



Research article

Epileptic seizure detection in EEG signals via an enhanced hybrid CNN with an integrated attention mechanism

Sakorn Mekruksavanich¹, Wikanda Phaphan² and Anuchit Jitpattanakul^{3,4,*}

¹ Department of Computer Engineering, School of Information and Communication Technology, University of Phayao, Phayao 56000, Thailand

² Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok 10800, Thailand

³ Department of Mathematics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok 10800, Thailand

⁴ Intelligent and Nonlinear Dynamic Innovations Research Center, Science and Technology Research Institute, King Mongkut's University of Technology North Bangkok, Bangkok 10800, Thailand

* **Correspondence:** Email: anuchit.j@sci.kmutnb.ac.th.

Abstract: Epileptic seizures, a prevalent neurological condition, necessitate precise and prompt identification for optimal care. Nevertheless, the intricate characteristics of electroencephalography (EEG) signals, noise, and the want for real-time analysis require enhancement in the creation of dependable detection approaches. Despite advances in machine learning and deep learning, capturing the intricate spatial and temporal patterns in EEG data remains challenging. This study introduced a novel deep learning framework combining a convolutional neural network (CNN), bidirectional gated recurrent unit (BiGRU), and convolutional block attention module (CBAM). The CNN extracts spatial features, the BiGRU captures long-term temporal dependencies, and the CBAM emphasizes critical spatial and temporal regions, creating a hybrid architecture optimized for EEG pattern recognition. Evaluation of a public EEG dataset revealed superior performance compared to existing methods. The model achieved 99.00% accuracy in binary classification, 96.20% in three-class tasks, 92.00% in four-class scenarios, and 89.00% in five-class classification. High sensitivity (89.00–99.00%) and specificity (89.63–99.00%) across all tasks highlighted the model's robust ability to identify diverse EEG patterns. This approach supports healthcare professionals in diagnosing epileptic seizures accurately and promptly, improving patient outcomes and quality of life.

Keywords: epileptic seizure detection; deep learning; CNN; residual network; bidirectional GRU; CBAM; EEG

1. Introduction

Epilepsy is a persistent condition of the brain marked by frequent and spontaneous seizures. It influences over fifty million individuals globally [1]. These epileptic episodes result from abnormal and inappropriate electrical reactions in the central nervous system, causing many signs, including unconsciousness, muscular contractions, and disruptions in sensory perception [2]. Precise and prompt designation of epileptic episodes is essential in the diagnosis, therapy, and control of epilepsy, eventually enhancing the quality of existence for individuals [3].

EEG is a commonly used method that does not need invasive procedures to observe neurological function and identify epileptic episodes [4]. By putting sensors on the top of the head, EEG detects the brain's electrical indications, providing crucial information about the timing and spatial distribution of neural activity [5]. However, manually interpreting EEG data is labor-intensive, demands considerable effort, and can yield inconsistent results among different evaluators [6]. Consequently, there is a growing demand for reliable automated methods to detect epileptic seizures through EEG analysis.

Deep learning algorithms have demonstrated considerable promise in various medical applications, including detecting epileptic seizures [7]. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are deep learning algorithms that can independently learn intricate features from unprocessed EEG data. These models can recognize spatial and temporal relationships within the data [8]. CNNs excel at identifying spatial patterns specific to a single region, while RNNs are particularly adept at capturing temporal connections that extend over prolonged periods [9].

Multiple research has used deep learning methods to identify epileptic episodes by analyzing EEG data. Shoeibi et al. [10] developed an approach using CNNs to categorize EEG patterns as epileptic or non-epileptic. This structure exhibited exceptional accuracy when assessed using a dataset that is readily available to anyone with access. Similarly, Xu et al. [11] integrated CNNs with long short-term memory (LSTM) networks to precisely express spatial and temporal characteristics. This technique demonstrated a higher level of effectiveness than conventional machine learning methodologies.

Although there has been advancement in using deep learning for seizure identification, more work still needs to be done in creating precise and resilient models that can successfully manage the intricate and diverse nature of EEG information. Residual networks, often known as ResNets, have been demonstrated to mitigate the issue of disappearing gradients and encourage more profound network training [12]. Nevertheless, the use of ResNets in the identification of epileptic seizures has been restricted.

Furthermore, attention mechanisms have gained popularity in deep learning due to their capability to selectively prioritize critical elements while disregarding unnecessary data [13]. One effective mechanism is the convolutional block attention module (CBAM) for refining feature maps by integrating spatial and channel-wise attention [14]. Integrating CBAM into deep learning models has led to advancements in various computer vision tasks [15]. While attention mechanisms have been applied to seizure detection, the specific integration of CBAM with CNN and residual BiGRU for EEG-based epileptic seizure detection represents a novel approach that has not been fully explored. Previous studies have not investigated how CBAM's dual attention mechanism (combining both channel and spatial attention) can enhance feature discrimination in the context of epileptic EEG signals when integrated with deep spatiotemporal architectures.

This research aims to overcome these constraints by introducing a hybrid deep learning method that combines a CNN and a residual BiGRU with a CBAM. The goal is to achieve accuracy identification of epileptic seizures. The main goals of this investigation are:

- 1) To develop a hybrid architecture combining CNN and residual BiGRU that efficiently captures spatial and temporal features from EEG data.
- 2) To boost the discriminative ability of the retrieved features, the proposed model will use the CBAM method.
- 3) To evaluate the model's performance; it will be tested using the publicly available EEG dataset and will be evaluated using the most advanced approaches currently available.

This research's primary contributions are as follows:

- Our work introduces a cutting-edge hybrid deep learning approach, CNN-ResBiGRU-CBAM, which combines a CNN and residual BiGRU with a CBAM mechanism to detect epileptic seizures.
- The study demonstrates the proposed model's effectiveness in accurately capturing spatial and temporal patterns from EEG data, improving seizure detection performance.
- We do significant research on the EEG dataset obtained from the University of Bonn, providing a thorough examination of the model's effectiveness, including conducting ablation tests and visualizing features.
- In addition, we evaluate our suggested model against the most advanced methodologies available, demonstrating its enhanced efficiency in terms of accuracy, sensitivity, and specificity.

Our research introduces an unexplored integration of advanced technologies, explicitly combining CNN with residual BiGRU and CBAM to address the complexities inherent in EEG data interpretation. Conventional approaches predominantly employ either CNN or RNN architectures in isolation. In contrast, our proposed framework leverages the complementary strengths of these models: the CNN is utilized to efficiently extract hierarchical spatial features while mitigating the vanishing gradient problem; the residual BiGRU captures intricate temporal dependencies by processing information in both forward and backward directions; and the CBAM is employed to provide a sophisticated dual-attention mechanism, which selectively emphasizes pertinent spatial and channel information.

The synergistic combination of these components offers a comprehensive approach to several pivotal challenges in analyzing EEG signals. Specifically, the residual connections facilitate the training of deeper networks, thereby enhancing the robustness and accuracy of feature extraction. The bidirectional processing capability of the BiGRU captures the complex evolution of seizure patterns over time, providing a holistic temporal representation. Additionally, the dual-attention mechanism of the CBAM plays a crucial role in attenuating irrelevant noise while accentuating features most indicative of seizure activity. Collectively, these elements form an integrated framework that improves the interpretability of EEG data and advances the accuracy and reliability of seizure detection and characterization.

The following parts of this paper are organized as follows: Section 2 provides a concise overview of prior studies on the use of deep learning in identifying epileptic seizures and the presently available EEG datasets for this purpose. Section 3 outlines the planned approach, encompassing the dataset, pre-processing techniques, and the hybrid residual CNN-BiGRU architecture with CBAM. Section

4 delineates the findings from the investigations. Section 5 provides an analysis of the outcomes. Section 6 serves as the last section of the work, providing a conclusion and presenting potential areas for additional exploration.

2. Related work

This section summarizes the current research on the topic, focusing specifically on three main categories: epileptic seizure detection (ESD), traditional machine-learning approaches, and deep learning-based methods for ESD.

2.1. ESD

ESD is a crucial field that seeks to precisely determine and classify seizure events from regular brain activity by analyzing EEG data. EEG is a noninvasive method that captures the electrical indications created by electrodes on the scalp. Because it can record the unique variations of activity related to seizures, it is commonly used to detect and treat epilepsy [16].

The main goal of epileptic identification is to develop automated techniques to identify seizures in actual time or offline scenarios accurately. Precise identification of seizures is crucial for multiple reasons. First, it facilitates prompt intervention and medical care, mitigating or reducing the possibility of damage inflicted by seizures. Furthermore, it aids in the impartial evaluation of the frequency and intensity of seizures, which is essential for arranging and tracking therapy. Moreover, it can notify caretakers or medical experts about the onset of seizures, notably when an individual cannot report it themselves [17].

Nevertheless, identifying epileptic seizures is a formidable challenge for many reasons. EEG signals exhibit complexity, lack stationarity, and regularly include diverse artifacts, including muscle contractions, eye movements, and electrical fluctuations. The occurrence of seizures can demonstrate substantial variations across people and even within the same person over time. In addition, some seizures could display mild or localized EEG indications that are challenging to differentiate from typical environment brain activity [18].

Conventional methods for detecting epileptic seizures include expert neurologists manually examining recordings of EEG activity. Nevertheless, this manual procedure is laborious, based on personal judgment, and susceptible to individual error. Due to the emergence of digital EEG recording technologies and developments in the processing of signals and machine-learning approaches, there is a growing fascination with creating computerized approaches for seizure identification [19].

Computerized epileptic seizure identification generally comprises four phases: (1) acquiring and pre-processing data, (2) extracting features, (3) selecting features, and (4) classifying. The pre-processing processes include decreasing noise, eliminating artifacts, and segmenting EEG data into appropriate epochs. Feature extraction approaches endeavor to identify the pertinent attributes of seizure activity, such as properties in the time domain, frequency domain, and time-frequency domain. Feature selection algorithms are used to identify the most distinctive attributes and reduce the dimensionality of the feature space. Classification techniques, including support vector machines and decision trees, ultimately distinguish seizure and non-seizure sequences. The choice of approach depends on the specific properties being considered [20].

2.2. Machine-learning techniques for ESD

The study focuses on algorithms that use machine learning designed to identify epileptic EEG patterns, which are becoming more common. The creation of the first globally acknowledged open EEG database for seizure forecasting in 2007 initiated a computational competition to compare achievements and promote ongoing developments in this domain [21]. Shoeb et al. [22] used scalp EEG data from the CHB-MIT dataset to implement machine-learning methods to identify seizures. Their efforts yielded impressive results. Tiwari et al. [23] employed a keypoint-based local binary pattern (LBP) technique with a support vector machine classifier to differentiate between seizures and seizure-free episodes. Al-Hadeethi et al. [24] attained positive outcomes with the AB-LS-SVM classification model by reducing EEG signal dimensions using a covariance matrix, extracting statistical features, and performing a non-parametric test to select the most significant features. Vicnesh et al. [25] categorized various types of epilepsy by analyzing non-linear EEG features and organizing them into a decision tree.

2.3. Deep-learning approaches for ESD

Machine learning has shown significant efficacy in detecting and identifying aberrant behaviors exhibited by individuals. Hence, automated detection and identification of anomalous behaviors are required. Various researchers have proposed distinct deep-learning techniques that significantly aid in automated identification. None of the abovementioned machine-learning approaches used a computerized approach for feature extraction; instead, they relied on manual methods. Hence, to circumvent the additional labor involved in manually extracting features, it is essential to use deep-learning methods to train deep-learning models using ample amounts of data. This section provides a compilation of researchers who used deep-learning methods to differentiate between epileptic and non-epileptic activities.

