International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-9 Issue-2, December, 2019

Exploring the Extreme Learning Machine for Classification of Brain MRIs

Pranati Satapathy, Sarbeswara Hota, Sateesh Kumar Pradhan

Abstract*: *Magnetic Resonance Imaging (MRI) technique of brain is the most important aspect of diagnosis of brain diseases. The manual analysis of MR images and identifying the brain diseases is tedious and error prone task for the radiologists and physicians. In this paper 2-Dimensional Discrete Wavelet Transformation (2D DWT) is used for feature extraction and Principal Component Analysis (PCA) is used for feature reduction. The three types of brain diseases i.e. Alzheimer, Glioma and Multiple Sclerosis are considered for this work. The Two Hidden layer Extreme learning Machine (TELM) is used for classification of samples into normal or pathological. The performance of the TELM is compared with basic ELM and the simulation results indicate that TELM outperformed the basic ELM method. Accuracy, Recall, Sensitivity and F-score are considered as the classification performance measures in this paper.

Keywords: Wavelet Transformation, Principal Component Analysis, Extreme Learning Machine, Magnetic Resonance Imaging.

I. INTRODUCTION

The human brain is prone to various diseases. The diagnosis of these disease require proper diagnosis from the scanning report of the radiologists through different imaging techniques [1]. Magnetic Resonance Imaging (MRI) provides a clear and proper neural tissue architecture of human brain. MRI is considered as one of the most important technique for the diagnosis of various brain disorders. This approach helps the physicians and technicians in identifying the brain abnormalities. But the manual process of MRI analysis is error prone and tedious. So the automated MRI analysis plays a vital role in medical community and research has been continued for the development of various software based image classification method using machine learning techniques [2]. The image classification process is broadly categorized as dataset collection, data preprocessing including feature extraction and feature reduction, and then classification [3-5]. From the literatures, it is found that neural network models are efficiently used for the brain image classification. The authors in [6] used K-NN and Artificial Neural Network (ANN) models to classify into normal and abnormal brain images. The authors applied 2-Dimensional Discrete Wavelet Transformation (DWT) and Principal Component Analysis

(PCA) to extract features from the images and to reduce the features respectively. Y. Zhang et al. [7] used the neural network based model for the classification of brain MR images. DWT and PCA were applied for the feature extraction and feature reduction tasks.

G. B. Huang *et al.* developed Extreme Learning Machine (ELM) that performed better than Multi-layer Perceptron [8-10].The ELM is associated with random weights and biases in the input layers. The output weights are mathematically determined using Moore-Penrose pseudo inverse. The authors in [11] used the ELM model for brain MRI classification. They proposed the modified Sine Cosine algorithm for the optimal determination of hidden layer parameters of ELM model. In [12], the authors used the ELM model for brain image classification. The randomness of the initial weights and hidden biases remain as one of the limitation of ELM. Various works have been conducted to improve the performances of ELM [13-15]. B. Que et *al.* in [16] proposed the two hidden layer ELM (TELM) model for the classification and regression problems. The algorithm and architecture of TELM are described in this work. The simulation study demonstrates that the proposed TELM model outperformed the basic ELM in different benchmark functions for the classification and regression task. Some authors have also used TELM and Multi hidden layer ELM for their works [17- 19]. The objective of this paper is to use the TELM model for the classification of brain images. In this paper, three brain MRI datasets i.e. Alzheimer, Glioma and Multiple sclerosis are used. The images are preprocessed with 2-D DWT and PCA, then TELM is used for classification into normal or pathological images. In this paper, section 2 describes the methodologies. The simulation study and the results analysis are discussed in section 3.Section 4 deals with the conclusion and future scope of this work.

II. METHODOLOGIES

This section discusses the working mechanisms of ELM and TELM models.

A. Extreme Learning Machine (ELM)

The literature study shows that Single Layer Feed forward Network (SLFN) has been used extensively in various applications [20]. But it has some limitations of expensive learning process due to the gradient based error and getting stuck in local optima. So the authors developed ELM as a relatively new kind of training method for SLFN [20].

The basic steps of ELM are as follows:

Let $M = \{(X_i, y_i) \mid \in R^d,\ y_i \in R, i = 1,2, \ldots, n\}$ be the dataset.

Revised Manuscript Received on December 30, 2019.

