Epileptic Seizure Detection using Deep Ensemble Network with Empirical Wavelet Transform

Sreelekha Panda¹, Abhishek Das², Satyasis Mishra¹, Mihir Narayan Mohanty²*

¹Dept. Of ECE, Centurion University of Technology and Management, Odisha, India
²Dept. Of ECE, ITER, Siksha ‘O’ Anusandhan (Deemed to be University), Bhubaneswar, Odisha, India
mihir.n.mohanty@gmail.com

Epileptic seizure attack is caused by abnormal brain activity of human subjects. Certain cases will lead to death. The detection and diagnosis is therefore an important task. It can be performed either by direct patient activity during seizure or by electroencephalogram (EEG) signal analysis by neurologists. EEG signal processing and detection of seizures using machine learning techniques make this task easier than manual detection. To overcome this problem related to a neurological disorder, we have proposed the ensemble learning technique for improved detection of epilepsy seizures from EEG signals. In the first stage, EEG signal decomposition is done by utilizing empirical wavelet transform (EWT) for smooth analysis in terms of sub-bands. Further, features are extracted from each sub. Time and frequency domain features are the two categories used to extract the statistical features. These features are used in a stacked ensemble of deep neural network (DNN) model along with multilayer Perceptron (MLP) for the detection and classification of ictal, inter-ictal, and pre-ictal (normal) signals. The proposed method is verified using two publicly available datasets provided by the University of Bonn (UoB dataset) and Neurology and Sleep Center - New Delhi (NSC-ND dataset). The proposed algorithm resulted in 98.93 % and 98 % accuracy for the UoB and NSC-ND datasets, respectively.

Keywords: Epilepsy, Empirical Wavelet Transform, DNN, Ensemble, EEG Classification.

1. INTRODUCTION

Epilepsy is considered a neurological disorder that causes instant seizures. The occurrence of repeated seizure attacks may cause severe damage to the person. The prediction of seizure attacks is somehow difficult due to their random nature. The cause of epilepsy seizure includes traumatic brain injury, central nervous system (CNS) infection, cerebrovascular disease, brain tumors, degenerative CNS disease, perinatal factors, familial and genetic factors [1]. The diagnosis process of epilepsy includes EEG monitoring or observing the patient through video. Manual monitoring of EEG signals for diagnosis may lead to delay in detection that may cause an increase in the risk on a patient’s life. Machine learning-based observations have proven to be automatic and accurate in comparison to classical detection methods for disease diagnosis.

In recent decades researchers have developed various methods in this field of research. Application of classic statistical methods like Fourier transform (FT) [2], wavelet transform (WT) [3], and principal component analysis (PCA) [4], etc., has been developed for epilepsy seizure detection. Classic machine learning techniques like the support vector machine (SVM) optimized using genetic algorithm and trained with features extracted by double density discrete wavelet transform (DWT) have shown a competitive performance in seizure detection [5]. Other machine learning methods like K-nearest neighbor (K-NN), decision tree (DT), Random forest, and naive Bayes are also utilized for epilepsy detection [6], [7]. Various deep learning-based methods are attracting researchers in EEG signal processing for epilepsy detection. Recurrent neural networks (RNN) have been used for epilepsy detection from EEG signals [8]. Long short-term memory-based recurrent models have been proven to be better in comparison to simple RNN models for epilepsy detection [9]-[11]. Convolutional neural networks have a higher attraction in comparison to other deep learning-based methods in epilepsy data detection for their automatic feature extraction characteristics [12]-[14].

Different decomposition methods have been utilized for sub-band generation from EEG signals in epilepsy detection. Empirical mode decomposition (EMD) has been considered in recent works for EEG signal decomposition and then for further processing by classifiers [15]-[17]. Improved EMD methods like ensemble-EMD (EEMD) and complete-EEMD (CEEMDAN) have also been applied for epilepsy detection from EEG signals [18], [19]. DWT-based EEG signal analysis for extracting envelopes is another way of decomposition [20]-[23]. EWT as a mode decomposition
method with various classifiers has also been chosen for EEG signal classification in epilepsy detection [24]-[26].

