

A Biologically Inspired ELM-based Framework for Classification of Brain MRIs

<https://doi.org/10.3991/ijoe.v16i10.15653>

Sarada Prasanna Pati, Debahuti Mishra
Siksha 'O' Anusandhan Deemed to be University, Odisha, India

Samarjeet Borah ^(✉)
Sikkim Manipal University, Sikkim, India
samarjeetborah@gmail.com

Abstract—Use of medical images for clinical analysis of various critical diseases have become increasingly predominant in modern health care systems. Application of machine learning technique in this context evolves as a potential solution in terms providing faster output with high diagnostic accuracy. In this work, we propose an Extreme Learning Machine (ELM) based classifier SFLA-ELM for detection of normal and pathological brain condition from brain Magnetic Resonance Images (MRIs). ELM is known for its speed and accuracy whereas the proposed method uses a swarm based evolutionary technique Shuffled Frog Leaping Algorithm (SFLA) and 10-fold cross validation method to optimally determine the network parameter of the ELM for better classification performance. The proposed model is experimented on three different brain MRI datasets of three different brain diseases. To get better approximation accuracy and generalization ability for the base ELM classifier, the suitable activation function and the appropriate number of hidden layer nodes are chosen. The performance validation of the proposed framework is done under two different network conditions, i.e. fixed network structure and varying network structure, by comparing its performance with two standard hybridized ELM classifiers, namely, PSO-ELM and ABC-ELM. The comparative performance analysis suggests that the proposed SFLA-ELM gives better classification performance in diagnosing the diseases in terms of accuracy, sensitivity, specificity, F-score and Area under ROC curve (AUC). Furthermore, the SFLA-ELM also found to offer better generalization ability and better stability with more compact network structure.

Keywords—Classification, Brain Image, Extreme Learning Machine, ELM, Shuffled Frog Leaping Algorithm, PSO-ELM, ABC-ELM, SFLA-ELM.

1 Introduction

Detection of normal and diseased brain conditions from Magnetic Resonance imaging (MRI) with high precision and accuracy has been a challenge for the health care professionals. High quality brain MRIs represent a large set of detailed information about the brain tissue anatomy and the condition of the patient's brain. In spite of the

knowledge and experience, the human vision system restricts the manual interpretation and analysis of MR images for the clinical experts. This is due to the large volume of information contained in an MR image which is hard to interpret by human vision [1-2]. This is the reason why use of automated image analysis methods utilizing machine learning and image processing techniques are of wide use in recent years in the field of MR image processing. These computers assisted diagnosis (CAD) techniques not only reduce burden on the radiologist and neurologists but also improve the accuracy and objectivity of diagnosis [3-7].

Extreme Learning Machine (ELM) [8-13], anticipated by Huang et al. [13], is an efficient and effective learning algorithm for Single Layer Feed Forward Networks (SLFNs), which is being used extensively for classification task. The input weights as well as the biases of the hidden layer neurons are selected randomly in ELM, from which the output weights are calculated mathematically using a simple inverse operation of the hidden layer output matrix. As compared to other traditional neural networks used for classification, ELM exhibits higher generalization ability, extremely faster training speed and smallest norm of weights [8]. ELM resolves several issues such as; local minima, stopping criterion, proper learning rate and the number of iterations in comparison to traditional neural networks trained using gradient descent methods. Literature study reveals that ELM shows better generalization and faster learning ability compared to contemporary classification techniques like Support Vector Machines (SVM) and deep learning [14].

However, ELM faces the challenge with respect to selection of input factors, e.g., hidden biases and the input weights, which are not optimal [15-16]. In ELM, while calculating the output weights, it tries to minimize the training error and finds the minimum norm of output weights. Because of the random selection of the input weights and hidden biases, sometimes the output matrix may not exhibit full rank leading to ill-conditioning of the network and may result in non-optimal solutions [17]. Hence, random selection of these parameters can ultimately affect the performance of ELM. Therefore, optimal selections of these parameters are crucial for the best performance of ELM and this implicates a complex optimization problem [18]. Research on enhancing the performance of ELM establish that evolutionary algorithms such as DE [19-20], PSO [21-22] and ABC [23] can be used to optimally determine the input weights and hidden biases of the SLFN by exploring the potential areas of the solution space. Use of such optimization techniques can also increase the generalization ability and stability of ELM with more compact network.

Shuffled Frog Leaping Algorithm (SFLA) is developed by Eusuff et al. [24], is a nature inspired meta-heuristic, based on population. SFLA comes under the class of memetic algorithms as it mimics the memetic evolution adapted by frog population while searching for location of good food sources available on discrete stones in a swamp. SFLA has been applied effectively to resolve several miscellaneous optimization problems in engineering [25-31]. This work makes use of SFLA to train ELM by optimally determining the network parameters.

Selecting an appropriate activation function is of great significance in ELM as it can affect the classification performance like learning accuracy and convergence speed of ELM. Literature on ELM reveal that ELM may produce different approximation

accuracy and generalization ability for same problem or training data using different activation functions [32] [33-34]. Activation functions are somehow dependent on the dataset at hand and there doesn't exist any provable relationship between these two. In essence, the suitable activation function can be chosen optimally either by trial or by tuning [34]. The network structure of the ELM, defined in terms of number of hidden layer nodes, also largely influences the classification performance of ELM classifier [32]. In this work, the classification performance of ELM has been evaluated using ten different commonly used activation functions at varying number of hidden layer nodes. Based on the accuracy rate obtained, the suitable activation function and the number of hidden layer nodes are chosen for the ELM classifier.