* Correspondence Author

Pranati Satapathy, Dept. of Computer Science and Applications, Utkal University, Bhubaneswar, India. satapathy.pranati@gmail.com

Sarbeswara Hota*, Computer Application Dept., Siksha O Anusandhan Deemed to be University, Bhubaneswar, India. sarbeswarahota@soa.ac.in.

Sateesh Kumar Pradhan, Computer Science and Applications, Utkal University, Bhubaneswar, India. Email: sateesh.cs@utkaluniversity.ac.in

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

Step-2: The output of the hidden layer is defined using the Equation (1).

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

Step-3: The output weights are determined as β, $where\ \beta = H^{\dagger}T$, where $H^{\dagger}$ is the MP generalized inverse of H and $T = (y_1, y_2, y_3, \dots \dots, y_n)^T$.

The structure of ELM is shown in Fig. 1.

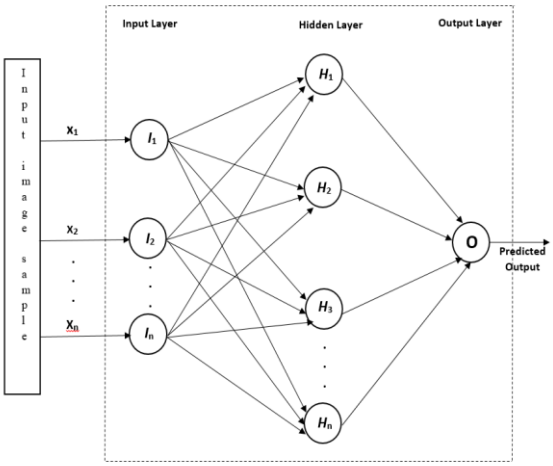


Fig.1. Basic ELM model

B. Two hidden layer Extreme Learning Machine

Due to the randomly generated hidden layer parameters i.e. weights between the input layer and hidden layer and biases in the hidden layer, the accuracy of the results produced using ELM are low. So research was continued to enhance the ELM performance. It leads to development of different extension to ELM i.e. Incremental ELM (I-ELM), Evolutionary ELM (E-ELM), online sequential ELM (OS-ELM) etc. The authors in [16] proposed a two hidden layer Extreme Learning Machine (TELM) that performed better as compared to basic ELM in some classification and regression problems. This TELM is described as below.

In TELM, an additional hidden layer is added between the hidden layer and output layer of the single layer ELM. The notations of TELM are explained in Table 1.

Table.1

Notation	Description
X	Input sample vector
T	Target vector (also known as labelled samples)
N_H	Number of hidden neurons in each hidden layer
W_{H1}	Weights between input layer and first hidden layer
B_1	Bias matrix of the first hidden layer
β	Weight matrix between the second hidden layer and the output layer.
W_{H2}	Weight matrix between the first and second hidden layer.
Y_1	Output of the first hidden layer
Y_2	Output of the Second hidden layer
B_2	Bias matrix of the second hidden layer
$\beta^{\dagger}$	MP Generalized inverse

Now the output of the first hidden layer is calculated using Equation (2).

$$Y_1 = f(W_{H1} \times X + B_1) \quad (2)$$

Similarly, the output of the second hidden layer is calculated using Equation (3).

$$Y_2 = f(W_{H2} \times Y_1 + B_2) \quad (3)$$

Also the expected output of the second hidden layer can be calculated using Equation (4).

$$Y_2 = T \times \beta^{\dagger} \quad (4)$$

The architecture of TELM is shown in Fig. 2.

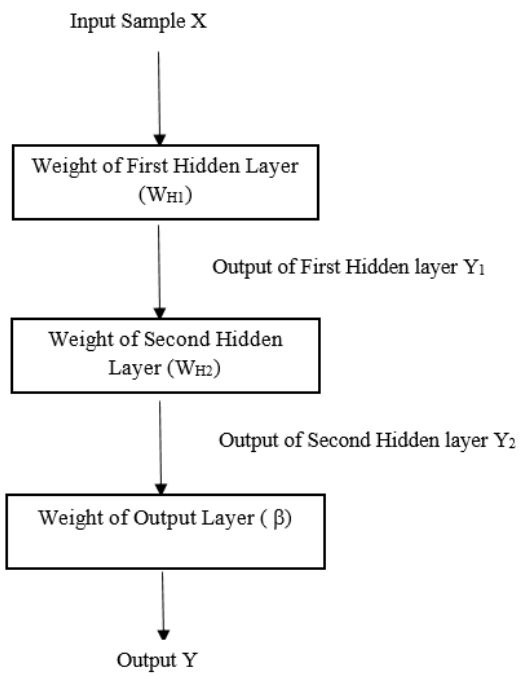


Fig. 2. TELM Architecture

The basic steps of TELM as proposed in [16] are expressed as follows:

Step 1: Generate randomly W_{H1} and B_1

Step 2: Determine Y_1 using Equation (2).

