Received: May 19, 2024. Revised: June 13, 2024.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Classification of Epileptic Seizure Using Hierarchical Long Short-Term Memory with Skip Connection Method

Nageshwar Voruganti[1]* **Vijaya Kumar Gurrala[2]**

[1]*Department of Electronics and Instrumentation Engineering,*
Vallurupalli Nageswara rao Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India
[2]*Department of Electronics and Communication,*
Vallurupalli Nageswara rao Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India
* Corresponding author's Email: nageshwar_v@vnrvjiet.in

Abstract: The neural activities of the brain are detected by Electroencephalography (EEG) that allows for analysis and classification of epileptic disease. The existing methods fail to capture high-dimensional data between adjacent sequences which make it difficult for the classifier to process and maximize the classification errors. This research proposes Hierarchical Long Short-Term Memory (H-LSTM) with a skip connection-based epileptic seizure classification method. The H- LSTM captures both short and long-term dependencies between adjacent sequences in the high-dimensional data. The skip connection is introduced between H- LSTM layers, facilitating the flow of data across adjacent sequences to improve the classification performance of epileptic seizures. The datasets used to evaluate the proposed H- LSTM with skip connection-based classification method are BONN-EEG and CHB-MIT EEG. The proposed H-LSTM with skip connection method attains 99.81% accuracy on BONN – EEG while attaining 99.34% accuracy on the CHB-MIT EEG dataset which is more effective than the existing methods namely, Bidirectional Gated Recurrent Unit (Bi-GRU) and Graph Convolutional Network (GCN).

Keywords: Electroencephalography, Hierarchical long short-term memory, High dimensional data, Long-term dependencies, Skip connection.

1. Introduction

Epileptic people suffer from unnecessary seizures with uncontrolled bursts of electrical activity that can cause death if not treated early [1]. Hence, the effective diagnosis and prediction of epileptic seizures is critical in recent times [2]. Epilepsy is generally diagnosed by Electroencephalogram (EEG) signals, which record brain waves that generally reflect electrical activity in the brain [3]. The EEG signals are generally diagnosed by experienced experts with the naked eye, but manual prediction of EEG signals leads to high cost and error [4]. Hence, the development of automatic EEG detection methods is highly significant in neuroscience [5]. The Machine Learning (ML) and Deep Learning (DL) techniques attain huge performance and high speed in the prediction of EEG signals and show the method's reliability [6]. The ML algorithms involve feature extraction as the initial stage, after which the classification is performed [7]. Feature extraction from EEG is a significant process in the classification of epileptic seizure techniques [8].

To identify the EEG seizure, various features have extracted from time, frequency, and time-frequency domains like power spectral density, spike rate, and energy of signals from wavelet transform [9]. The DL algorithms provides automatic feature extraction from EEG signals, without manual feature extraction [10-12]. In recent times, DL algorithms obtain features through the learning of model without manual feature selection, and exhibit an optimum performance in EEG recognition and classification. [13-15] The DL algorithms include Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) networks that attain an effective performance in sequence EEG data [16]. The existing methods have drawbacks like difficulty in capturing high-dimensional data between adjacent sequences that minimizes the sample density and data representation. To overcome these limitations, this research proposes a Hierarchical LSTM with a skip connection-based epileptic seizure classification method. The Hierarchical LSTM captures both low and high-dimensional data between adjacent sequences and skips the connection between LSTM layers, enabling the network facilitate the flow of information across adjacent sequences. The essential contributions of this research are given as follows:

- The Hierarchical Long Short-Term Memory (LSTM) with skip connection technique is proposed for classifying epileptic EEG seizures, which captures both low and high-dimensional data between adjacent sequences to enhance the classification performance.
- The skip connection between LSTM layers facilitates gradient flow of data across adjacent sequences, which mitigates the issue of gradient vanishing during training, thereby enhancing the classification performance of Epileptic EEF seizure.
- The Short Time Fourier Transform (STFT) and Continuous Wavelet Transform (CWT) based feature extraction techniques are utilized in this research to capture necessary temporal and spectral characteristics of EEG to differentiate variance between seizure and non-seizure.

This rest of the research paper is organized as follows: Section 2 analyzes the literature review of recent DL algorithms for epileptic seizure classification, while Section 3 describes the process of the proposed method with detailed information, Section 4 gives results and discussion and finally, the conclusion of this research is given in Section 5.

2. Literature review

Hassan [17] presented a novel epileptic detection technique with the integration of Empirical Mode Decomposition (EMD) and Mutual Information Best Individual Feature (MIBIF) selection technique and Multi-layered Perceptron Neural Network (MLPNN). At first, fixed length of EEG epochs was decomposed to amplitude, which is known as Intrinsic Mode Functions (IMFs). The important features were chosen from the measured feature through MIBIF technique for producing the last feature subset. Then, the produced feature subset was given to the MLPNN classifier. The presented method effectively detected and classified epileptic seizures. However, the problem of gradient vanishing occurred during training which affected the classification performance.

Abdulwahhab [18] suggested a DL algorithm that had two simultaneous methods to detect the activity of epileptic seizures by EEG signals. The image of time-frequency in EEG and raw waves were crucial for input elements of Convolutional Neural Network (CNN), RNN with LSTM. Further, two signal processing techniques, Short-Time Fourier Transform (STFT) and Continuous Wavelet Transformation (CWT) were employed to produce spectrogram and scalogram images. Nonetheless, the method had difficulty in capturing the high dimensional data between adjacent sequences.

