Computers in Biology and Medicine 176 (2024) 108565

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Parallel Dual-Branch Fusion Network for Epileptic Seizure Prediction

Hongcheng Ma [a,b,1], Yajing Wu [b,1,*], Yongqiang Tang [b,*], Rui Chen [a,b], Tao Xu [c], Wensheng Zhang [a,b,d]

[a] School of Information and Communication Engineering, Hainan University, Haikou, China
[b] State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China
[c] Shanxi Key Laboratory of Big Data Analysis and Parallel Computing, Taiyuan University of Science and Technology, Taiyuan, China
[d] School of Computer Science and Cyber Engineering, Guangzhou University, Guangzhou, China

ARTICLE INFO

Keywords:
Seizure prediction
Electroencephalography signal
Deep learning
Transformer
Convolutional neural network

ABSTRACT

Epilepsy is a prevalent chronic disorder of the central nervous system. The timely and accurate seizure prediction using the scalp Electroencephalography (EEG) signal can make patients adopt reasonable preventive measures before seizures occur and thus reduce harm to patients. In recent years, deep learning-based methods have made significant progress in solving the problem of epileptic seizure prediction. However, most current methods mainly focus on modeling short- or long-term dependence in EEG, while neglecting to consider both. In this study, we propose a Parallel Dual-Branch Fusion Network (PDBFusNet) which aims to combine the complementary advantages of Convolutional Neural Network (CNN) and Transformer. Specifically, the features of the EEG signal are first extracted using Mel Frequency Cepstral Coefficients (MFCC). Then, the extracted features are delivered into the parallel dual-branches to simultaneously capture the short- and long-term dependencies of EEG signal. Further, regarding the Transformer branch, a novel feature fusion module is developed to enhance the ability of utilizing time, frequency, and channel information. To evaluate our proposal, we perform sufficient experiments on the public epileptic EEG dataset CHB-MIT, where the accuracy, sensitivity, specificity and precision are 95.76%, 95.81%, 95.71% and 95.71%, respectively. PDBFusNet shows superior performance compared to state-of-the-art competitors, which confirms the effectiveness of our proposal.

1. Introduction

Epilepsy is a prevalent neurological disorder resulting from abnormal neuronal discharges in the brain [1]. The World Health Organization (WHO) estimates that approximately 50 million individuals worldwide are affected by epilepsy [2]. As seizures can occur at any time, the duration and severity can affect the patient's normal life and can increase the patient's mental stress, sometimes to the point of being life-threatening. Therefore, predicting seizures is of paramount importance. Early prediction of seizures before they occur and taking preventive measures can reduce the harm and improve the patient's quality of life.

Currently, epilepsy can be diagnosed based on the amplitude and waveform of Electroencephalography (EEG), which is an effective tool for monitoring brain activity [3–6]. Based on EEG signals, two distinct stages of brain activity in epilepsy patients (*i.e.*, preictal and interictal state) are given particular attention in the seizure prediction task [7]. Concretely, the preictal state refers to the period just before the seizure. The occurrence of the preictal state implies the imminent arrival of epileptic seizures. The period just preceding the preictal state is termed interictal, indicating a normal brain state. Seizure prediction aims to differentiate between the preictal and the interictal brain stages of a subject [8].

The majority of seizure prediction algorithms consist of two main steps: EEG feature extraction and classifier training. Typically, these methods commence by extracting epilepsy-related EEG features in either the time-domain (such as Willison amplitude [11] and Hjorth parameters [12]), frequency-domain (such as Power Spectral Density (PSD) [13]), or time–frequency domain (such as Short-Time Fourier Transform (STFT) [14] and Discrete Wavelet Transformation (DWT) [15]). The classifiers fall into either machine learning-based or deep learning-based categories. Commonly employed machine learning techniques encompass Support Vector Machines (SVM) [16], the K-Nearest

* Corresponding authors.
E-mail addresses: mahongcheng@hainanu.edu.cn (H. Ma), yajing.wu@ia.ac.cn (Y. Wu), yongqiang.tang@ia.ac.cn (Y. Tang), chenrui_1216@163.com (R. Chen), S202120110710@stu.tyust.edu.cn (T. Xu), zhangwenshengia@hotmail.com (W. Zhang).
[1] Hongcheng Ma and Yajing Wu contributed equally to this work.

