Undergraduate Conference on Intelligent Computing and Systems (UCICS 2025)
26–27 February, 2025; Varendra University, Rajshahi, Bangladesh

Attention-Based LSTM System for Epileptic Seizure Detection from EEG Signals

Sakir Ahamed
Computer Science and Engineering
University of Rajshahi
Rajshahi, Bangladesh
sakir.ahamed96@gmail.com

Nuhan Intekhab
Computer Science and Engineering
Varendra University
Rajshahi, Bangladesh
intekhabnuhan@gmail.com

Samira Tareque
Computer Science and Engineering
Varendra University
Rajshahi, Bangladesh
samira@vu.edu.bd

Sabina Yasmin
Computer Science and
Engineering Varendra University
Rajshahi, Bangladesh
sabina@vu.edu.bd

MSU Zaman
Computer Science and Engineering
Rajshahi University of Engineering and
Technology Rajshahi, Bangladesh
zaman@cse.ruet.ac.bd

Md. Khademul Islam Molla
Computer Science and Engineering
University of Rajshahi
Rajshahi, Bangladesh
khademul.cse@ru.ac.bd

***Abstract*—The electroencephalogram (EEG) has become one of the most important tools for clinicians to detect seizures and other neurological irregularities of the human brain over the past few decades. An accurate diagnosis of epilepsy is essential due to its unique characteristics and the adverse consequences of epileptic seizures on individuals. An urgent need exists for an automated epilepsy detection system utilizing electroencephalography (EEG) for clinical use. This paper employs the Discrete Wavelet Transform (DWT) to decompose EEG signals into multiple subbands. Various features used to discriminate spike events and extracted from each subband signal of an EEG trial. The attention mechanism augments the network's capacity to concentrate on discriminative features and temporal steps within the feature sequence, thereby enhancing interpretability and detection precision. The weighted number of features effectively distinguishes the underlying characteristics of EEG signals indicative of seizure and non-seizure events. The attention mechanism utilizing Long Short-Term Memory (LSTM) is employed to classify seizure and non-seizure EEG signals. The effectiveness of the proposed method is assessed through multiple experiments utilizing a public dataset acquired from the University of Bonn. The experimental findings indicate that the proposed seizure detection method obtains a classification accuracy of 99.65%, outperforming the performance of current techniques. The efficacy of the LSTM with attention model is compared with support vector machine classifiers, which exhibit a classification accuracy of 98.52%. Thus, the proposed method is validated as a potential indicator for EEG-based seizure detection.**

***Index Terms*—electroencephalography, epilepsy, discrete wavelet transform, attention, LSTM, seizure**

I. INTRODUCTION

A seizure is considered a neurological malady of the central nervous system in which abnormal brain activity disrupts electrical signals. Seizures cause noticeable symptoms, unusual behaviors, sensations, or even loss of consciousness. Seizures may occur without visible symptoms [1]. Electroencephalography (EEG) is a useful tool for diagnosing epilepsy seizures. A robust algorithm for detecting seizures from these signals would greatly assist specialists. EEG captures the electrical brain activity that includes valuable pathological data. In identifying epileptic seizures, the identification of spikes in EEG patterns is employed [2]. In the analysis of brain disorders, particularly epileptic seizures, this EEG pattern acts as an essential character [3]. However, an automatic seizure detection system helps to monitor, diagnose, and rehabilitation in the long term [4]. Many methods improve classification by decomposing EEG signals into multiple levels. In this work, EEG signals are divided into narrow-band components. Discriminative features are extracted and dominant features are weighted using the attention mechanism for seizure detection using the model Long-Short-Term Memory (LSTM) improved version of recurrent neural network (RNN). This work introduced an automated system for seizure detection that utilizes the BONN EEG data set.

