

RESEARCH

Open Access



Epileptic seizure prediction via multidimensional transformer and recurrent neural network fusion

Rong Zhu^{1,2}, Wen-xin Pan^{1,2*} , Jin-xing Liu^{1,2} and Jun-liang Shang^{1,2}

Abstract

Background Epilepsy is a prevalent neurological disorder in which seizures cause recurrent episodes of unconsciousness or muscle convulsions, seriously affecting the patient's work, quality of life, and health and safety. Timely prediction of seizures is critical for patients to take appropriate therapeutic measures. Accurate prediction of seizures remains a challenge due to the complex and variable nature of EEG signals. The study proposes an epileptic seizure model based on a multidimensional Transformer with recurrent neural network (LSTM-GRU) fusion for seizure classification of EEG signals.

Methodology Firstly, a short-time Fourier transform was employed in the extraction of time-frequency features from EEG signals. Second, the extracted time-frequency features are learned using the Multidimensional Transformer model. Then, LSTM and GRU are then used for further learning of the time and frequency characteristics of the EEG signals. Next, the output features of LSTM and GRU are spliced and categorized using the gating mechanism. Subsequently, seizure prediction is conducted.

Results The model was tested on two datasets: the Bonn EEG dataset and the CHB-MIT dataset. On the CHB-MIT dataset, the average sensitivity and average specificity of the model were 98.24% and 97.27%, respectively. On the Bonn dataset, the model obtained about 99% and about 98% accuracy on the binary classification task and the tertiary upper classification task, respectively.

Conclusion The findings of the experimental investigation demonstrate that our model is capable of exploiting the temporal and frequency characteristics present within EEG signals.

Keywords Epilepsy prediction, Transformer, Recurrent neural network, EEG signals

Introduction

Epilepsy affects approximately 50 million people with epilepsy worldwide. Moreover, age, sex, or race can be affected by epilepsy, which is mainly characterized by

recurrent episodes of momentary brain dysfunction [1]. During epileptic seizures, abnormal discharges of brain electrical activity occur, leading to various symptoms, such as loss of consciousness and muscle spasms. Epileptic seizures are unpredictable, causing great inconvenience and hidden problems in patients' work and lives [2]. A number of techniques have been developed with the objective of enabling the detection and prediction of epileptic seizures. Amongst these techniques, those based on magnetic resonance imaging (MRI), magnetoencephalography (MEG) and electroencephalography

*Correspondence:

Wen-xin Pan
panwenxin1006@163.com

¹ School of Computer Science, Qufu Normal University, Rizhao 276826, Shandong, China

² Rizhao-Qufu Normal University Joint Technology Transfer Center, Qufu Normal University, Rizhao 276826, Shandong, China



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

(EEG) have proven particularly effective. Of these, EEG has been more widely studied [3].

EEG-based epilepsy prediction methods are among the most widely studied techniques. EEG is a technique that records the electrical activity of neurons. It has been employed extensively in the detection and diagnostic treatment of epilepsy [4]. For EEG, researchers have processed EEG signals by signal processing methods such as short-time Fourier transform (STFT), wavelet transform (WT), and continuous wavelet transform (CWT). Saly Abd-Elateif El-Gindy et al. [5] predicted epileptic seizures by using different attributes of wavelet-transformed EEG signals, such as amplitude, localized mean, derivative, and entropy. In a study by Yan Jianjun and colleagues [6], the authors employed the STFT to extract time-frequency information from electroencephalography data to develop a classification prediction model for epilepsy.

In light of the accelerated pace of technological advancement, machine learning has been employed in the domain of epilepsy classification, and a plethora of techniques for epilepsy classification have been devised. In the field of machine learning, triple classification or binary classification of epileptic seizures has been studied using random forests [7, 8], support vector machines [9, 10], k-nearest neighbors [11, 12], and decision trees [13, 14]. Similarly, there are many ways to improve machine learning; for example, Al-Hussaini et al. [15] developed a new seizure detection framework through machine learning, which was used to improve the sensitivity of wearable devices. Mary et al. [16] used random forests and least squares support vector machines to classify EEG signals to improve the efficiency of clinical epilepsy EEG signal analysis. He Jiaxiu et al. [17] used SVM and gradient-enhanced decision tree (GBDT) classifiers for classifying epilepsy, and the classification accuracy reached 90.00%.

Although machine learning is widely used in epilepsy detection and classification [18], it has poor representation learning capability and inadequate fitting ability. For epilepsy detection with the application of EEG signals, machine learning is less capable of modeling time-series data and is computationally inefficient, and it is difficult to scale the model. To solve these problems, epilepsy researchers have proposed various deep learning methods for improving the deficiencies in machine learning models. Huang Zixuan et al. [19] proposed a novel graph-regularized fuzzy generalized learning system (GFBLs). GFBLs performs IED detection by feature input and obtains an accuracy rate of up to 91%, greatly decreasing the training time. EEGs contain rich time series and frequency information, and EEGs can be categorized in several ways using recurrent neural networks (RNNs) as well as improved RNNs, long

short-term memory networks (LSTMs) and gated recurrent units (GRUs). Aliyu et al. [20] proposed, among others, the LSTM network for the classification of epileptic EEG signals, and the results obtained are superior to those of machine learning methods such as logistic regression (LR) and SVM. Ma Yahong et al. [21] designed a multichannel feature fusion model that combines a CNN and Bi-LSTM, extracts spatial features through a CNN and temporal features through Bi-LSTM, and finally fuses the classification through an attention mechanism. The average accuracy of this model reached 94.83%, and the prediction of epileptic seizures was accurate. Xin et al. [22] proposed a DRSN-GRU approach for epileptic seizure prediction that combines deep residual shrinkage networks (DRSNs) and GRUs to provide new ideas for epileptic prediction research. Although RNNs have many advantages in dealing with long sequence data, they still suffer from problems such as gradient vanishing and explosion. In contrast, the self-attention mechanism of the Transformer can capture long-distance dependencies more effectively. Moreover, transformers can also address time series problems, so researchers in epilepsy detection have also applied transformers to epilepsy detection. Shu et al. [23] proposed a transformer-based network model for epilepsy detection, called EpilepsyNet. This model demonstrated the efficacy of transformers in detecting epilepsy and produced stable and reliable results.

However, while the Transformer is good at capturing global dependencies, it may be less adept at dealing with local temporal dependencies in time-series data relative to LSTM or GRU. This may be a drawback when dealing with EEG data, as some important physiological features in EEG signals may be highly localized in time. LSTM and GRU require iterative computation on a time-step-by-time-step basis, which restricts their parallel processing capabilities. For large-scale EEG datasets containing a large number of time points, this may lead to slow processing, especially in application scenarios that require fast processing.

Transformers have global modeling capabilities and parallel computing advantages. LSTM and GRUs are suitable for extracting local time-series features, and their combination not only captures complex dependencies in sequence data but also improves the processing efficiency and enhances the robustness and adaptability of the model. In order to address the aforementioned issues, this paper proposes a transformer-based RNN model for seizure prediction in epilepsy patients, which combines a transformer, an LSTM, and a GRU.

This work contributes to the field of seizure detection in the following ways:

1. We propose an epileptic seizure model based on a multidimensional transformer with LSTM-GRU fusion for epilepsy prediction from EEG signals. The proposed method is effective for epilepsy prediction.
2. We used a method that combines the time and frequency information of EEG signals for data extraction and training via a multidimensional transformer encoder and an LSTM-GRU network for more efficient processing and learning of complex patterns and dependencies in EEG time series data.
3. We conducted extensive experiments using 5 cases from the combined Bonn EEG dataset and 20 patients from the CHB-MIT dataset. The findings of the experiment demonstrate the superiority of our method in several key metrics, including accuracy, sensitivity, and specificity, and confirm its significant performance advantage over other competing methods, thereby providing additional validation of its effectiveness and reliability in seizure prediction.