Zhang [19] developed a Bidirectional Gated Recurrent Unit (Bi-GRU) neural network for seizure detection. The developed technique facilitated the treatment and diagnosis of epileptics. Initially, wavelet transforms were assigned for EEG recordings to filter in the pre-processing phase. Next, related signal energies in various specific frequency bands were measured and given to the Bi-GRU method. The developed BiGRU method captured long-term dependencies in EEG signals in both positive and negative directions. Nevertheless, developed technique had a high classification error rate that caused incorrect classification of epileptic seizures.

Jia [20] introduced a Graph Convolutional Network (GCN) method to predict seizures for resolving the issue of oversize seizure prediction methods depending on graph architecture of EEG signals. In the graph classification, network structure included graph convolution layers which extracted node features with single-hop neighbors, pooling layers summarized the node features, and fully connected layers were implemented for classification. The introduced method resulted in effective prediction and lesser network size. However, the method captured only the time-based features because the method did not consider rhythmic oscillations of EGG, leading to less discriminative power in differentiating between seizure and non-seizure EEG.

Islam [21] implemented a highly heterogenous and included Dense Convolutional Blocks (DCB), Feature Attention Modules (FAM), Residual Blocks (RB), and Hypercolumn Technique (HT). Initially, DCB was utilized for providing discriminative features from EEG samples. Next, FAM extracted significant features, following which RB learned many important parts as a whole and utilized data in convolutional layer. At last, HT retained effective local features extracted from layers placed at various

Received: May 19, 2024. Revised: June 13, 2024.

phases of methods. Nevertheless, the method had a lesser effective flow of information across adjacent sequences, resulting in the loss of significant temporal context in classifying seizure patterns.

Ra and Li [22] suggested a Synchro Extracting Transformation and Singular Value Decomposition (SET-SVD) method for enhancing the resolution of time-frequency. The SET was much energy focused on representation of TF than the traditional analysis techniques. Then, the classification of the pre-seizure technique was employed as one-Dimensional (1D-CNN). The suggested SET-SVD method displayed effective performance in EEG prediction. Nonetheless, the extracted hand-crafted features were not effective when classifying the signals because of the limited representation of features. From the above analysis, the existing methods have drawbacks such as the problem of gradient vanishing, difficulty in capturing high dimensional data between adjacent sequences, high classification error rate, while capturing only time-based features, less effective flow of information across adjacent sequences, and extraction of the hand-crafted features only. In order to tackle these limitations, this research proposes a Hierarchical LSTM with skip connection method. The proposed method captures high-dimensional data and minimizes the error rate in classification. The CNN-based feature extraction is performed to capture the deep features with hand-crafted features.

3. Proposed methodology

The Hierarchical LSTM with skip connection-based classification method is proposed for epileptic EEG seizure classification. The BONN-EEG and CHB-MIT EEG datasets are used in this research and is pre-processed by using Least Mean Square (LMS) and Z-score normalization techniques. Then, the time, frequency, and time-frequency features are extracted by using STFT and CWT based feature extraction methods. Further, it is classified by using Hierarchical LSTM with a skip connection-based classification method. Fig. 1 represents the process of epileptic seizure classification.

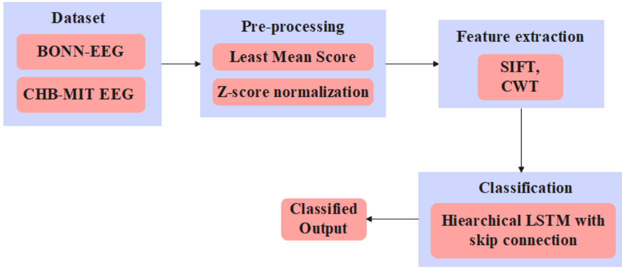

Figure. 1 Process of Epileptic Seizure Classification

3.1 Dataset

The two publicly available EEG datasets utilized in epileptic seizure prediction are BONN-EEG [23] created through the University of Bonn Germany and CHB-MIT EEG [24] by the Children's Hospital Boston, USA dataset. A brief explanation of these datasets is given as follows:

3.1.1. BONN-EEG dataset

The BONN-EEG dataset contains 5 subsets labeled as A, B, C, D, and E. Every subset has single-channel EEG data with particular characteristics. The subset A and B have scalp EEG data from healthy volunteers and subsets, while C and D have intracranial EEG data from non-focal and focal epileptic patients. At last, E has seizure-relevant intracranial EEG signals. Every subset has 100 files and every file has 4096 samples, every file recording 23.6 at a sampling rate of 173.61 Hz. Table 1 displays the dataset description of BONN-EEG.

