaluation. Here figures (a), (f) and (j) represented the Guided filtered images. Figure (b), (g) and (l) represented the magnitude of Gabor filter then figures (c), (h) and (l) are the real part of Gabor filter image. After the classification, the abnormal images are segmented and the exact tumor portions are detected with the WDWWO algorithm and which are given in figures (e), (i) and (m).
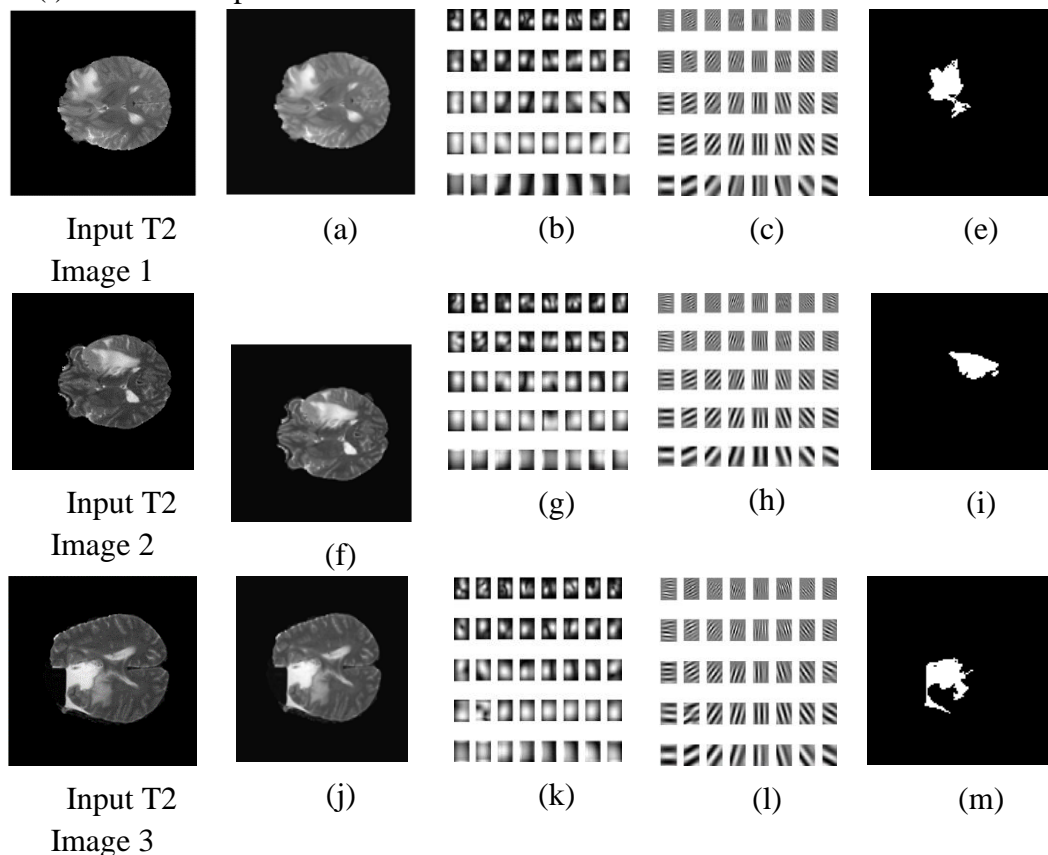

Input T2   (a)   (b)   (c)   (e)
Image 1

Input T2   (g)   (h)   (i)
Image 2   (f)

Input T2   (j)   (k)   (l)   (m)
Image 3

Figure 6: BRATS image dataset registered with T2 Images

In this above figure (6), T2 images from the BRATS data set are taken for the evaluation. Here figures (a), (f) and (j) represented the Guided filtered images. Figure (b), (g) and (l) represented the magnitude of Gabor filter then figures (c), (h) and (l) are the real part of Gabor filter image. After the classification the abnormal images are segmented and the exact tumor portions are detected with the WDWWO algorithm and which are given in figures (e), (i) and (m).

Table 3: Severity detection of brain images

(a)	This corresponding FLAIR image should be suggested to be a severe treatment.
(b)	This corresponding T1C image should be suggested to be a severe treatment.
(c)	This corresponding FLAIR image are not needed severe treatment.
(d)	This corresponding T2 image are not needed severe treatment.

In table 3, based upon the convex [23] estimation the tumor volume is measured and the severity of the tumor is identified. In this above table, (a) represented the patient with severe treatment, (b) gives the result as severe treatment is necessary, (c) mentions the result as FLAIR image with slightly affected tumor and thus treatment is not needed, (d) gives the result as it needs severe treatment and finally (e) represents the result as the corresponding patient requires severe treatment.