Ensemble learning has proven to be a better option for obtaining increased performance in comparison to single classifiers in various fields of the image as well as in single-dimensional signal processing [27]-[30]. Ensemble learning has also shown its importance in epilepsy seizure detection and classification including the combination of various statistical machine learning as well as deep learning approaches [31]-[33]. Still, research is going on to predict the seizure attack by studying the frequency of seizure occurrence signals, which can be studied in both the time domain and frequency domain by considering features in both domains for training purposes [34]. The authors in this work have provided details of features that can be used for further analysis. This paper mainly focuses on the following few points:

- The proposed classification model is trained directly with the EEG dataset to observe the performance.
- The epilepsy prediction is done using EWT as a feature extractor and ensemble of DNNs for classification.
- The decomposed signals generated by EWT with a different number of modes (2 to 8) are then processed to extract statistical features both in the time domain and frequency domain.
- The frequency-domain features include AM and FM bandwidths, spectral entropy, spectral power, and spectral centroid, whereas the time domain features include skewness, kurtosis, Hjorth activity, Hjorth mobility, and Hjorth complexity. These features are then divided into two groups to train the classification model.
- The classification results are compared in terms of results obtained from the two groups of features.
- The ensemble of DNNs with MLP as meta classifier is used for the classification of ictal, inter-ictal, and normal (pre-ictal) epilepsy EEG signals.
- The DNN is designed with 128, 64, 32, 64, 128 nodes activated by the ReLU activation function and at the end 2 number of nodes are taken for base level classification activated by the Softmax activation function. Each model of DNN is optimized with Adam optimizer. Binary cross entropy is used to calculate the loss.
- The proposed method is verified on two publicly available datasets, i.e. the UoB dataset and the NSC-ND dataset.

The remaining part of the paper is organized as follows: Proposed algorithms adopted for EEG signal analysis for epilepsy detection are discussed in section 2; section 3 provides detailed information about the results obtained from the proposed method, and the conclusion is drawn in section 4.

2. PROPOSED METHOD

Instead of directly feeding the EEG signals of various kinds like ictal, inter-ictal, and normal (pre-ictal), to the classification model, signals from each category are decomposed into their corresponding modes using EWT [35].

The statistical features are then extracted into two groups to train the stacked ensemble model for the detection of epilepsy seizures. The workflow diagram of the proposed algorithm is summarized in Fig.1.

![Workflow diagram of the proposed method](image)

**2.1. Empirical Wavelet Transform (EWT)**

EWT is designed to extract the AM and FM components of a signal adaptively using wavelet filter banks. The decomposition is possible due to the compact Fourier supports of such components following the assumptions: (1) the signal is real-valued to maintain the symmetry property, and (2) periodic with a period $2\pi$. But the signal analysis is performed in the range $[0, \pi]$ to satisfy Shannon’s sampling criterion. The number of modes (N) to which the signal will be decomposed is set before the decomposition that makes the method different from the previously developed empirical mode decomposition (EMD) designed by Huang et al.[36] in 1998.

**2.2. Feature extraction**

The proposed ensemble model is trained with 10 statistical features extracted from the decomposed signals of each model. The features are divided into two groups depending upon the signal domain. Frequency domain and time domain features are the two main groups each including five different features. This work is mainly focusing on epilepsy seizure detection as well as the importance of features for this purpose.

**2.2.1 Frequency domain method**

The features extracted from the decomposed signals under this category are AM and FM bandwidths, spectral entropy, spectral power, and spectral centroid. AM and FM bandwidths are evaluated using the equations (1) and (2) as mentioned in [37].

$$B_{AM} = \sqrt{\frac{1}{E} \int \left(\frac{dA(t)}{dt}\right)^2 dt}$$

$$B_{FM} = \sqrt{\frac{1}{E} \int \left(\frac{dA(t)}{dt}\right)^2 dt}$$
Where $A$ represents the amplitude of the EEG signal, $E$ is the energy of the signal and $\langle \omega \rangle$ represents the central frequency corresponding to each mode. The third and fourth features are the power and entropy of the frequency domain representation of the signal, i.e. spectral power ($S_{pow}$) and spectral entropy ($S_{Ent}$) are calculated as equations (3) and (4).

$$S_{pow} = \frac{f_2}{N} \sum_{f=0}^{f_2} P_{XX}(f)$$ (3)

$$S_{Ent} = -\sum_{f=0}^{f_2} P_{XX} \log[P_{XX}(f)]$$ (4)

Where, $N$ represents the count of spectral coefficients, and $P_{XX}$ is the power spectral density calculated by using Welch’s method [38], and $P_{XX}$ represents the normalized power spectral density.