Table 1. BONN-EEG dataset description

<table>
<tr><th>Subsets</th><th>Subject details</th><th>File name</th><th>Duration in seconds</th><th>Description</th></tr>
<tr><td>A</td><td rowspan="2">5 normal subjects (healthy)</td><td>Z001.txt – Z100.txt</td><td rowspan="5">100 (23.6)</td><td>EEG records with eyes open</td></tr>
<tr><td>B</td><td>O001.txt – O100.txt</td><td>EEG records with eyes closed</td></tr>
<tr><td>C</td><td rowspan="3">5 epilepsy patients</td><td>N001.txt – N100.txt</td><td>EEG records of hippocampal formation in hemisphere opposite to epileptogenic zone. It is recorded in seizure-free periods.</td></tr>
<tr><td>D</td><td>F001.txt – F100.txt</td><td>EEG records of epileptogenic zone. It is recorded in seizure-free periods</td></tr>
<tr><td>E</td><td>S001.txt – S100.txt</td><td>EEG records of epileptic seizure activity from hippocampal focus.</td></tr>
</table>

3.1.2. CHB-MIT EEG dataset

The CHB-MIT EEG dataset includes data acquired from seizure and interictal periods, following 10-20 international standard electrode placement system. The dataset has multiple channel EEG records with sampling rate of 256 Hz with a total of 23 records from 22 subjects. The N/A in the table represents the not specified, while Table 2 displays dataset description of the CHB-MIT EEG dataset.

3.2 Pre-processing

The EEG signal is pre-processed to obtain important features with high possibility of ictal and interictal portions of correlation. The pre-processed techniques used in this research are Least Mean Square (LMS) and z-score normalization. A brief explanation of pre-processed techniques is given below:

3.2.1. Noise removal using least mean square (LMS)

The quality of EEG signals is highly influenced by noise that degrades the performance of EEG epileptic classification. The frequency of noise is a critical examination of EEG signals. The frequency of noise and unnecessary data in EEG signals is eliminated by using LMS [4]. The LMS method eliminates noises like the Gaussian noise which arises from random fluctuations in EEG signals by adjusting the filter coefficients, so as to reduce the mean squared error between the filtered result and the desired signal.

3.2.2. Z-score normalization

Normalization is performed to carry two signals for the same or predefined series. The distinctive sample of the pre-defines series is a statistical discernment of normalization which converts the signal where the value of mean is 1, and that of the standard deviation is 1. In this research, z-score method [4] is performed for normalization and the z-score method reveals the classification performance through signal flattening. The numerical expression for the z-score value is given as Eq. (1).

$$z - score = \frac{score - mean}{standard\ deviation} \tag{1}$$

The score is the data point, mean is the average of all data points and standard deviation is the amount of variation in data. The normalization preserved correlation among normalized and actual EEG signals reduce the selection bias. The z-score normalization standardizes EEG data and enhances the feature discrimination.

3.3 Feature extraction

The pre-processed signals are given as input to the feature extraction phase for extracting the time

Table 2. CHB-MIT EEG dataset description

Records number	Patient's ID	Age	Gender	Count of seizures	Duration in hours
Chb01	1-1	11	F	7	45.00
Chb02	2	11	M	3	39.57
Chb03	3	14	F	7	57.87
Chb04	4	22	M	4	154.41
Chb05	5	7	F	5	38.09
Chb06	6	1.5	F	10	89.25
Chb07	7	14.5	F	3	67.23
Chb08	8	3.5	M	5	26.38
Chb09	9	10	F	4	65.92
Chb10	10	3	M	7	72.49
Chb11	11	12	F	3	73.30
Chb12	12	2	F	40	N/A
Chb14	14	9	F	8	41.50
Chb15	15	16	M	20	62.29
Chb16	16	7	F	10	17.03
Chb17	17	12	F	3	34.11
Chb18	18	18	F	6	62.29
Chb19	19	19	F	3	61.58
Chb20	20	6	F	8	41.43
Chb21	1-2	13	F	4	55.71
Chb22	21	9	F	3	75.93
Chb23	22	6	F	7	70.90

and frequency features. The Short-Time Frequency Transform (STFT) and Continuous Wavelet Transform (CWT) techniques are used in this research to extract the time and frequency domains. The brief explanation of these techniques is explained below:

3.3.1. Short time-frequency transform (STFT)

The STFT extracts the representations of time-frequency in EEG signals which gives significant features for epileptic seizure classification. The features of time-frequency capture variations in spectra over short time intervals and provide discriminative features that differentiate between seizure and non-seizure. In STFT, the non-stationary signals are separated to little segments and those segments are taken as sequential, and therefore FT is used for every portion. These portions are acquired through utilizing the windowing function and the technique is known as signals windowing. By using STFT, the time-dependent signals are stated in time and frequency axes. The numerical expression for STFT ($\gamma(w,\tau)$)is given as Eq. (2),

$$\gamma(w,\tau) = STFT\{f(t) = \int f(t)W(t-\tau)e^{-jwt}dt \quad (2)$$

The $f(t)$ denotes time domain signal, W denotes function of windowing, w denotes the parameter of frequency, t denotes the parameter of time, $\gamma(w,\tau)$ denotes the outcome of SIFT, e^{-jwt} represents the exponential function and τ denotes the parameter of slow time. In this research, the hamming is utilized as a windowing function in STFT. For BONN-EEG and CHB-MIT EEG datasets, the size of window is decided as 4,128.64 and parameter has number of points for overlapping among windows is utilized and determined as 2,64,32.