Deep-learning technology has experienced rapid and substantial progress, with neural network models widely adopted in various fields. Rosas-Romero et al. [26] demonstrated that utilizing functional near-infrared spectroscopy (fNIRS) for detecting epileptic seizures produces more effective results compared to traditional EEG. Additionally, the study highlights that applying deep-learning techniques to this issue is appropriate due to the specific characteristics of fNIRS data. Zhang et al. [27] applied wavelet packet decomposition and conventional spatial pattern methods to extract distinctive features of scalp EEG signals in both time and frequency domains, using a shallow CNN to classify pre-seizure and inter-seizure periods. Ma et al. [28] employed an RNN with LSTM to predict epileptic seizures by inputting statistical features of the EEG data into the LSTM architecture, marking the first use of this method for seizure prediction. Daoud et al. [29] developed an LSTM-based algorithm for seizure prediction tailored to individual patients. Jana et al. [30] used a one-dimensional CNN to detect epileptic seizures from EEG recordings by inputting the generated spectral graph matrix. Hu et al. [31] proposed an innovative approach utilizing a deep bidirectional LSTM (BiLSTM) network for seizure detection. Tsiouris et al. [32] improved the accuracy of predicting seizures by using a two-layer LSTM network that used four pre-seizure windows of varying lengths.

Recent developments in hybrid deep-learning structures have demonstrated encouraging outcomes in EEG signal processing. Wang et al. [33] introduced a hybrid approach combining 2D CNN and

LSTM to analyze time series portions of EEG data for motor imagery classification. Their methodology illustrates the efficacy of integrating CNN's spatial feature extraction with LSTM's temporal modeling, enhancing classification accuracy. Roy [34] developed a multiscale feature combined CNN based on adaptive transfer learning for EEG motor imagery classification, tackling inter-subject variability and intricate signal characteristics through innovative integration of convolution scales and transfer learning methodologies. This study attained notable enhancements in classification accuracy by efficient feature extraction from various frequency bands and adaptive learning methodologies. Although these studies concentrated on motor imagery classification, they offer essential perspectives into the efficacy of hybrid architectures for EEG signal evaluation. They reinforce our methodology of integrating spatial and temporal processing inside the CNN-ResBiGRU-CBAM model for epileptic seizure identification.

2.4. Comparison with existing approaches

Despite the significant advancements achieved by current methodologies in epileptic seizure identification, our proposed CNN-ResBiGRU-CBAM model offers a series of transformative features that distinguish it markedly from the state-of-the-art methods. Conventional CNN approaches focus on spatial feature extraction, while traditional RNN models are limited to capturing temporal dependencies. In contrast, our hybrid architecture uniquely integrates three distinct components: CNN for comprehensive spatial feature extraction, residual BiGRU for enhanced temporal representation, and CBAM for selective attention to spatial and channel dimensions. This multifaceted integration enables a more thorough spatiotemporal analysis of EEG signals, significantly surpassing the scope of current methodologies, which tend to adopt a limited, single-focus perspective.

Contemporary techniques often need to improve on notable limitations such as the vanishing gradient issue in deep networks, deterioration of temporal information, and suboptimal feature selection mechanisms. Our approach overcomes these constraints through several strategic enhancements. Implementing residual connections within BiGRU mitigates the vanishing gradient problem and allows for the efficient preservation of temporal gradients across deeper layers. Bidirectional processing within the BiGRU captures the temporal evolution of seizure dynamics in both forward and backward directions, thereby providing a more complete temporal context. Moreover, the dual-attention mechanism of CBAM ensures that the model selectively emphasizes pertinent features, effectively suppressing irrelevant noise and refining the focus on seizure-relevant patterns. These integrated advancements enable our model to discern intricate EEG signal patterns with heightened precision while maintaining computational efficiency.

In addition to the architectural improvements, our proposed model demonstrates significant advantages in terms of computational efficiency compared to current methodologies. Our approach achieves a superior balance between model depth and operational efficiency by promoting enhanced parameter sharing and reducing computational complexity without compromising performance. This equilibrium is facilitated by a novel feature fusion mechanism, which ensures that our model is computationally feasible for practical applications, especially in scenarios where hardware resources are constrained, such as in wearable or mobile healthcare devices.

Furthermore, the robustness and generalizability of our model offer substantial benefits in practical clinical applications. Unlike previous models, which often exhibit strong performance exclusively

in highly controlled or specific experimental settings, our CNN-ResBiGRU-CBAM model maintains consistent performance across binary and multi-class classification tasks. It effectively handles diverse EEG patterns and generalizes to multiple seizure types, demonstrating resilience across varied patient populations and seizure characteristics. The model also yields more balanced performance metrics, increasing its reliability and applicability in clinical contexts where precision, recall, and balanced accuracy are critical for patient outcomes.

In summary, the proposed CNN-ResBiGRU-CBAM architecture significantly advances automated epileptic seizure detection. By addressing critical limitations in existing methodologies and introducing a holistic, computationally efficient solution for spatiotemporal feature extraction and attention, our model sets a new benchmark for accuracy and practical feasibility in analyzing EEG signals. These contributions underscore its potential as a highly effective tool in clinical environments, facilitating more accurate and timely seizure diagnosis and monitoring.

3. Research methodology

This section outlines the specific procedures used in our investigation, detailing the use of the EEG dataset, the data pre-processing steps, and the development of the suggested CNN-ResBiGRU-CBAM model. The dataset comprises monochannel EEG data obtained from both healthy subjects and patients diagnosed with epilepsy. It is divided into five distinct subsets labeled A through E. We discuss data-preparation techniques, such as linear interpolation, filtering, and normalization, to ensure the data's quality and consistency. The CNN-ResBiGRU-CBAM model is a robust deep-learning structure specifically built to forecast epileptic seizures by effectively collecting both spatial and temporal patterns in EEG data. In addition, we provide a description of the assessment criteria utilized to assess the predictive model's effectiveness, which includes accuracy, sensitivity, specificity, and F1-score.

3.1. Bonn EEG dataset

The Bonn EEG dataset included in this study is derived from the University of Bonn, consisting of single-channel EEG signals extracted from multi-channel EEG recordings of both healthy subjects and epileptic patients [35]. The EEG dataset consists of two categories of recordings: scalp EEG from individuals in good health and intracranial EEG from epileptic sufferers. Although these recordings offer significant data for seizure identification investigation, it is crucial to acknowledge that the dataset description needs more details regarding the overall number of collected seizure episodes and needs to clarify the precise EEG channel configurations employed during the gathering of information.

It is divided into five subsets (A–E), each containing 100 recordings from five subjects. The signals were sampled at a rate of 173.61 Hz, and the bandpass was filtered within the frequency range of 0.53–40 Hz. Subsets A and B consist of signals from awake, healthy individuals, with subset A recorded with eyes closed and subset B with eyes open. Subsets C and D include interictal signals from non-epileptogenic and epileptogenic areas of epilepsy patients who were not experiencing seizures at the time. Subset E contains ictal recordings taken during epileptic seizures. The Bonn EEG dataset provides sample EEG signals, illustrated in Figure 1, facilitating the comparison of seizure detection algorithms with EEG data representing normal, interictal, and ictal states from various brain regions. Table 1 lists all five subsets along with their respective sample counts.

Table 1. Details of the EEG dataset from the University of Bonn.

Subset ID	Type	Description	No. of samples
A (Z)	Healthy	Five individuals in a state of relaxation, with their eyes closed, and in good physical condition.	100
B (O)	Healthy	Five individuals in a state of relaxation, with their eyes open, and in good physical condition.	100
C (F)	Seizure-Free Epilepsy	Five individuals with epilepsy, who had not experienced seizures, had EEG recordings taken from the hippocampal formation in the brain's opposite hemisphere	100
D (N)	Seizure-Free Epilepsy	Five individuals with epilepsy who are free from seizures and have no abnormal brain activity recorded in the epileptogenic zone.	100
E (S)	Under the seizure attack with epilepsy	Five individuals with pathological conditions characterized by epileptic activity related to seizures.	100

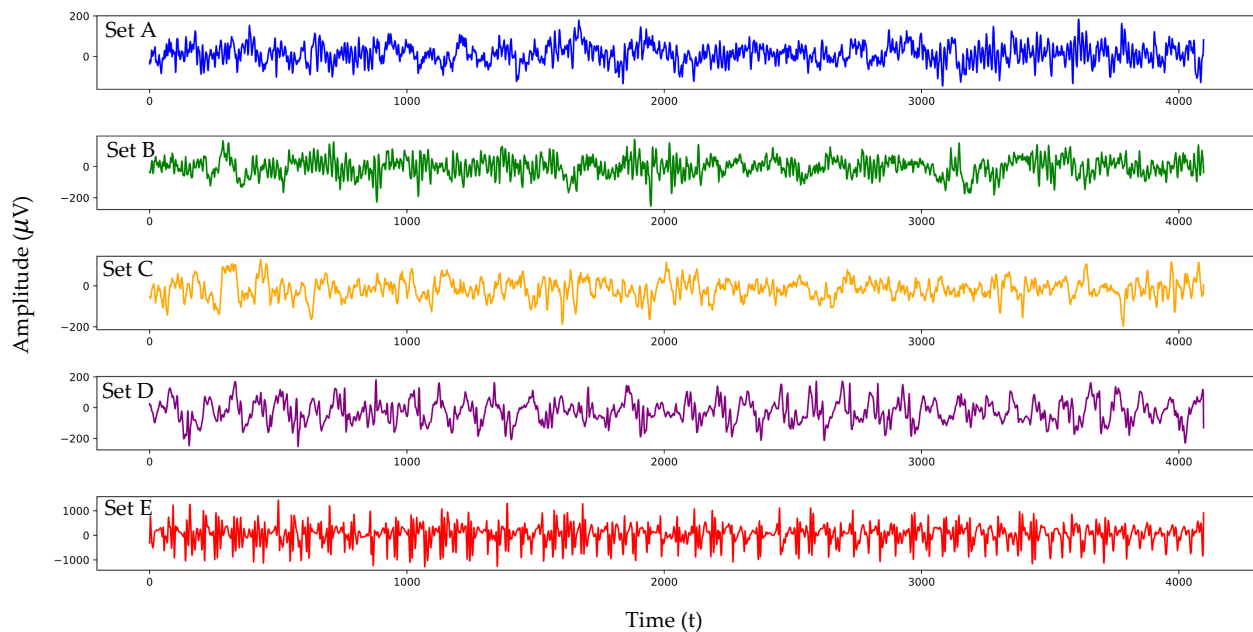


Figure 1. Example signals of EEG dataset: (a) Set A—wide open eyes; (b) Set B—typical closed eyes; (c) Set C—hippocampal growth during the interictal period; (d) Set D—epileptogenesis during the interictal period; (e) Set E—ictal (disorder characterized by seizures).

Figure 1 presents EEG example signals from the dataset, showcasing the varied characteristics of EEG recordings across different brain states and conditions. The distinctive patterns and features of EEG signals in subsets A to E highlight brain activity's complex and diverse nature during normal, interictal, and ictal states. These differences form the basis for developing automated algorithms to detect epileptic seizures. The goal is distinguishing between EEG signal types and reliably identifying seizure occurrences.

3.2. Data pre-processing

During this phase, the EEG data is subjected to filtering and normalizing procedures to guarantee uniformity and appropriateness when developing a model for detection. This technique entails the elimination of any insufficient or abnormal information values by the given protocol:

- We employed linear interpolation to address missing sensor data values, eliminating any existing noise. We applied sequential low-pass and mid-pass filtering techniques to reduce the impact of irrelevant EEG signals. Initially, a third-order Butterworth filter was used to eliminate high-frequency artifacts above 20 Hz. Following this, a third-order median filter replaced each data point with the median value of its neighboring points. Focusing on the median helps disregard occasional outliers and anomalies, favoring the less skewed central trend. This approach smooths out irregular peaks, providing waveforms that reflect the underlying dynamics more accurately. It removes extraneous variations, preserving essential characteristics needed for model training and avoiding distortions caused by random points that falsify normal or abnormal brain activity.
- Furthermore, we used a normalization technique to normalize each distinct segment of EEG data by computing the mean and standard deviation. This phase was essential in guaranteeing uniformity and the capacity to make meaningful comparisons across the data.