II. LITERATURE REVIEW

A large number of researchers have developed feature extraction techniques to automate epileptic seizures detection. Reviewing these methods is fundamental for understanding current approaches' strengths, seeing where improvements can be made, and identifying their limitations. An efficient feature vector can identify epileptic seizure events in the EEG signal in an acceptable timeframe, supported by a discriminative approach of feature section [5]. EEG signals are non-linear and non-stationary that is why it is very difficult to analyze [6]. EEG signals are non-linear and non-stationary, that is why it is very crucial to analyze [6]. For seizure detection, a hybrid system has been developed that integrates feature extraction using the Fast Fourier Transform and decision-making through a decision tree classifier, operating under the assumption that EEG signals remain stationary for short periods [7]. The method does not match the EEG characteristics, although it achieved accuracy of classification of 98.72%. For epileptic EEG detection, a lightweight PCNN-BiLSTM hybrid network has been developed, integrating a small window segmentation technique and SMOTE-based data augmentation for balanced classification, achieving 98.52% precision with only 9371 training parameters [8]. The deep learning approach is very computational complexity. To discriminate the non-stationary signal, the entropy-based features derived from EEG signals, various techniques have been created for epilepsy detection [9]. Another machine learning-based approach has been developed that uses fractal parameters, spectral entropy, and power spectral density, demonstrating high accuracy across multiple public datasets with a minimal feature set, but it has higher time complexity because of the machine learning model to get the better accuracy [10]. To address these problems, the discrete wavelet transform is employed to break down EEG signals into subbands for feature extraction and seizure classification. To evaluate the intricacy of EEG signals and enhance seizure detection, several spike-based features are extracted. The system effectively differentiates seizure signals from normal signals. Furthermore, adding an LSTM with attention increased the classification accuracy.

III. DATA DESCRIPTION AND PREPROCESSING

A. Dataset

The BONN dataset [11], presented by the Department of Epileptology, University of Bonn, Germany, is applied in the evaluation of this proposed method. This dataset is accessible to the public. It includes five sets of EEG data, labeled as A, B, C, D, and E. Each holds 100 single-channel EEG segments of 23.6 s long. Continuous multichannel recordings were used to derive the segments. To remove artifacts from eye movements and muscle activity, the segments were derived with prior visual.Sets A and B were acquired from surface EEG recordings of five healthy participants. The recordings were made using a standardized electrode placement, with set A captured during an awake state with eyes open and set B with eyes closed. Sets C, D, and E were extracted from presurgical diagnostic EEGs of five patients. These patients, who had confirmed epileptogenic zones, achieved full seizure control following hippocampal resection. Sets C and D were recorded during seizure-free intervals from the hippocampal formation of the opposite hemisphere and within the epileptogenic zone, respectively. Set E contained only seizure activity. All EEG signals were recorded using a consistent 128-channel amplification system with a standard common reference. The data were sampled at 173.61 Hz with 12-bit analog-to-digital conversion.

B. Data Preprocessing

Each set in this dataset consists of 100 segments, each segment lasting 23.6 seconds. The frame of each EEG segment is carried out with an 75% overlap of a feasible portion. To meet the requirements of the medical application, the frame size is set to 10 seconds [5].

IV. METHODOLOGY

The Subband feature extraction method played a vital role in extracting effective discriminative features [5]. Effective feature extraction methods, LSTM and Attention mechanism are utilized in this work. The introduced block diagram is demonstrated in Figure 1.

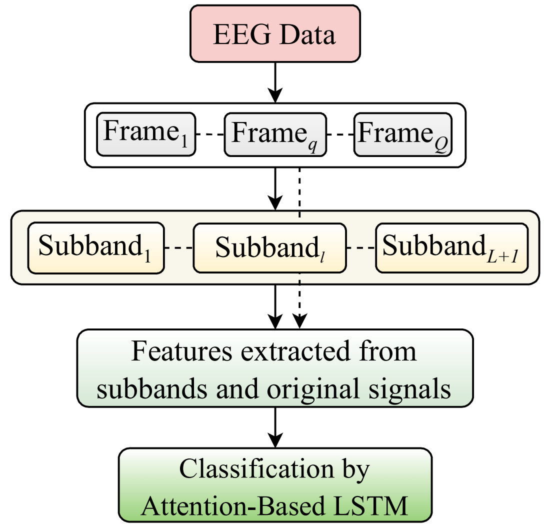

Figure 1. The block diagram of the proposed method.