The other components of this paper are listed below. Section 2 describes the experimental dataset used in this study and the setup of this dataset for the experiments, as well as the proposed methodology and method setup for the EEG data; Sect. 3 presents the experimental results of the proposed model on two datasets and the comparison results with similar experimental models; Sect. 4 discusses the comparison of the proposed model with the current state-of-the-art models, the interpretability of the proposed model, as well as the current shortcomings of the proposed model and its future development; and Sect. 5 gives the conclusions of our experiments.

Materials and methods

Dataset

CHB-MIT dataset

The CHB-MIT dataset [24] contains epileptic EEG data from Children's Hospital Boston included in the MIT EEG database, which is available at <https://archive.physionet.org/physiobank/database/chbmit/>. The CHB-MIT dataset contains 23 cases from 22 pediatric epilepsy patients, where the 1st and 21st EEG recordings are from the same patient and the EEG recordings are separated by 1.5 years. The EEG recordings were conducted in accordance with the internationally recognized 10–20 electrode arrangement system. Each EEG recording encompassed a range of 9–42 EEG files, with the exception of a subset of EEG files from chb04, chb06, chb07, chb09, chb10, and chb23, which exhibited a predominant recording duration of 2 or 4 h. The recording times were all 1 h. The database comprises approximately 844 h of EEG recordings, encompassing 163 seizures. The onset and offset of each seizure are clearly delineated.

In performing the experiments, we reconfigured the CHB-MIT dataset for experimental purposes. Among the 24 data points in the CHB-MIT dataset, 23 electrodes were used, except for the chb12, chb14, chb15, chb16, chb17, and chb18 data, for which 18 electrodes were used; in addition, chb24 had a serious lack of information and an unknown number of electrodes. To standardize the number of electrodes, we unified the EEG data of the following 18 electrodes: P3-O1, FP1-F3, FP2-F4, F3-C3, F4-C4, F7-T7, T8-P8, F8-T8, FP2-F8, T7-P7, C3-P3, P4-O2, FP1-F7, P7-O1, P8-O2, C4-P4, CZ-PZ, and FZ-CZ. where the individual electrode [25] letters represent the following: FP is the anterior pole, F is the frontal lobe, T is the temporal lobe, O is the occipital lobe, C is the central lobe, and P is the parietal lobe. The Chb12 and Chb13 channels change frequently during recording, which may lead to contamination of the EEG recordings. The frequency of seizures in Chb24 patients is high, so we lacked sufficient seizure interval data for training the model [6]. The amount of data in Chb04 was very large and beyond the processing capacity of our model. Therefore, we deleted the EEG data of these four patients in our experiment.

Bonn EEG dataset

The Bonn dataset [26, 27], the Bonn University EEG dataset, is an EEG dataset widely used for epilepsy diagnosis and classification. The dataset is from the Institute for Epileptic Electroencephalography at the University of Bonn, Germany, and contains five subsets of data from five healthy individuals and five epileptic patients. The datasets comprise 100 single-channel EEG segments, with a sample frequency of 173.61 Hz and an average duration of 23.6 s, containing 4097 data points per segment. Subset A and Subset B were obtained from scalp EEGs of 5 healthy individuals. subset A is the EEGs of subjects with eyes open and subset B is the EEGs of subjects with eyes closed. Subset C, subset D, and subset E were obtained from intracranial EEGs of 5 epileptic patients. subset C and subset D recorded interictal EEGs. Subset C consisted of EEGs obtained from the contralateral side of the epileptic foci, whereas subset D comprised EEGs taken from the epileptic foci themselves. EEG acquired from the epileptic focus. Subset F records seizure-phase EEG data and contains all data captured by intracranial electrodes. Segments of the Bonn dataset were manually clipped from long-range multichannel EEGs, and possible interferences such as myokinetic artifacts and oculomotor artifacts were removed during the clipping process.

In the experiment, we divided these five cases into five cases based on clinical correlation studies. Case 1 (AB-CD-E) and Case 2 (A-D-E) studied the tri-categorical

EEG signals: ictal, interictal, and normal. Case 3 (ABCD-E), Case 4 (AB-E) and Case 5 (AB-CD) studied dichotomized EEG signals. Cases 3 and 4 study epileptic seizures and non-epileptic seizures, and the epileptic seizure data for both cases are reported from subset E. The non-epileptic seizure data for Case 3 are from subsets A, B, C, and D, and include normal signals and signals from the inter-seizure interval as non-epileptic seizure signals. The non-epileptic seizure data for Case 4 comes from Subset A and Subset B and contains only normal EEG signals. Case 5 examines the EEG signals of epileptic unseizure and inter-seizure periods, with the unseizure EEG data coming from subsets 1 and 2, and the inter-seizure EEG data coming from subsets C and D. Because this dataset has already been processed, we apply it by processing it only as spectrograms as input data to the neural network.

Methods

Figure 1 depicts the flow of the model proposed in pursuance of this study. First, the EEG signals present

within the preictal and interictal stages of the original EEG signals were extracted and labelled. Second, the temporal and frequency characteristics of the EEGs in the preictal and interictal periods were extracted using the STFT, and the frequency spectra were plotted. Then, the spectral information is input to the corresponding encoder to learn the extracted time-frequency features. Subsequently, the temporal characteristics of the EEG signals are further elucidated through the input of the time and frequency features processed by the encoder into the LSTM and GRU, respectively. Fifth, simple feature fusion and classification of the outputs of the LSTM and GRU are performed using a gating mechanism. Finally, postprocessing is used to predict epileptic seizures.

Due to the limited size of the Bonn dataset, we performed spectralization solely to assess the model’s performance in the classification task. To provide a more detailed illustration of our methodology, we used the CHB-MIT dataset as a representative example.