1.1 Aim of the work

The aim of this work is to design an efficient classifier model that can distinguish the brain MRI samples as normal and pathological, with more accuracy and effectively ranks both the positive and negative instances with better success rate. In this work, a hybrid ELM based classifier, called SFLA-ELM has been proposed for classifying brain MRI data to normal and pathological brain. SFLA-ELM applies SFLA for enhancing the performance of basic ELM classifier by optimally determining the input weight and hidden bias of the SLFN. The proposed model is experimented on three different brain MRI datasets of three different brain diseases. To get better approximation accuracy and generalization ability for the base ELM classifier, the suitable activation function and the appropriate number of hidden layer nodes are chosen based on evaluating the classification accuracy of ELM on ten standard activation functions. The classification performance of the SFLA-ELM is validated by comparing its performance with two standard hybridized ELM classifiers, namely, PSO-ELM and ABC-ELM. The performance validation is done under two different network conditions, i.e. fixed network structure and varying network structure. 10-fold cross validation is used and average estimation is considered in all the experiments in this work to reduce the variance in performance estimation.

1.2 Organization of the paper

The rest of the paper is arranged as follows – Section 2 defines the methodologies adopted for experimentation; experimental design has been discussed in Section 3; the experimentation and result analysis has been given in Section 4. Section 5 discusses the findings of this study and finally Section 6 concludes the work highlighting future scope.

2 Methodologies Adopted for the Experiment

2.1 Extreme Learning Machine (ELM)

ELM is a very fast, simple and efficient learning algorithm for training the SLFNs. Unlike that of in traditional learning methods used to train SLFNs, in ELM the output weight matrix of the network is calculated mathematically in a non-iterative way from the randomly chosen input weights and hidden biases using the Moore–Penrose generalized inverse [8]. In ELM, the training is carried out in batch mode by presenting all the training data to the SLFN before computing the weights in a single iteration. ELM exhibits several striking features that makes it superior compared to other learning methods for SLFN such as; extremely fast learning speed with better learning performance, better generalization ability, free from getting stuck in local minima, no network tuning, no control parameter setup, simple and fixed network structure and ease of implementation [14]. While the gradient-based learning algorithms strive to reach minimum training error without considering the magnitude of weights, ELM tries to reach smallest training error with smallest norm of weights. The learning steps of the ELM algorithm can be summarized in the following three steps [8]. Given a training set $N = \{(X_i, t_i) \mid X_i \in R^d, t_i \in R, i = 1, 2, \dots, n\}$, hidden node number N and activation function g :

Step 1: Assign the input weight w_i and bias b_i ($i = 1, 2, \dots, N$) randomly.

Step 2: Calculate the hidden layer output matrix H using the equation (1).

$$H = \begin{pmatrix} g(W_1 \cdot X_1 + b_1)g(W_2 \cdot X_1 + b_2) \cdots g(W_N \cdot X_1 + b_N) \\ g(W_1 \cdot X_2 + b_1)g(W_2 \cdot X_2 + b_2) \cdots g(W_N \cdot X_2 + b_N) \\ \vdots \cdots \vdots \\ g(W_1 \cdot X_n + b_1)g(W_2 \cdot X_n + b_2) \cdots g(W_N \cdot X_n + b_N) \end{pmatrix}_{n \times N} \quad (1)$$

Step 3: Calculate the output weight β , where $\beta = H^\dagger T$, here H^\dagger is the Moore–Penrose (MP) generalized inverse of H and $T = (t_1, t_2, t_3, \dots, t_n)^T$.

2.2 Shuffled Frog Leaping Algorithm (SFLA)

SFLA [24] is a population based meta-heuristic for solving hard combinatorial optimization problems. SFLA comes under the class of memetic algorithms as it mimics the memetic evolution adapted by frog population while searching for location of food sources available on discrete stones in a swamp. The algorithm comprises of three major steps namely partitioning, local search and shuffling, apart from initialization of population. While the local searches enable evolution of intellect possessed by frogs (meme) within each partition independently, shuffling allows global interchange of intellect among frogs in the entire population. The process of partitioning, local exploration and global shuffling continue until the optimal solution is achieved or the stopping criteria are met. The detail explanation on SFLA can be found in [24-25]

2.3 Working Principle of Proposed SFLA-ELM Classifier

In this section, a novel learning framework for SLFNs, called SFLA-ELM has been proposed. This method uses SFLA and the k -fold cross validation scheme to find the optimal input weights and hidden biases. Further, Moore-Penrose generalized inverse is applied to mathematically calculate the output weights. Fitness function used by SFLA for optimization is the minimization of the *mean square error* (MSE). The optimization problem thus, is based on minimizing the function stated in equation (5).

$$\varphi = E_{MSE}(y_k, \hat{y}_k), E_{MSE}(y_k, \hat{y}_k) = \frac{1}{N} \sum_{k=1}^N (y_k - \hat{y}_k)^2 \quad (5)$$

Here, N = Number of samples, y_k = Desired output and \hat{y}_k = Obtained output
The algorithm of SFLA-ELM for classification of MRI dataset is as follows:

Algorithm: Brain Image Classification

Input: Brain image dataset; Population size (P); Size of hidden layer (H_c); Number of memplexes (m); number of folds for n -fold cross validation (n).

Output: Classification accuracy

Description: Global best (X_g); Local best (X_b); Local worst (X_w).