A. Subband Decomposition

Subband decomposition is a signal processing method that is used to divide a signal into multiple frequency bands. Effective discriminative insights for seizure detection are offered by narrowband features [12]. Fourier transformation is not suitable for EEG signals analysis, since EEG signals are inherently nonstationary. EEG signals are more effectively analyzed using DWT, which is a widely utilized method for subband decomposition of EEG frames. This technique employs a filter bank subband decomposition approach based on an orthogonal basis, accomplished by convolving the original signal $x(n)$ with the filter. There are highpass filter h and lowpass filter g. The high-pass filter and lowpass filter compute the detail coefficient by $d^l(n) = \sum_k g(k) * a^{l-1}(2n-k)$c and the approximate coefficient by $a^l(n) = \sum_k h(k) * a^{l-1}(2n-k)$, where l is the current wavelet decomposition level, n is the number of time observations, $g(k)$ is the filter coefficient of a low-pass filter and $h(k)$ are the filter coefficients of high-pass filters. At level l, both the approximation and detail coefficients are exclusively derived from the approximation coefficients at the preceding level $l-1$. The db4 wavelet function is used to decompose EEG signals into subbands through L levels of DWT, producing one approximation and L detail coefficients. This results in $L+1$ reconstructed subband signals $(s_1, s_2, \ldots, s_{L+1})$. In this paper, five levels of wavelet decomposition is applied and six subbands are obtained (Sb_1 to Sb_6) with cutoff frequencies 43–86 Hz, 22–43 Hz, 11–22 Hz, 6–11 Hz, 3–6 Hz, and 0–3 Hz. Except for Sb_1, the other subbands combine components from multiple rhythms. However, Gamma rhythms ($> 30Hz$) which have played a vital role for seizure detection, other rhythms also help to detect the same task.

B. Feature Extraction

For classification of nonstationary EEG signals, extracting effective features is the most challenging step in signal processing. It helps extract discriminative features that help to improve the model accuracy. The variance [13], kurtosis [13] are calculated to capture the variability and distribution of the EEG data; the root mean square(rms) [13] is used to determine the effective value of the signal; the coefficient of variation (CV) is calculated as a quantitative measure of fluctuations in the amplitude of a signal [5]. The fluctuation index (Fi) is measured for rapid variations in the amplitude of a signal [5]. A feature vector is created that concatenates all features of each frame.

C. Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) is a specialized type of recurrent neural network (RNN) created to overcome the problem of dissolving gradient issue common to standard RNNs, enabling it to learn information across multiple time steps [14]. LSTM has input, forget, and output gates.

$$i_t = \delta(W_i[h_{t-1}, X_t] + b_i) \tag{1}$$

$$f_t = \delta(W_f[h_{t-1}, X_t] + b_f) \tag{2}$$

$$O_t = \delta(W_o[h_{t-1}, X_t] + b_o) \tag{3}$$

Here, i_t is the activation of the input gate at the time interval t, controlling which input information is updated in the cell state. f_t is the activation of the forget gate at the time step t, deciding which part of the previous cell state is discarded. O_t is the activation of the output gate at the time step t, determining the output and the contribution of the cell state to the hidden state. W_i, W_f, W_o are the weight matrices for the input gate, the forget gate, and the output gate, sequentially. b_i, b_f, b_o are the Bias vectors for the input, forget, and output gates, respectively. h_{t-1} is the hidden state of the previous time interval. X_t is the input at the current time interval t. δ is the activation function of the sigmoid. The lstm model captures long-term dependencies in sequential data that help to analyze the EEG signal, where patterns over time are critical for classification.

D. Attention

The encoder-decoder model was created for tasks such as text translation and question answering. It has been utilized in time series forecasting but faces limitations with long sequences due to its fixed-length vector representation. To address this limitation, the attention mechanism was introduced that mitigates the constraints and leverages information from the entire sequence [15]. The mechanism to calculate attention weight is shown below.

$$e_{ij} = \mathrm{a}\left(h_{i-1}, h_j\right) \tag{4}$$

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{T_x} \exp(e_{ik})} = \mathrm{softmax}(e_{ij}) \tag{5}$$

$$C_i = \sum_{j=1}^{t} \alpha_{ij} h_j \tag{6}$$

Here, The attentional weight α_{ij} denotes the allocation of attention, whereas e_{ij} measures the attentional correlation from moment j to moment i [15]. The non-linear function a assesses the hidden state h_j of the input sequence X_j at time j against the hidden state h_{i-1} of the generated output y_i at the preceding time in the decoder, quantifying the degree of correspondence between X_j and y_i. A greater match degree leads to increased e_{ij} and α_{ij}, representing a stronger impact of moment j on i. Thus, the attention mechanism significantly upgrades the encoder-decoder model.