https://doi.org/10.1016/j.compbiomed.2024.108565
Received 22 December 2023; Received in revised form 9 April 2024; Accepted 5 May 2024
Available online 8 May 2024

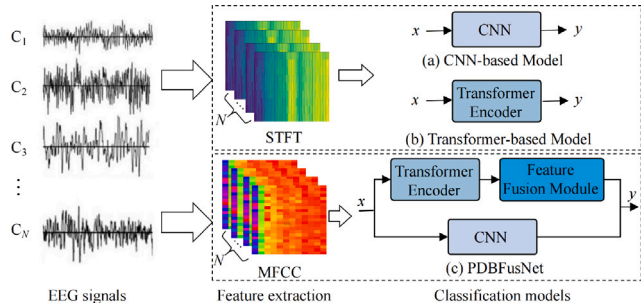


Fig. 1. The comparison of (a) CNN-based model [9], (b) Transformer-based model [10], and the proposed (c) PDBFusNet. N represents the number of channels of the EEG signal. x denotes the input features of the model, and y is the prediction result.

Neighbour algorithm (KNN) [17], Random Forest (RF) [18], Decision Tree (DT) [19], and Linear Discriminant Analysis (LDA) [20], among others.

With the rapid development of deep learning technology, significant achievements have been made in the fields of computer vision [21,22], natural language processing [23], and automatic control systems [24]. Certainly, it also further promotes the application of deep learning technology in medical research. For example, the segmentation of medical images [25,26] helps to achieve the diagnosis and treatment of diseases, and it helps to achieve early and accurate prediction of epileptic seizures in the task of seizure prediction. Commonly used deep learning algorithms include: Multi-Layer Perceptrons (MLP) [27], Convolutional Neural Networks (CNN) [28,29], Long Short-Term Memory (LSTM) network [30] and Transformer [10]. Among them, CNN is particularly popular in the past few years. However, since CNN extracts the deep feature within the local receptive field, this would limit the perception ability of global context information (or saying that long-term dependence) for EEG [31,32]. Benefited by the achievement in computer vision [21], natural language processing [23] and speech recognition [33], more recently, Transformer has been introduced into epilepsy prediction [34] and achieved impressive performance. Nevertheless, Transformer-based methods mainly capture long-term dependence through position coding and self-attention mechanism, while not sensitive to local structure [35].

Actually, in the seizure prediction task, the segmented EEG slice may have two types of status epilepticus. Short-term dependency of the EEG signal within current slice and long-term dependency between two consecutive slices are both important for accurate prediction [36]. To overcome the limitations of a single model, we propose a Parallel Dual-Branch Fusion Network (PDBFusNet) to simultaneously handle the short- and long-term dependencies. As shown in Fig. 1(c), the framework of PDBFusNet is composed of a feature extraction and a classifier. Specifically, Mel Frequency Cepstral Coefficient (MFCC) features are extracted from EEG signals and help differentiate between interictal and preictal shifts [37]. The CNN branch is used to extract local features from short-term dependence, while the Transformer branch obtains global features from long-term dependence. In the Transformer branch, considering that the information from the time, frequency and channel dimensions are not fully utilized through concatenate or average pooling operations, a feature fusion module is further designed to assign different weights and fuse EEG information from different dimensions. To evaluate the effectiveness of our proposed framework, we use a Leave-One-Out Cross-Validation (LOOCV), a more strict setup, for experimental verification. The experimental results show that it outperforms most state-of-the-art epilepsy prediction methods. The main contributions of this paper are summarized as follows:

- We propose a novel deep learning framework named PDBFusNet to handle the EEG-based seizure prediction task. PDBFusNet harnesses the synergistic advantages of both CNN and Transformer architectures for simultaneously modeling local and global features from short- and long-term dependencies, respectively.
- For the output of the single-branch Transformer model, we design a feature fusion module that is capable of fusing information in three different dimensions: time, frequency, and channel.
- The widely used public data is applied to our method. Extensive experiments are conducted to validate the effectiveness of PDBFusNet. Experimental results also show that the method proposed in this paper outperforms existing methods.