Step 3: Calculate $\beta = T \times Y_1^{\dagger}$

Step 4: Now evaluate the output of second hidden layer Y_2 using Equation (4).

Step 5: Calculate the weights of the second hidden layer W_{H2} using Equation (5).

$$W_{H2} = f^{-1}(Y_2) \times Y_1^{\dagger} \quad (5)$$

Step 6: Determine the actual output of the second hidden layer using Equation (3).

Step 7: Recalculate the output weight matrix using Equation (6).

$$\beta' = T \times Y_2^{\dagger} \qquad (6)$$

Step 8: The final output is determined using Equation (7).

$$Y = f(W_{H2} f(W \times X + B_1) + B_2) \times \beta' \qquad (7)$$

III. SIMULATION RESULTS

This section discusses the simulation study of this paper.

In this paper, three brain MRI datasets i.e. Alzheimer, Glioma ad Multiple Sclerosis are collected which are axial and T2-weighted brain MRI images. These datasets were taken from the Medical school of Harvard University. 2-D DWT technique is used to extract the features from the brain images. All the extracted features are not beneficial in classification. So PCA is used for feature reduction. In this work 2D DWT and PCA techniques are used for data preprocessing. Table 2 describes the features after applying PCA.

Table 2. Feature reduction using PCA

Name of the Datasets	No. of original features	No. of reduced features
Alzheimer	1296	34
Glioma	1296	24
Multiple Sclerosis	1296	39

The preprocessed datasets are now taken for training and testing the ELM model and TELM model. The datasets are divided into train and test datasets randomly. After training the models, the test data are used for testing the model. The Accuracy, F-score, Recall and Precision are used for performance comparison. These four metrics are explained in Equation (8) to Equation (11).

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \qquad (8)$$

$$Precision = \frac{TP}{TP + FP} \qquad (9)$$

$$Recall = \frac{TP}{TP + FN} \qquad (10)$$

$$F - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \qquad (11)$$

Table 3 shows the performance values for the three datasets for ELM and TELM.

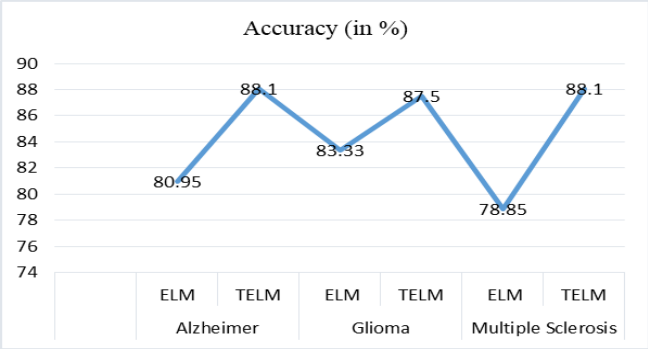
Table 3. Accuracy and Precision values of two models

Name of the Dataset	Name of the Model	Accuracy	Precision
Alzheimer	ELM	80.95	85.71
	TELM	88.1	92.85
Glioma	ELM	83.33	85.71
	TELM	87.5	88.57
Multiple Sclerosis	ELM	78.85	85.71
	TELM	88.1	92.85

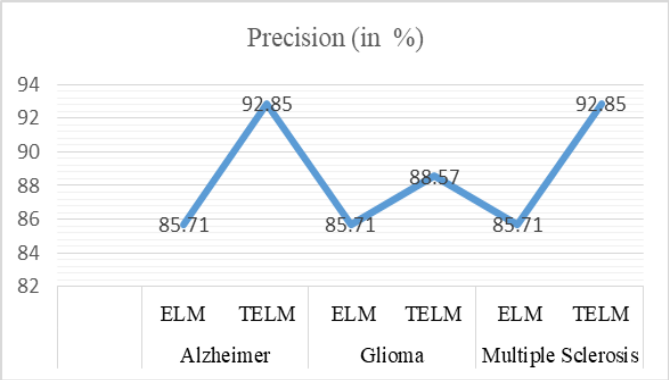
Table 4. Recall and F-score values of two models