The normalization method linearly scales the raw EEG data using the min-max strategy. Once the data has been cleansed and standardized, it can be used as input for further data preparation and classification processes. To streamline the classifier's training process, the data is divided into many segments according to the selected approach. Subsequently, the following subset is used as a test set to evaluate the algorithm's effectiveness.

3.3. The proposed CNN-ResBiGRU-CBAM model

The suggested model combines convolution and residual BiGRU blocks inside and throughout its entire deep-learning architecture. The proposed model's overall design is shown in Figure 2.

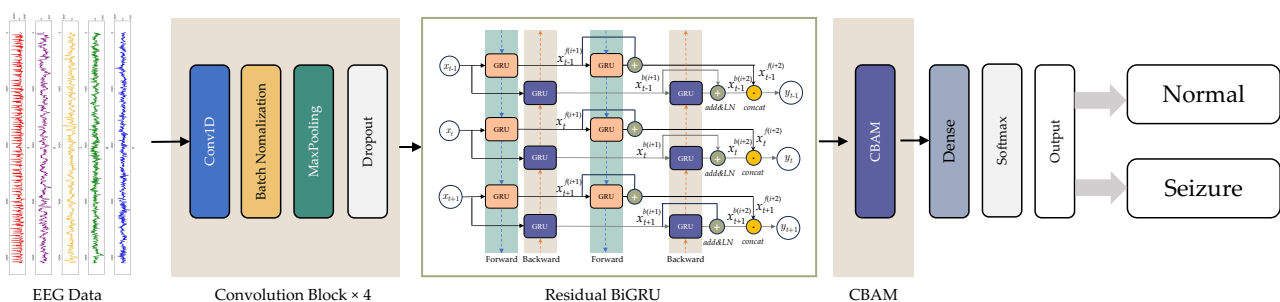


Figure 2. Detailed architecture of the proposed CNN-ResBiGRU model.

The first element, known as the convolution block, is tasked with extracting spatial features from the pre-processed input. By adjusting the convolution kernel's stride, the time series length is significantly reduced, resulting in a faster identification operation. Subsequently, the BiGRU network collects temporal patterns from the data the convolution block has handled. Including this feature improves the model's capacity to recognize long-term relationships in the time series data by

leveraging the advantages of the BiGRU. This combination supports the model's comprehension of intricate temporal patterns and contributes to its ability to recognize them. Behavioral input is classified through the fully linked layer and SoftMax function. The outcomes of this categorization procedure act as the identification results, offering a forecast of the particular activity being executed. We will comprehensively describe each element in the following parts, clarifying its functions and accomplishments within our suggested framework.

3.3.1. Convolution block

A CNN often utilizes a distinct collection of elements. CNNs are often employed when supervised learning is employed. Typically, these networks establish links between each neuron and all the neurons in the subsequent layers. The activation function of a neural network transforms the input value of neurons into the matching output value. The efficacy of the activation function is determined by two crucial factors: the paucity of data and the capacity of the lower network layers to sustain gradient flow. CNNs frequently utilize pooling techniques to decrease the dimensionality of data, typically using max-pooling and average-pooling as the conventional approaches.

This work utilizes convolutional blocks (ConvB) to extract essential information from unprocessed sensor data. Figure 2 illustrates the structure of ConvB, which consists of four layers: 1D-convolutional (Conv1D), batch normalization (BN), a max-pooling layer (MP), and a dropout layer. Conv1D employs several trainable convolutional kernels to identify different characteristics, with each kernel generating a feature map. The batch normalization layer enhances performance and accelerates the training process.

3.3.2. Residual BiGRU block

Human behaviors are temporal in nature, meaning that relying just on the convolution block to retrieve spatial characteristics is insufficient for accurately detecting activities. It is crucial to consider the chronological sequence of the occurrence. RNNs have superior capabilities in effectively processing and analyzing time series data. However, these models can experience problems like gradient vanishing and data loss as the time series grows.

Hochreiter et al. [36] pioneered the introduction of an LSTM model. LSTM, unlike simple RNNs, is a kind of RNN that effectively maintains time-related data over long periods using gating processes. Moreover, it outperforms ordinary RNNs in efficiently handling larger time series. Biological data is influenced not just by past occurrences but also by future ones.

Although LSTM has effectively resolved the vanishing gradient in RNNs, its memory cells lead to increased memory consumption. Cho et al. [37] introduced the GRU network, a novel model that builds upon the concept of RNNs. The GRU is a modified version of the LSTM with no separate memory cell in its structure [38]. In a GRU network, adjustments and reset gates control the degree to which each hidden state is altered. It determines what details should be provided to the next step and which should be excluded. A BiLSTM is a neural network design that integrates forward and backward signals by employing two GRU networks. BiGRU enhances extracting features from time series data by capturing and retaining information about both forward and backward relationships, which sets it apart from the GRU network. Therefore, utilizing a BiGRU network to identify temporal patterns from behavioral data is appropriate.

While the BiLSTM network excels at catching temporal patterns, it might be more efficient in acquiring spatial information. Furthermore, as the amount of stacked layers rises, the issue of gradient vanishing becomes more significant throughout the training process. In 2015, the Microsoft Research team launched ResNet, a residual network, to tackle the issue of gradient vanishing [12]. The depth of this neural network is 152 layers, and each residual block could be represented as follows:

$$x^{i+1} = x^i + F(x^i, W_i) \quad (3.1)$$

The residual blocks are comprised of two components: x_i , which represents a direct mapping, and $F(x^i, W_i)$, which represents the residual portion.

The framework mentioned above is also used to construct the encoder element in the transformer model. Our study uses its advantages to present a residual design that integrates the BiGRU network. The BiGRU network could additionally be subjected to normalization procedures. Layer normalization (LN) is more beneficial for RNNs than batch normalization (BN). LN is calculated similarly to BN and can be represented by the following equation:

$$\widehat{x}^i = \frac{x^i - E(x^i)}{\sqrt{\text{var}(x^i)}} \quad (3.2)$$

The input vector of size i is denoted as x^i , whereas the output after layer normalization is denoted as \widehat{x}^i .

This study presents a new method called ResBiGRU, which combines residual structure with layer normalization in a BiGRU network. The structure of ResBiGRU can be seen in Figure 3. The recursive feature data, denoted as y , could be precisely specified as:

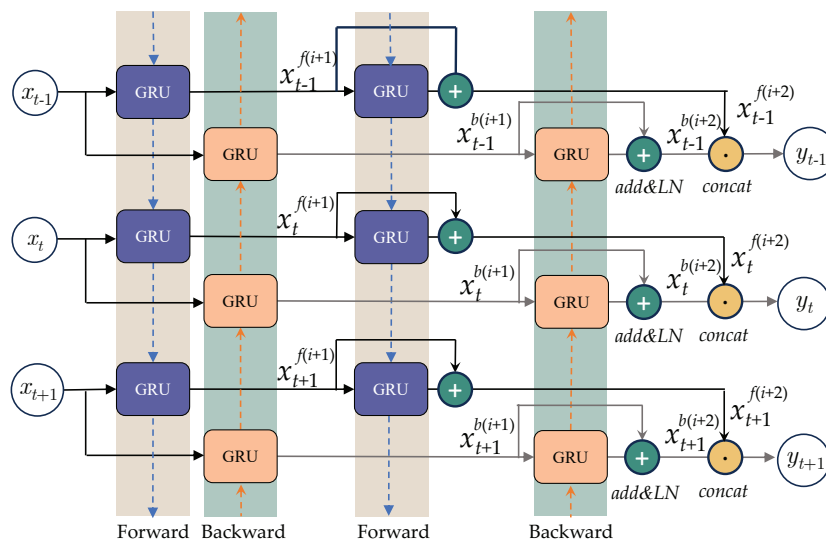


Figure 3. Structure of the ResBiGRU.

$$x_t^{f(i+1)} = LN(x_t^{f(i)} + GRU(x_t^{f(i)}, W_i)) \quad (3.3)$$

$$x_t^{b(i+1)} = LN(x_t^{b(i)} + GRU(x_t^{b(i)}, W_i)) \quad (3.4)$$

$$y^t = \text{concat}(x_t^f, x_t^b) \quad (3.5)$$

LN stands for layer normalization, whereas *G* represents the computation of input states in the GRU network. The subscript t in $x_t^{f(i)}$ denotes the t -th instant in the time series. The superscript f denotes the forward state, b represents the backward state, and $(i+1)$ shows the number of stacked layers. The data encoded y_t at time t is formed by combining the forward and backward states.

3.4. CBAM block

The CBAM [14] is an attention mechanism designed to boost the effectiveness of CNNs. It achieves this by emphasizing relevant channels and important spatial areas within feature maps. CBAM consists of two sequential submodules: the channel attention module (CAM) and the spatial attention module (SAM), shown in Figure 4.

The CAM employs adaptive learning to evaluate the significance of each channel. This assessment is carried out through a combination of maximum and average pooling. Subsequently, a shared multilayer perceptron regulated by a reduction ratio (r) is used to balance computational efficiency with attention precision. Feature maps processed by the CAM are then passed to the SAM.

The SAM creates a spatial attention mask using maximum and average pooling across channels, concatenation, and a convolutional layer. This mask is then applied to each element of the input feature maps individually. As a result, the network can focus on the most informative spatial locations. CBAM improves the network's perceptual capability and robustness by incorporating both channel and spatial attention [39].

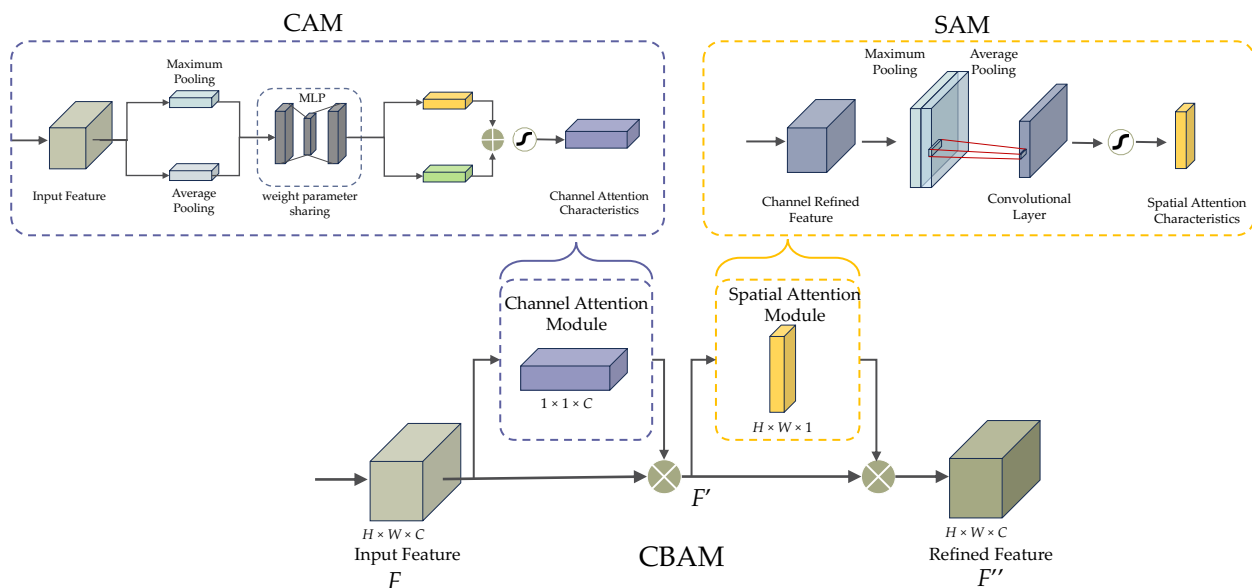


Figure 4. The convolutional block attention module (CBAM).