3.3.2. Continuous wavelet transform (CWT)

The CWT gives effective localization of time-frequency features that allow simultaneous analysis of frequency and time characteristics of EEG signals. The CWT is used in this research to minimize the loss of resolution produced through selection of window size in STFT. The CWT utilizes a windowing function known as mother wavelet and variance among this windowing process, while windowing is deployed in STFT and is scalable. When processing the CWT, wavelet function is shifted in time and scale, and its mathematical expression is given as Eq. (3),

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}}\Psi\left(\frac{t-b}{a}\right) \; a,b \in R, a \neq 0 \quad (3)$$

In the above Eq. (3), the parameter a refers to the scaling, b refers to the translation, $\Psi(t)$represents the mother function. The parameters of low-scale compress signals when parameters of huge-scale expand the signals. The parameters of high-scale capture lesser frequencies, whereas the parameters of low-scale capture the high frequencies. The CWT is determined as signal integral to be examined with difficult conjugate of wavelet function and its mathematical formulation is given as Eq. (4),

$$CWT\{f(t), a, b\} = \int_{-\infty}^{+\infty} f(t)\Psi^{*}{}_{a,b}(t)\,dt \quad (4)$$

Where, the $\Psi^{*}{}_{a,b}(t)$ represents complex conjugate in scaled and translated mother wavelet. The low-frequency data is taken for EEG signals due to delta band of EEG signals having less frequencies.

3.4 Classification using hierarchical long short-term memory

The extracted time and frequency features are given as input to the classification stage to classify the epileptic seizure using hierarchical LSTM. The adjacent sequences define the segments of EEG data which are consecutive and temporal. The LSTM is used for classification which captures long-term dependencies and works well on the sequence data (i.e. EGG signals). In hierarchical LSTM, the multi-layers of LSTM are organized hierarchically, with every layer processing the input sequence at various layers. The traditional LSTM concentrates on capturing the temporal dependencies with individual sequences, but the hierarchical LSTM extends its ability to capture hierarchical temporal dependency across multiple levels of representation. The hierarchical LSTM organizes input into multi-levels of representations, where every representation corresponds to various temporal scales. At every level of the hierarchy, an individual LSTM layer processes the input sequence and learns the temporal dependency with that level. The output representation from low levels is integrated to develop an input sequence for high levels, intern allowing the hierarchical LSTM to capture high-level temporal dependencies between adjacent sequences. In this research, hierarchical LSTM is developed for capturing two insights into EEG sequences. Initially, the method captures the correlation of local temporal context, alongside the channel correlation of every sample. The initial layer of hierarchical LSTM is

Received: May 19, 2024. Revised: June 13, 2024.

sample encoder layer, that captures local temporal correlations among samples with epochs. Additionally, the skip connections are introduced between hierarchical LSTM layers to facilitate the information flow across adjacent sequences. By using a skip connection, the learned features are transferred from one layer of the encoder to the respective layer of the decoder. The skip connection allows the hierarchical LSTM network to bypass certain layers and propagate data much more effectively. This process enhances the ability of the network to capture high-level temporal dependencies between adjacent sequences and gradients during training.

To formulate sample encoder for j^{th} epoch, $X_j = \{x_{j1}, x_{j2}, \ldots, x_{jT}\}$ describes EEG sample in j epoch, $x_{jt} \in R^{k\times1}, t \in \{1,2, \ldots T\}$. The hierarchical LSTM introduced with every time stage method for representing the hidden state is $h_{jt} \in R^{L\times1}$ and is executed as $h_{jt} = f\big(W_{jxh}x_{jt} + W_{jhh}h_{j(t-1)} + b_{jh}\big)$. Here, $W_{jxh} \in R^{T\times k}$ and $W_{jhh} \in R^{T\times T}$ represent the weight matrix integrated with x_{jt} input vector, where $h_{j(t-1)}$ represents the previous time step. For obtaining attention, a layer Multi-Layer Perceptron (MLP) is initially utilized for hidden representation r_{jt} for h_{jt} at every time step, and the numerical expression is given as Eq. (5),

$$r_{jt} = \tanh\big(W_{local}h_{jt} + b_{local}\big) \tag{5}$$

In the above eq (5), the $W_{local} \in R^{L\times1}$ represents the weight vector, h_{jt} represents hidden state and b_{local} represents the bias vector. The normalized weight α_{jt} is measured through comparing similarities with r_{local} context vector and numerical expression is given as Eq. (6),

$$\alpha_{jt} = \frac{\exp\,(r_{jt}^T r_{local}}{\sum_t \exp\big(r_{jt}^T r_{local}\big)} \tag{6}$$

The numerical expression for e_j is the last representation of epoch j as given as Eq. (7),

$$e_j = \sum_t \alpha_{jt} h_{jt} \tag{7}$$

The inputs for the epoch encoder layer are represented as p outcomes from the encoder layer $e_1, e_2, \ldots, e_p$. In that, $e_j \in R^{L\times1}$ for $j \in \{1,2, \ldots, P\}$. Next, another LSTM layer is utilized to encode the correlation of temporal context in epoch representation to produce the hidden unit $h_j \in R^{L\times1}\forall j$ as outcome. The completed representation of a fully connected layer of sequence is given as Eqs. (8) – (10),

$$r_j = \tanh\big(W_{global}h_j + b_{global}\big) \tag{8}$$

$$\alpha_j = \frac{\exp\big(r_j^T r_{global}\big)}{\sum_j \exp\big(r_j^T r_{global}\big)} \tag{9}$$

$$v = \sum_i \alpha_j h_j \tag{10}$$

In above eq (8)-(10), the r_j represents the output of neuron, h_j represents the input to that neuron, $W_{global} \in R^{L\times1}$ and b_{global} represent weights and bias vectors, while the α_j represents the corresponding weight, v represents the final representation of the sequence.