The remainder of this paper is organized as follows. In Section 2 we discuss the most related work in epileptic seizure prediction. Section 3 describes the dataset of the epileptic seizure prediction and the methods in our study. Section 4 presents the experimental results, analysis, and compares them with those from other advanced methods. Finally, Section 5 draws conclusions.

2. Related work

Artificial intelligence algorithms for epileptic seizure prediction have been investigated for several decades. Currently, these methods can be categorized into traditional Machine Learning (ML)-based methods and Deep Learning (DL)-based methods.

2.1. Traditional ML for seizure prediction

The process of traditional ML methods to solve the problem of epileptic seizure prediction includes data pre-processing, feature extraction, and classification. Data pre-processing aims at filtering out noise and artifact to reduce interference to the EEG signal. In the feature extraction stage, the information of each state is extracted from the EEG signals of epilepsy patients. There are several ways that can be tracked, including time-domain features such as Willison amplitude [11] and Hjorth parameters [12], frequency-domain features such as PSD [13], time–frequency domain features such as STFT [38], Wavelet Transform (WT) [39] and Stockwell Transform (ST) [40], nonlinear features such as entropy [41]. In the classification stage, various ML algorithms are employed to epileptic seizure prediction tasks, such as SVM, KNN, and DT.

At present, there have been many efforts that attempt to combine different feature extraction methods with ML algorithms for epileptic seizure prediction. Current research route follows two lines, i.e., using SVM as classifier based on different feature extraction methods, and using the same feature extraction method with different classifiers. For the first line, Subasi et al. [42] proposed to make use of DWT for EEG signal analysis, where they first extracted features through DWT, then different classical methods such as principal component analysis, linear discriminant analysis, and independent component analysis were used to reduce the dimensionality of the features. These low-dimensional features were used as inputs to SVM to classify EEG signals. Patidar et al. [43] extracted the Kraskov entropy as the only feature to get the nonlinear values of EEG signals and then used Least Squares Support Vector Machines (LS-SVM) for seizure prediction. Given the nonlinear and non-stationary nature of EEG signals, researchers have increasingly focused on investigating nonlinear dynamic characteristics in recent years. Chen et al. [44] applied information entropy theory to extract nonlinear dynamic characteristics, where one-way analysis of variance and forward sequential feature selection technique are jointly utilized to identify features that effectively characterize epilepsy. For the second line, considering that the model fitting ability of the RF method is stronger than that of SVM, Tzimourta et al. [45] used RF as classifier using DWT-based time–frequency features. Yuan et al. [46] decomposed the signal into multiple frequency bands by discrete wavelet decomposition, extracted the diffusion distance between the bands as a multi-scale feature, and performed seizure prediction based on Bayesian Linear Discriminant Analysis (BLDA).

In general, traditional ML methods have made many achievements in epileptic seizure prediction task. However, their performance highly relies on the selection of features and classifiers.

2.2. DL-based methods for seizure prediction

Recently, with the rapid development of DL, it has been applied to many fields. DL can automatically extract features from data. Therefore, many researchers have applied DL technology to automatic feature extraction and classification for epileptic seizure prediction, and have achieved impressive results.

In DL, CNN can not only be used as a feature extractor but also as a classifier. Hisham et al. [8] used CNN as a feature extractor, followed by Bi-directional Long Short-Term Memory (Bi-LSTM) as a classifier, and achieved an accuracy of 99.66% and 99.72% sensitivity. Truong et al. [47] extracted time–frequency domain features from EEG time-domain signals by STFT and then used CNN for classification, finally achieving a sensitivity of 81.2% and a False Prediction Rate (FPR) of 0.16/h. Usman et al. [48] proposed an ensemble learning method for epileptic seizure prediction. Handcrafted features and CNN-based features are used for feature extraction, and Pearson Correlation Coefficient (PCC) and Particle Swarm Optimization (PSO) are combined to select the extracted features. The classifier then combines the three-part output of an SVM, CNN, and LSTM using model-agnostic meta-learning. Finally, they achieved a sensitivity of 96.28%, an accuracy of 96.05%, and a specificity of 95.65%. However, these methods are based on CNN to extract local features of EEG signals, ignoring contextual information or long-range dependencies.