Step 1 Divide the dataset to *train_set* (*train_input*, *train_output*) and *test_set* (*test_input*, *test_output*) as per n - fold cross validation

Step 2 For $j = 1$ to n

j^{th} set is used as *test_set* and $n - j$ sets are as *train_set*

Step 3 Set K random weight population; each of size $1 \times N_c$

$P_i = \{V_1^i, V_2^i, V_3^i, \dots, V_{N_c}^i\}$ for $i = 1, 2, 3, \dots, K$

Step 4 For each population P_i find the error value in ELM using Step 5

Step 5 For each *train_input* find

$H_i = \text{train}_{input_i} \times P_i$

$\beta_i = \text{pseudo}_{inverser}(H) \times \text{train}_{output}$

$obt_{output} = (\text{test}_{input} \times P_i) \times \beta_i$

$O_i = E_{MSE}(obt_{output}, \text{test}_{output})$

Step 6 Sort each population in ascending order with respect to maximum fitness considering minimum MSE.

Step 7 Find X_g , where O_i is with minimum value.

Step 8 Partition the populations into m number of memplexes.

The distribution of the sorted population to m number of memplexes is done in such a way, that, the first population is assigned to first memplex and second population is assigned to second memplex and the same process is continued till the m^{th} memplex. Then the $m+1$ population will be assigned to first memplex and so on till all population are distributed.

Step 9 Within each memplex, generate X_b and X_w through the objective function repeating Step 5 with respect to fitness value.

Step 10 Update position of X_w with respect to X_b using the eq. (2) and eq. (3)

Step 11 If Fitness of $X_w(\text{new})$ is better than X_w , then,

- Replace X_w with $X_w(new)$
Else
Update X_w with respect to X_g using equation (4) and equation (3)
a. If
Fitness of $X_w(new)$ is better than X_w , then replace X_w with $X_w(new)$
Else
Generate a new position of X_w randomly
- Step 12** Merge the memplexes and find the fitness value of the new population using Step 5
- Step 13** Repeat Step 6 to Step 12 until the termination condition is satisfied
- Step 14** X_g is considered as final weight and accordingly the $\beta_{global_{best}}$ is generated and used for the classification of test data.
- Step 15** Repeat the process for all the value of j , find the classification accuracy at each fold to generate the average accuracy.

3 Experiment Design

Three ELM based classification frameworks, namely, SFLA-ELM, ABC-ELM and PSO-ELM were implemented for classification of the brain MRI datasets. The classification performance of proposed SFLA-ELM is evaluated, compared and validated through experimentation results. The validation is done under two different network conditions of ELM.

- a) *Fixed network structure*: Keeping the network structure of the ELM classifier fixed, after determining the suitable activation function and the number of hidden layer nodes [32] [35]. The activation function and the number of hidden layer nodes are finalised after experimentally analysing the classification performance of 10 commonly used activation functions with respect to varying number of nodes.
- b) *Varying network structure*: Varying the number of hidden layer nodes while using the same activation function for ELM as determined in Phase I.

3.1 Experimental setup

All experiments have been conducted using MATLAB (R2016) on a personal computer with 3.30 GHz Core-i5 processor having 4GB RAM running under Windows 10 operating system. Throughout the experiments, all data were scaled within the range of [0, 1]. The proposed method for brain image classification employs ELM as the base classifier. In all experiments, 10-fold cross validation were used for training the ELM to ensure bias free and reliable performance estimates by nullifying the influence of training and testing set on the classification outcomes. In addition, for each fold, 20 trials were carried out in order to nullify the effect of random inputs to the network and the average values of the results are considered. For the optimization techniques used

in this work applying SFLA, ABC and PSO, the maximum number of iterations considered was 100 and the population size taken was 20.

3.2 Datasets used for experimentation

The proposed classifier model is experimented on three brain image datasets collected from [36], supported and maintained by *Medical School of Harvard University*. These three datasets contain T2-weighted brain MR images in *gif* format with size 256×256 . All brain images are in *transaxial plane*. For our purpose we have collected three different brain image datasets for three different brain diseases, namely, *Alzheimer*, *Glioma* and *Multiple Sclerosis*. Feature extraction is the first step to represent the image in its compact and high level form to prepare the image dataset for classification. The datasets used in this work are prepared by applying Discrete Wavelet Transformation (DWT) using bi-orthogonal wavelet function for image decomposition and feature extraction from the MRI images. The images are decomposed by using bi-orthogonal 1.3 wavelet function up to level 3 and thereby reducing the image size to 32×32 . The detail description of datasets is shown in **Table 1**. Each of these datasets comprises of both normal as well as diseased instances for brain images pertaining to that particular disease.

Table 1. Description of Image Dataset

Type of Brain Image	Num. of Images	Image Size
Alzheimer	101	256×256
Glioma	123	256×256
Multiple Sclerosis	127	256×256

4 Experimentation, Result Analysis and Validation

This section presents the details of the three separate sets of experiments that were conducted, the analysis of the results obtained and validation of the proposed classification method. The first set of experiments were conducted with an objective to choose the suitable activation function and optimally decide the number of nodes for the base ELM classifier for each dataset. In second set of experiments the classification of the three MRI datasets was carried out using SFLA-ELM

4.1 Experiment I: Determining the activation function and network structure

The network structure of the SLFN, i.e. the activation function and the number of hidden layer nodes, has large influence on the learning efficiency and classification accuracy of ELM. In order to decide the most suitable activation function and the optimal number of hidden nodes of the base classifier ELM, a comprehensive study considering 10 commonly used activation function on ELM has been performed. The activation functions used are *sigmoid*, *tanh*, *sine*, *bipolar*, *hardlimit*, *Gaussian*, *tribas*, *radbas*, *multi-quadratic* and *inverse multi-quadratic*. Series of simulations were carried

out for each dataset to analyze the performance of ELM at different number of nodes by considering the activation functions and 10-fold cross validation is used for bias free assessment. In addition, for each fold, 20 trials are taken with the same number of nodes and activation function on a dataset in order to nullify the effect of random inputs to the network. Finally, the average values of the result are considered for performance measure.