3.4.1. Skip connections

Skip connections are introduced between hierarchical LSTM layers to facilitate the information flow across adjacent sequences. The skip connection allows the hierarchical LSTM network to bypass certain layers and propagate data much more effectively. This process enhances the ability of the network to capture high-level temporal dependencies between adjacent sequences and gradients during the training. The numerical expression for skip connection is given as Eq. (11),

$$skip = Multiply()([skip, gate]) \tag{11}$$

In the above Eq. (11), the $Multiply()$ represents the element-wise multiplication between skip connection and gate output, the $skip$ represents the previous layer and $gate$ represents the corresponding element.

3.4.2. Network training

The v vector of sequence describes many essential and robust features. The hierarchical LSTM with skip connection is trained by N training samples $\{(X_1, y_1), \ldots, (X_N, y_N)\}$, the y_n represents the label of EEG sequence. The training process is done through reducing cross-entropy error over training samples. The numerical expression is given as Eq. (12),

$$E(\theta) = \frac{1}{N}\sum_{n=1}^{N} y_n \log\big(y_n(X_n, \theta)\big) + \frac{\lambda}{2}\|\theta\|_2^2 \tag{12}$$

In the above Eq. (12), θ represents collection network parameters and hyper-parameters, $E(\theta)$ represents cost function, N represents total number of training samples, X_n represents input features, y_n

corresponding label and λ represents the Lagrange multiplier. The Adam optimizer is deployed for performing the optimization. The LMS method and z-score normalization are used in the pre-processing stage to remove the noise and standardize the data. Then, the features are extracted by using the SIFT and CWT techniques to differentiate the variance between seizure and non-seizure. Next, the extracted features are classified by using the hierarchical LSTM with a skip connection-based classification method. By using the skip connection method in hierarchical LSTM, it facilitates the flow of data across adjacent sequences which maximizes the performance of seizure classification.

4. Experimental analysis

The proposed Hierarchical LSTM with skip connection method is simulated with MATLAB 2020a environment and the system requirements are Windows 10, i7 processor, and 16 GB RAM. The evaluation measures utilized to analyze performance of proposed Hierarchical LSTM with skip connection are accuracy, sensitivity, specificity, f1-score, and false detection rate (FDR). The numerical expressions for evaluation measures are given as Eqs. (13) – (15),

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN} \times 100 \tag{13}$$

$$Sensitivity = \frac{TP}{TP+FN} \times 100 \tag{14}$$

$$Specificity = \frac{TN}{TN+FP} \times 100 \tag{15}$$

In the above equations (13) – (15), the TP represents truly positive and TN represents true negative that refers to number of seizure and non-seizure segments that are accurately classified through the proposed Hierarchical LSTM with the skip connection method. The FP represents false positive which states that number of non-seizure EEG segments incorrectly classified through the proposed method and FN represents false negative that incorrectly labels the seizure segments. For every patient, FDR is referred to as mean amount of false detections per hour in non-seizure periods.

4.1 Quantitative and qualitative analysis

Table 3 represent the performance of the Hierarchical LSTM method on the BONN-EEG dataset with different evaluation measures. The existing neural networks are considered to analyze the proposed Hierarchical LSTM without skip connection method are Convolutional Neural Network (CNN), Multi-Layer Perceptron (MLP), Recurrent Neural Network (RNN), and LSTM. The hierarchical LSTM attains 98.27% accuracy, 97.62% sensitivity, 97.17% specificity, 97.45% f1-score and

Table 3. Performance of Hierarchical LSTM method on BONN-EEG dataset

Methods	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-score (%)	FDR (%)
CNN	95.04	94.67	94.27	93.90	0.84
MLP	95.82	95.16	95.16	94.73	0.77
RNN	96.49	96.03	95.71	95.36	0.73
LSTM	97.32	96.76	96.56	96.26	0.61
Hierarchical LSTM	98.27	97.62	97.17	97.45	0.54

Table 4. Performance of proposed Hierarchical LSTM with skip connection method on BONN-EEG dataset

Methods	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-score (%)	FDR (%)
CNN	96.05	95.72	95.37	95.54	0.75
MLP	96.40	96.26	96.01	96.12	0.62
RNN	97.49	97.18	96.68	96.93	0.51
LSTM	98.34	98.04	98.46	98.26	0.44
Hierarchical LSTM with skip connection	99.81	99.87	99.75	99.79	0.33