Transformer has been widely applied in many fields due to their excellent ability to capture long-distance dependencies. Bhattacharya et al. [34] introduced the Transformer model into epileptic seizure prediction for the first time. They used STFT to extract the time–frequency features of EEG signals, and then used the Transformer model for classification, achieving a sensitivity of 98.46%. Yan et al. [10] applied an STFT to the EEG signal and then expanded the resulting third-order tensor into three matrices based on different dimensions. These matrices were used as inputs to their model for epilepsy prediction. This approach effectively leveraged time, frequency, and channel information within the EEG signal, resulting in an impressive sensitivity of 96.01% and a low FPR of 0.047/h. Godoy et al. [49] focused on the information from different EEG channels and proposed two Transformer-based neural network models called Temporal Multi-Channel Transformer (TMC-T) and Temporal Multi-Channel Visual Transformer (TMC-ViT). The results showed that both models achieved the best performance when the preictal duration was set to 60 min and the sample segment was 20 s. Among them, the TMC-T model achieved an accuracy of 93.74% and a sensitivity of 93.87%, and the TMC-ViT model achieved an accuracy of 95.73% and a sensitivity of 96.46%. Gao et al. [50] proposed a general sample-weighted framework that uses genetic algorithms to optimize the sample weights of the training set. Then weighted and unweighted comparative experiments were conducted on SVM, CNN and Transformer respectively to evaluate their effectiveness. However, these works are based on Transformer to model the global characteristics of EEG signals, ignoring the influence of local characteristics.

There are some differences between previous studies and ours. In terms of input features, directly inputting features into the model does not take into account the difference in contributions of different dimensions of the features. The feature fusion module we designed can fully fuse features by giving corresponding weights of different dimensions to the features. In terms of model structure, most of the above-mentioned works adopt a single-branch structure. We adopt a parallel dual-branch structure and take advantage of the synergistic advantages of Transformer and CNN to extract global and local features from EEG signals respectively.

Table 1
Detailed description of the CHB-MIT dataset used in this study.

Patient ID	Gender	Age	Record time (h)	Duration of seizure (s)	No. of seizures
chb01	F	11	40.6	499	7
chb02	M	11	25.3	175	3
chb03	F	14	28.0	409	7
chb05	F	7	39.0	563	5
chb06	F	1.5	66.7	147	10
chb07	F	14.5	68.1	328	3
chb08	M	3.5	20.0	924	5
chb10	M	3	50.0	454	7
chb11	F	12	34.8	809	3
chb14	F	9	26.0	117	7
chb19	F	19	30.0	239	3
chb20	F	6	29.0	302	8
chb21	F	13	33.0	303	4
chb22	F	9	31.0	207	3
chb23	F	6	28.0	431	7

Gender: Female(F) and Male(M). No. of seizures: The number of seizures.

Fig. 2. The segmentation of epileptic EEG data. The preictal period is defined as one-hour before the seizure, and the interictal period is defined as four-hour before or after the end of the seizure.

3. Materials and methodology

In this section, we first introduce the data used in the study. Then the framework of PDBFusNet is described.

3.1. Data description

The CHB-MIT dataset [51] is a scalp EEG dataset collected by Boston Children's Hospital, which includes EEG recordings of pediatric individuals with intractable epileptic seizures and is open access at PhysioNet.org.[2] The scalp EEG recordings involve 22 participants, comprising 5 males aged between 3 and 22 and 17 females aged between 1.5 and 19. The recordings were obtained at a sampling frequency of 256 Hz using scalp EEG data from 22 electrodes following the international bipolar 10–20 system. Because there are differences in the electrodes used for each patient across multiple experiments, it is difficult to analyze without selecting their common channels. Therefore, we chose 15 patients for the experiment who shared 23 channels, namely: 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 and T8-P8. Details are shown in Table 1.

As stated earlier, the task of epileptic seizure prediction concerns the preictal and interictal stages. The segmentation of these two stages is depicted in Fig. 2. Following [8,37], the preictal period is defined as the one-hour interval preceding a seizure, while the interictal period encompasses the four-hour timeframe both before and after the seizure episode. The EEG signal is divided into non-overlapping 5-second segments, and the number of samples in the interictal period is selected to be equal to the number of samples in the preictal period.