The CBAM module in our architecture processes feature maps arranged as three-dimensional matrices with dimensions of $H \times W \times C$. In this setup, H (height) corresponds to the temporal dimension derived from the BiGRU output, W (width) denotes the feature dimension at each time step, and C (channel) indicates the number of feature maps generated by the BiGRU block.

This configuration enables CBAM to focus on channel attention within the C dimension, interpret temporal information as spatial relationships within the $H \times W$ dimension, and retain the temporal consistency of the BiGRU's output. Consequently, this structure ensures that the temporal patterns captured by the BiGRU block are preserved while facilitating an effective attention mechanism.

3.5. Evaluation metrics

Given the critical role of medical diagnostics, assessing systems for detecting epileptic seizures requires thorough consideration of multiple performance criteria. We use a thorough assessment methodology that considers clinical relevance and statistical reliability. The metrics listed below are specially selected to evaluate various facets of the detection system's functionality:

- Accuracy – The total correctness of classification across all classes is measured by accuracy. It symbolizes the model's overall dependability in differentiating between various EEG states in the setting of epileptic seizure identification. The definition of the measurement is:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.6)$$

- Sensitivity (Recall) – Since sensitivity assesses the model's capacity for correctly identifying actual seizure episodes, it is essential in epileptic seizure identification. For the safety of patients, a high sensitivity means fewer seizures are ignored. The definition of the measurement is:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (3.7)$$

From a clinical standpoint, this statistic is particularly crucial because failing to detect a seizure occurrence could harm the treatment and care of patients.

- Specificity – This score assesses how well the model can detect non-seizure occurrences. This measure is essential for avoiding false alarms that can cause patients to get anxious or require needless procedures. The definition of the measurement is:

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (3.8)$$

The identification system's practical application in medical environments depends on high specificity, which guarantees that ordinary neurological activity is not mistakenly identified as seizure activity.

- F1-score – This metric integrates precision and recall into one measure to offer a fair assessment of the model's effectiveness. Because seizure events are less common than non-seizure events in EEG datasets, which are generally unbalanced, this is especially important for epileptic seizure identification. The definition of the indicator is:

$$\text{F1-score} = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \quad (3.9)$$

where, true positive (TP) represents the number of seizure signals that are correctly identified as seizures. False negative (FN) refers to the number of seizure signals that are incorrectly classified as non-seizures. True negative (TN) is the count of non-seizure signals accurately classified as

non-seizures. Finally, false positive (*FP*) represents the number of non-seizure signals that are mistakenly identified as seizures.

These assessment criteria are related to one another and offer complimentary information about how well the model performs from various clinical viewpoints. The F1-score provides a balanced metric that considers both sensitivity and specificity, which is essential in clinical settings because missed seizures and false alarms can substantially influence patient care. Sensitivity guarantees accurate identification of actual seizures, while specificity avoids false alarms. When combined with metrics, these measures offer a thorough evaluation of the model's clinical value, striking a balance between statistical effectiveness and real-world patient care implications. This allows for a comprehensive evaluation of the system's efficacy in actual epileptic seizure detection settings.

4. Experiments

This section presents the experimental results that evaluate the performance of our hybrid residual CNN-BiGRU model, which incorporates the CBAM mechanism, for detecting epileptic seizures. The experiments utilized the EEG dataset. The primary goal was to test the model's effectiveness in distinguishing EEG segments that indicate epileptic activity from those that do not. The experimentation also evaluated the model's ability to distinguish between different kinds of EEG data in binary and multi-class classification situations.

4.1. Experimental settings

The experimental setup utilized to assess our suggested CNN-ResBiGRU-CBAM model for epileptic seizure detection is described in depth in this part. To ensure the reproducibility of our findings, we offer thorough details regarding the implementation libraries, training methods, and hardware and software environment.

4.1.1. Hardware and software infrastructure

The experiments were performed on Google Colab Pro+ using a Tesla V100-SXM2-16GB GPU (Hewlett Packard Enterprise, Los Angeles, USA) to speed up the training process for the deep-learning models. This powerful computational setup efficiently handled EEG data from the Bonn EEG dataset. The implementation was executed in Python version 3.6.9, with TensorFlow 2.2.0 as the main framework for building and training the models. GPU computations were optimized using CUDA 10.2 as the backend, improving overall training efficiency.

4.1.2. Software libraries

Our implementation utilized many Python libraries, each fulfilling distinct model-building roles. We employed NumPy and Pandas for efficient data processing and EEG signal evaluation, while SciPy offered crucial signal processing and statistical calculation functionalities. The model was developed and trained using Keras with a TensorFlow backend, enabling the creation of our CNN-ResBiGRU-CBAM architecture. Scikit-learn was utilized for data preparation, cross-validation processes, and effectiveness measurement computations. We employed CUDA libraries to enhance efficiency when computing using GPU acceleration. We utilized Matplotlib and Seaborn libraries to provide detailed

effectiveness visualizations and outcome assessments, supplemented by bespoke evaluation criteria tailored for evaluations of recognizing seizures.

4.1.3. Training protocol

The training approach was meticulously crafted to ensure potent model effectiveness and dependable assessment outcomes. We employed a 5-fold cross-validation method to evaluate model robustness and generalization capabilities. The training utilized a batch size of 64 samples across 200 epochs, with a dropout probability of 0.25 for optimal regularization. We deployed the Adam optimizer with an initial learning rate of 0.001 and implemented a cross-entropy loss function for the classification assignments. We established an early terminating mechanism for 50 epochs to avert overfitting and enhance training productiveness. This extensive training methodology guarantees the reproducibility of our findings and establishes a robust basis for significant comparisons with alternative methodologies in the area.

4.2. Experiments, training, and validation details

For this research, we performed a sequence of tests to evaluate the precision of our suggested model by using different combinations of EEG data from the dataset. As stated in Section 3.1, the dataset comprises five separate sets of EEG signals: A, B, C, D, and E. Set A and Set B were obtained from the scalps of five subjects who were in good health. Set A was recorded with the subjects' eyes closed, while Set B was recorded with their eyes open. The recordings were made using the usual 10–20 electrode placement technique. Three sets, C, D, and E, consist of EEG data obtained from five individuals diagnosed with epilepsy. Set D consists of EEG signals acquired from the epileptogenic zone. Set C consists of signals obtained from the hippocampus formation in the contralateral hemisphere during times without seizures, referred to as the interictal state. Set E comprises EEG data recorded over seizures or the ictal phase.

The primary aim of the initial experiment was to examine the binary classification challenge in the context of detecting epileptic seizures. The present study investigated four binary classification scenarios (A-E, B-E, C-E, and D-E) using 200 EEG signal samples for model training and evaluation.

The second experiment aimed to examine a three-class classification (AB-CD-E). This involved distinguishing between healthy states (AB), interictal states (CD), and ictal states (E). The third experiment focused on four-class classification (AB-C-D-E). It separated healthy states (AB), interictal states in the hippocampal formation (C), interictal states in the epileptogenic zone (D), and ictal states (E). Lastly, the fourth experiment explored five-class classification (A-B-C-D-E), covering all distinct categories of EEG signals—the experimental procedures used all 500 samples from the EEG dataset.

The Adam optimizer was utilized for training the recommended models by optimizing the categorical cross-entropy cost function. Because it incorporates the benefits of the AdaGrad and RMSProp algorithms, the optimizer was implemented. Throughout the deep-learning model-training process, the optimizer improves the efficiency of computation [40,41].

Training and assessing the proposed models is accomplished via five-fold cross-validation. The EEG signals are partitioned arbitrarily into five equal-sized segments. One-fold cross-validation was utilized for evaluating, and four-fold cross-validation was employed for instruction.

The training data details for binary and multi-class classification are meticulously designed to ensure the model's accuracy. For binary classification, 160 EEG signals with a duration of 23.6 seconds are used, with 40 signals dedicated for testing in each fold. For three-, four-, and five-class classifications, 400 EEG signals with a duration of 23.6 seconds are trained on, with 100 signals used for testing, providing a comprehensive and thorough assessment of the model's performance.

4.3. Experimental results of binary classification

Table 2 presents the interpretation metrics of the CNN-ResBiGRU-CBAM model, designed for binary classification tasks such as A-E, B-E, C-E, and D-E. The model demonstrates exceptional results across all these binary classification tasks, reflected in its high values for F1-score, sensitivity, specificity, and accuracy.

Table 2. Performance measures of the proposed model for binary classification.

Binary Classification	Performance (mean(\pm standard deviation))			
	Accuracy	Sensitivity	Specificity	F1-score
A-E	99.00%(\pm 2.00%)	99.00%(\pm 2.00%)	99.00%(\pm 2.00%)	99.00%(\pm 2.01%)
B-E	99.00%(\pm 1.22%)	99.00%(\pm 1.22%)	99.00%(\pm 1.22%)	99.00%(\pm 1.23%)
C-E	97.50%(\pm 2.24%)	97.50%(\pm 2.24%)	97.50%(\pm 2.24%)	97.49%(\pm 2.24%)
D-E	98.50%(\pm 1.22%)	98.50%(\pm 1.22%)	98.50%(\pm 1.22%)	98.50%(\pm 1.23%)

For the A-E categorization, the model achieves an average accuracy of 99.00% with a standard deviation of 2.00%. The sensitivity and specificity are both 99.00% (\pm 2.00%), indicating the model's balanced capability to differentiate between epileptic seizure (E) and non-seizure (A) EEG data. The F1-score, representing the harmonic mean of precision and recall, is 99.00% with a standard deviation of 2.01%.

In the B-E classification, the model similarly attains an average accuracy, sensitivity, and specificity of 99.00%, with a standard deviation of 1.22%. The F1-score for this classification is 99.00% with a standard deviation of 1.23%.

The model's performance in the C-E classification is slightly inferior to the A-E and B-E classifications. It achieves an average accuracy, sensitivity, and specificity of 97.50% with a standard deviation of 2.24%. The F1-score is 97.49%, with a standard deviation of 2.24%. This suggests that distinguishing between interictal (C) and ictal (E) EEG signals is more challenging than differentiating between healthy (A or B) and ictal (E) data.

The model's performance for the D-E classification is comparable to the A-E and B-E classifications. It attains an average accuracy, sensitivity, and specificity of 98.50% with a standard deviation of 1.22%. The F1-score for this classification is 98.50%, with a standard deviation of 1.23%.

The minimal standard deviations observed in all measures suggest that the model's effectiveness remains constant and stable throughout many cross-validation folds.

The CNN-ResBiGRU-CBAM model demonstrates exceptional results in binary classification tests. It accurately differentiates between healthy and ictal EEG data (A-E and B-E), and while its ability to distinguish between interictal and ictal signals (C-E and D-E) is slightly reduced, it remains highly accurate. These findings underscore the efficacy of the CNN-ResBiGRU-CBAM design, which

incorporates the CBAM mechanism, in precisely detecting epileptic seizures by capturing spatial and temporal patterns.

4.4. Experimental results of multi-class classification

Table 3 interprets the proposed model's performance metrics for multi-class classification tasks, including 3-class, 4-class, and 5-class issues. The model distinguishes satisfactorily among various EEG signal categories, attaining high accuracy, sensitivity, specificity, and F1-score.

Table 3. Performance measures of the proposed model for multi-class classification.

Experiment	Performance (mean(\pm standard deviation))			
	Accuracy	Sensitivity	Specificity	F1-score
(3-Class Problem)				
AB-CD-E	96.20%(\pm 2.86%)	96.50%(\pm 2.76%)	96.92%(\pm 2.23%)	96.66%(\pm 2.57%)
(4-Class Problem)				
AB-C-D-E	92.00%(\pm 2.53%)	90.38%(\pm 3.34%)	96.75%(\pm 1.07%)	90.94%(\pm 3.34%)
(5-Class Problem)				
A-B-C-D-E	89.00%(\pm 3.63%)	89.00%(\pm 3.63%)	89.63%(\pm 4.87%)	88.94%(\pm 3.64%)

The model achieves an average accuracy of 96.20% with a standard deviation of 2.86% for the 3-class problem (AB-CD-E). Its sensitivity stands at 96.50%(\pm 2.76%), indicating a high capability to correctly identify each class. The model's specificity is 96.92%(\pm 2.23%), demonstrating its effectiveness in distinguishing between different classes. Additionally, the F1-score, which balances precision and recall, is 96.66%(\pm 2.57%).