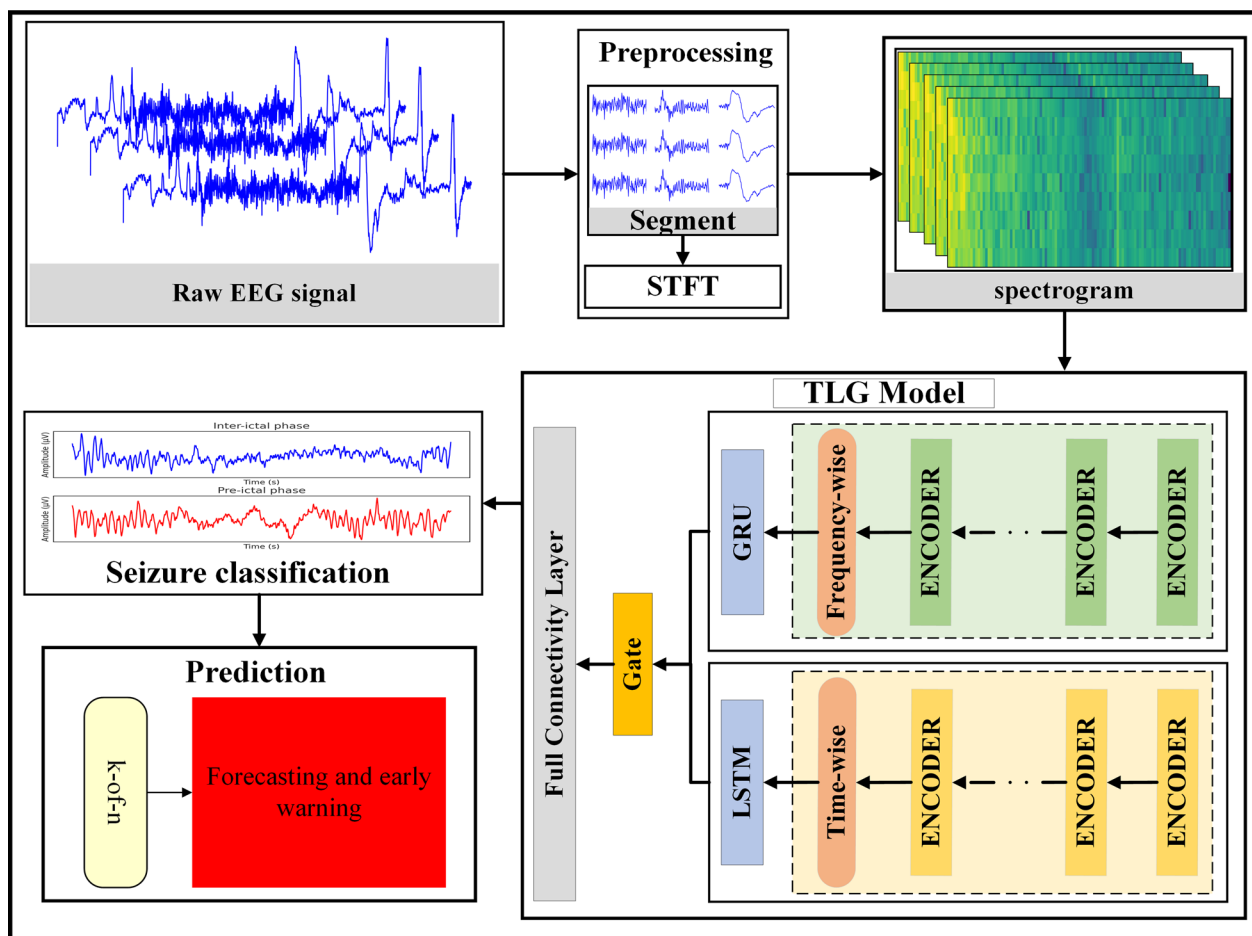


Fig. 1 Flowchart of the epilepsy prediction model

Preprocessing

In epilepsy prediction, we mainly consider the EEG signals of the interictal and preictal phases, so we discard the EEG signal segments from the ictal phase in the CHB-MIT dataset and transform epileptic seizure prediction from a three-classification problem to a two-classification problem. Due to the large amount of CHB-MIT data, at the beginning of preprocessing, we segmented the original EEG signals and labeled the segmented data. However, in the CHB-MIT dataset, because some patients have a low number of seizures, very unbalanced data between the interictal and preictal periods are problematic. In our experiments, we found that before the data balancing process, the training metrics would be distorted during the experimental training, and the model would not be able to perform calculations. Due to the very small amount of data in the pre-episode period, there is also an inability to capture the patterns of a few classes in the training. This issue is addressed by utilizing overlapping sliding windows, which results in the acquisition of a greater quantity of preictal data. To ensure that the ratio of interictal data to preictal data was approximately one to one, the preictal window size was set in accordance with the specifications outlined in Eq. (1):

$$W_l = \frac{x_l}{x_i} W, \tag{1}$$

where x_l represents the number of preictal electroencephalogram segments per patient and x_i represents the number of interictal EEG segments for delicious patients. W represents the size of the EEG window.

After segmenting the original data, we transform the segmented data into spectral data for model training. The present study employs the use of the STFT to convert electroencephalographic signals into spectrograms, which are composed of both temporal and spectral domains, and to analyse these EEG signals within the context of the time-frequency domain. We chose the cosine analysis window for the 5-second samples for the STFT. During the data analysis, we found that the EEG data recordings used were disturbed by 60 Hz power line noise. Consequently, in the course of our experiments, components were rejected in the frequency ranges of 55–65 Hz and 115–125 Hz, utilizing band-stop filters and DC components at 1 Hz, respectively, to facilitate the effective removal of power line noise and DC component interference. The use of spectral data may result in the emergence of extreme values or a considerable degree of variability within the signal data, which could subsequently lead to issues such as gradient explosion or gradient vanishing during model training. This could have a detrimental impact on the overall processing efficiency of the data. To address this issue, this study also normalized the spectral data using a combination of logarithmic transformation and min–max range scaling.

Model

Transformers have very powerful modeling capabilities in dealing with time series data. In recent years, the Transformer has been used for epilepsy detection and classification research. As illustrated in Fig. 2, we present a novel network architecture for the analysis and prediction of EEG signals associated with epileptic seizures.

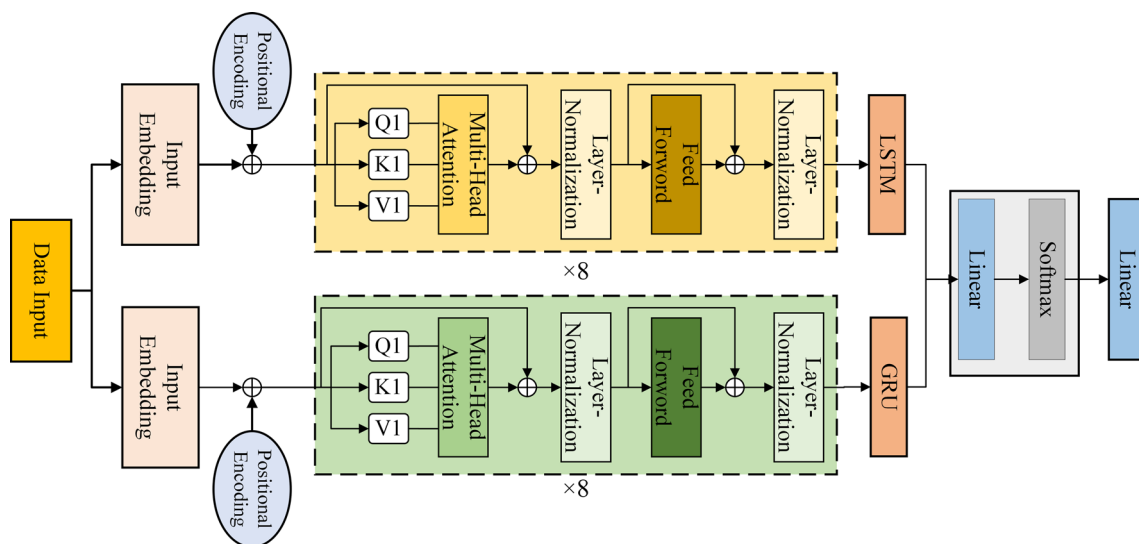


Fig. 2 General architecture diagram of the proposed network model

This approach integrates models such as Transformer encoder, LSTM, and GRU networks.

The preprocessed EEG data, after the time-frequency domain transformer, yield EEG segments with three-dimensional features (S, T, F), where S represents the time dimension of the EEG, T represents the channel dimension of the EEG, and F represents the frequency dimension of the EEG. However, the traditional transformer architecture is used to process text, and text data typically exist in a two-dimensional format (L, D). Therefore, we flatten the matrices $\delta \in \mathbb{R}^{S \times T \times F}$ of two different dimensions, time and frequency, into two-dimensional matrices $\delta_S \in \mathbb{R}^{S \times (T \cdot F)}$, $\delta_F \in \mathbb{R}^{F \times (S \cdot T)}$. These two matrices were used as inputs to the model, while positional coding was added to the input embedding of frequency and time at the model inputs. To comprehensively capture the serial correlation within our data, we implemented two sets of encoders, each tailored to a specific dimension. This approach enabled us to analyse the correlation patterns in terms of both frequency and time step independently for each dimension. The two sets of encoders have the same structure, with each layer including a multihead self-attention mechanism and a subsequent feedforward layer. In the multihead attention mechanism, the output matrix of each attention mechanism is calculated by Eq. (2):

$$\text{Attention}(Q, K, V) = \text{soft max} \left(\frac{QK^T}{\sqrt{d_k}} \right) V, \quad (2)$$

where Q stands for Query, K stands for Key and V stands for Value. The multihead self-attention mechanism [28] enables the entire model to concentrate on distinct representation subspaces of the input, thereby enhancing the diversity of the learned representations. The combination of information from different subspaces enables the model to achieve a more comprehensive feature representation, as specifically manifested in Eq. (3).

$$\text{MultiHead}(Q, K, V) = \text{Concat}(h_1, \dots, h_m), \quad (3)$$

where m is the number of attention heads, W_i^Q , W_i^K and W_i^V are the learned projection matrices, and $h_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)$.