Table 2 shows the highest average training accuracy achieved for each activation function. The training accuracies obtained at different number of hidden layer nodes in these experiments are represented in **Fig 1 to Fig 3**. All experiments are conducted by varying the number of nodes with an increment of 5 nodes, up to 60 nodes. That activation function is chosen as the most suitable activation function which shows best accuracy at lesser number of nodes compared to other activation functions tested here. By inspecting the accuracy results obtained using different activation function it is observed that the *Multi-quadratic* activation function shows better classification accuracies for all the three datasets. In fact, we limit the maximum number of nodes to 60 for our comparative study as in all three datasets it is found that by 60 nodes the training accuracies obtained reaches the maximum.

Table 2. Recognition of Performance of Activation Functions for the Three Datasets

Activation functions	Alzheimer		Glioma		Multi Sclerosis	
	Training Accuracy	No. of Nodes	Training Accuracy	No. of Nodes	Training Accuracy	No. of Nodes
Sine	0.918	60	0.908	60	0.918	60
Sigmoid	0.99	60	0.996	60	0.975	60
Hardlim	0.99	60	0.996	60	0.974	60
Tribas	0.5	5	0.503	60	0.5	05
Radbas	0.515	25	0.518	60	0.512	60
Tanh	0.993	60	0.995	60	0.972	60
Gaussian	0.507	35	0.517	60	0.523	60
Inv. Multi-quadratic	0.979	60	0.979	60	0.942	60
Bipolar	0.989	60	0.995	60	0.979	60
Multi-quadratic	0.999	55	1	40	0.999	55

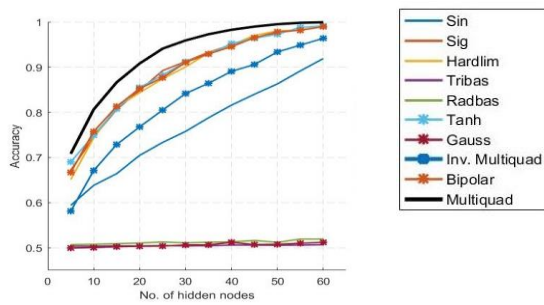


Fig. 1. Training accuracy for activation functions in *Alzheimer* dataset

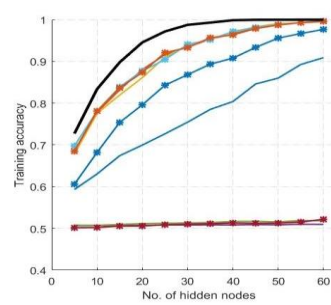


Fig. 2. Training accuracy for activation functions in *Glioma* dataset

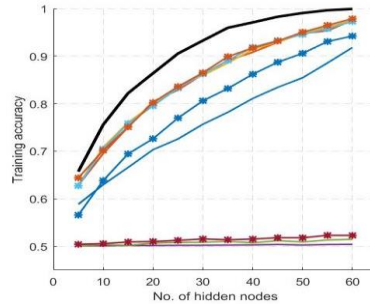


Fig. 3. Training accuracy for activation functions in Multiple Sclerosis dataset

After choosing the activation function with best performance, the number of nodes that led to this best performance had been considered as the optimal number of nodes for the SLFN. The network structure that we choose for the ELM classifier which gives best classification performance in all three datasets is given in **Table 3**.

Table 3. The Network Structure of ELM with best Classification Performance

Dataset	Activation Function	No. of nodes
Alzheimer	Multi-quadratic	55
Glioma	Multi-quadratic	40
Multi Sclerosis	Multi-quadratic	55

4.2 Experiment-II: Classification and performance analysis at fixed network structure

In order to study the classification performance and effectiveness of the SFLA-ELM method, it was applied on three MRI datasets, namely, *Alzheimer*, *Glioma* and *Multiple Sclerosis*. Two of the established hybrid ELM models such as; ABC-ELM and PSO-ELM were also implemented for classification of the MRI datasets to have a comparative study of the classification performances and validate the effectiveness of the model.

Table 4. Performance Comparison of Proposed Model Based on Accuracy, Sensitivity, Specificity

Methods	Alzheimer			Glioma			Multi Sclerosis		
	Acc	Sen	Spe	Acc	Sen	Spe	Acc	Sen	Spe
SFLA-ELM	0.979	1.000	0.917	0.997	0.944	1.000	0.988	0.968	1.000
PSO-ELM	0.957	0.963	0.887	0.947	0.867	0.985	0.976	0.988	0.833
ABC-ELM	0.910	0.933	0.853	0.895	0.866	0.833	0.908	0.917	0.831

*Acc: Accuracy, Sen: Sensitivity, Spe: Specificity

For the classification of each of the dataset, the network structure for the ELM and the activation function used in the experiments for all three methods were taken as per **Table 3**. Classification accuracy and Receiver Operating Characteristics (ROC) curve are the two important measures considered primarily to evaluate the performance of the classifiers. Moreover, other performance measures like sensitivity, specificity and AUC

were evaluated in order to compare the effectiveness of the classifier. The average classification accuracy, sensitivity, specificity values obtained for SFLA-ELM, PSO-ELM and ABC-ELM in case of all the three datasets is presented in **Table 4**. SFLA-ELM is found to achieve highest accuracy of 0.947, 0.944 and 0.988 for *Alzheimer*, *Glioma* and *Multiple Sclerosis* dataset respectively. it is evident that SFLA-ELM better scores for specificity and sensitivity as compared to PSO-ELM and ABC-ELM in case of all the three datasets except the case of *Multiple Sclerosis* dataset, where PSO-ELM scores better than SFLA-ELM in terms of the sensitivity value.