For the 4-class problem (AB-C-D-E), the model's performance declines slightly, achieving a mean accuracy of 92.00%(\pm 2.53%). With a sensitivity of 90.38%(\pm 3.34%), the model's ability to correctly identify the interictal epileptogenic zone (D), ictal (E), interictal hippocampal formation (C), and healthy (AB) states is somewhat reduced. However, the specificity remains high at 96.75%(\pm 1.07%), indicating continued effectiveness in differentiating between classes. The F1-score is 90.94%(\pm 3.37%).

The model's performance drops further in the 5-class problem (A-B-C-D-E), achieving an average accuracy of 89.00%(\pm 3.63%). The sensitivity and specificity are 89.00%(\pm 3.63%) and 89.63%(\pm 4.87%), respectively. The F1-score is 88.94%(\pm 3.64%). These results suggest that distinguishing among all five EEG signal categories is the most challenging task, requiring the model to detect more nuanced differences between the classes.

Compared to binary classification tasks, the standard deviations for multi-class issues are slightly higher. This indicates the model's adaptability to different tasks, with more significant variability in its performance across different cross-validation folds. This variability is expected, as binary classification tasks are inherently less complex than multi-class problems, highlighting the model's versatility.

The proposed model excels in multi-class classification tasks, with the highest accuracy observed in the 3-class problem (AB-CD-E). The 5-class problem (A-B-C-D-E) is the most challenging, with the model's performance decreasing slightly as the number of classes increases. Nonetheless, the model maintains high sensitivity, specificity, accuracy, and F1-score, demonstrating its ability to capture the distinct characteristics that separate various EEG signal categories. Our findings further confirm the robustness of the hybrid residual CNN-BiGRU architecture with the CBAM mechanism in epileptic

seizure detection.

4.5. Experimental results using multi-channel EEG data

This part presents supplementary investigations utilizing the CHB-MIT benchmark EEG dataset to assess the effectiveness of the suggested CNN-ResBiGRU-CBAM model with multi-channel EEG data.

The CHB-MIT dataset comprises EEG recordings of recurrent seizures from 24 patients aged between 1.5 and 22 years. The individuals were monitored for several days following the cessation of anti-epileptic medication to assess their candidacy for surgery and to determine the intensity of their seizures. The worldwide 10–20 system was employed for electrode positioning and nomenclature with a sampling frequency of 256 Hz [42]. Each individual's data comprises 23 dispersed channels, specifically: FP1-F7, F7-T7, T7-P7, P7-O1, FP1-F3, F3-C3, C3-P3, P3-O1, FP2-F4, F4-C4, C4-P4, P4-O2, FP2-F8, F8-T8, T8-P8, P8-O2, FZ-CZ, CZ-PZ, P7-T7, T7-FT9, FT9-FT10, FT10-T8, T8-P8. Furthermore, in conjunction with the EEG recordings, certain people possess other signal recordings, including vagus nerve stimulation (VNS) and electrocardiogram (ECG) signals. This study employs EEG data from 14 patients, maintaining consistent channel allocation for result standardization [43]. Table 4 presents comprehensive facts regarding the subjects with seizure onsets employed in this research.

Table 4. Detailed information on the subjects with seizure onsets included in this study.

Model	Performance (mean(\pm standard deviation))			
	Accuracy	Sensitivity	Specificity	F1-score
CNN-ResBiGRU-CBAM	98.60%(\pm 1.44%)	98.59%(\pm 1.43%)	98.59%(\pm 1.42%)	98.60%(\pm 1.43%)

5. Discussion

5.1. Comparison results with baseline deep-learning models

To assess the CNN-ResBiGRU-CBAM model's effectiveness, we performed additional investigations to compare it with five widely employed deep-learning models (CNN, LSTM, BiLSTM, GRU, and BiGRU) in arrangement modeling operations, specifically in the context of epileptic seizure identification [8, 44]. The detailed hyperparameters of the baseline deep-learning model are shown in Appendix A.

- 1) CNNs are very proficient at capturing spatial characteristics from sensor data and have shown significant promise in challenges related to recognizing human activities. They excel in acquiring knowledge of specific patterns in the local area and understanding the relationships between different elements in the supplied data.
- 2) LSTM is a specific sort of RNN intended to collect long-term dependencies in time series data. It employs a system of gates to regulate the transmission of data across time, which makes it very suitable for representing temporal patterns in sensor data.
- 3) BiLSTM is an extension of LSTMs that processes the input sequence in both directions (forward and backward). This model's bidirectional computing enables it to consider both the past and

potential future environment, which helps it collect more information and enhance its effectiveness in sequence modeling.

- 4) GRU is a kind of RNN with a less-complex structure than LSTM networks. Although simple, GRUs are capable of successfully capturing temporal relationships in sequences via a gating mechanism that regulates the flow of data.
- 5) BiGRU operates by processing the input sequence in both directions, forward and backward, similar to BiLSTMs. The bidirectional computation enables the model to gather a more extensive range of temporal data and consider both past and future contexts.

For a thorough and equitable comparison, all conventional models were subjected to the same rigorous training and testing using identical portions of the EEG dataset, as outlined in Section 4.2. This systematic training and evaluation process provides a solid foundation for an objective assessment of the CNN-ResBiGRU-CBAM model's effectiveness compared to standard deep-learning models.

Table 5 presents the impressive performance of the CNN-ResBiGRU-CBAM model when pitted against other core deep-learning models. The evaluation covers both binary and multi-classification challenges using the EEG dataset. The results unequivocally demonstrate the suggested model's exceptional ability to handle spatial and temporal dependencies, surpassing the standard models in terms of accuracy, sensitivity, specificity, and F1-score.

The CNN-ResBiGRU-CBAM model consistently outperforms all baseline models in accuracy, sensitivity, specificity, and F1-score for binary classification tasks A-E, B-E, C-E, and D-E. Among all models, the CNN model shows the weakest performance. While the LSTM, BiLSTM, GRU, and BiGRU models perform better than the CNN model, they still fall short of the proposed model's standards. The CNN-ResBiGRU-CBAM model achieves superior accuracy, sensitivity, and specificity, reaching 99.00% for both A-E and B-E classifications and 97.50% and 98.50% for C-E and D-E classifications.

The CNN-ResBiGRU-CBAM model excels in multi-class classification tasks, outperforming baseline models in the 3-class, 4-class, and 5-class problems. In the 3-class scenario (AB-CD-E), the model shows outstanding performance with an accuracy of 96.20%, sensitivity of 96.50%, specificity of 96.92%, and an F1-score of 96.66%. These metrics significantly surpass those of the baseline models. Among them, the BiGRU model comes closest, with an accuracy of 88.80%, sensitivity of 88.83%, specificity of 90.08%, and an F1-score of 89.19%.

The CNN-ResBiGRU-CBAM model's performance in the 4-class issue (AB-C-D-E) further demonstrates its adaptability and potential for clinical use. With an accuracy of 92.00%, sensitivity of 90.38%, specificity of 96.75%, and an F1-score of 90.94%, the model outperforms the BiGRU model, which achieves an accuracy of 75.00%, sensitivity of 70.63%, specificity of 91.75%, and an F1-score of 70.96%. The CNN-ResBiGRU-CBAM model's performance in the 5-class problem (A-B-C-D-E) underscores its versatility and potential for clinical applications. With an accuracy of 89.00%, sensitivity of 89.00%, specificity of 89.63%, and an F1-score of 88.94%, the model outperforms the GRU model, which achieves an accuracy of 65.00%, sensitivity of 65.00%, specificity of 76.38%, and an F1-score of 64.43%.

Table 5. Comparison results of baseline deep-learning models and the proposed CNN-ResBiGRU-CBAM.

Experiment	Model	Performance (mean%)			
		Accuracy	Sensitivity	Specificity	F1-score
(2-Class Problem)					
A-E	CNN	60.00%	60.00%	60.00%	49.31%
	LSTM	96.00%	96.00%	96.00%	95.99%
	BiLSTM	98.00%	98.00%	98.00%	98.00%
	GRU	98.00%	98.00%	98.00%	98.00%
	BiGRU	94.50%	94.50%	94.50%	94.49%
	CNN-ResBiGRU-CBAM	99.00%	99.00%	99.00%	99.00%
B-E	CNN	61.50%	61.50%	61.50%	51.29%
	LSTM	94.00%	94.00%	94.00%	93.98%
	BiLSTM	92.50%	92.50%	92.50%	92.48%
	GRU	92.50%	92.50%	92.50%	92.48%
	BiGRU	89.50%	89.50%	89.50%	89.49%
	CNN-ResBiGRU-CBAM	99.00%	99.00%	99.00%	99.00%
C-E	CNN	90.00%	90.00%	90.00%	89.77%
	LSTM	96.50%	96.50%	96.50%	96.49%
	BiLSTM	96.00%	96.00%	96.00%	96.00%
	GRU	97.00%	97.00%	97.00%	97.00%
	BiGRU	96.50%	96.50%	96.50%	96.49%
	CNN-ResBiGRU-CBAM	97.50%	97.50%	97.50%	97.49%
D-E	CNN	73.50%	73.50%	73.50%	66.75%
	LSTM	93.50%	93.50%	93.50%	93.48%
	BiLSTM	94.50%	94.50%	94.50%	94.49%
	GRU	93.50%	93.50%	93.50%	93.48%
	BiGRU	94.50%	94.50%	94.50%	94.49%
	CNN-ResBiGRU-CBAM	98.50%	98.50%	98.50%	98.50%
(3-Class Problem)					
AB-CD-E	CNN	65.80%	60.33%	74.25%	58.43%
	LSTM	87.60%	87.00%	90.75%	87.69%
	BiLSTM	87.80%	87.50%	89.92%	87.97%
	GRU	82.80%	83.83%	85.58%	83.67%
	BiGRU	88.80%	88.83%	90.08%	89.19%
	CNN-ResBiGRU-CBAM	96.20%	96.50%	96.92%	96.66%
(4-Class Problem)					
AB-C-D-E	CNN	65.20%	59.75%	72.83%	60.07%
	LSTM	69.60%	65.38%	86.58%	65.02%
	BiLSTM	73.20%	68.75%	90.33%	68.72%
	GRU	72.00%	66.88%	90.92%	66.62%
	BiGRU	75.00%	70.63%	91.75%	70.96%
	CNN-ResBiGRU-CBAM	92.00%	90.38%	96.75%	90.94%
(5-Class Problem)					
A-B-C-D-E	CNN	58.60%	58.60%	65.50%	58.13%
	LSTM	59.20%	59.20%	70.88%	59.17%
	BiLSTM	62.20%	62.20%	74.63%	62.05%
	GRU	65.00%	65.00%	76.38%	64.43%
	BiGRU	61.20%	61.20%	73.25%	61.01%
	CNN-ResBiGRU-CBAM	89.00%	89.00%	89.63%	88.94%

These findings underscore the efficiency of the CNN-ResBiGRU-CBAM model in accurately identifying epileptic seizures by capturing both spatial and temporal patterns from EEG signals. Integrating residual CNN, BiGRU, and CBAM methods enables the model to extract more distinct features, resulting in superior performance compared to baseline deep-learning models. The results

highlight the potential of the proposed model for clinical applications in epilepsy diagnosis and monitoring.

5.2. Comparison results with state-of-the-art work

Table 6 compares the effectiveness of the suggested CNN-ResBiGRU-CBAM model to other state-of-the-art models from prior research for binary and multi-class classification challenges utilizing the Bonn EEG dataset.