The output matrix of the multihead attention layer is fed into the feed-forward neural network, which can operate independently at each position, so it can process the vectors at all positions in parallel, thus improving training efficiency. The Transformer encoder is capable of efficiently capturing contextual information and handling long-distance dependencies through a self-attention mechanism. GRU and LSTM are good at modeling local features and long-distance dependencies

in time series. Accessing the GRU and LSTM after the Transformer encoder can further refine the local temporal features and frequency features. Therefore, after the encoder, we add two recurrent neural networks, LSTM [29] and GRU [30]. Because GRU efficiently handles short-term dependencies through a gating mechanism that can quickly capture frequency changes in EEG signals, the complex gating mechanism of LSTM provides an advantage in processing long time sequences, and it can efficiently remember and utilize long-term information. Therefore, we use LSTM for further learning features in the time dimension and GRU for further learning features in the frequency dimension.

Then, after the GRU and LSTM, the feature information learned by the LSTM and GRU is spliced using a gating mechanism. We set the outputs of LSTM and GRU as M and N , respectively, connect them into vectors, project them to H through a linear layer, and then assign gating weights to each output through softmax μ_1 and μ_2 . In this process, μ_1 is assigned to the output matrix of the LSTM, and μ_2 is assigned to the output matrix of the GRU. Through this mapping, the feature vector τ is finally obtained. The specific process can be expressed by Eqs. (4), (5) and (6):

$$H = W \cdot \text{Concat}(M, N) + b, \quad (4)$$

$$\mu_1, \mu_2 = \text{Soft max}(H), \quad (5)$$

$$\tau = \text{Concat}(M \cdot \mu_1, N \cdot \mu_2). \quad (6)$$

Finally, the feature vectors τ are passed through the fully connected layer into vectors of dimension 2, and then the classification results are obtained.

Postprocessing

To assess the model's ability to predict seizures, we introduced two metrics, the seizure prediction time period (SOP) and the seizure alarm time period (SPH). SOP denotes the time period during which the model predicts that a seizure will occur at a certain time in the future, and SPH denotes the time period from the time the model issues an alarm to the start of the SOP. A real seizure is considered a successful prediction if it occurs within the SOP or after SPH, and a prediction is considered a failure if a seizure occurs within the SPH or if there is no seizure within the SOP. Considering that too long of a seizure prediction time increases patient anxiety, we refer to [6] and set the SPH to 5 min and the SOP to 30 min to achieve a reasonable warning time.

It should be noted that there are certain subjective assumptions in the setting of SOP and SPH, which may not exactly match the EEG characteristics of some

patients before an actual seizure, which may have had some impact on the accuracy of the predictions [6]. In addition, since the model is mainly used to identify preictal and interictal periods, sporadic false alarms in the interictal period may also generate false alarms. To reduce false alarms, we postprocessed the model results. Using the k-of-n method [31], if at least 24 out of 30 consecutive predicted segments are predicted to be positive, the final output is a seizure alert [6]. The application of this postprocessing method can lead to a reduction in the rate of false alarms and an increase in the prediction performance of the model in practical applications.

Result

In this experiment, we employed a comprehensive evaluation strategy based on the metrics of accuracy, sensitivity, F1 score, specificity, precision and false positive rate (FPR), to assess the performance of our designed seizure prediction model.

CHB-MIT dataset

We evaluated the proposed model through an empirical investigation in comparison to the Gated Transformer Network (GTN) model [32] and the Triple Transformer Tower (hereafter referred to as TTT) model [6]. In order to evaluate the predictive efficacy of the models, we utilize sensitivity, precision, and specificity as metrics for assessing the performance of the three distinct models. The results of the experiments conducted on the proposed model are presented in Table 1. Table 2 contains the results of the experiments carried out on the GTN and TTT models. As the GTN and TTT models were not included in the proposed models, the associated confidence intervals were not calculated.

The model is based on the TTT model and the GTN model. The GTN model employs two pairs of identical transformer encoders to process the time-domain and frequency features of the EEG signals. The TTT model uses three transducers to process the time, frequency, and channel characteristics of EEG signals. We believe that compared to using only the transformer encoder to capture the feature information of the EEG signal, adding the LSTM layer and GRU layer after the transformer encoder can capture richer and more complex feature information. The robustness and generalizability of the network model can also be improved by the fusion of different networks.

A comparison of the three indicators for the three models in Tables 1 and 2 shows that the average sensitivity, specificity, and precision of the TTT model were greater than those of the GTN model for the 20 patient samples. However, the proposed model achieved a mean sensitivity of 98.24% and a mean specificity of

Table 1 Sensitivity, specificity and precision of the proposed models

Patient	Sen (%)	Spe (%)	Pre (%)
Chb01	99.07 ± 1.52	99.84 ± 1.24	99.87 ± 1.17
Chb02	99.26 ± 1.37	99.67 ± 1.37	99.69 ± 1.32
Chb03	99.23 ± 1.32	92.81 ± 3.52	93.73 ± 3.67
Chb05	99.91 ± 0.96	99.9 ± 0.61	98.95 ± 0.73
Chb06	97.87 ± 1.93	98.38 ± 1.12	98.76 ± 1.07
Chb07	89.18 ± 3.25	94.89 ± 2.36	95.16 ± 2.25
Chb08	94.00 ± 2.03	91.67 ± 3.08	98.95 ± 1.89
Chb09	98.43 ± 1.72	96.73 ± 1.99	96.65 ± 2.01
Chb10	99.58 ± 1.04	99.69 ± 0.96	99.67 ± 0.97
Chb11	100.00 ± 0.00	92.78 ± 3.52	92.02 ± 3.67
Chb14	100.00 ± 0.00	99.29 ± 0.76	98.36 ± 1.29
Chb15	98.00 ± 1.24	93.67 ± 2.35	83.76 ± 3.74
Chb16	95.87 ± 1.98	97.56 ± 1.42	97.33 ± 1.45
Chb17	97.32 ± 1.67	98.63 ± 1.54	98.61 ± 1.55
Chb18	99.91 ± 0.92	99.86 ± 1.03	99.92 ± 0.96
Chb19	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
Chb20	100.00 ± 0.00	99.97 ± 0.21	99.95 ± 0.29
Chb21	97.76 ± 1.63	96.00 ± 1.93	95.84 ± 2.05
Chb22	99.44 ± 1.25	94.94 ± 2.46	94.3 ± 2.62
Chb23	100.00 ± 0.00	99.94 ± 0.67	99.96 ± 0.53
Average	98.24 ± 1.16	97.27 ± 1.26	97.03 ± 1.72

97.27%, both of which were superior to those of the TTT model. This demonstrates that our model outperforms the other two models for both preictal and interictal detection. In addition, the accuracy of our model was also better than that of the TTT model. To more effectively convey the distinction in average performance between our model and the other models, we have presented the results graphically, as illustrated in Fig. 3.