Spectral centroid (SC) is determined using equation (5).

$$SC = \frac{\sum_{f=0}^{f_2} \omega(f)M(f)}{\sum_{f=0}^{f_2} M(f)}$$ (5)

Where $f$ represents the frequency segment of the original signal. $\omega(f)$ and $M(f)$ are the central frequency and magnitude of PSD of the segment $f$, respectively.

2.2.2. Time-domain method

For time-domain feature extraction, we have considered the skewness, kurtosis, and Hjorth parameters. Skewness and Kurtosis calculation are done as per equations (6) and (7).

$$Skewness = E \left[ \left( \frac{f(t) - \mu}{\sigma} \right)^3 \right]$$ (6)

$$Kurtosis = E \left[ \left( \frac{f(t) - \mu}{\sigma} \right)^4 \right]$$ (7)

Where, $\sigma$ and $\mu$ represent the standard deviation and mean of the signal $f(t)$.

Hjorth parameters include activity, mobility, and complexity. These parameters are calculated using equations (8)-(10).

$$H_{Activity} = \text{var}(f(t))$$ (8)

$$H_{Mob} = \sqrt{\frac{\text{var} \left( \frac{df(t)}{dt} \right)}{\text{var}(f(t))}}$$ (9)

$$H_{Comp} = \frac{\text{Mob} \left( \frac{df(t)}{dt} \right)}{\text{Mob}(f(t))}$$ (10)

Where $\text{var}$) is the variance of the signal.

These features are well studied by providing them to the classification model and the results obtained from these two groups are discussed in the results section.

2.3. Classification

Once the features are extracted, these features are dived into two groups as described above, depending upon the time and frequency domain representation. The classification model proposed in this work is a stacked ensemble model. The ensemble part consists of four DNN models with the same number of layers and nodes in each model. We have analyzed the number of base learners by varying it from 2 to 4. Considering the two numbers of base learners is not advisable as their results may bias the final decision when both the base classifiers will provide results opposite to each other. So, we have not considered two numbers of DNNs. The training accuracy obtained with three and four numbers of DNNs are 96.54 % and 100 %, respectively. The number of base classifiers also increased to five. The result obtained with five base classifiers was deteriorating due to overfitting of the data and provided less accuracy than that of four-DNNs model. Therefore, the proposed model is designed with four DNNs as base learners. The prediction outcomes of each base model are combined to form the metadata that is passed to the meta classifier MLP for training and final classification. The MLP model has 8 nodes in the input layer to accept the 8 outcomes from the four base models. The hidden layer consists of 32 numbers of nodes activated using the ReLU activation function and the final output layer again consists of two nodes activated by the Softmax activation function for two-class classification, i.e. for ictal versus pre-ictal (normal) and ictal versus inter-ictal classification. Algorithm 1 is used for model training, and classification. The meta classifier is the final classifier that performs the linear stacking on the classification results by base learners. The outputs of base learners, i.e. $B_{1}, B_{2}, ..., B_{4}$, are linearly combined with weights $W_{i}, i \in 1,...,n$, learned by the MLP, which is given by equation (11).

$$C_{\beta}(x) = \sum_{i=1}^{n} W_{i} B_{i}(x)$$ (11)
### Algorithm 1: The proposed method

1. **Input:** Dataset $\mathcal{D} = \{x_i,y_i\}_{i=1}^m$
2. **Output:** Class
3. **Step 1:** Train the Base Learners
   4. For $n = 1$ to $4$
      - Train $B_{L_n}$ with $\mathcal{D}$
   5. End for
4. **Step 2:** generate the input for Meta Learner
   7. for $i = 1$ to $n$
      - $D_{\text{meta}} = \{x_i',y_i\}$
      - Where $x_i' = \{B_{L_1}(x_i), B_{L_2}(x_i), \ldots, B_{L_4}(x_i)\}$
   8. end for
5. **Step 3:** Train Meta Learner ($M_{\text{L}}$)
   10. Train $M_{\text{L}}$ with $D_{\text{meta}}$
6. **Return** Class

Where $x_i$: features, $y_i$: Labels of features, $m$: number of features, $n$: number of base learners

### 3. RESULTS

The proposed method was implemented in Python 3.7 on the Google Collaboratory platform utilizing the online GPU provided by Google. The dataset and the results obtained using the proposed method are discussed in this section. A comparative analysis is also done based on the two types of features in the time and frequency domain.