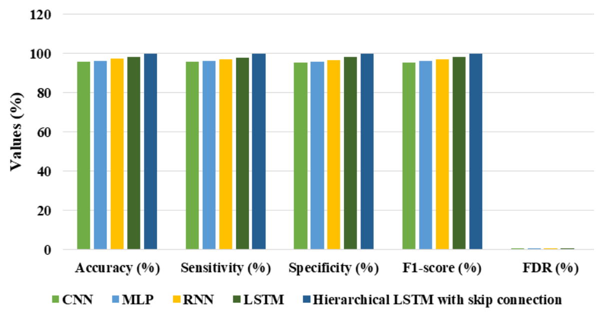

Figure. 2 Performance of proposed Hierarchical LSTM with skip connection method on BONN-EEG dataset

Table 5. Performance of Hierarchical LSTM method on the CHB-MIT EEG dataset

Methods	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-score (%)	FDR (%)
CNN	94.03	93.78	93.45	93.02	0.79
MLP	94.68	94.32	94.04	93.82	0.73
RNN	95.57	95.42	95.17	95.01	0.66
LSTM	96.46	96.34	96.02	95.67	0.57
Hierarchical LSTM	97.53	97.02	96.65	96.82	0.45

Table 6. Performance of proposed Hierarchical LSTM with skip connection method on CHB-MIT EEG dataset

Methods	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-score (%)	FDR (%)
CNN	96.17	94.16	96.03	95.58	0.63
MLP	96.67	95.02	96.63	95.67	0.58
RNN	97.23	95.56	97.47	96.65	0.47
LSTM	98.78	96.23	98.03	97.27	0.35
Hierarchical LSTM with skip connection	99.34	96.86	98.62	97.54	0.27

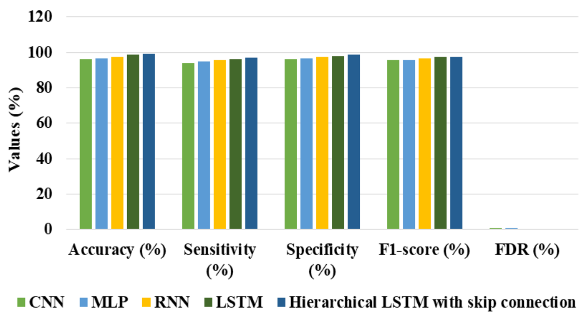

Figure. 3 Performance of proposed Hierarchical LSTM with skip connection method on CHB-MIT EEG dataset

0.54% of FDR. The hierarchical LSTM performs better on BONN-EEG dataset than other existing neural networks. By using the hierarchical LSTM-based classification method, it captures both the low and high-level features between adjacent sequences to increase the performance of epileptic seizure classification performance. Then, the skip connection is introduced between Hierarchical LSTM method to facilitate flow of data between adjacent sequences which increases the performance of the hierarchical LSTM method for epileptic seizure classification.

Table 4 and Fig. 2 display the performance of the proposed Hierarchical LSTM with skip connection method on the BONN-EEG dataset with different evaluation measures. The existing neural networks considered to analyze the proposed Hierarchical LSTM with the skip connection method are CNN, MLP, RNN, and LSTM. The hierarchical LSTM with the skip connection method attains 99.81% accuracy, 99.87% sensitivity, 99.75% specificity, 99.79% f1-score and 0.33% of FDR. The hierarchical LSTM with the skip connection method performs well on the BONN-EEG dataset than other existing neural networks.

Table 5 exhibit the outcomes of Hierarchical LSTM method on the CHB-MIT EEG dataset with different evaluation measures. The existing neural networks considered to analyze the proposed Hierarchical LSTM without the skip connection method are CNN, MLP, RNN, and LSTM. The hierarchical LSTM method achieves 97.53% accuracy, 97.02% sensitivity, 96.65% specificity, 96.82% f1-score and 0.45% of FDR. The hierarchical LSTM performs well on the CHB-MIT EEG dataset than other existing neural networks.

Table 6 and Fig. 3 represent performance of the proposed Hierarchical LSTM with skip connection method on the CHB-MIT EEG dataset with different evaluation measures. The existing neural networks considered to analyze the proposed Hierarchical LSTM with the skip connection method are CNN, MLP, RNN, and LSTM. The hierarchical LSTM with the skip connection method achieves 99.34% accuracy, 99.86% sensitivity, 98.62% specificity, 97.54% f1-score and 0.27% of FDR. The hierarchical LSTM with the skip connection method performs in a superior manner on the CHB-MIT EEG dataset than other existing neural networks.

4.2 Comparative analysis

The proposed Hierarchical LSTM with skip connection method's performance is compared to other existing methods like EMD+MIBIF+MLPNN [17], PCNN-LSTM [18], Bi-GRU [19] and GCN [20] on BONN-EEG and CHB-MIT EEG datasets. The performance of proposed Hierarchical LSTM with the skip connection method is evaluated with evaluation measures of accuracy, sensitivity, specificity, and FDR.