It is obvious from Fig. 3 that our model outperforms the other two models in all the indicators; in particular, the improvement in sensitivity and specificity is more obvious. This fully verifies that the design of the model has achieved significant improvement and can better recognize the different stages of epileptic seizures. To evaluate the proposed model more comprehensively, we further compared it with the TTT model and the GTN model in terms of F1 score and accuracy. The results of the comparison are presented in Fig. 4.

As shown in Fig. 4, compared with the GTN model, our model performs well on both the accuracy and F1 score metrics, with a large increase. Compared to the TTT model, our model also shows an increase in accuracy and F1 score. Although the magnitude is small, it still shows some improvement. Based on the comparison results, the new model we designed shows some improvement in

Table 2 Sensitivity, specificity and precision of GTN and TTT models

Patient	GTN			TTT		
	Sen (%)	Spe (%)	Pre (%)	Sen (%)	Spe (%)	Pre (%)
Chb01	98.62	98.99	98.96	100.00	99.88	99.9
Chb02	92.61	94.56	94.49	96.88	98.64	98.58
Chb03	90.65	91.01	89.87	97.82	96.87	94.31
Chb05	92.34	90.40	90.6	94.00	99.45	98.95
Chb06	91.26	87.93	88.29	93.51	92.1	92.18
Chb07	89.04	93.99	93.73	99.11	97.12	97.2
Chb08	94.44	91.72	91.89	97.22	95.86	95.89
Chb09	87.59	92.00	91.71	94.53	97.15	97.1
Chb10	95.65	90.33	90.53	96.48	98.2	98.11
Chb11	89.5	86.25	81.14	97.48	96.39	94.69
Chb14	98.59	97.88	97.89	99.29	99.65	99.65
Chb15	91.60	88.87	86.67	93.22	97.43	96.63
Chb16	85.85	88.00	87.74	89.85	92.31	92.11
Chb17	93.19	95.12	95.04	98.54	97.24	97.28
Chb18	94.25	96.06	95.35	98.24	99.67	99.61
Chb19	93.56	91.59	89.05	95.56	94.21	92.35
Chb20	95.20	94.58	94.43	98.64	97.55	97.49
Chb21	89.58	88.63	88.81	93.50	91.55	91.77
Chb22	91.20	91.21	90.35	95.53	96.21	95.79
Chb23	88.25	89.01	88.81	93.33	91.21	91.29
Average	92.10	91.90	91.20	96.13	96.43	96.044

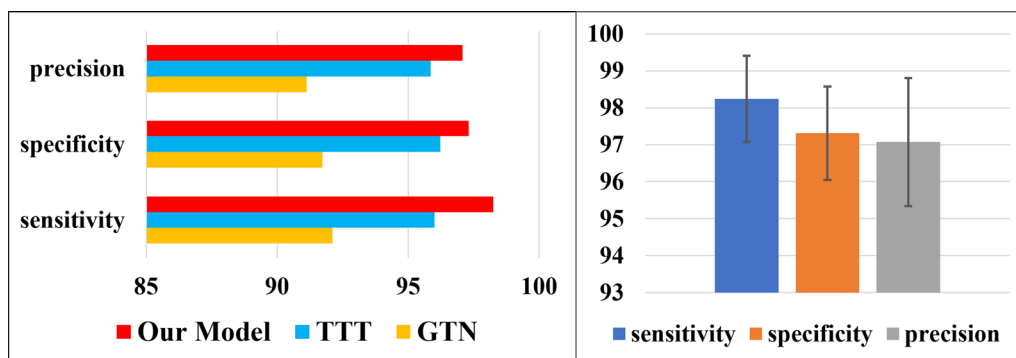


Fig. 3 Comparison of the performance of the proposed model with the TTT model and the GTN model

overall recognition performance compared to the GTN model and the TTT model.

We further compared our model with the TTT model in terms of FPR metrics, and Fig. 5 shows the comparison results. Our proposed model outperforms the TTT model in most cases. Specifically, Fig. 5a shows that the FPR metric curve of our model is always lower than that of the TTT model in most cases. This indicates that our model can effectively reduce the occurrence of false

alarms. This is also confirmed by the quantitative results in Fig. 5b. Our model reduces the average false alarm rate from 0.038/h in the TTT model to 0.033/h, which is a 13.2% reduction. Overall, our model greatly reduces the likelihood of a normal EEG being misclassified as a seizure. Through these quantitative analyses and comparisons, we can see that our proposed model optimizes the FPR metrics and accomplishes the seizure prediction task compared to previous TTT models.

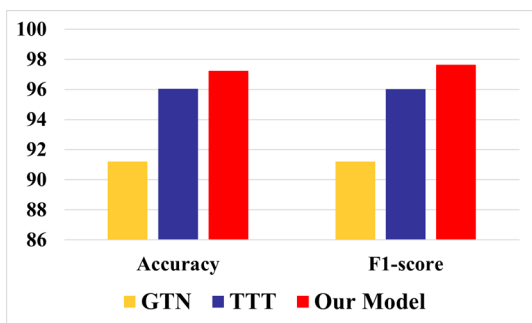


Fig. 4 Comparison of accuracy and F1 scores between the proposed model and the TTT and GTN models

Bonn dataset

In order to verify the ability of the proposed model to collect signals under different conditions, We performed further performance checks on the model using the Bonn dataset. We did three experiments on the Bonn dataset, the first one is the experiment of seizure and normal periods of epileptic, the second one is the experiment of normal, interictal and seizure periods of epileptic, and the third one is the experiment of normal and interictal periods of epileptic. Tables 3 and 4 show the dichotomized and trichotomized data respectively

Cases 3 and 4 in Table 3 were used to validate the normal and ictal periods of seizures, while Case 5 was used to validate the normal and interictal periods of seizures. Case 3 (ABCD-E), which was used to distinguish between normal and interictal periods, performed well in all performance metrics with 99.75% accuracy, 98.75% sensitivity and specificity, and close to 100% precision and F1 score. These results indicate that the model

Table 3 Binary classification performance on the Bonn dataset

Case	Acc (%)	Sen (%)	Spe (%)	Pre (%)	F1 score (%)
Case 3	99.75 ± 2.43	98.75 ± 4.67	98.75 ± 4.67	99.75 ± 2.43	99.74 ± 2.45
Case 4	99.33 ± 3.24	98.50 ± 4.95	98.50 ± 4/95	99.33 ± 3.24	99.32 ± 3.26
Case 5	98.75 ± 2.43	98.24 ± 1.88	98.50 ± 2.13	98.32 ± 2.99	98.24 ± 2.26

Table 4 The triple classification performance on the Bonn dataset

Case	Acc (%)	Sen (%)	Spe (%)	Pre (%)	F1 score (%)
Case 1	98.75 ± 2.43	98.33 ± 2.32	99.17 ± 1.76	98.75 ± 2.43	98.74 ± 2.43
Case 2	98.33 ± 3.24	98.33 ± 2.97	99.17 ± 1.95	98.33 ± 2.97	98.33 ± 2.97

is extremely reliable and stable in recognizing the two states of epileptic seizures. Case 4 (AB-E) was also used to differentiate between normal and seizure phases, but the data used were different from Case 3. The normal phase data in Case 4 contains only two subsets, whereas Case 3 includes four subsets. Although the accuracy, sensitivity, specificity, precision, and F1 score of Case 4 were slightly lower than those of Case 3, its performance was still quite good, indicating that the model is still effective in distinguishing between seizure states. Case 5 (AB-CD) was used to differentiate between normal and interictal periods and its performance was also excellent. The accuracy was 98.75%, sensitivity was 98.24%, specificity was 98.50%, and both precision and F1 score were around 98%. This indicates that the model has good recognition ability in identifying normal and interictal periods.