In binary classification (ABCD-E), our CNN-ResBiGRU-CBAM model attains an accuracy of 99.00%, surpassing current methodologies. Our model demonstrates enhancements of 1.83% in accuracy and 4.39% in sensitivity relative to the stacking ensemble-based deep-learning method proposed by Akyol [45]. Our technique exhibits an 8% enhancement in accuracy relative to the deep neural network model put forward by Thara et al. [46]. The most notable enhancement is shown in comparison to the LSTM model of Shekokar et al. [47], with our model demonstrating a 14% increase in accuracy and an 11.50% gain in sensitivity. Our model attains an F1-score of 98.37%, signifying a balanced achievement in precision and recall.

In the complex five-class classification problem (A-B-C-D-E), our model achieves a commendable accuracy of 89.00%, reflecting a 1.80% enhancement compared to the multivariate empirical mode decomposition method proposed by Zahra et al. [48]. Furthermore, our model offers extensive indicators of achievement, including sensitivity (89.00%), specificity (89.63%), and F1-score (88.94%), indicating balanced categorization proficiency across all five categories.

The comparisons indicate that our CNN-ResBiGRU-CBAM model not only attains state-of-the-art effectiveness in binary classification but also sustains outstanding outcomes in the more complex multi-class scenario, underscoring the efficacy of our combined strategy that incorporates CNN, residual BiGRU, and attention mechanisms.

Table 6. Comparison results of models in previous works and the proposed CNN-ResBiGRU-CBAM.

Experiment	Model	Performance (mean%)			
		Accuracy	Sensitivity	Specificity	F1-score
(2-Class Problem)					
ABCD-E	Stacking ensemble-based deep-learning approach [45]	97.17%	93.11%	98.18%	-
	Deep neural network [46]	91.00%	-	-	91.00%
	LSTM [47]	85.00%	86.00%	85.00%	-
	CNN-ResBiGRU-CBAM	99.00%	97.50%	97.50%	98.37%
(5-Class Problem)					
A-B-C-D-E	Multivariate empirical mode decomposition [48]	87.20%	-	-	-
	CNN-ResBiGRU-CBAM	89.00%	89.00%	89.63%	88.94%

5.3. Ablation studies

Ablation studies are predominantly utilized in neural networks [49] to evaluate the effectiveness of a model by examining the impact of modifying specific elements [50]. Consequently, the impacts of ablation are examined on the model we have created through three case studies by altering various blocks and layers to assess their impact on the suggested design [51]. Upon concluding all study

instances, the optimal configuration of our suggested CNN-BiGRU-CBAM model could be attained with the optimum recognition efficiency.

5.3.1. Impact of the convolution block

We performed ablation research utilizing the Bonn EEG dataset to examine the effect of the convolutional block on the model's effectiveness. We established a standard structure by excluding the convolutional block from the suggested CNN-ResBiGRU-CBAM architecture while retaining the residual BiGRU and CBAM blocks. This baseline model inputs the raw sensor data specifically into the ResBiGRU block, omitting any spatial feature extraction.

Table 7 illustrates the essential significance of the convolutional block inside our design. In binary classification tasks (A-E, B-E, C-E, and D-E), eliminating the CNN component particularly impaired the model's efficiency. The accuracy declined from 97.50–99.00% to roughly 51.50–52.00% in all binary classification contexts. This significant decline in efficiency suggests that the CNN block is essential for extracting relevant spatial characteristics from the raw EEG signals.

Table 7. Impact of the convolution block.

Experiment	Model	Performance (mean %)			
		Accuracy	Sensitivity	Specificity	F1-score
(2-Class Problem)					
A-E	The Proposed Model without CNN	51.50%	51.50%	51.50%	36.53%
	The Proposed CNN-ResBiGRU-CBAM	99.00%	99.00%	99.00%	99.00%
B-E	The Proposed Model without CNN	51.50%	51.50%	51.50%	36.53%
	The Proposed CNN-ResBiGRU-CBAM	99.00%	99.00%	99.00%	99.00%
C-E	The Proposed Model without CNN	52.00%	52.00%	52.00%	37.51%
	The Proposed CNN-ResBiGRU-CBAM	97.50%	97.50%	97.50%	97.49%
D-E	The Proposed Model without CNN	51.50%	52.00%	52.00%	36.45%
	The Proposed CNN-ResBiGRU-CBAM	98.50%	98.50%	98.50%	98.50%
(3-Class Problem)					
AB-CD-E	The Proposed Model without CNN	40.00%	33.33%	50.00%	19.05%
	The Proposed CNN-ResBiGRU-CBAM	96.20%	96.50%	96.92%	96.66%
(4-Class Problem)					
AB-C-D-E	The Proposed Model without CNN	40.00%	25.00%	50.00%	14.29%
	The Proposed CNN-ResBiGRU-CBAM	92.00%	90.38%	96.75%	90.94%
(5-Class Problem)					
A-B-C-D-E	The Proposed Model without CNN	20.20%	20.20%	50.00%	7.11%
	The Proposed CNN-ResBiGRU-CBAM	89.00%	89.00%	89.63%	88.94%

The effect is further amplified in multi-class classification problems. In the three-class issue (AB-CD-E), the elimination of the CNN block resulted in a drastic decrease in accuracy from 96.20% to 40.00%, accompanied by a decline in the F1-score from 96.66% to merely 19.05%. The four-class classification (AB-C-D-E) exhibited comparable degradation, with accuracy declining from 92.00% to 40.00% and the F1-score lowering to 14.29%. The most significant impact occurred in the five-class issue (A-B-C-D-E), with accuracy declining from 89.00% to 20.20% and the F1-score plummeting to only 7.11%.

The findings unequivocally indicate that the convolutional block is crucial for efficient feature extraction from EEG signals. In the absence of the CNN part, the model encounters difficulties in differentiating various EEG patterns, leading to nearly uncontrolled effectiveness in classification. The significant decline in effectiveness across all categorization scenarios underscores that spatial

feature extraction via convolutional processes is essential for the efficacy of our suggested design in detecting epileptic seizures.

5.3.2. Impact of the BiGRU block

The BiGRU component in our suggested design is intended to identify temporal dependencies in EEG signals by analyzing sequences in both forward and backward orientations. We performed ablation research to assess its efficacy by omitting the BiGRU block while retaining the CNN and CBAM aspects.

Table 8 presents intriguing patterns across several classification circumstances. In binary classification tasks (A-E, B-E), the model exhibited excellent results (99.00% across all measures) regardless of including the BiGRU block. The BiGRU block provided a marginal enhancement in the C-E classification, elevating accuracy and F1-score from 97.00% to 97.50%. In the D-E scenario, both models attained comparable outcomes (98.50% across all criteria).

Table 8. Impact of the ResBiGRU block.

Experiment	Model	Performance (mean%)			
		Accuracy	Sensitivity	Specificity	F1-score
(2-Class Problem)					
A-E	The Proposed Model without ResBiGRU	99.00%	99.00%	99.00%	99.00%
	The Proposed CNN-ResBiGRU-CBAM	99.00%	99.00%	99.00%	99.00%
B-E	The Proposed Model without ResBiGRU	99.00%	99.00%	99.00%	99.00%
	The Proposed CNN-ResBiGRU-CBAM	99.00%	99.00%	99.00%	99.00%
C-E	The Proposed Model without ResBiGRU	97.00%	97.00%	97.00%	96.99%
	The Proposed CNN-ResBiGRU-CBAM	97.50%	97.50%	97.50%	97.49%
D-E	The Proposed Model without ResBiGRU	98.50%	98.50%	98.50%	98.50%
	The Proposed CNN-ResBiGRU-CBAM	98.50%	98.50%	98.50%	98.50%
(3-Class Problem)					
AB-CD-E	The Proposed Model without ResBiGRU	95.40%	95.50%	96.08%	95.80%
	The Proposed CNN-ResBiGRU-CBAM	96.20%	96.50%	96.92%	96.66%
(4-Class Problem)					
AB-C-D-E	The Proposed Model without ResBiGRU	92.60%	91.13%	97.25%	91.58%
	The Proposed CNN-ResBiGRU-CBAM	92.00%	90.38%	96.75%	90.94%
(5-Class Problem)					
A-B-C-D-E	The Proposed Model without ResBiGRU	88.20%	88.20%	93.00%	88.16%
	The Proposed CNN-ResBiGRU-CBAM	89.00%	89.00%	89.63%	88.94%

The BiGRU block is significant because it offers more evidence in intricate multi-class situations. In the three-class issue (AB-CD-E), the incorporation of the BiGRU block enhanced the model's accuracy from 95.40% to 96.20%, alongside comparable enhancements in sensitivity (95.50% to 96.50%), specificity (96.08% to 96.92%), and F1-score (95.80% to 96.66%).

Notably, in the four-class issue (AB-C-D-E), the model devoid of BiGRU exhibited marginally more effective results in specific indicators, attaining 92.60% accuracy and 91.58% F1-score, in contrast to 92.00% accuracy and 90.94% F1-score with BiGRU. In the most difficult five-class scenario (A-B-C-D-E), the BiGRU block proved its efficacy by enhancing accuracy from 88.20% to 89.00%, but with a minor reduction in specificity (from 93.00% to 89.63%).

The findings indicate that the BiGRU block has a negligible effect on more straightforward binary classification work but enhances efficiency in more intricate multi-class situations. The bidirectional handling of temporal information is especially advantageous for differentiating between various EEG

signal categories, underscoring the significance of collecting temporal relationships in thorough epileptic seizure detection techniques.

5.3.3. Impact of the CBAM block

The CBAM in our concept aims to improve feature representation through the use of channel and spatial attention techniques. To assess its contribution, we performed an ablation research by eliminating the CBAM block while preserving the CNN and BiGRU components.

Table 9 presents diverse effects across several classification circumstances. In binary classification tasks, the impact of CBAM was negligible to moderate. Incorporating CBAM in the A-E classification resulted in a marginal performance enhancement, elevating all measures from 98.50% to 99.00%. In the B-E, C-E, and D-E situations, the models exhibited identical performance with and without CBAM, attaining 99.00%, 97.50%, and 98.50% accordingly across all criteria.

Table 9. Impact of the CBAM block.

Experiment	Model	Performance (mean %)			
		Accuracy	Sensitivity	Specificity	F1-score
(2-Class Problem)					
A-E	The Proposed Model without CBAM	98.50%	98.50%	98.50%	98.50%
	The Proposed CNN-ResBiGRU-CBAM	99.00%	99.00%	99.00%	99.00%
B-E	The Proposed Model without CBAM	99.00%	99.00%	99.00%	99.00%
	The Proposed CNN-ResBiGRU-CBAM	99.00%	99.00%	99.00%	99.00%
C-E	The Proposed Model without CBAM	97.50%	97.50%	97.50%	97.49%
	The Proposed CNN-ResBiGRU-CBAM	97.50%	97.50%	97.50%	97.49%
D-E	The Proposed Model without CBAM	98.50%	98.50%	98.50%	98.50%
	The Proposed CNN-ResBiGRU-CBAM	98.50%	98.50%	98.50%	98.50%
(3-Class Problem)					
AB-CD-E	The Proposed Model without CBAM	93.40%	93.67%	95.00%	93.93%
	The Proposed CNN-ResBiGRU-CBAM	96.20%	96.50%	96.92%	96.66%
(4-Class Problem)					
AB-C-D-E	The Proposed Model without CBAM	91.60%	89.75%	97.00%	90.43%
	The Proposed CNN-ResBiGRU-CBAM	92.00%	90.38%	96.75%	90.94%
(5-Class Problem)					
A-B-C-D-E	The Proposed Model without CBAM	89.00%	89.00%	92.88%	88.98%
	The Proposed CNN-ResBiGRU-CBAM	89.00%	89.00%	89.63%	88.94%

The contribution of the CBAM is undeniable in multi-class classification jobs. The incorporation of CBAM in the three-class problem (AB-CD-E) resulted in significant enhancements, elevating accuracy from 93.40% to 96.20%, sensitivity from 93.67% to 96.50%, specificity from 95.00% to 96.92%, and F1-score from 93.93% to 96.66%. This signifies one of the most substantial enhancements seen in our ablation research.