Table 7. Comparative analysis of proposed Hierarchical LSTM with skip connection method on two datasets

Dataset	Methods	Accuracy (%)	Sensitivity (%)	Specificity (%)	FDR (%)
Bonn-EEG	EMD+MIBIF+MLPNN [17]	99.54	N/A	N/A	N/A
	PCNN-LSTM [18]	99.75	99.83	99.62	N/A
	Proposed Hierarchical LSTM with skip connection	99.81	99.87	99.75	0.33
CHB-MIT EEG	PCNN-LSTM [18]	99.12	96.75	97.49	N/A
	Bi-GRU [19]	98.49	93.89	98.49	0.31
	GCN [20]	N/A	96.51	N/A	N/A
	Proposed Hierarchical LSTM with skip connection	99.34	96.86	98.62	0.27

The proposed Hierarchical LSTM with skip connection method attains 99.81% accuracy on BONN – EEG and 99.34% accuracy on the CHB-MIT EEG dataset. The Hierarchical LSTM captures the high-dimensional data between adjacent sequences and skips connection between LSTM layers, thus facilitating the flow of information across adjacent sequences which helps improve the performance of epileptic seizure classification. Table 7 represents the comparative analysis of Hierarchical LSTM with the skip connection method.

4.3 Discussion

The section explains the results observed from Hierarchical LSTM with skip connection method for improving the classification process of epileptic EEG seizures. The EMD+MIBIF+MLPNN [17] method has the issue of gradient vanishing. The PCNN-LSTM [18] method has difficulty in capturing the high dimensional data between adjacent sequences of EEG. The Bi-GRU [19] method suffers from high classification error, while the GCN [20] method suffers from lesser effective flow of information across adjacent sequences. To overcome these limitations, the proposed Hierarchical LSTM with skip connection method exhibits commendable classification performance than the previous methods. The hierarchical LSTM-based classification method is used to capture both low and high-level features between adjacent sequences for maximizing the classification performance of epileptic seizures. The skip connection is introduced between the Hierarchical LSTM method to facilitate the data flow between adjacent sequences for maximizing the performance of the hierarchical LSTM method.

5. Conclusion

This research proposes a Hierarchical LSTM with a skip connection-based classification method to capture the high dimensional data between adjacent sequences. The hierarchical LSTM captures both short and long-term dependencies between adjacent sequences in high-dimensional data. The skip connection is introduced between hierarchical LSTM layers, therefore facilitating the flow of data across adjacent sequences to improve the classification performance of hierarchical LSTM. The datasets used to evaluate the proposed Hierarchical LSTM with the skip connection method are BONN-EEG and CHB-MIT EEG datasets. The proposed Hierarchical LSTM with skip connection method attains 99.81% accuracy, 99.87% sensitivity, 99.75% specificity and 0.33% FDR on BONN – EEG and 99.34% accuracy, 96.86% sensitivity, 98.62% specificity and 0.27% FDR on the CHB-MIT EEG dataset which is more effective than the existing methods of Bi-GRU and GCN. In the future, an optimization-based feature selection method can be used to eliminate the irrelevant features to further improve the classification performance.

Notation

Notations	Description
$f(t)$	Time Domain Signal
W	Windowing Function
w	Parameter Of Frequency
$\gamma(w,\tau)$	Outcome Of SIFT
$X_j = \{x_{j1}, x_{j2}, \dots, x_{jT}\}$	EEG Sample In j Epoch
$h_{jt} \in R^{L\times 1}$	Hidden State
$W_{jxh} \in R^{T\times k}$	Weight Matrix Integrated With x_{jt} Input Vector
$h_{j(t-1)}$	Previous Time Step
$W_{local} \in R^{L\times 1}$	Weight Matrix
b_{local}	Bias Vector
e_j	Last Representation of Epoch j
$Multiply()$	Element-Wise Multiplication
y_n	Label Of EEG Sequence
θ	Network Parameters
λ	Lagrange Multiplier
e^{-jwt}	Exponential Function
$\Psi(t)$	Mother Function
$\Psi^*{}_{a,b}(t)$	Complex Conjugate
α_{jt}	Normalized Weight
$E(\theta)$	Cost Function
X_n	Input Features
y_n	Corresponding Labels

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization, NV; methodology, VKG; software, VKG; validation, VKG; formal analysis, NV; investigation, NV; resources, VKG; data curation, NV; writing—original draft preparation, NV; writing—review and editing, VKG; visualization, NV; supervision, VKG; project administration, VKG.

References

[1] F. Hassan, S.F. Hussain, and S.M. Qaisar, "Epileptic seizure detection using a hybrid 1D CNN-machine learning approach from EEG data", *Journal of Healthcare Engineering*, Vol. 2022, p. 9579422, 2022.

[2] X. Liu, J. Wang, J. Shang, J. Liu, L. Dai, and S. Yuan, "Epileptic Seizure Detection Based on Variational Mode Decomposition and Deep Forest Using EEG Signals", *Brain Sciences*, Vol. 12, p. 1275, 2022.

[3] C. Wang, L. Liu, W. Zhuo, and Y. Xie, "An epileptic EEG detection method based on data augmentation and lightweight neural network", *IEEE Journal of Translational Engineering in Health and Medicine*, Vol. 22, pp. 22-31, 2024.

[4] I. Palanichamy, and F.B.B. Ahamed, "Prediction of Seizure in the EEG Signal with Time Aware Recurrent Neural Network", *Revue d'Intelligence Artificielle*, Vol. 36, No. 5, pp. 717-724, 2022.

[5] X. Jiang, X. Liu, Y. Liu, Q. Wang, B. Li, and L. Zhang, "Epileptic seizures detection and the analysis of optimal seizure prediction horizon based on frequency and phase analysis", *Frontiers in Neuroscience*, Vol. 17, p. 1191683, 2023.