The ACC, SEN, PRE, SPE, and FM results provided by the proposed Shannon entropy based segmentation is found superior to other alternatives. Further, the potential BRATS dataset are tested for the WDWWO optimization and the maximum results provided by similarity and statistical measures are provided in the above table. The results attained from these tables indicate that DNN-EHO approach attains better sensitivity of 99.09% specificity of 88.80%, F-measure 98.78%, precision of 98.47% and accuracy of 93.80% values with the BRATS dataset.

4.2 Evaluation metrics with the state of the art comparison

Accuracy, sensitivity, precision, specificity, and F-measure are the few metrics considered in this work to evaluate the effectiveness of the proposed approach. The comparison is performed between the proposed and existing works to show that the performance of this proposed method was better than the other methods. In our work, the current performance metrics are compared with three different existing methods namely TLBO [24], SGO [25] and DNN [26]. In case of specificity and sensitivity, the proposed work is compared with the

existing works namely TLBO, SGO, and DCNN [27]. The comparative results for

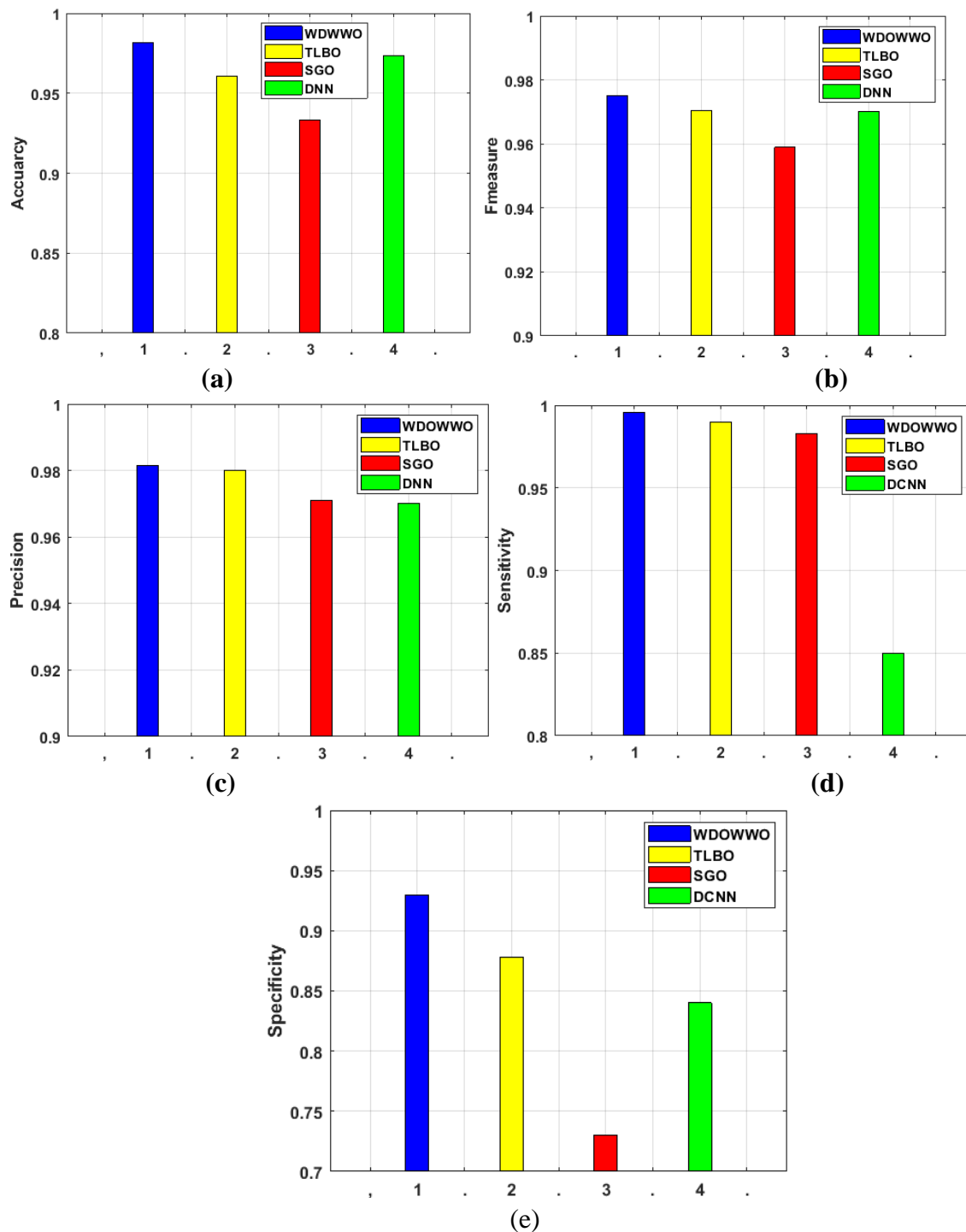this performance assessment factors are shown in the following figures,