3.2. Methodology

The details of the proposed PDBFusNet are shown in Fig. 4. First, MFCC features are extracted from EEG signals. Second, MFCC features

[2] https://www.physionet.org/content/chbmit/1.0.0/

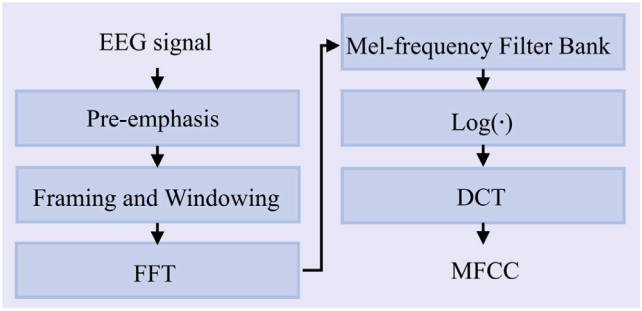


Fig. 3. The process of MFCC feature extraction. First, the time domain EEG signal is preprocessed by pre-emphasis, framing and windowing. Then, FFT is applied to convert it to frequency domain signal and Mel filter energy is obtained by Mel filter bank. Finally, logarithmic transformation and DCT are used to obtain the MFCC.

serve as inputs to both the Transformer and CNN branches. Additionally, feature fusion of different dimensions occurs at the output of the Transformer branch. Finally, after concatenating the outputs of the two branches, the results of epileptic seizure prediction are output. In this paper, matrices are represented in bold capital letters, vectors in bold lowercase letters, and scalars in regular lowercase letters for clarity. The key notations and their corresponding descriptions are outlined in the Table 2.

3.2.1. Feature extraction

The proposed framework uses MFCC features as the input to the following network to extract the pattern from the Mel frequency cepstrum domain. MFCC has shown a wide range of applicability in DL tasks associated with biomedical signals [52] and contains predictive biomarkers related to interictal and preictal brain states [37]. It captures more low-frequency information through nonlinear scaling, which makes it very effective when processing non-stationary signals [53]. MFCC features are obtained by mapping the signal frequency to the nonlinear Mel frequency cepstrum domain. Formally, the Mel scale frequency $Mel(\cdot)$ is defined as follows:

$$Mel(f) = 2595 \times \log\left(1 + \frac{f}{700}\right) \tag{1}$$

where f is the actual frequency of the EEG signal.

In this study, the concrete process in MFCC feature extraction is shown in Fig. 3. After pre-processing operations including pre-emphasis, framing, and windowing, the segmented EEG signal sequence without noise is obtained. Then, the Fast Fourier Transform (FFT) is performed to convert the time-domain signal into a frequency-domain signal. Next, the spectral line energy of each frame of the frequency-domain signal is calculated. Sequentially, the obtained discrete power spectrum is filtered through the Mel filter to obtain the corresponding Mel filter energy. Finally, the Mel filter energy is logarithmically transformed to obtain the logarithmic spectrum, and the MFCC is obtained after the Discrete Cosine Transform (DCT). The shape of the calculated MFCC feature map is $(N \times C \times T)$, in which N is the channel dimension, C is the order of MFCC and T is the time scale. For a single frame EEG, the ith MFCC coefficient $MFCC_i$ is calculated as follows:

$$MFCC_i = \sum_{m=1}^{M} S(m) \cos\left[\frac{\pi i (m - 0.5)}{M}\right], i = 1, 2, \ldots, C \tag{2}$$

where $S(m)$ is the logarithmic energy output by the filter bank, cos represents the cosine function, and m represents the index of the Mel filter with M Mel filters.

3.2.2. Parallel dual-branch fusion network

Our PDBFusNet uses a parallel architecture that harnesses the complementary strengths of both CNN and Transformer networks as shown in Fig. 4. The Transformer can capture long correlations when analyzing and predicting sequential data to obtain global information, while CNN is adept at extracting local features.

Table 2
The key notations and the corresponding descriptions.