[6] G. Kumar, S. Chander, and A. Almadhor, "An intelligent epilepsy seizure detection system using adaptive mode decomposition of EEG signals", *Physical and Engineering Sciences in Medicine*, Vol. 45, No. 1, pp. 261-272, 2022.

[7] M. Varlı, and H. Yılmaz, "Multiple classification of EEG signals and epileptic seizure diagnosis with combined deep learning", *Journal of Computational Science*, Vol. 67, p. 101943, 2023.

[8] Y. Gao, X. Chen, A. Liu, D. Liang, L. Wu, R. Qian, H. Xie, and Y. Zhang, "Pediatric seizure prediction in scalp EEG using a multi-scale neural network with dilated convolutions", *IEEE Journal of Translational Engineering in Health and Medicine*, Vol. 10, p. 4900209, 2022.

[9] W. Chen, Y. Wang, Y. Ren, H. Jiang, G. Du, J. Zhang, and J. Li, "An automated detection of epileptic seizures EEG using CNN classifier based on feature fusion with high accuracy", *BMC Medical Informatics and Decision Making*, Vol. 23, No. 1, p. 96, 2023.

[10] P. Kunekar, M.K. Gupta, and P. Gaur, "Detection of epileptic seizure in EEG signals using machine learning and deep learning techniques", *Journal of Engineering and Applied Science*, Vol. 71, No. 1, p. 21, 2024.

[11] Y. Zaid, M. Sah, and C. Direkoglu, "Pre-processed and combined EEG data for epileptic seizure classification using deep learning", *Biomedical Signal Processing and Control*, Vol. 84, p. 104738, 2023.

[12] N. Abderrahim, A. Echtioui, R. Khemakhem, W. Zouch, M. Ghorbel, and A. Ben Hamida, "Epileptic Seizures Detection Using iEEG Signals and Deep Learning Models", *Circuits, Systems, and Signal Processing*, Vol. 43, No. 3, pp. 1597-1626, 2024.

[13] W. Zeng, L. Shan, B. Su, and S. Du, "Epileptic seizure detection with deep EEG features by convolutional neural network and shallow classifiers", *Frontiers in Neuroscience*, Vol. 17, p. 1145526, 2023.

[14] A. K. R. Britto, S. Srinivasan, S. K. Mathivanan, M. Venkatesan, B. A. M. B. Malar, S. Mallik, and H. Qin, "A multi-dimensional hybrid CNN-BiLSTM framework for epileptic seizure detection using electroencephalogram signal scrutiny", *Systems and Soft Computing*, Vol. 5, p. 200062, 2023.

[15] G. Yogarajan, N. Alsubaie, G. Rajasekaran, T. Revathi, M. S. Alqahtani, M. Abbas, M. M. Alshahrani, and B. O. Soufiene, "EEG-based epileptic seizure detection using binary dragonfly algorithm and deep neural network", *Scientific Reports*, Vol. 13, No. 1, p. 17710, 2023.

[16] A. A. E. Shoka, M. M. Dessouky, A. El-Sayed, and E. E. D. Hemdan, "An efficient CNN based epileptic seizures detection framework using encrypted EEG signals for secure telemedicine applications", *Alexandria Engineering Journal*, Vol. 65, pp. 399-412, 2023.

[17] K. M. Hassan, M. R. Islam, T. T. Nguyen, and M. K. I. Molla, "Epileptic seizure detection in EEG using mutual information-based best individual feature selection", *Expert Systems with Applications*, Vol. 193, p. 116414, 2022.

[18] A. H. Abdulwahhab, A. H. Abdulaal, A. H. T. Al-Ghrairi, A. A. Mohammed, and M. Valizadeh, "Detection of epileptic seizure using EEG signals analysis based on deep learning techniques", *Chaos, Solitons & Fractals*, Vol. 181, p. 114700, 2024.

[19] Y. Zhang, S. Yao, R. Yang, X. Liu, W. Qiu, L. Han, W. Zhou, and W. Shang, "Epileptic seizure detection based on bidirectional gated recurrent unit network", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 30, pp. 135-145, 2022.

[20] M. Jia, W. Liu, J. Duan, L. Chen, C. L. P. Chen, Q. Wang, and Z. Zhou, "Efficient graph convolutional networks for seizure prediction using scalp EEG", *Frontiers in Neuroscience*, Vol. 16, p. 967116, 2022.

[21] M. S. Islam, K. Thapa, and S.-H. Yang, "Epileptic-Net: An Improved Epileptic Seizure Detection System Using Dense Convolutional

Block with Attention Network from EEG", *Sensors*, Vol. 22, p. 728, 2022.

[22] J. S. Ra, and T. Li, "A novel epileptic seizure prediction method based on synchroextracting transform and 1-dimensional convolutional neural network", *Computer Methods and Programs in Biomedicine*, Vol. 240, p. 107678, 2023.

[23] BONN-EEG dataset: https://www.upf.edu/web/ntsa/downloads/-/asset_publisher/xvT6E4pczrBw/content/2001-indications-of-nonlinear-deterministic-and-finite-dimensional-structures-in-time-series-of-brain-electrical-activity-dependence-on-recording-regi

[24] CHB-MIT EEG dataset: https://physionet.org/content/chbmit/1.0.0/