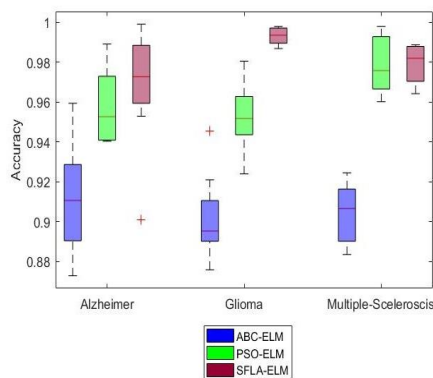


Fig. 4. Boxplots Showing the Variation in Classification Accuracy Values

Figure 4 shows the box-plots related to the distribution of the classification accuracy values. Each box in the plot corresponds to the accuracy distribution range obtained over the 10-fold cross validation performed for classification using the algorithm. As we can figure out, SFLA-ELM generates more compact boxes in all three datasets compared to PSO-ELM and ABC-ELM, and this confirms that SFLA-ELM gives better stability in terms of classification results compared to the other two methods.

Figure 5 to **Figure 7** present the ROC curves for the classification methods on the three datasets. A close observation of the ROC curves establishes that, SFLA-ELM outperforms PSO-ELM and ABC-ELM in case of all the three datasets.

Figure 8 presents the comparison of the AUC values obtained for the three methods. The results clearly show that SFLA-ELM exhibits better classification performance compared to both PSO-ELM and ABC-ELM in all the three datasets. AUC values obtained for SFLA-ELM are 0.97, 0.94 and 0.95 for *Alzheimer*, *Glioma* and *Multiple Sclerosis* datasets respectively.

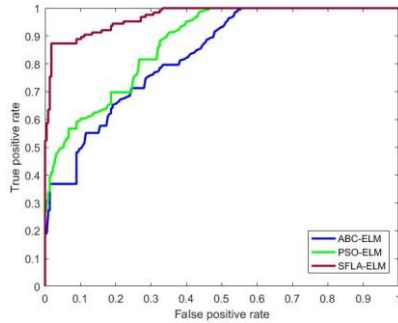


Fig. 5. ROC curve for *Alzheimer*

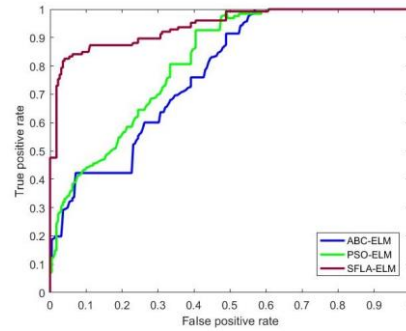


Fig. 6. ROC curve for *Glioma*

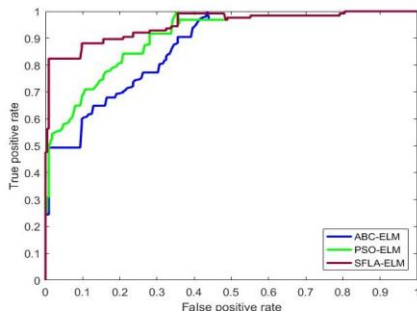


Fig. 7. ROC curve for *Multiple Sclerosis*

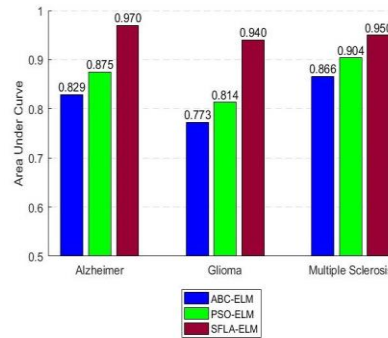


Fig. 8. Comparison plot of AUC values for the three datasets

4.3 Experiment III: Classification and Performance Analysis at Varying Network Structure

Experiments were conducted to investigate the influence of varying the number of nodes of the SLFN on the classification performance of ELM, PSO-ELM, ABC-ELM and SFLA-ELM, with an objective to compare and determine the most compact network that yields best classification accuracy in each of the datasets. The numbers of nodes are varied from 5 to 80, with an increment of 5 nodes. 10-fold cross-validation is used and for each fold 20 trials were carried out. The plots showing the variation in classification accuracy with respect to increase in number of hidden nodes are given in **Figure 9**.