E. Attention-Based LSTM

Long Short-Term Memory (LSTM) utilizes recurrent structures to handle sequential data, eliminating the need for external memory and reducing computational complexity. LSTM learning occurs in two phases. One of the phases is structure learning. Structure learning generates rules based on firing thresholds. The other phase is parameter learning, which minimizes error to optimize performance. To assign spatial and temporal firing strengths, Gaussian membership functions are used. Combining an attention mechanism with LSTM increases the model's ability to focus on critical EEG features and improves its ability to identify temporal patterns linked to seizures. The attention mechanism allows the model to emphasize and prioritize significant portions of the EEG signal, enhancing both its accuracy and reliability. The BONN dataset is used to evaluate this method, which has proven to be an effective and efficient solution for detecting seizure events. The proposed structure of this model is shown in Figure 2.

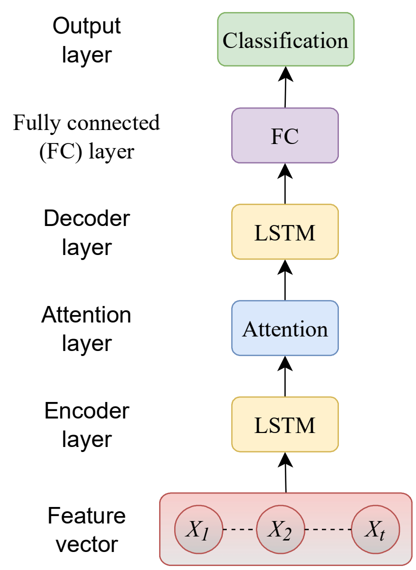

Figure 2. Proposed structure of LSTM-attention-LSTM model.

V. EXPERIMENTAL RESULTS AND DISCUSSION

The proposed LSTM with Attention model produces higher-level performance with an average accuracy of 99.65%. In addition, this proposed method achieved an average accuracy 98.52% using an SVM [16]. The dataset is categorized into four cases to evaluate the proposed method. A feature vector is created for each instance and assessed individually. The precision for each case is shown in Table I.

TABLE I
DIFFERENT CASES ACCURACY FOR EPILEPTIC SEIZURE DETECTION

Case	Class1 - Class2	Accuracy (%)
1	A-E	99.92
2	B-E	99.67
3	C-E	99.75
4	D-E	99.25

The proposed method surpasses other advanced methodologies, such as LS-SVM model with 99.56% [17], convolution neural network-gated recurrent unit-attention mechanism (CNN-GRU-AM) with 99.35% [18], Bi-directional long short-term memory (Bi-LSTM) with 99.18% [19], and combined features with graph eigen decomposition (CFGED) with 99.55% [5]. Its ability to leverage both temporal EEG patterns and attention mechanisms without extensive preprocessing distinguishes it from conventional models, which often require significant computational resources and also feature selection methods. Hence, this proposed method advances the field of seizure detection. The comparison with recent studies on the Bonn dataset shown in Table II.

TABLE II
PERFORMANCE COMPARISON WITH OTHER SYSTEMS ON BONN DATASET

Model	**Accuracy (%)**
Bi-directional LSTM [19]	99.18
CNN-GRU-AM [18]	99.35
CFGED [5]	99.55
LS-SVM [17]	99.56
Proposed LSTM-attention-LSTM	**99.65**

VI. CONCLUSIONS

Detecting epileptic seizures automatically remains a challenge due to subject dependency and limited training data. In this paper, the BONN dataset is utilized to assess the performance of the introduced approach. The primary contribution is the creation of an efficient feature vector and a discriminative feature attention scheme for accurate seizure detection. Our proposed methodology allows the model to effectively learn temporal patterns from EEG signals while focusing on the most relevant features by merging LSTM and attention mechanisms. This approach eliminates the need for extensive preprocessing and feature selection, as well as reduces computational requirements compared to traditional models. Effective feature extraction techniques are used to ensure that only the most significant features are utilized and optimize classification performance. Although the suggested approach attains the highest accuracy in this benchmark dataset, its generalizability in diverse patient populations is limited due to subject dependency and lack of clinical validation. Future work includes addressing these limitations by conducting clinical trials and extending the LSTM with attention model to a wider range of datasets to enhance generalization capabilities. Furthermore, improving the model's scalability by integrating it into a more comprehensive deep neural network architecture is also considered future work.