Figure 7: Evaluation metrics

The accuracy of the current work is 98% and compared with TLBO, the accuracy is 2% varied. In case SGO, the

accuracy is varied with nearly 5% and in case of DNN, the value has varied into nearly 1%. The value of F-measure in the proposed work is 0.975. The difference in

the proposed work with the existing TLBO is nearly 0.5% and the value of SGO is 15%. In case of DNN, it has changed to 5%. The precision of the current work is 0.983 and the range of precision is high while comparing to the existing works. In case of TLBO, the value of precision is 0.98, it has 3% variation while comparing to the current work. Then the value of SGO is 0.97% and in case of DNN, the value is 0.969. So the variation of SGO and DNN in case of precision is 12%, 11% respectively.

The sensitivity of the current work is 0.995 and the range of sensitivity is high while comparing to the existing works. In case of TLBO, the value of sensitivity is 0.989, it has 3% variation while comparing to the current work. Then the value of SGO is 0.978% and in case of DNN, the value is 0.85. So the variation of SGO and DNN in case of precision is 2%, 14% respectively. The specificity of the current work is 0.925 and the range of sensitivity is high while comparing to the existing works. In case of TLBO, the value of sensitivity is 0.89, it has 3% variation while comparing to the current work. Then the value of SGO is 0.725% and in case of DNN, the value is 0.84. So the variation of SGO and DNN in case of precision is 15%, 8% respectively.

5. CONCLUSION

In this approach, the WDWWO algorithm assisted segmentation and DNN-EHO based classification techniques are analyzed over well-known brain MRI to show the effectiveness of this proposed optimization based segmentation and classification process. The proposed technique is found more efficient as it successfully extracts the mass of tumor from the available MRI, for such extraction different modalities like Flair, T1, and T2 are used. This entire classification process is categorized into two groups named pre-processing and post-processing region. The experimental analysis proved that the involvement of Shannon's entropy based thresholding achieves higher results for this considered dataset. To test the effectiveness of this proposed segmentation process, the BRATS 2012 dataset is utilized in this method. Moreover, this segmentation process also achieved higher performance for the modalities Flair, T1, and T2 of the MRI dataset. The accuracy, precision, sensitivity, F-measure, specificity values attained by this proposed segmentation algorithm illustrate that this proposed method is clinically significant.

REFERENCES

[1] Anand S, Riccardo C, and Pavone FS: Optical Fiber-Probe Spectroscopy of Brain Tumors. In Neuro photonics and Biomedical Spectroscopy, Elsevier 1-23, 2019

[2] Rehman ZU, Naqvi SS, Khan TM, Khan MA, and Bashir T: Fully automated multi-parametric brain tumor segmentation using superpixel based classification. Expert Systems with Applications 118: 598-613, 2019

[3] Manogaran G, Shakeel PM, Hassanein AS, Kumar PM, and Babu GC: Machine Learning Approach-Based Gamma Distribution for Brain Tumor Detection and Data Sample Imbalance Analysis. IEEE Access 7: 12-19, 2019

[4] Quail DF, and Joyce JA: The micro environmental landscape of brain

tumors. Cancer cell 31(3): 326-341, 2017

[5]Polepaka S, Rao CS, and Mohan MC: A Brain Tumor: Localization Using Bounding Box and Classification Using SVM. In Innovations in Electronics and Communication Engineering, Springer, Singapore 61-70, 2019

[6]Li YQ, Chiu KS, Liu XR, Hsiao TY, Zhao G, Li SJ, Lin CP and Sun CW: Polarization-Sensitive Optical Coherence Tomography for Brain Tumor Characterization. IEEE Journal of Selected Topics in Quantum Electronics 25(1): 1-7, 2019

[7]Shakeel PM, Tobely TEE, Al-Feel H, Manogaran G, and Baskar S: Neural Network Based Brain Tumor Detection Using Wireless Infrared Imaging Sensor. IEEE Access 7: 5577-5588, 2019