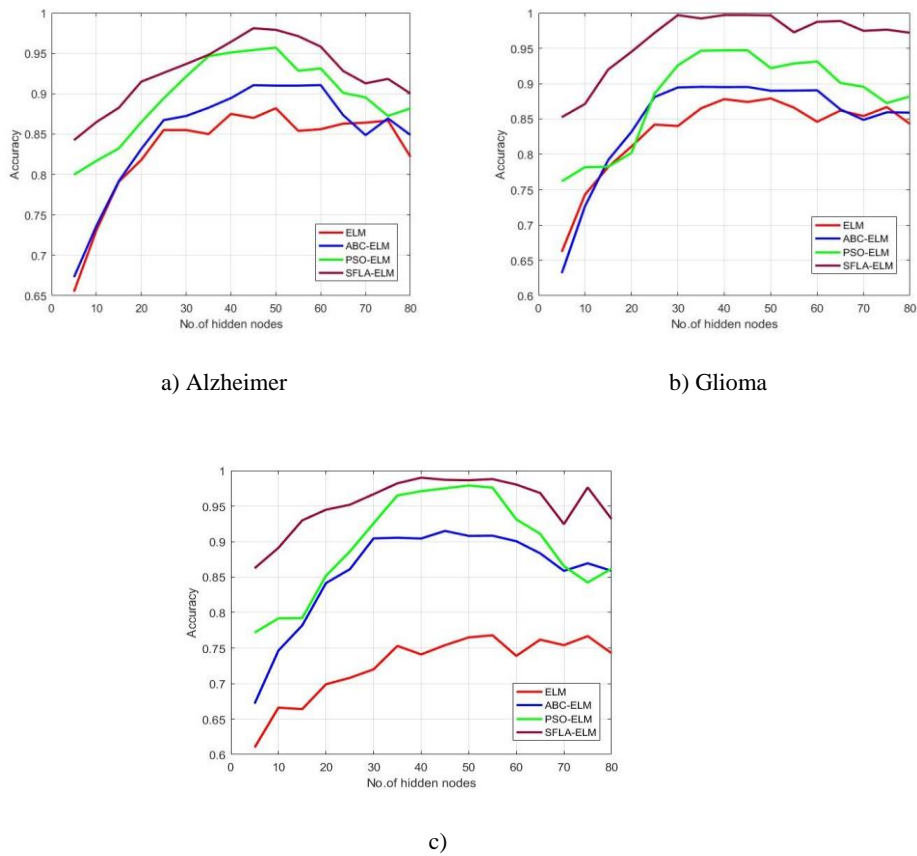


Fig. 9. Classification Accuracy Vs. Number of Nodes for the three Datasets

Table 5 reports the highest average testing accuracy reached by the three methods and the number of nodes at which the highest accuracy is reached. Examining the results in **Table 5**, it is evident that SFLA-ELM achieves highest accuracy at minimum number of nodes as compared to ELM, PSO-ELM and SFLA-ELM in case of all three datasets. This experimental result establishes that SFLA-ELM shows better classification accuracy with a relatively compact and robust network.

Table 5. Highest Accuracy Achieved while Varying the Number of Nodes

Datasets	Classifier	Highest Accuracy	No. of Nodes
Alzheimer	SFLA-ELM	0.991	40
	PSO-ELM	0.973	40
	ABC-ELM	0.912	45
	ELM	0.882	50
Glioma	SFLA-ELM	0.998	30
	PSO-ELM	0.948	35
	ABC-ELM	0.885	30
	ELM	0.871	40
Multi Sclerosis	SFLA-ELM	0.99	40
	PSO-ELM	0.972	50
	ABC-ELM	0.918	45
	ELM	0.768	55

5 Findings

The analysis of the of the experiments results suggested that the proposed SFLA-ELM classifier exhibits better overall performance in successfully discriminating the diseased and normal brain MRI samples with better accuracy in comparison to the existing hybridized ELM methods, PSO-ELM and ABC-ELM. Some of the important findings of these experimental studies are as follows.

- a) The proposed SFLA-ELM method achieves better accuracy in all three MRI datasets compared to the other two hybrid ELM based methods, i.e. ABC-ELM and PSO-ELM.
- b) Analysis of the box plot results for accuracy further confirms that SFLA-ELM gives better stability in terms of classification results compared to the other two methods.
- c) The ROC curve analysis of the three hybrid methods shows that the curve for SFLA-ELM is superior to ABC-ELM and PSO-ELM. The higher AUC values obtained for SFLA-ELM in case of all three datasets further prove that SFLA-ELM is more successful in ranking both the positive and negative instances.
- d) SFLA-ELM achieves better sensitivity and specificity than the other two hybrid methods and thereby making it evident that, the proposed SFLA-ELM method is more appropriate for diagnosis of brain diseases from MRI data.
- e) While evaluating the performance of the three-classification method over varying number of hidden layer nodes, SFLA-ELM achieves highest accuracy at minimum number of nodes as compared to ELM, PSO-ELM and SFLA-ELM in case of all three datasets. This experimental result establishes that SFLA-ELM shows better classification accuracy with a relatively compact network.

6 Conclusion and Future Scope

This paper presents a novel hybrid classification method SFLA-ELM to classify MR brain image data for diagnosis of pathological brain conditions. SFLA-ELM applies SFLA for enhancing the performance of basic ELM classifier by optimally determining the input weight and hidden bias of the SLFN. The performance validation was done under both fixed network structure as well as varying network structure of ELM. Experimental results reveal that the SFLA-ELM method outperforms the other two hybrid methods PSO-ELM and ABC-ELM and achieves better performance in terms of various performance measures like accuracy, ROC, AUC, sensitivity, specificity. Furthermore, SFLA-ELM is found to be more successful in ranking both the positive and negative instances by achieving higher AUC values and also shows better classification accuracy with a relatively compact network with better stability and thereby making it evident that, the proposed SFLA-ELM method can be more appropriate for diagnosis of brain diseases from MRI data. Future work could focus on implementing this model for multi-class classification of brain MR images and also on other type medical images for diagnosis of other diseases. Larger image dataset can also be used to enhance the training process of the classifier.