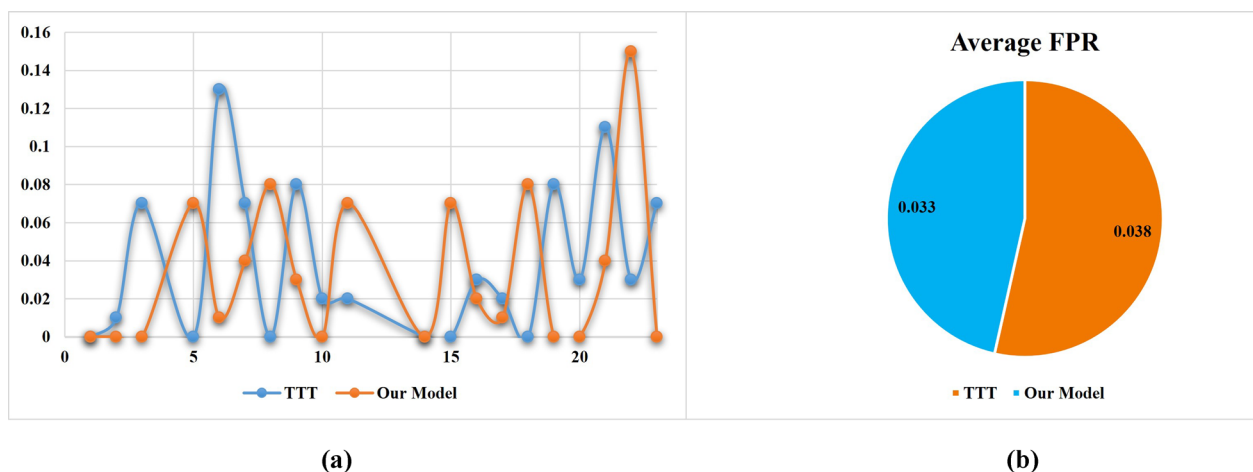


Fig. 5 Comparison of the FPRs between the proposed model and the TTT model. **a** Comparison of FPRs per patient. **b** Comparison of the mean FPRs

Overall, in different status epilepticus classification tasks, although the performance of the models may vary depending on the difficulty of the task and the nature of the dataset, in general, they show a good ability to recognize different states of epileptic seizures.

In Table 4, both Case 1 and Case 2 performed the three classification tasks for seizures (ictal, interictal, and normal). The metrics for Case 1 show that the model performed very well on this dataset, with an accuracy and precision of 98.75%, a sensitivity of 98.33%, an F1 score of 98.74%, and a specificity of 99.17%. This indicates that the model has high accuracy and stability in distinguishing between the three states of epilepsy, and performs particularly well in identifying normal-phase samples. In contrast, Case 2, although the specificity was also 99.17%, the accuracy, sensitivity, precision, and F1 score were 98.33% with a large standard deviation, indicating that the model's performance fluctuated a lot on this dataset, but the performance of all the indexes was high, which also demonstrated the stability of the model.

Discussion

Comparison with advanced work

The CHB-MIT dataset contains rich and diverse EEG data, providing a valuable resource for the study of epilepsy, a neurological disorder. Lately, the Transformer architecture, a new natural language processing architecture, has been used extensively across different bioinformatics domains. In epilepsy research, many researchers have also utilized the CHB-MIT dataset and Transformer to study seizure patterns and classify epilepsy. To further enhance the research results, we present a comparison between the proposed model and the state-of-the-art model in Table 5.

In Table 5, the literature [33] proposes a hybrid epilepsy prediction framework based on hybrid modular feature extraction and a deep transformer, where features are extracted by Fourier transform, and then, epilepsy prediction is performed by a deep transformer, obtaining an average sensitivity of 95.2% and an FPR of 0.02 per hour.

Although the model proposed in the literature [33] outperforms our model in terms of FPR metrics, our model achieves a significant improvement in terms of average sensitivity. Specifically, the model of literature [33] has an FPR of 0.02 per hour, which is lower than our model's FPR of 0.33 per hour. However, the average sensitivity of our model reaches 98.24%, which is much greater than the 86% reported in the literature [33]. This shows that although there is still room for optimization of our model in terms of reducing false positives, our model achieved a substantial improvement in its ability to identify real seizures. This indicates that the modeling framework and technical route we designed are effective in improving the sensitivity of epilepsy recognition.

Wu et al. [34] combined successive variational modal decomposition (SVMD) and a transformer for time-frequency analysis and epilepsy prediction. SVMD decomposes the data into multiple modes and then selects the relevant sequences for processing. Li et al. [35] combined the advantages of CNNs and transformers and proposed the TGCNN model for seizure prediction. Chen et al. [36] proposed a patient-specific approach by preprocessing transformed EEG data to extract epileptic seizure features, which improved the prediction performance to a certain extent. Deng et al. [37] proposed a hybrid visual transformer architecture to address the possible blurring and noisy representations of single EEG sample embedding. Busia et al. [38] proposed a transformer model called the EEGformer model, which reduces the epilepsy detection delay by more than 20% compared to other models.

Compared with the above proposed models, our model has advantages in both average sensitivity and FPR. Among the above studies, the most prominent results were obtained by Saketh Maddineni et al. However, compared to the proposed model, they obtained less sensitivity and FPR. On the CHB-MIT dataset, the proposed model achieved an average sensitivity of 98.24% and an FPR of 0.033 events per hour, both of which were better than those of the models mentioned in the above

Table 5 Comparison with advanced models

Authors	Feature	Classifier	Sen (%)	FPR (/h)
Saketh Maddineni [33]	STFT	Deep Transformer	95.2	0.02
Xiao Wu [34]	SVMD	BERT	86	0.18
Chang Li [35]	STFT	TGCNN	91.5	0.145
Ryan Chen [36]	STN+STFT	CNN	82	0.38
Zhiwei Deng [37]	-	HViT	87.9± 2.2	0.056
Paola Busia [38]	Wavelet Transform	EEGformer	-	-
Our Works	STFT	TLG	98.24	0.033

studies. This indicates that the structure of the model we designed is more conducive to improving the epilepsy detection ability while controlling the FPR within an acceptable range, which fully demonstrates the effectiveness of our proposed model.

In terms of the overall results, our method has a lower FPR, a higher true positive rate, and better overall results in the seizure prediction task. This suggests that our proposed model is superior for extracting key features of epileptic EEG signals and for making accurate predictions. This is mainly attributed to the network structure we adopted and the deep optimization of the model to improve the understanding and processing of complex EEG signals. Overall, our study is an important step forward in the prediction of epileptic seizures.

Interpretability of the proposed model

In order to visualize the interpretability [39, 40] and plausibility of the proposed model more intuitively, we used the t-SNE [41] dimensionality reduction technique for each module of the model. This technique elucidates the adaptive learning ability of the proposed model for seizure EEG signals. We take Case 1 in the Bonn dataset as an example and extract three different types of EEG signal features from each module. The

results are shown in Fig. 6, where these feature vectors are simplified to 2D space for visualization.