Notation	Description
C	Order of MFCC
N	Number of EEG channels
T	Time scale
d_i	Feature dimension of the output of the ith Encoder branch
w_1, w_2, w_3	Weights for the three features
$\mathbf{X}_N$	Channel-wise feature
$\mathbf{X}_C$	Coefficient-wise feature
$\mathbf{X}_T$	Time scale-wise feature
$\mathbf{Z}_i$	Feature representation after encoder
$\mathbf{Z}_i^L$	Feature representation after linear layer mapping
$\mathbf{Z}^C$	Representation of features after concatenation
$\mathbf{Z}^P$	Feature representation after pooling
$\mathbf{Z}^{Trans}$	Output features of the Transformer branch
$\mathbf{Z}^{CNN}$	Output features of the CNN branch

(1) Transformer with a fusion module in PDBFusNet: After feature extraction, a three-order tensor $\mathcal{X}$ with shape $N \times C \times T$ is obtained. To explore more useful knowledge hidden in such a high-order feature, the tensor is transformed into three feature matrices, namely, channel-wise feature $\mathbf{X}_N \in \mathbb{R}^{N \times (C \times T)}$, coefficient-wise feature $\mathbf{X}_C \in \mathbb{R}^{C \times (N \times T)}$, and time scale-wise feature $\mathbf{X}_T \in \mathbb{R}^{T \times (C \times N)}$, which are deemed as different learning branches of our Transformer module, as depicted in Fig. 4(a)(1). Considering that the EEG signal is continuous, in the Transformer module, a fully connected layer is used to replace the embedding layer of the original Transformer, for projecting each feature matrix as a preliminary abstract representation. To utilize the order of EEG signal sequences, relative or absolute position information of the markers needs to be further involved in the sequence. Hence, we propose to add positional encoding in the coefficient-wise and time scale-wise branches following their preliminary representations. Thus far, three different representations are in our hands to be input into the encoders for further representation refinement. Not hard to find that, three encoders are cooperating with each other to capture the correlation of each dimensional sequence in terms of time, frequency, and channel information. Our encoder contains D encoder layers, and each of whom mainly includes a fully connected layer and a multi-head attention layer, as shown in Fig. 4(b). More specifically, the multi-head attention layer consists of multiple self attention layers, which map different output vectors to multiple new subspaces to learn features from different positions. Self-attention is calculated as follows:

$$Attention(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = softmax\left(\frac{\mathbf{Q}\mathbf{K}^{\mathbf{T}}}{\sqrt{d_k}}\right)\mathbf{V} \tag{3}$$

where $\mathbf{Q}$, $\mathbf{K}$, $\mathbf{V}$ represent query, key, value matrices and T is transposition, d_k is the dimension of $\mathbf{K}$, $softmax$ refers to the activation function. The formula of multi-head attention is given as follows:

$$\begin{aligned} MultiHead(\mathbf{Q}, \mathbf{K}, \mathbf{V}) &= Concat\left(head_1, \ldots, head_h\right)\mathbf{W}^O \\ head_i &= Attention\left(\mathbf{Q}\mathbf{W}_i^Q, \mathbf{K}\mathbf{W}_i^K, \mathbf{V}\mathbf{W}_i^V\right) \end{aligned} \tag{4}$$

where h is the number of attention heads, i represents the ith attention head, $Concat$ denotes the concatenation operation, $\mathbf{W}_i^Q$, $\mathbf{W}_i^K$, $\mathbf{W}_i^V$ and $\mathbf{W}^O$ are the trainable weight matrices.

As we can see, the Transformer encoders can generate three kinds of feature representations $\{\mathbf{Z}_i \in \mathbb{R}^{B \times d_i}\}_{i=1}^{3}$, where B is the batch size and d_i is the output feature dimension of the ith encoder branch. Inspired by the simplex weighting strategy in multi-view learning [54–57] and multi-modal fusion [58,59], we claim that these feature representations should present different degrees of importance, thus a feature fusion