Name of the Dataset	Name of the Model	Recall	F-Score
Alzheimer	ELM	85.71	85.71
	TELM	89.65	91.22
Glioma	ELM	90.9	88.23
	TELM	93.93	91.17
Multiple Sclerosis	ELM	87.5	83.58
	TELM	93.75	89.55

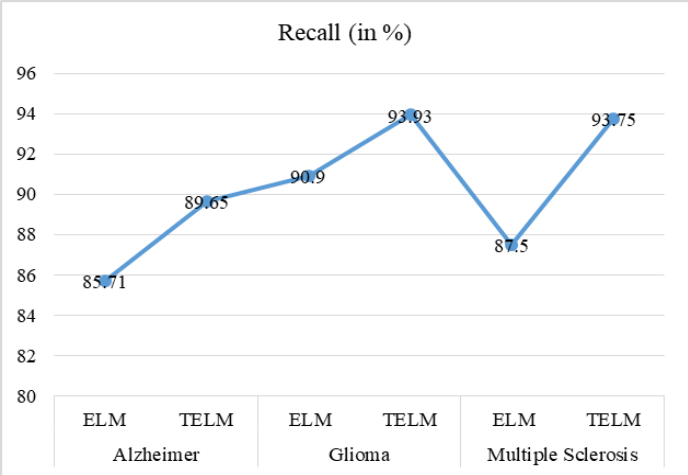
The above values are shown graphically in Fig.3.



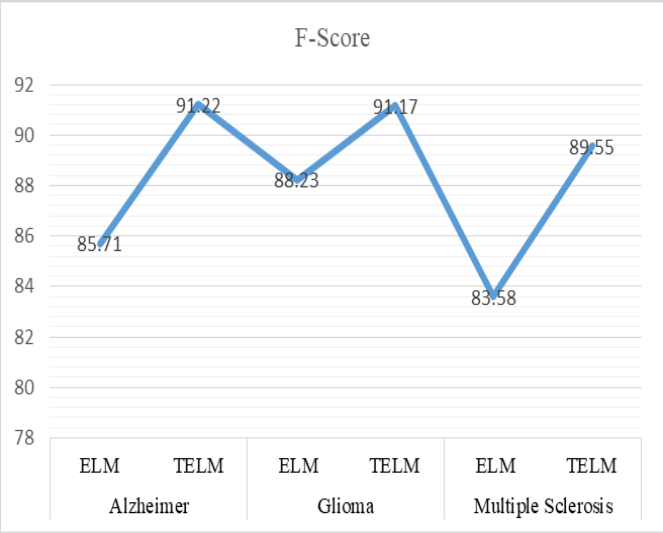
(a)



(b)



(c)



(d)

Fig. 3. Graphical Represetation of (a) accuracy (b) precision (c) recall (d) F-score for three datasets using ELM and TELM model

The performance measures are evaluated for the three brain MRI datasets using both ELM and TELM model. From the above simulation study, it is found that the accuracy values of TELM model are 88.1, 87.5 and 88.1 for Alzheimer, Glioma and Multiple Sclerosis respectively. Similarly, for the other three measures, the values of TELM model are better than ELM. So it is concluded that TELM outperformed ELM in classification of brain images into normal or diseased for three of the brain MRI datasets.

IV. CCONCLUSION

This paper discusses the TELM model in which an extra hidden layer is appended with the single hidden layer ELM. This model is used for the classification of brain MRI datasets in this work. The three categories of brain diseases i.e. Alzheimer, Glioma and Multiple Sclerosis are considered and these datasets are preprocessed before classification. 2D DWT is used for feature extraction and PCA is used for feature reduction. 5-fold cross validation scheme is used for training and testing the models. The simulation results indicates that TELM model outperformed the ELM model with respect to accuracy, precision, recall and F-score for these three brain MRI datasets.