In the four-class classification (AB-C-D-E), the CBAM block demonstrated minor enhancements, with accuracy rising from 91.60% to 92.00%, sensitivity growing from 89.75% to 90.38%, and F1-score improving from 90.43% to 90.94%, but with a minor reduction in specificity from 97.00% to 96.75%. In the five-class scenario (A-B-C-D-E), both models attained the same accuracy (89.00%), although the model devoid of CBAM exhibited marginally superior specificity (92.88% compared to 89.63%).

The outcomes demonstrate that the attention mechanism of the CBAM block is highly proficient in managing tasks of moderate complexity, notably in three-class categorization. Although its effect

may be diminished in binary classification or very intricate multi-class situations, CBAM's overall contribution to the model's effectiveness warrants its incorporation into the design, especially for its capacity to improve feature differentiation in multi-class contexts.

5.4. Practical applications

Our CNN-ResBiGRU-CBAM model demonstrates varying levels of improvement across different classification tasks, which should be interpreted with practical applications in mind. Although the five-class and four-class classification tasks achieve lower overall accuracy than binary classification, they hold greater clinical importance for several reasons.

Doctors must differentiate between multiple EEG states in real-world clinical settings rather than simply identifying seizure versus non-seizure conditions. Classifying normal, interictal, and ictal patterns is crucial for monitoring disease progression, assessing treatment effectiveness, understanding seizure onset patterns, and identifying various forms of epileptic movement.

While binary classification (seizure vs. non-seizure) is valuable for alert procedures, multi-class classification offers richer diagnostic insights. It enhances understanding of seizure dynamics, supports better treatment planning, and improves the monitoring of patient responses to interventions. Although the model shows only modest improvements in binary classification compared to baseline models (all exceeding 90% accuracy), its significant advancements in multi-class classification are noteworthy. These tasks are inherently more complex and offer greater clinical relevance, making the improvements in this area particularly impactful.

5.5. Limitations

Although our suggested CNN-ResBiGRU-CBAM model exhibits robust efficacy in epileptic seizure identification, it is crucial to recognize several constraints of the present investigation. The principal limitation pertains to dataset constraints. The assessment was performed exclusively on the Bonn EEG dataset, which, despite its widespread utilization, has a minimal sample size. The dataset comprises pristine, pre-selected EEG segments, while real-world EEG data frequently encompasses several artifacts and noise. The data is derived from a restricted sample of participants, which may not adequately reflect the range of epilepsy symptoms among different categories of patients.

From a signal-processing standpoint, numerous constraints are present. Our proposed methodology utilizes a 20 Hz cutoff filter to mitigate high-frequency noise, which may inadvertently omit pertinent brain activity in the higher frequency range (20–50 Hz) that could provide critical insights for seizure identification. Advanced noise reduction methods, such as independent component analysis (ICA), may more effectively distinguish noise from significant brain signals while maintaining pertinent high-frequency information. Furthermore, our model analyzes single-channel EEG data, whereas clinical environments generally utilize multi-channel recordings. The method does not explicitly account for the variety in seizure patterns that may arise within the same individual over time. Moreover, the model's efficacy in continuous, long-term EEG recordings remains unassessed, which is essential for practical applications.

The clinical application must consider several limits. The existing approach lacks real-time functionality, which is essential for clinical applications. The model's efficacy across various seizure types (focal, generalized, etc.) has not been thoroughly evaluated. Furthermore, the model's

decision-making process involves enhanced interpretability to foster clinician trust and adoption.

Our present technique also faces technical restrictions. The computational demands of the complete architecture may need to be improved for resource-limited devices, thus restricting its implementation in portable or wearable applications. The existing approach needs to tackle the issue of early seizure forecasting, concentrating instead on identification. Furthermore, the model's resilience to adversarial noise and artifacts requires additional examination to guarantee dependable performance in practical applications.

6. Conclusions and future works

This study enhances epileptic seizure identification by presenting an innovative architecture for deep learning that uniquely combines three robust aspects: CNN, residual BiGRU, and CBAM. Our methodology offers multiple innovative contributions to the existing body of investigations: (1) The particular amalgamation of CNN with residual BiGRU facilitates more efficient spatiotemporal feature extraction from EEG signals while mitigating the vanishing gradient issue, (2) the incorporation of CBAM's dual attention mechanism allows for selective emphasis on both channel and spatial features, thereby augmenting the model's capacity to differentiate between various seizure patterns, and (3) the holistic design exhibits outstanding effectiveness in both binary and multi-class classification contexts, attaining up to 99.00% accuracy in binary classification endeavors. These developments signify a substantial progression in automated epileptic seizure identification, providing enhanced accuracy and resilience relative to current methodologies.

The CNN-ResBiGRU-CBAM model surpassed other deep-learning models and prior state-of-the-art methods in binary and multi-class classification evaluations. The empirical findings revealed the efficacy of the hybrid framework in identifying spatial and temporal patterns from EEG signals, resulting in enhanced seizure detection. The CNN component effectively retrieved intricate spatial characteristics, but the residual BiGRU component assimilated long-term temporal relationships. Incorporating the CBAM mechanism improved the model's capacity to differentiate various traits by emphasizing the most informative geographical and temporal areas.

The suggested model demonstrated outstanding outcomes, exceeding all prior models, with outstanding accuracy, sensitivity, specificity, and F1-score across multiple classification tests. It attained an accuracy of up to 99.00% in binary classification tasks, demonstrating its capacity to distinguish between epileptic and non-epileptic EEG segments. In multi-class scenarios, the model exhibited exceptional performance, with accuracies between 89.00% and 96.20%, indicating its proficiency in differentiating several types of EEG data. The comparison investigation established that the CNN-ResBiGRU-CBAM model outperforms baseline models and previous studies in detecting epileptic seizures. It frequently surpassed other deep-learning architectures and manually created feature-based methods, underscoring its potential for practical application in diagnosing and monitoring epilepsy.

Despite these promising results, there are areas for further research to enhance the proposed approach. Future work should include evaluating the model's effectiveness on more extensive and diverse EEG datasets, integrating multi-channel EEG data, expanding its capabilities to detect and predict seizure onset, improving the model's interpretability, and validating its performance in real-time, online seizure detection scenarios. Addressing these aspects will enhance the model's

applicability, clinical relevance, transparency, and practical implementation, facilitating its incorporation into wearable devices or monitoring systems for continuous seizure detection.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This research was funded by the University of Phayao; Thailand Science Research and Innovation Fund (Fundamental Fund); National Science, Research and Innovation Fund (NSRF); and King Mongkut's University of Technology North Bangkok with contract no. KMUTNB-FF-67-B-09.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. A. Guekht, M. Brodie, M. Secco, S. Li, N. Volkers, S. Wiebe, The road to a world health organization global action plan on epilepsy and other neurological disorders, *Epilepsia*, **62** (2021), 1057–1063. <https://doi.org/10.1111/epi.16856>
2. I. E. Scheffer, S. Berkovic, G. Capovilla, M. B. Connolly, J. French, L. Guilhoto, et al., Ilae classification of the epilepsies: Position paper of the ilae commission for classification and terminology, *Epilepsia*, **58** (2017), 512–521. <https://doi.org/10.1111/epi.13709>
3. K. M. Fiest, K. M. Sauro, S. Wiebe, S. B. Patten, C. S. Kwon, J. Dykeman, et al., Prevalence and incidence of epilepsy, *Neurology*, **88** (2017), 296–303. <https://doi.org/10.1212/WNL.000000000000350>
4. M. K. Alharthi, K. M. Moria, D. M. Alghazzawi, H. O. Tayeb, Epileptic disorder detection of seizures using eeg signals, *Sensors*, **22** (2022), 6592. <https://doi.org/10.3390/s22176592>
5. Y. Tang, Q. Wu, H. Mao, L. Guo, Epileptic seizure detection based on path signature and Bi-LSTM network with attention mechanism, *IEEE Trans. Neural Syst. Rehabil. Eng.*, **32** (2024), 304–313. <https://doi.org/10.1109/TNSRE.2024.3350074>
6. J. Jing, H. Sun, J. A. Kim, A. Herlopian, I. Karakis, M. Ng, et al., Development of expert-level automated detection of epileptiform discharges during electroencephalogram interpretation, *JAMA Neurol.*, **77** (2020), 103–108. <https://doi.org/10.1001/jamaneurol.2019.3485>
7. Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, J. Faubert, Deep learning-based electroencephalography analysis: A systematic review, *J. Neural Eng.*, **16** (2019), 051001. <https://doi.org/10.1088/1741-2552/ab260c>
8. A. Shoeibi, M. Khodatars, N. Ghassemi, M. Jafari, P. Moridian, R. Alizadehsani, et al., Epileptic seizures detection using deep learning techniques: A review, *Int. J. Environ. Res. Public Health*, **18** (2021), 5780. <https://doi.org/10.3390/ijerph18115780>

9. M. Kaseris, I. Kostavelis, S. Malassiotis, A comprehensive survey on deep learning methods in human activity recognition, *Mach. Learn. Knowl. Extr.*, **6** (2024), 842–876. <https://doi.org/10.3390/make6020040>
10. A. Shoeibi, N. Ghassemi, R. Alizadehsani, M. Rouhani, H. Hosseini-Nejad, A. Khosravi, et al., A comprehensive comparison of handcrafted features and convolutional autoencoders for epileptic seizures detection in eeg signals, *Expert Syst. Appl.*, **163** (2021), 113788. <https://doi.org/10.1016/j.eswa.2020.113788>
11. G. Xu, T. Ren, Y. Chen, W. Che, A one-dimensional cnn-lstm model for epileptic seizure recognition using eeg signal analysis, *Front. Neurosci.*, **14** (2020), 1–9. <https://doi.org/10.3389/fnins.2020.578126>
12. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2016), 770–778. <https://doi.org/10.1109/CVPR.2016.90>
13. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, et al., Attention is all you need, in *31st International Conference on Neural Information Processing Systems (NIPS'17)*, (2017), 6000–6010.
14. S. Woo, J. Park, J. Y. Lee, I. S. Kweon, Cbam: Convolutional block attention module, in *European Conference on Computer Vision (ECCV)* (eds. V. Ferrari, M. Hebert, C. Sminchisescu and Y. Weiss), Springer International Publishing, Cham, (2018), 3–19.
15. A. Ju, Z. Wang, Convolutional block attention module based on visual mechanism for robot image edge detection, *EAI Endorsed Trans. Scalable Inf. Syst.*, **9** (2021), 1–9. <https://doi.org/10.4108/eai.19-11-2021.172214>
16. S. J. M. Smith, EEG in the diagnosis, classification, and management of patients with epilepsy, *J. Neurol. Neurosurg. Psychiatry*, **76** (2005), ii2–ii7. <https://doi.org/10.1136/jnnp.2005.069245>
17. A. T. Tzallas, M. G. Tsipouras, D. I. Fotiadis, Epileptic seizure detection in eegs using time–frequency analysis, *IEEE Trans. Inf. Technol. Biomed.*, **13** (2009), 703–710. <https://doi.org/10.1109/TITB.2009.2017939>
18. U. R. Acharya, S. V. Sree, G. Swapna, R. J. Martis, J. S. Suri, Automated eeg analysis of epilepsy: A review, *Knowl.-Based Syst.*, **45** (2013), 147–165. <https://doi.org/10.1016/j.knosys.2013.02.014>
19. T. N. Alotaiby, S. A. Alshebeili, T. Alshawi, I. Ahmad, F. E. A. El-Samie, Eeg seizure detection and prediction algorithms: A survey, *EURASIP J. Adv. Signal Process.*, **2014** (2014), 183. <https://doi.org/10.1186/1687-6180-2014-183>
20. S. Ramgopal, S. Thome-Souza, M. Jackson, N. E. Kadish, I. Fernández, J. Klehm, et al., Seizure detection, seizure prediction, and closed-loop warning systems in epilepsy, *Epilepsy Behav.*, **37** (2014), 291–307. <https://doi.org/10.1016/j.yebeh.2014.06.023>
21. B. Maimaiti, H. Meng, Y. Lv, J. Qiu, Z. Zhu, Y. Xie, et al., An overview of eeg-based machine learning methods in seizure prediction and opportunities for neurologists in this field, *Neuroscience*, **481** (2022), 197–218. <https://doi.org/10.1016/j.neuroscience.2021.11.017>