REFERENCES

[1] L. Guo, D. Rivero, J. Dorado, J. R. Rabunal, and A. Pazos, "Automatic epileptic seizure detection in eegs based on line length feature and artificial neural networks," *Journal of neuroscience methods*, vol. 191, no. 1, pp. 101–109, 2010.

[2] S. Mukhopadhyay and G. Ray, "A new interpretation of nonlinear energy operator and its efficacy in spike detection," *IEEE Transactions on biomedical engineering*, vol. 45, no. 2, pp. 180–187, 1998.

[3] S. Li, W. Zhou, Q. Yuan, S. Geng, and D. Cai, "Feature extraction and recognition of ictal eeg using emd and svm," *Computers in biology and medicine*, vol. 43, no. 7, pp. 807–816, 2013.

[4] F. Mormann, R. G. Andrzejak, C. E. Elger, and K. Lehnertz, "Seizure prediction: the long and winding road," *Brain*, vol. 130, no. 2, pp. 314–333, 2007.

[5] M. K. I. Molla, K. M. Hassan, M. R. Islam, and T. Tanaka, "Graph eigen decomposition-based feature-selection method for epileptic seizure detection using electroencephalography," *Sensors*, vol. 20, no. 16, p. 4639, 2020.

[6] R. B. Pachori and S. Patidar, "Epileptic seizure classification in eeg signals using second-order difference plot of intrinsic mode functions," *Computer methods and programs in biomedicine*, vol. 113, no. 2, pp. 494–502, 2014.

[7] K. Polat and S. Güneş, "Classification of epileptiform eeg using a hybrid system based on decision tree classifier and fast fourier transform," *Applied Mathematics and Computation*, vol. 187, no. 2, pp. 1017–1026, 2007.

[8] 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*, 2023.

[9] N. S. Tawfik, S. M. Youssef, and M. Kholief, "A hybrid automated detection of epileptic seizures in eeg records," *Computers & Electrical Engineering*, vol. 53, pp. 177–190, 2016.

[10] S. Abhishek, S. Kumar, N. Mohan, and K. Soman, "Eeg based automated detection of seizure using machine learning approach and traditional features," *Expert Systems with Applications*, vol. 251, p. 123991, 2024.

[11] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Physical Review E*, vol. 64, no. 6, p. 061907, 2001.

[12] D. Wang, D. Miao, and C. Xie, "Best basis-based wavelet packet entropy feature extraction and hierarchical eeg classification for epileptic detection," *Expert Systems with Applications*, vol. 38, no. 11, pp. 14314–14320, 2011.

[13] D. Zwillinger and S. Kokoska, *CRC standard probability and statistics tables and formulae*. Crc Press, 1999.

[14] A. Sherstinsky, "Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network," *Physica D: Nonlinear Phenomena*, vol. 404, p. 132306, 2020.

[15] D. Bahdanau, "Neural machine translation by jointly learning to align and translate," *arXiv preprint arXiv:1409.0473*, 2014.

[16] M. Nandan, S. S. Talathi, S. Myers, W. L. Ditto, P. P. Khargonekar, and P. R. Carney, "Support vector machines for seizure detection in an animal model of chronic epilepsy," *Journal of Neural Engineering*, vol. 7, no. 3, p. 036001, 2010.

[17] E. D. Übeyli, "Least squares support vector machine employing model-based methods coefficients for analysis of eeg signals," *Expert Systems with Applications*, vol. 37, no. 1, pp. 233–239, 2010.

[18] J. Zhang, S. Zheng, W. Chen, G. Du, Q. Fu, and H. Jiang, "A scheme combining feature fusion and hybrid deep learning models for epileptic seizure detection and prediction," *Scientific Reports*, vol. 14, no. 1, p. 16916, 2024.

[19] S. Srinivasan, S. Dayalane, S. k. Mathivanan, H. Rajadurai, P. Jayagopal, and G. T. Dalu, "Detection and classification of adult epilepsy using hybrid deep learning approach," *Scientific Reports*, vol. 13, no. 1, p. 17574, 2023.