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 physicists. 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 \( M = \{ (x_1, y_1), \ldots, (x_n, y_n) \} \in \mathbb{R}^d, y_i \in \mathbb{R}, i = 1, 2, \ldots, n \) be the dataset.

Revised Manuscript Received on December 15, 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.

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

Published By:
Blue Eyes Intelligence Engineering & Sciences Publication

Retrieval Number: A19091091192019©BEIESP
DOI: 10.35940/ijeat.A1909.129219

3654
Step-1: Assign the weight $w_i$ and bias $b_i(i = 1, 2, ..., N)$ between the input and hidden layer randomly.

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

$$H = \left( g(W_{1i}X_i + b_i), g(W_{2i}X_i + b_i), ..., g(W_{Ni}X_i + b_i) \right)$$

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

The structure of ELM is shown in Fig. 1.

![Fig.1. Basic ELM model](image)

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>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X$</td>
<td>Input sample vector</td>
</tr>
<tr>
<td>$T$</td>
<td>Target sample vector also known as labelled samples</td>
</tr>
<tr>
<td>$N_H$</td>
<td>Number of hidden neurons in each hidden layer</td>
</tr>
<tr>
<td>$W_{H1}$</td>
<td>Weights between input layer and first hidden layer</td>
</tr>
<tr>
<td>$B_1$</td>
<td>Bias matrix of the first hidden layer</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Weight matrix between the first and second hidden layer</td>
</tr>
<tr>
<td>$W_{H2}$</td>
<td>Weight matrix between the second hidden layer and output layer</td>
</tr>
</tbody>
</table>

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

$$Y_1 = f(W_{H1}X + B_1)$$

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

$$Y_2 = f(W_{H2}Y_1 + B_2)$$

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

$$Y_2 = T \times \beta^\top$$

The architecture of TELM is shown in Fig. 2.

![Fig. 2. TELM Architecture](image)
Step 8: The final output is determined using Equation (7).
\[ Y = f(W_{H2}f(W \times \beta_2) + B_2) \times \beta'' \quad (7) \]

### III. SIMULATION RESULTS

This section discusses the simulation study of this paper.

In this paper, three brain MRI datasets i.e. Alzheimer, Glioma and 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>
<thead>
<tr>
<th>Name of the Datasets</th>
<th>No. of original features</th>
<th>No. of reduced features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alzheimer</td>
<td>1296</td>
<td>34</td>
</tr>
<tr>
<td>Glioma</td>
<td>1296</td>
<td>24</td>
</tr>
<tr>
<td>Multiple Sclerosis</td>
<td>1296</td>
<td>39</td>
</tr>
</tbody>
</table>

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).

\[
\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (8)
\]

\[
\text{Precision} = \frac{TP}{TP + FP} \quad (9)
\]

\[
\text{Recall} = \frac{TP}{TP + FN} \quad (10)
\]

\[
F\text{-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (11)
\]

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

<table>
<thead>
<tr>
<th>Name of the Dataset</th>
<th>Name of the Model</th>
<th>Accuracy</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alzheimer</td>
<td>ELM</td>
<td>80.95</td>
<td>85.71</td>
</tr>
<tr>
<td></td>
<td>TELM</td>
<td>88.1</td>
<td>92.85</td>
</tr>
<tr>
<td>Glioma</td>
<td>ELM</td>
<td>83.33</td>
<td>85.71</td>
</tr>
<tr>
<td></td>
<td>TELM</td>
<td>87.5</td>
<td>88.57</td>
</tr>
<tr>
<td>Multiple Sclerosis</td>
<td>ELM</td>
<td>78.85</td>
<td>85.71</td>
</tr>
<tr>
<td></td>
<td>TELM</td>
<td>88.1</td>
<td>92.85</td>
</tr>
</tbody>
</table>

Table 4 shows the recall and F-score values for the two models.

<table>
<thead>
<tr>
<th>Name of the Dataset</th>
<th>Name of the Model</th>
<th>Recall</th>
<th>F-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alzheimer</td>
<td>ELM</td>
<td>85.71</td>
<td>85.71</td>
</tr>
<tr>
<td></td>
<td>TELM</td>
<td>89.65</td>
<td>91.22</td>
</tr>
<tr>
<td>Glioma</td>
<td>ELM</td>
<td>90.9</td>
<td>88.23</td>
</tr>
<tr>
<td></td>
<td>TELM</td>
<td>93.93</td>
<td>91.17</td>
</tr>
<tr>
<td>Multiple Sclerosis</td>
<td>ELM</td>
<td>87.5</td>
<td>83.58</td>
</tr>
<tr>
<td></td>
<td>TELM</td>
<td>93.75</td>
<td>89.55</td>
</tr>
</tbody>
</table>

The above values are shown graphically in Fig. 3.
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


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 at 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.