22. A. Shoeb, A. Kharbouch, J. Soegaard, S. Schachter, J. Guttag, A machine-learning algorithm for detecting seizure termination in scalp eeg, *Epilepsy Behav.*, **22** (2011), S36–S43. <https://doi.org/10.1016/j.yebeh.2011.08.040>
23. A. K. Tiwari, R. B. Pachori, V. Kanhangad, B. K. Panigrahi, Automated diagnosis of epilepsy using key-point-based local binary pattern of eeg signals, *IEEE J. Biomed. Health Inf.*, **21** (2017), 888–896. <https://doi.org/10.1109/JBHI.2016.2589971>
24. H. Al-Hadeethi, S. Abdulla, M. Diykh, R. C. Deo, J. H. Green, Adaptive boost ls-svm classification approach for time-series signal classification in epileptic seizure diagnosis applications, *Expert Syst. Appl.*, **161** (2020), 113676. <https://doi.org/10.1016/j.eswa.2020.113676>
25. J. Vicnesh, Y. Hagiwara, Accurate detection of seizure using nonlinear parameters extracted from eeg signals, *J. Mech. Med. Biol.*, **19** (2019), 1940004. <https://doi.org/10.1142/S0219519419400049>
26. R. Rosas-Romero, E. Guevara, K. Peng, D. K. Nguyen, F. Lesage, P. Pouliot, et al., Prediction of epileptic seizures with convolutional neural networks and functional near-infrared spectroscopy signals, *Comput. Biol. Med.*, **111** (2019), 103355. <https://doi.org/10.1016/j.compbiomed.2019.103355>
27. Y. Zhang, Y. Guo, P. Yang, W. Chen, B. Lo, Epilepsy seizure prediction on eeg using common spatial pattern and convolutional neural network, *IEEE J. Biomed. Health Inf.*, **24** (2020), 465–474. <https://doi.org/10.1109/JBHI.2019.2933046>
28. X. Ma, S. Qiu, Y. Zhang, X. Lian, H. He, Predicting epileptic seizures from intracranial eeg using lstm-based multi-task learning, in *Pattern Recognition and Computer Vision* (eds. J. H. Lai, C. L. Liu, X. Chen, J. Zhou, T. Tan, N. Zheng and H. Zha), Springer International Publishing, Cham, (2018), 157–167. https://doi.org/10.1007/978-3-030-03335-4_14
29. H. Daoud, M. A. Bayoumi, Efficient epileptic seizure prediction based on deep learning, *IEEE Trans. Biomed. Circuits Syst.*, **13** (2019), 804–813. <https://doi.org/10.1109/TBCAS.2019.2929053>
30. G. C. Jana, R. Sharma, A. Agrawal, A 1D-CNN-spectrogram based approach for seizure detection from eeg signal, *Procedia Comput. Sci.*, **167** (2020), 403–412. <https://doi.org/10.1016/j.procs.2020.03.248>
31. X. Hu, S. Yuan, F. Xu, Y. Leng, K. Yuan, Q. Yuan, Scalp eeg classification using deep Bi-LSTM network for seizure detection, *Comput. Biol. Med.*, **124** (2020), 103919. <https://doi.org/10.1016/j.compbiomed.2020.103919>
32. K. Tsiouris, V. Pezoulas, M. Zervakis, S. Konitsiotis, D. Koutsouris, D. Fotiadis, A long short-term memory deep learning network for the prediction of epileptic seizures using eeg signals, *Comput. Biol. Med.*, **99** (2018), 24–37. <https://doi.org/10.1016/j.compbiomed.2018.05.019>
33. J. Wang, S. Cheng, J. Tian, Y. Gao, A 2D CNN-LSTM hybrid algorithm using time series segments of eeg data for motor imagery classification, *Biomed. Signal Process. Control.*, **83** (2023), 104627. <https://doi.org/10.1016/j.bspc.2023.104627>
34. A. M. Roy, Adaptive transfer learning-based multiscale feature fused deep convolutional neural network for eeg mi multiclassification in brain–computer interface, *Eng. Appl. Artif. Intell.*, **116** (2022), 105347. <https://doi.org/10.1016/j.engappai.2022.105347>

35. R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, 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, *Phys. Rev. E*, **64** (2001), 061907. <https://doi.org/10.1103/PhysRevE.64.061907>
36. S. Hochreiter, The vanishing gradient problem during learning recurrent neural nets and problem solutions, *Int. J. Uncertain. Fuzziness Knowl. Based Syst.*, **6** (1998), 107–116. <https://doi.org/10.1142/S0218488598000094>
37. K. Cho, B. van Merriënboer, D. Bahdanau, Y. Bengio, On the properties of neural machine translation: Encoder–decoder approaches, in *Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation (SSST-8)*, Association for Computational Linguistics, Doha, Qatar, (2014), 103–111. <https://doi.org/10.3115/v1/W14-4012>
38. J. Chung, C. Gulcehre, K. Cho, Y. Bengio, Empirical evaluation of gated recurrent neural networks on sequence modeling, in *NIPS 2014 Deep Learning and Representation Learning Workshop*, (2014), 1–9. <https://doi.org/10.48550/arXiv.1412.3555>
39. S. Agac, O. D. Incel, On the use of a convolutional block attention module in deep learning-based human activity recognition with motion sensors, *Diagnostics*, **13** (2023), 1861. <https://doi.org/10.3390/diagnostics13111861>
40. P. Nagabushanam, S. T. George, S. Radha, Eeg signal classification using LSTM and improved neural network algorithms, *Soft Comput.*, **24** (2020), 9981–10003. <https://doi.org/10.1007/s00500-019-04515-0>
41. S. Mallick, V. Baths, Novel deep learning framework for detection of epileptic seizures using eeg signals, *Front. Comput. Neurosci.*, **18** (2024), 1–17. <https://doi.org/10.3389/fncom.2024.1340251>
42. A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, et al., Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals, *Circulation*, **101** (2000), e215–e220. <https://doi.org/10.1161/01.CIR.101.23.e215>
43. Y. Song, C. Fan, X. Mao, Optimization of epilepsy detection method based on dynamic eeg channel screening, *Neural Networks*, **172** (2024), 106119. <https://doi.org/10.1016/j.neunet.2024.106119>
44. H. I. Fawaz, G. Forestier, J. Weber, L. Idoumghar, P. A. Muller, Deep learning for time series classification: A review, *Data Min. Knowl. Discovery*, **33** (2019), 917–963. <https://doi.org/10.1007/s10618-019-00619-1>
45. K. Akyol, Stacking ensemble based deep neural networks modeling for effective epileptic seizure detection, *Expert Syst. Appl.*, **148** (2020), 113239. <https://doi.org/10.1016/j.eswa.2020.113239>
46. D. K. Thara, B. G. PremaSudha, F. Xiong, Epileptic seizure detection and prediction using stacked bidirectional long short term memory, *Pattern Recognit. Lett.*, **128** (2019), 529–535. <https://doi.org/10.1016/j.patrec.2019.10.034>
47. K. Shekocar, S. Dour, G. Ahmad, Epileptic seizure classification using lstm, in *2021 8th International Conference on Signal Processing and Integrated Networks (SPIN)*, (2021), 591–594. <https://doi.org/10.1109/SPIN52536.2021.9566118>

48. A. Zahra, N. Kanwal, N. ur Rehman, S. Ehsan, K. D. McDonald-Maier, Seizure detection from eeg signals using multivariate empirical mode decomposition, *Comput. Biol. Med.*, **88** (2017), 132–141. <https://doi.org/10.1016/j.compbiomed.2017.07.010>
49. S. Montaha, S. Azam, A. K. M. R. H. Rafid, P. Ghosh, M. Z. Hasan, M. Jonkman, et al., Breastnet18: A high accuracy fine-tuned vgg16 model evaluated using ablation study for diagnosing breast cancer from enhanced mammography images, *Biology*, **10** (2021), 1347. <https://doi.org/10.3390/biology10121347>
50. C. de Vente, L. H. Boulogne, K. V. Venkadesh, C. Sital, N. Lessmann, C. Jacobs, et al., Improving automated covid-19 grading with convolutional neural networks in computed tomography scans: An ablation study, preprint, arXiv:2009.09725.
51. R. Meyes, M. Lu, C. W. de Puiseau, T. Meisen, Ablation studies in artificial neural networks, preprint, arXiv:1901.08644.

Appendix

Table A.1. The summary of the hyperparameters for the CNN network used in this work.

Stage	Hyperparameters	Values
Architecture	1D-Convolution	Kernel Size Stride Filters
		5 1 32
	Dropout	0.25
	Max Pooling	2
	Flatten Dense	- 128
Training	Loss Function	Cross-entropy
	Optimizer	Adam
	Batch Size	64
	Number of Epochs	200

Table A.2. The summary of the hyperparameters for the LSTM network used in this work.

Stage	Hyperparameters	Values
Architecture	LSTM Unit	128
	Dropout	0.25
	Dense	128
Training	Loss Function	Cross-entropy
	Optimizer	Adam
	Batch Size	64
	Number of Epochs	200

Table A.3. The summary of the hyperparameters for the BiLSTM network used in this work.

Stage	Hyperparameters	Values
Architecture	BiLSTM Unit	128
	Dropout	0.25
	Dense	128
Training	Loss Function	Cross-entropy
	Optimizer	Adam
	Batch Size	64
	Number of Epochs	200

Table A.4. The summary of the hyperparameters for the GRU network used in this work.

Stage	Hyperparameters	Values
Architecture	GRU Unit	128
	Dropout	0.25
	Dense	128
Training	Loss Function	Cross-entropy
	Optimizer	Adam
	Batch Size	64
	Number of Epochs	200

Table A.5. The summary of the hyperparameters for the BiGRU network used in this work.

Stage	Hyperparameters	Values
Architecture	BiGRU Unit	128
	Dropout	0.25
	Dense	128
Training	Loss Function	Cross-entropy
	Optimizer	Adam
	Batch Size	64
	Number of Epochs	200

Table A.6. The summary of the hyperparameters for the CNN-ResBiGRU-CBAM network used in this work.

Stage	Hyperparameters	Values	
Architecture	<u>Convolution Block</u>		
	1D-Convolution	Kernel Size	5
		Stride	1
		Filters	256
	Batch Normalization		-
	Activation		Smish
	Max Pooling		2
	Dropout		0.25
	1D-Convolution	Kernel Size	5
		Stride	1
		Filters	128
	Batch Normalization		-
	Activation		Smish
	Max Pooling		2
	Dropout		0.25
	1D-Convolution	Kernel Size	5
		Stride	1
		Filters	64
	Batch Normalization		-
	Activation		Smish
Max Pooling		2	
Dropout		0.25	
1D-Convolution	Kernel Size	5	
	Stride	1	
	Filters	32	
Batch Normalization		-	
Activation		Smish	
Max Pooling		2	
Dropout		0.25	
<u>ResBiGRU Block</u>			
ResBiGRU_1	Neural	128	
ResBiGRU_2	Neural	64	
<u>Attention Block</u>			
CBAM			
	Dropout	0.25	
	Dense	128	
	Activation	SoftMax	
Training	Loss Function	Cross-entropy	
	Optimizer	Adam	
	Batch Size	64	
	Number of Epochs	200	



AIMS Press

© 2025 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>)