#### 3.1. Dataset

The proposed stacked ensemble model is verified using two publicly available datasets: (i) the UoB dataset, provided by the University of Bonn [39], and (ii) the NSC-ND dataset provided by Neurology and Sleep Center, Hauz Khas, New Delhi [40]. The UoB dataset has five subfolders with names F, N, O, S, and Z. Each subfolder contains 100 numbers of EEG data sampled at 173.6 Hz for the duration of 23.6 seconds. The signals under F and N categories are the EEG signals taken within the seizure intervals by applying electrodes on the epileptogenic zone and just opposite to hippocampus. The EEG signals under the category S are taken during the seizure attack. O and Z are the normal EEG signals taken from five healthy subjects with eyes closed and open. There are a total of 500 signals in the dataset with 100 numbers of signals in each category.

The NSC-ND dataset contains three different types of EEG signals sampled at 200 Hz for the duration of 5.12 seconds. The dataset contains three folders named as ictal, inter-ictal, and pre-ictal. Each category of these folders contains 50 signals. The datasets were divided subject wise. EEG data from each subject are divided to training and testing sets with 80:20 ratios. The test data is used as validation set for hyperparameter estimation.

The samples of decomposed signals obtained from the NSC-ND dataset for ictal, pre-ictal, and inter-ictal signal with $N = 4$, are provided in Fig.2, to Fig.4.

#### 3.2. Classification results

The performance of the proposed classification model is verified first by directly providing the EEG signals from the dataset without any decomposition and statistical feature extraction. The DNN models have extracted the features automatically for training. The performance of the proposed classification model is tested by applying the decomposition method and statistical features prior to training. It is observed that the performance is better in statistical feature-based training approach in comparison to direct use of datasets. The comparative analysis is provided in Table 1.
Table 1. Results analysis with and without statistical features.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Classification</th>
<th>Accuracy (%)</th>
<th>Without EWT and statistical features</th>
<th>With EWT and Statistical features</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Time-Domain</td>
<td>Frequency-Domain</td>
</tr>
<tr>
<td>NSC-ND</td>
<td>Ictal vs pre-ictal (Normal)</td>
<td>94.73</td>
<td>0.9600</td>
<td>0.9800</td>
</tr>
<tr>
<td></td>
<td>Ictal vs inter-ictal</td>
<td>95.16</td>
<td>0.9600</td>
<td>0.9700</td>
</tr>
<tr>
<td>UoB</td>
<td>Ictal vs pre-ictal (Normal)</td>
<td>94.43</td>
<td>0.9702</td>
<td>0.9782</td>
</tr>
<tr>
<td></td>
<td>Ictal vs inter-ictal</td>
<td>95.01</td>
<td>0.965</td>
<td>0.9893</td>
</tr>
</tbody>
</table>

Table 2. Validation results obtained for Ictal vs pre-ictal (Normal).

<table>
<thead>
<tr>
<th></th>
<th>Time Domain Features</th>
<th>Frequency Domain Features</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UoB Dataset</td>
<td>NSC-ND dataset</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.9702</td>
<td>0.9600</td>
</tr>
<tr>
<td>Precision</td>
<td>1.0000</td>
<td>0.9500</td>
</tr>
<tr>
<td>Recall</td>
<td>0.9300</td>
<td>0.9600</td>
</tr>
<tr>
<td>F1-Score</td>
<td>0.9600</td>
<td>0.9549</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.9238</td>
<td>0.9000</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.9721</td>
<td>0.9300</td>
</tr>
</tbody>
</table>

Table 3. Validation results obtained for ictal vs inter-ictal.