[8]Nabizadeh N, and Kubat M: Brain tumors detection and segmentation in MR images: Gabor wavelet vs. statistical features. Computers & Electrical Engineering 45: 286-301, 2015

[9]Sajjad M, Khan S, Muhammad K, Wu W, Ullah A, and Baik SW. Multi-grade brain tumor classification using deep CNN with extensive data augmentation. Journal of computational science 30: 174-182, 2019

[10] Louis DN, Perry A, Reifenberger G, von Deimling A, FigarellaBranger D, Cavenee WK, Ohgaki H, Wiestler OD, Kleihues P, Ellison DW: The World Health Organization classification of tumors of the central nervous system: a summary. Acta Neuropathol. 131 (6): 803–820, 2016

[11]Anaraki AK, Ayati M, and Kazemi F: Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms. Bio

cybernetics and Biomedical Engineering 39(1): 63-74, 2019

[12]Amin J, Sharif M, Yasmin M, and Fernandes SL. A distinctive approach in brain tumor detection and classification using MRI. Pattern Recognition Letters 2017

[13]Talo M, Baloglu UB, Yıldırım O, and Acharya UR. Application of deep transfer learning for automated brain abnormality classification using MR images. Cognitive Systems Research 54: 176-188, 2019

[14]Sharma M., Purohit GN, Mukherjee S. Information Retrieves from Brain MRI Images for Tumor Detection Using Hybrid Technique K-means and Artificial Neural Network (KMANN), Networking Communication and Data Knowledge Engineering, Springer, Singapore 4, 2018

[15]Panda A, Mishra TK, and Phaniharam VG: Automated Brain Tumor Detection Using Discriminative Clustering Based MRI Segmentation. In Smart Innovations in Communication and Computational Sciences, pp. 117-126. Springer, Singapore 2019

[16]Bhakat S, and Periannan S: Brain Tumor Detection Using Cuckoo Search Algorithm and Histogram Thresholding for MR Images. In Smart Innovations in Communication and Computational Sciences, Springer, Singapore 85-95, 2019

[17]Shivhare SN, Sharma S, and Singh N: An Efficient Brain Tumor Detection and Segmentation in MRI Using Parameter-Free Clustering. In Machine Intelligence and Signal Analysis, Springer, Singapore 485-495, 2019

[18]Srinivas B, and Rao GS: Performance Evaluation of Fuzzy C Means Segmentation and Support Vector Machine Classification for MRI Brain Tumor. In Soft Computing for Problem

Solving, pp. 355-367. Springer, Singapore 2019.

[19] He K, Sun J, Tang X: Guided image filtering. IEEE transactions on pattern analysis and machine intelligence. 35(6):1397-409, 2013 Jun

[20] Wang GG, Deb S, and Coelho LDS: Elephant herding optimization. In 2015 3rd International Symposium on Computational and Business Intelligence (ISCBI) IEEE 1-5, 2015, December

[21] Zhang J, Zhou Y, Luo Q: Nature-inspired approach: a wind-driven water wave optimization algorithm. Applied Intelligence 49(1):233-52, 2019 Jan 15

[22] Harrabi R, Braiek EB: Isotropic and anisotropic filtering techniques for image denoising: A comparative study with classification. In2012 16th IEEE Mediterranean Electro technical Conference IEEE. 370-374, 2012 Mar 25

[23] Shally HR, and Chitharanjan K: Tumor volume calculation of brain from MRI slices. International Journal of Computer Science & Engineering Technology (IJCSET) 4(8): 1126-1132, 2013

[24] Rajinikanth V, Satapathy SC, Fernandes SL, and Nachiappan S: Entropy based segmentation of tumor from brain MR images–a study with teaching learning based optimization. Pattern Recognition Letters 94: 87-95, 2017

[25] Rajinikanth V, and Satapathy SC: Segmentation of ischemic stroke lesion in brain MRI based on social group optimization and Fuzzy-Tsallis entropy. Arabian Journal for Science and Engineering 43(8): 4365-4378, 2018

[26] Mohsen H, El-Dahshan ESA, El-Horbaty ESM, and Salem ABM: Classification using deep learning neural networks for brain tumors. Future Computing and Informatics Journal 3(1): 68-71, 2018

[27] Hussain S, Anwar SM, Majid M: Segmentation of glioma tumors in brain using deep convolutional neural network. Neuro computing 282: 248-61, 2018 Mar 22