7 References

- [1] A. Ortiz, J.M. Gorriz, J. Ramirez, D. Salas-Gonzalez, J.M. Llamas-Elvira, Two fully-unsupervised methods for MR brain image segmentation using SOM-based strategies, *Appl. Soft Comput. J.*, 13 (2013) 2668–2682. <https://doi.org/10.1016/j.asoc.2012.11.020>
- [2] D. Jude Hemanth, C. KeziSelva Vijila, A. Immanuel Selvakumar, J. Anitha, Performance Improved Iteration-Free Artificial Neural Networks for Abnormal Magnetic Resonance Brain Image Classification, *Neurocomputing*, 130 (2014) 98–107. <https://doi.org/10.1016/j.neucom.2011.12.066>
- [3] Geethu Mohan, M. Monica Subashini, MRI based medical image analysis: Survey on brain tumor grade classification, *Biomedical Signal Processing and Control*, 39 (2018) 139–161. <https://doi.org/10.1016/j.bspc.2017.07.007>
- [4] Amr S. Mady, Samir Abou El-Seoud, An Overview of Volume Rendering Techniques for Medical Imaging, *iJOE – Vol. 16, No. 6, 2020*, pp. 95-106, <https://doi.org/10.3991/ijoe.v16i06.13627>
- [5] Ahmed Elnakib, Hanan M. Amer, Fatma E.Z. Abou-Chadi, Early Lung Cancer Detection using Deep Learning Optimization, *iJOE – Vol. 16, No. 6, 2020*, pp. 82-94, <https://doi.org/10.3991/ijoe.v16i06.13657>
- [6] Hewawasam Puwakpitiyage C. A., Paramesura Rao V. R., Muhammad Azizi M. S. A., Tee W. J., Murugesan R. K., Hamzah M. D., A Proposed Web Based Real Time Brain Computer Interface (BCI) System for Usability Testing, *iJOE – Vol. 15, No. 8, 2019*, pp. 108-119, <https://doi.org/10.3991/ijoe.v15i08.10406>
- [7] Mohammad Subhi Al-Batah, Belal Mohammad Zaqibeh, Saleh Ali Alomari, Mowafaq Saleem Alzboon, Gene Microarray Cancer Classification using Correlation Based Feature Selection Algorithm and Rules Classifiers, *iJOE – Vol. 15, No. 8, 2019*, pp. 62-73, <https://doi.org/10.3991/ijoe.v15i08.10617>

- [8] G. Huang, Y. Qin, C.K. Siew, Extreme learning machine: theory and applications, *Neurocomputing*, 70(1) (2006) 489–501. <https://doi.org/10.1016/j.neucom.2005.12.126>
- [9] G. Huang, L. Chen, C. K. Siew, Universal approximation using incremental constructive feed forward networks with random hidden nodes, *IEEE Trans. Neural Networks*, 17(4) (2006)870–892. <https://doi.org/10.1109/TNN.2006.875977>
- [10] G. Huang, L. Chen, Convex Incremental extreme learning machine, *Neurocomputing*, 70(16) (2007) 3056–3062. <https://doi.org/10.1016/j.neucom.2007.02.009>
- [11] G. Huang, L.Chen, Enhanced random search based incremental extreme learning machine, *Neurocomputing*, 71(16) (2008) 3460–3468. <https://doi.org/10.1016/j.neucom.2007.10.008>
- [12] G. Huang, L.Chen, Optimization method based extreme learning machine for classification, *Neurocomputing*, 74(1) (2010) 155–163. <https://doi.org/10.1016/j.neucom.2010.02.019>
- [13] G. Huang, Q. Zhu, C.K. Siew, Extreme learning machine: a new learning scheme of feed forward neural networks, *Proceedings of the IEEE International Joint Conference on Neural Networks*, 2004.
- [14] Gao Huang, Guang-Bin Huang, Shiji Song, KeyouYoua, Trends in extreme learning machines: A review, *Neural Networks*, 61 (2015) 32–48. <https://doi.org/10.1016/j.neunet.2014.10.001>
- [15] Bin Li, Yibin Li, Xuewen Rong, The extreme learning machine learning algorithm with tunable activation function, *Neural Comput & Applic*, (2013) 22:531–539. <https://doi.org/10.1007/s00521-012-0858-9>
- [16] Yongshan Zhang, Jia Wu, Zhihua Cai, Peng Zhang, Ling Chen, Memetic Extreme Learning Machine, *Pattern Recognition*, 58(2016)135–148. <https://doi.org/10.1016/j.patcog.2016.04.003>
- [17] P. Mohapatra, S. Chakravarty, P. K. Dash, an improved cuckoo search based extreme learning machine for medical data classification, *Swarm and Evolutionary Computation*, 24 (2015) 25–49. <https://doi.org/10.1016/j.swevo.2015.05.003>
- [18] S. Suresh, S. Saraswathi, N. Sundararajan, Performance enhancement of extreme learning machine for multi category sparse data classification problems, *Engineering Applications of Artificial Intelligence*, 23 (2010) 1149–1157. <https://doi.org/10.1016/j.engappai.2010.06.009>
- [19] J. Sánchez-Monedero, C. Hervás-Martínez, P.Gutiérrez, M. C. Ruz, M. R. Moreno, M. Cruz-Ramírez, Evaluating the performance of evolutionary extreme learning machines by a combination of sensitivity and accuracy measures, *Neural Network World*, 20 (7) (2010) 899.
- [20] Q.-Y. Zhu, A. K. Qin, P. N. Suganthan, G.-B. Huang, Evolutionary extreme learning machine, *Pattern Recognition*, 38 (10) (2005) 1759–1763. <https://doi.org/10.1016/j.patcog.2005.03.028>
- [21] F. Han, H.-F. Yao, Q.-H. Ling, an improved evolutionary extreme learning machine based on particle swarm optimization, *Neurocomputing*, 116 (2013) 87–93. <https://doi.org/10.1016/j.neucom.2011.12.062>
- [22] Y. Xu, Y. Shu, Evolutionary extreme learning machine-based on particle swarm optimization, *Proc. of the 2006 international symposium on neural networks*, 2006, pp. 644–652. https://doi.org/10.1007/11759966_95
- [23] Chao Ma, An Efficient Optimization Method for Extreme Learning Machine Using Artificial Bee Colony, *Journal of Digital Information Management*, 15(3) (2017).
- [24] M. Eusuff, K. Lansey, Optimization of water distribution network design using the shuffled frog leaping algorithm, *J Water Resour Plann Manage*, 129(3) (2003) 210–25. [https://doi.org/10.1061/\(ASCE\)0733-9496\(2003\)129:3\(210\)](https://doi.org/10.1061/(ASCE)0733-9496(2003)129:3(210))