Figure 6 shows that as the depth of the model is increased, the borders of the data feature distributions between the different types of EEG signals become more distinct. Initially, the two categories of raw signals were randomly distributed in the two-dimensional plane, mainly due to the large amount of redundancy in the raw signals, when the raw data of epileptic seizure EEG signals were input into the model. With the increase of model depth, the separability of the proposed model for different features gradually increases. In particular, the ability of the model to distinguish seizure EEG signals is significantly improved after learning through the transformer encoder module. This is mainly due to the ability of the Transformer Encoder to deep mine and catch the overall temporal features of the epileptic seizure EEG signals, which effectively focuses on the important features in the EEG signals through the self-attention mechanism to improve the classification accuracy. EEG data after LSTM and GRU module, the same epileptic seizure EEG signal feature information can achieve good clustering, the reason is that LSTM and GRU can effectively capture time-dependent and complex nonlinear patterns in EEG signal processing, which makes the clustering

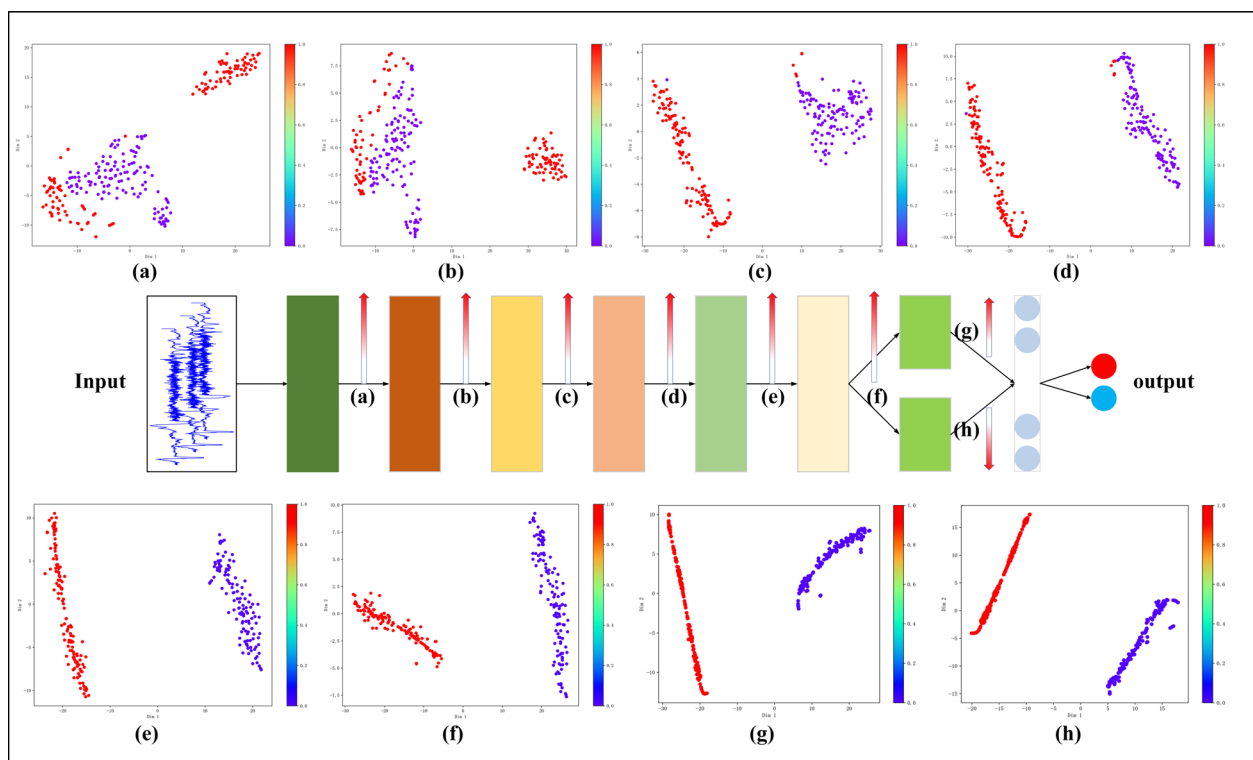


Fig. 6 Visualization of the results of the modules in the proposed model

effect better. Finally, the classification of seizure EEG signals is accomplished by the GATE mechanism.

The findings show that the suggested model is able to identify valid seizure EEG signal features from the raw signals, which fully demonstrates the strong adaptive ability of the model in seizure EEG signal classification and further confirms the reliability of the proposed model in seizure prediction.

Present limitations and future work

Despite the good performance of our work in both epilepsy datasets, there are several limitations that can still be improved. First, due to the experimental conditions, we were not able to acquire the seizure EEG signals ourselves, so we were unable to apply our model to real-time EEG signals or to biosignals, which resulted in our inability to validate the performance of the model in wearable devices. At the same time, we are unable to conduct clinical experiments using the proposed model in the real world, and due to the inability to conduct clinical experiments, we are unable to adequately validate the performance and reliability of the model in a real-world environment. This limitation means that although the model may perform well on offline datasets, it may encounter unforeseen problems in practical applications. Secondly, the proposed model is structurally complex, containing components such as LSTM, GRU and multi-group encoder, which increases the difficulty of debugging and training. Finally, our pre-processing of EEG signals is not good enough to handle certain EEG signals that are difficult to analyse.

In our future work, we plan to improve our research through the following measures: first, we will seek collaboration with hospitals and research institutions to collect more diverse real-time EEG signal data to improve the generalization ability of the model. Second, we will explore the application of the model in real-time signal processing and validate it on wearable devices to evaluate its actual performance. Meanwhile, we will investigate simplifying and optimizing the model structure to reduce debugging difficulty and computational overhead. In addition, we will improve the pre-processing techniques for EEG signals and develop more advanced denoising and feature extraction methods to enhance the robustness of signal processing. Finally, with the improvement of the model, we plan to conduct small-scale clinical trials in the future to evaluate its performance in real clinical settings. These improvements will help overcome the limitations of existing studies, further enhance the utility and reliability of the model, and provide a more effective tool for early detection and treatment of epilepsy.

Conclusion

Seizures often plague epileptic patients, not only causing physical pain but also bringing great pressure to their work and life. How to accurately predict epileptic seizures has become an urgent problem for clinicians and researchers. This paper presents a hybrid deep learning model for the prediction of epileptic seizures. The model integrates the Transformer and recurrent neural networks. We employ the Transformer module to identify the long-range dependent features of electroencephalogram signals, while the GRU and LSTM modules are utilized to learn the temporal characteristics of EEG signals. Then, the learned feature representations from the two models are spliced, fully connected and softmax classified to predict epileptic seizures. The proposed model achieves excellent results on both datasets, performing both prediction and classification tasks. This suggests that compared to a single recurrent neural network or attention model, the model in this paper combines the advantages of both approaches and significantly improves the modeling and prediction of complex EEG signals.

Although some progress has been made, more accurate and reliable seizure predictions still require in-depth research. In the future, we will refine the model structure and investigate the optimal integration of disparate models. Moreover, the development of new time-frequency domain feature extraction and data enhancement techniques is also a direction worth exploring. We expect to continuously improve the model prediction performance, minimize the pain of epilepsy patients and improve the quality of life of epilepsy patients.

Acknowledgements

This work is supported in part by the National Natural Science Foundation of China under Grant No. 62472250, 62172254, 62473179.

Data availability

The data underlying this article will be shared on reasonable request corresponding author.

Received: 12 June 2024 Accepted: 4 September 2024

Published online: 04 October 2024

References

1. WHO: Epilepsy. 2023. <https://www.who.int/news-room/fact-sheets/detail/epilepsy>. Accessed 20 Sept 2023.
2. Katz DI, Bernick C, Dodick DW, Mez J, Stern RA. National institute of neurological disorders and stroke consensus diagnostic criteria for traumatic encephalopathy syndrome. *Neurology*. 2023;96(18):848–63.
3. Hanming Z, Jingang M, Ningning Z, Zhenzhen Z, Ming L. Advances, in deep learning for epilepsy detection. *J Comput Eng Appl*. 2023;59(10):35–47.
4. Yuan Q, Zhou W, Zhang L, Zhang F, Xu F, Leng Y, Wei D, Chen M. Epileptic seizure detection based on imbalanced classification and wavelet packet transform. *Seizure*. 2017;50(10):99.