REFERENCES

1. D. Jude Hemanth, C. KeziSelvaVijila, A. Immanuel Selvakumar, J. Anitha, "Performance Improved Iteration-Free Artificial Neural Networks for Abnormal Magnetic Resonance Brain Image Classification", Neurocomputing, 130 , 98–107, 2014.
2. C. A. Cocosco, A. P. Zijdenbos, A. C. Evans, "A fully automatic and robust brain MRI tissue classification method",Medical image analysis, 7(4), 2 513-527,2003.
3. . S. Chaplot, L.M.Patnaik, N. Jagannathan, "Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network. Biomedical signal processing and control, 1(1), 86-92,2006.
4. N. V. Shree, T.N.R. Kumar, "Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network", Brain informatics, 5(1), 23-30, 2018.
5. V. Kumar, J. Sachdeva, I. Gupta, N. Khandelwal, C. K. Ahuja, "Classification of brain tumors using PCA-ANN", In: 2011 IEEE World Congress on Information and Communication Technologies, 1079-1083, 2011.
6. E. Dahshan, T. Hosny, A. Salem, "Hybrid intelligent techniques for MRI brain images classification", Digital Signal Processing, 20(2), 433-441, 2010.
7. Y. Zhang, Z. Dong, L. Wu, S. Wang. "A hybrid method for MRI brain image classification", Expert Systems with Applications, 38(8), 10049-10053, 2011.
8. G.B. Huang, Q. Zhu, C. Siew, "Extreme Learning Machine: Theory and applications", Neurocomputing, 70, 489-501, 2006.
9. G. B.Huang, Q. Y. Zhu, C. K. Siew,"Extreme learning machine: a new learning scheme of feedforward neural networks", Neural networks, 2, 985-990, 2004.
10. H. J. Rong, Y. S. Ong, A. H. Tan, Z. Zhu, "A fast pruned-extreme learning machine for classification problem", Neurocomputing, 72(1-3), 359-366,2008.
11. D. R. Nayak, R. Dash, B. Majhi, S. Wang, "Combining extreme learning machine with modified sine cosine algorithm for detection of pathological brain", Computers & Electrical Engineering, 68, 366-380, 2018.
12. R. K. Lama, J. Gwak, J. S. Park, S. W. Lee, "Diagnosis of Alzheimer's disease based on structural MRI images using a regularized extreme learning machine and PCA features", Journal of healthcare engineering, 2017.
13. G. B. Huang, N. Y. Liang,H. J. Rong, P. Saratchandran, N. Sundararajan, "On-line sequential extreme learning machine", Computational Intelligence, 232-237, 2005.
14. Q. Y. Zhu, A. K. Qin, P. N. Suganthan, G. B. Huang, Evolutionary extreme learning machine, "Pattern recognition", 38(10), 1759-1763, 2005.
15. Y. Miche, A. Sorjamaa, P. Bas, O. Simula, C. Jutten, A. Lendasse, "OP-ELM: optimally pruned extreme learning machine", IEEE transactions on neural networks, 21(1), 158-162, 2009.
16. B. Y. Qu, B. F. Lang, J. J. Liang, A. K. Qin, O. D. Crisalle, "Two-hidden-layer extreme learning machine for regression and classification", Neurocomputing, 175, 826-834, 2016.
17. C. T. Yue, J. Liang, B. Lang, B. Qu, "Two-hidden-layer extreme learning machine based wrist vein recognition system", Big Data & Information Analytics, 2(1), 59-68, 2017.
18. J. Tang, C. Deng, G. B. Huang, "Extreme learning machine for multilayer perceptron", IEEE transactions on neural networks and learning systems, 27(4), 809-821, 2015.
19. G.B.Huang, "Learning capability and storage capacity of two-hidden-layer feed forward networks, IEEE Transaction on. Neural Network, 14(2), 274–281, 2003.
20. G. B. Huang, D. Wang, Y. Lan, "Extreme learning machines: a survey", International journal of machine learning and cybernetics, 2(2), 107-122, 2011.

AUTHORS PROFILE

Pranati Satapathy: She is pursuing Ph. D. in the department of Computer Science ad Applications of Utkal University. Her research area includes medical image processing and soft computing. She has 8 years of teaching experience.

Sateesh Kumar Pradhan: He is working as a Professor in the department of Computer Science and Applications in Utkal University. His research area includes computer networking, parallel computing, soft computing etc. He has published more than 50 research papers in various journals and conferences.

Sarbeswara Hota: He is working as an Assistant Professor in the department of Computer application in SOA deemed to be university. He has 17 years of teaching experience. His research area includes financial computing, soft computing and machine learning.