<table>
<thead>
<tr>
<th></th>
<th>Time Domain Features</th>
<th>Frequency Domain Features</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UoB Dataset</td>
<td>NSC-ND dataset</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.9600</td>
<td>0.965</td>
</tr>
<tr>
<td>Precision</td>
<td>0.9710</td>
<td>0.9500</td>
</tr>
<tr>
<td>Recall</td>
<td>0.9600</td>
<td>0.9632</td>
</tr>
<tr>
<td>F1-Score</td>
<td>0.9649</td>
<td>0.9565</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.9738</td>
<td>0.9502</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.9424</td>
<td>0.9730</td>
</tr>
</tbody>
</table>

From Table 1, it is observed that the classification model is providing better results while trained with statistical features extracted from decomposed EWT signals in comparison to direct training on EEG datasets.

The proposed model is trained with time-domain features as well as frequency-domain features. The results are evaluated in terms of accuracy, sensitivity, specificity, precision, recall, and F1 score. The accuracy plots for comparison purposes to show the improvement in results by stacked ensemble learning are given in Fig.5. for ictal versus inter-ictal classification using the UoB dataset. The validation results obtained by the proposed method are provided in Table 2 and Table 3 for ictal versus pre-ictal (normal) and ictal versus inter-ictal signal, respectively.

A comparative analysis with respect to related works is provided in Table 4.

Table 4. Comparative analysis with respect to related works.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Method Used</th>
<th>Accuracy (%)</th>
<th>F1-score (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[2]</td>
<td>DNN</td>
<td>95</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[3]</td>
<td>DWT</td>
<td>88</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[4]</td>
<td>FT + Sparse Denoising Autoencoder</td>
<td>93.82</td>
<td>96.05</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[8]</td>
<td>RNN</td>
<td>85</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[9]</td>
<td>LSTM</td>
<td>-</td>
<td>-</td>
<td>99.86</td>
<td>-</td>
</tr>
<tr>
<td>[10]</td>
<td>LSTM</td>
<td>97.78</td>
<td>-</td>
<td>98.85</td>
<td>-</td>
</tr>
<tr>
<td>[12]</td>
<td>WT+DNN</td>
<td>97.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[13]</td>
<td>CNN</td>
<td>-</td>
<td>-</td>
<td>87.8</td>
<td>-</td>
</tr>
<tr>
<td>[14]</td>
<td>CNN</td>
<td>97.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[15]</td>
<td>EMD-DWT + KNN</td>
<td>89.4</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[18]</td>
<td>EEMD + K-means Clustering</td>
<td>98%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[19]</td>
<td>CEEMD</td>
<td>98.67</td>
<td>-</td>
<td>98.67</td>
<td>98.72</td>
</tr>
<tr>
<td>[20]</td>
<td>DWT + Ensemble-NN</td>
<td>98.78</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>This work</td>
<td>EWT + Stacked ensemble-based DNN model</td>
<td>98.93</td>
<td>98.99</td>
<td>99.5</td>
<td>100</td>
</tr>
</tbody>
</table>

3.3. Discussion

From Fig.5, it can be observed that the ensemble of four DNNs with stacked MLP performs better in comparison to single DNN models. The statistical features extracted from each mode are sufficient to train the model to obtain the objective of the work. From Table 1 and Table 2, it can be observed that the proposed model is performing better in the case of frequency domain features in comparison to time-domain features. The accuracy in epilepsy seizure detection
reached the highest value of 98.93 in ictal versus inter-ictal classification with frequency-domain features that is competitive with the state-of-the-art methods. The proposed model is also providing a promising result of 98% accuracy in ictal versus pre-ictal classification.

4. CONCLUSIONS

The remarkable characteristics of ensemble learning are verified with epilepsy seizure detection by preprocessing the EEG datasets using empirical wavelet transform. The decomposition of EEG signals provides detailed information hidden in the multi-resolution arrays of each signal. The time domain and frequency domain analysis of signal are well studied by training the classification model in two different steps. A comparative analysis is provided along with improved seizure detection. The accuracy obtained in epilepsy seizure detection is 98.93% for ictal versus inter-ictal classification with frequency-domain features that are competitive with state-of-the-art methods. In the future, the work is planned to be carried out with other decomposition models with improved accuracy for detection and classification.

REFERENCES


Received April 24, 2021
Accepted August 05, 2021