5. El-Gindy AE, Hamad A, El-Shafai W, Khalaf AAM, El-Samie FEA. Efficient communication and eeg signal classification in wavelet domain for epilepsy patients. *J Ambient Intell Humaniz Comput.* 2021;12(6):1–16.
6. Yan J, Li J, Xu H, Yu Y, Xu T. Seizure prediction based on transformer using scalp electroencephalogram. *Appl Sci.* 2022;12(9):4158.
7. Ben Messaoud R, Chavez M. Random forest classifier for eeg-based seizure prediction. *arXiv e-prints* (2021)
8. Fang Y, Zeng T, Song T. Classification method of eeg based on evolutionary algorithm and random forest for detection of epilepsy. *J Med Imaging Health Inform.* 2020;10(5):979–83.
9. Tapani KT, Nevalainen P, Vanhatalo S, Stevenson NJ. Validating an svm-based neonatal seizure detection algorithm for generalizability, non-inferiority and clinical efficacy. *Comput Biol Med.* 2022;145: 105399.
10. Janga V, Edara SR. Epilepsy and seizure detection using jltn based icffa and multiclass svm classifier. *Traitement du Signal.* 2021;38(3):883–93.
11. Dash DP, Kolekar MH, Jha K. Surface eeg based epileptic seizure detection using wavelet based features and dynamic mode decomposition power along with knn classifier. *Multimed Tools Appl.* 2021;81(29):42057–77.
12. Ghaderyan P, Abbasi A, Sedaaghi MH. An efficient seizure prediction method using knn-based undersampling and linear frequency measures. *J Neurosci Methods.* 2014;232:134–42.
13. Xu X, Lin M, Xu T. Epilepsy seizures prediction based on nonlinear features of eeg signal and gradient boosting decision tree. *Int J Environ Res Public Health.* 2022;19(18):11326.
14. Shoaran M, Farivar M, Emami A. Hardware-friendly seizure detection with a boosted ensemble of shallow decision trees. In: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Orlando, FL, USA. 2016. p. 1826–29.
15. Al-Hussaini I, Mitchell CS. Seizft: interpretable machine learning for seizure detection using wearables. *Bioengineering.* 2023;10(8):918.
16. Mary G, Chitti S, Vallabhaneni RB, Renuka N, et al. Eeg signal classification automation using novel modified random forest approach. *J Sci Ind Res.* 2023;82(1):101–8.
17. He J, Yang L, Liu D, Song Z. Automatic recognition of high-density epileptic eeg using support vector machine and gradient-boosting decision tree. *Brain Sci.* 2022;12(9):1197.
18. Farooq MS, Zulfiqar A, Riaz S. Epileptic seizure detection using machine learning: taxonomy, opportunities, and challenges. *Diagnostics.* 2023;13(6):1058.
19. Huang Z, Duan J. Gfbls: graph-regularized fuzzy broad learning system for detection of interictal epileptic discharges. *Eng Appl Artif Intell.* 2023;125: 106763.
20. Aliyu I, Lim CG. Selection of optimal wavelet features for epileptic eeg signal classification with lstm. *Neural Comput Appl.* 2023;35(2):1077–97.
21. Ma Y, Huang Z, Su J, Shi H, Wang D, Jia S, Li W. A multi-channel feature fusion cnn-bi-lstm epilepsy eeg classification and prediction model based on attention mechanism. *IEEE Access.* 2023;11:62855–64.
22. Xu X, Zhang Y, Zhang R, Xu T. Patient-specific method for predicting epileptic seizures based on drsn-gru. *Biomed Signal Process Control.* 2023;81: 104449.
23. Lih OS, Jahmunah V, Palmer EE, Barua PD, Dogan S, Tuncer T, Garcia S, Molinari F, Acharya UR. Epilepsynet: novel automated detection of epilepsy using transformer model with eeg signals from 121 patient population. *Comput Biol Med.* 2023;164: 107312.
24. Shoeb AH, Guttag JV. Application of machine learning to epileptic seizure detection. In: *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*. 2010. p. 975–82.
25. Ding X, Nie W, Liu X, Wang X, Yuan Q. Compact convolutional neural network with multi-headed attention mechanism for seizure prediction. *Int J Neural Syst.* 2023;33(3):2350014.
26. Thuwajit P, Rangpong P, Sawangjai P, Autthasan P, Chaisaen R, Banluesombatkul N, Boonchit P, Tatsaringkansakul N, Sudhawiyangkul T, Wilaiprasitporn T. Eegwavenet: multiscale cnn-based spatiotemporal feature extraction for eeg seizure detection. *IEEE Trans Ind Inform.* 2021;18(8):5547–57.
27. Rahman R, Varnosfaderani S.M, Makke O, Sarhan N.J, Asano E, Luat A, Alhawari M. Comprehensive analysis of eeg datasets for epileptic seizure prediction. In: 2021 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE; 2021. p. 1–5.
28. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput.* 1997;9(8):1735–80.
29. Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y. Learning phrase representations using rnn encoder-decoder for statistical machine translation. 2014. *arXiv preprint arXiv:1406.1078*.
30. Truong ND, Nguyen AD, Kuhlmann L, Bonyadi MR, Yang J, Ippolito S, Kavehei O. Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram. *Neural Netw.* 2018;105:104–11.
31. Feravich SM, Keller CM. Application and use of prime electrodes and eye leads. *Neurodiagn J.* 2014;54(1):48–67.
32. Liu M, Ren S, Ma S, Jiao J, Chen Y, Wang Z, Song W. Gated transformer networks for multivariate time series classification. 2021. *arXiv preprint arXiv:2103.14438*.
33. Maddinesi S, Janapati S, Kosana V, Teeparthi K. A hybrid deep transformer model for epileptic seizure prediction. In: 2022 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC). IEEE; 2022. p. 1–6.
34. Wu X, Zhang T, Zhang L, Qiao L. Epileptic seizure prediction using successive variational mode decomposition and transformers deep learning network. *Front Neurosci.* 2022;16: 982541.
35. Li C, Huang X, Song R, Qian R, Liu X, Chen X. Eeg-based seizure prediction via transformer guided cnn. *Measurement.* 2022;203: 111948.
36. Chen R, Parhi KK. Seizure prediction using convolutional neural networks and sequence transformer networks. In: 2021 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE; 2021. p. 6483–86.
37. Deng Z, Li C, Song R, Liu X, Qian R, Chen X. Eeg-based seizure prediction via hybrid vision transformer and data uncertainty learning. *Eng Appl Artif Intell.* 2023;123: 106401.
38. Busia P, Cossettini A, Ingolfsson T.M, Benatti S, Burrello A, Scherer M, Scrugli M.A, Meloni P, Benini L. Eegformer: Transformer-based epilepsy detection on raw eeg traces for low-channel-count wearable continuous monitoring devices. In: 2022 IEEE Biomedical Circuits and Systems Conference (BioCAS). IEEE; 2022. p. 640–44.
39. Couplet E, Lambert P, Verleysen M, Mulders D, Lee JA, De Bodt C. Natively interpretable t-sne. In: *Proceedings of AIMLAI Workshop*, vol. 1, 2023. p. 1.
40. Bibal A, Vu VM, Nanfack G, Frénay B. Explaining t-sne embeddings locally by adapting lime. In: 28th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning: ESANN2020. ESANN (i6doc. com); 2020. p. 393–98.
41. Maaten L, Hinton G. Visualizing data using t-sne. *J Mach Learn Res.* 2008;9(86):2579–605.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.