- [25] Muzaffar Eusuff, Kevin Lansey, Fayzul Pasha, Shuffled frog-leaping algorithm: a memetic meta-heuristic for discrete optimization, *Engineering Optimization*, 38(2) (2006) 129–154. <https://doi.org/10.1080/03052150500384759>
- [26] Parmeet Kaur, Shikha Mehta, Resource provisioning and work flow scheduling in clouds using augmented Shuffled Frog Leaping Algorithm, *J. Parallel Distrib. Comput.*, 101 (2017) 41–50. <https://doi.org/10.1016/j.jpdc.2016.11.003>
- [27] Sunita Panda, Archana Sarangi, Siba Prasada Panigrahi, A new training strategy for neural network using shuffled frog-leaping algorithm and application to channel equalization, *Int. J. Electron. Commun.*, 68 (2014) 1031–1036. <https://doi.org/10.1016/j.aeue.2014.05.005>
- [28] Xinmin Cheng, Xiaodan Zhang, Li Zhao, Aideng Deng, Yongqiang Bao, Yong Liu, Yunliang Jiang, The application of shuffled frog leaping algorithm to wavelet neural networks for acoustic emission source location, *C. R. Mecanique*, 342 (2014) 229–233. <https://doi.org/10.1016/j.crme.2013.12.006>
- [29] Jianping Luo, Xia Li, Min-Rong Chen, Hongwei Liu, A novel hybrid shuffled frog leaping algorithm for vehicle routing problem with time windows, *Information Sciences*, 316 (2015) 266–292. <https://doi.org/10.1016/j.ins.2015.04.001>
- [30] Zhaokun Li, Jingtai Cao, Xiaohui Zhao, Wei Liu, Swarm intelligence for atmospheric compensation in free space optical communication—Modified shuffled frog leaping algorithm, *Optics & Laser Technology*, 66 (2015) 89–97. <https://doi.org/10.1016/j.optlastec.2014.08.012>
- [31] Jian-ping Luo, Xia Li, Min-rong Chen, Hybrid shuffled frog leaping algorithm for energy-efficient dynamic consolidation of virtual machines in cloud data centers, *Expert Systems with Applications*, 41 (2014) 5804–5816. <https://doi.org/10.1016/j.eswa.2014.03.039>
- [32] Jianfu Xia, Huiling Chen, Qiang Li, Minda Zhou, Limin Chen, ZhennaoCai, Yang Fang, Hong Zhou, Ultrasound-based differentiation of malignant and benign thyroid Nodules: An extreme learning machine approach, *Computer Methods and Programs in Biomedicine*, 147 (2017) 37–49. <https://doi.org/10.1016/j.cmpb.2017.06.005>
- [33] Bin Li, Yibin Li, Xuewen Rong, The extreme learning machine learning algorithm with tunable activation function, *Neural Comput & Applic*, 22 (2013) 531–539. <https://doi.org/10.1007/s00521-012-0858-9>
- [34] Ömer Faruk Ertuğrul, A novel type of activation function in artificial neural networks: Trained activation function, *Neural Networks*, 99 (2018) 148–157. <https://doi.org/10.1016/j.neunet.2018.01.007>
- [35] E. Malar, A. Kandaswamy, D. Chakravarthy, A. GiriDharan, A novel approach for detection and classification of mammographic microcalcifications using wavelet analysis and extreme learning machine, *Computers in Biology and Medicine*, 42 (2012) 898–905. <https://doi.org/10.1016/j.compbiomed.2012.07.001>
- [36] <http://www.med.harvard.edu/AANLIB>

8 Authors

Sarada Prasanna Pati has completed his Ph.D. in Computer Science & Engineering and works as an Associate Professor in the Computer Science & Engineering Department of Siksha ‘O’ Anusandhan Deemed to be University. His research area includes medical image analysis, machine learning and soft computing.

Debahuti Mishra is working as Professor and Head of the Computer Science & Engineering Department of Siksha ‘O’ Anusandhan Deemed to be University. She has guided 32 number of M. Tech. students and 9 number of PhD students. She has

published 165 number of research papers in reputed journals and conferences. Her research area includes Data Mining and soft computing techniques.

Samarjeet Borah is currently working as Professor in the Department of Computer Applications, Sikkim Manipal University (SMU), Sikkim, India. He Borah handles various academics, research and administrative activities such as curriculum development, BoS, DRC, IT Infrastructure Management etc. in SMU. Dr. Borah is involved with various funded projects from AICTE (Govt. of India), DST-CSRI (Govt. of India) etc. as PI/Co-PI. Dr. Borah is involved with various journals of repute and book volumes as Guest Editor, Editor and Series Editor. samarjeetborah@gmail.com

Article submitted 2020-05-20. Resubmitted 2020-06-21. Final acceptance 2020-06-25. Final version published as submitted by the authors.