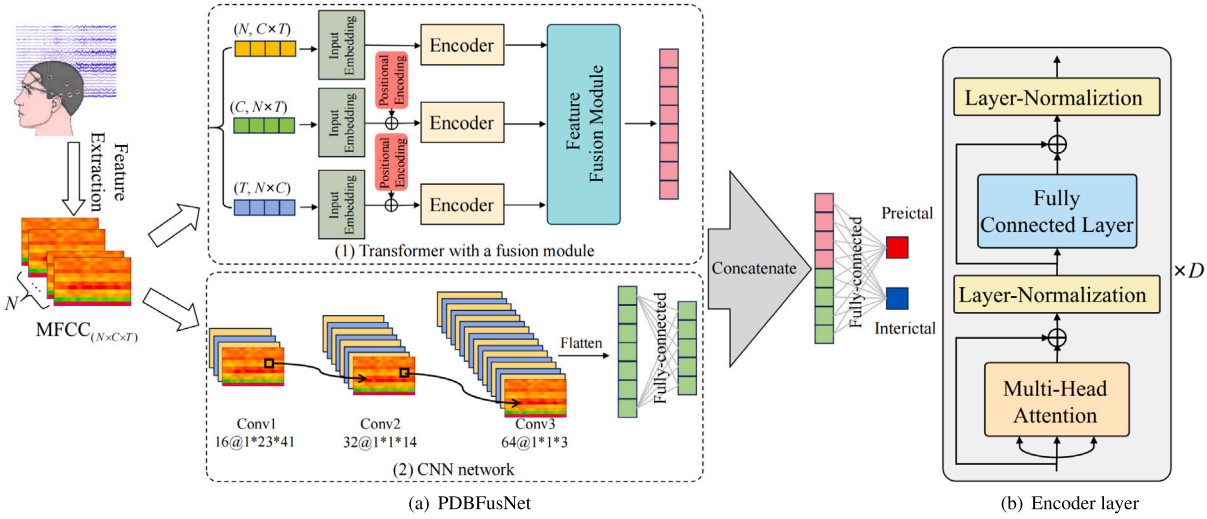


Fig. 4. The architecture of our proposal. Our framework consists of MFCC feature extraction and a classifier to distinguish the preictal and the interictal brain stages. Classification results are output by fusion based on the Transformer branch and CNN branch. In the Transformer branch, the encoder consists of D Transformer encoder layers.

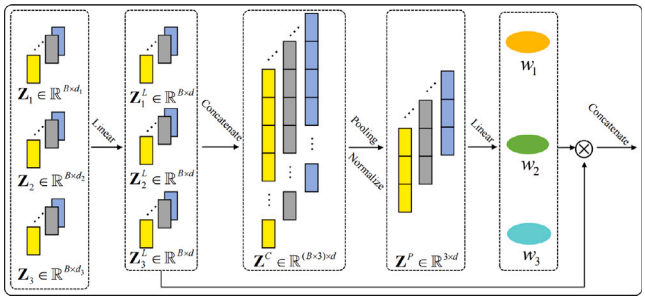


Fig. 5. The structure of the proposed feature fusion module. Weights are obtained and multiplied with the corresponding features to achieve feature fusion.

module (as shown in Fig. 5) is introduced to provide different attention weights for different feature representations. In detail, we project $\{\mathbf{Z}_i \in \mathbb{R}^{B\times d_i}\}_{i=1}^{3}$ onto a shared d-dimensional feature space to produce $\{\mathbf{Z}_i^L \in \mathbb{R}^{B\times d}\}_{i=1}^{3}$. Then the matrix $\mathbf{Z}^C \in \mathbb{R}^{(B\times 3)\times d}$ can be obtained by the concatenation operation $Concat$:

$$\mathbf{Z}^C = Concat\left(\mathbf{Z}_1^L, \mathbf{Z}_2^L, \mathbf{Z}_3^L\right) \tag{5}$$

Next, the matrix $\mathbf{Z}^C$ is normalized and an average pooling layer is performed on the batch size dimension to get $\mathbf{Z}^P \in \mathbb{R}^{3\times d}$. We send $\mathbf{Z}^P$ to a fully connected layer and pass the $Softmax$ function to obtain the attention weights w_1, w_2 and w_3 of the three features:

$$\left[w_1, w_2, w_3\right]^{\mathrm{T}} = Softmax\left(\mathbf{W}\mathbf{Z}^P + \mathbf{b}\right) \tag{6}$$

where $\mathbf{W}$ represents the weight matrix, $\mathbf{b}$ represents the bias vector. Finally, a weighted aggregation layer is used to combine the fusion representation of multiple features:

$$\mathbf{Z}^{Trans} = Concat\left(w_1\mathbf{Z}_1, w_2\mathbf{Z}_2, w_3\mathbf{Z}_3\right) \tag{7}$$

Following the Transformer branch's processing, the long-term dependencies in the EEG signal can be captured, and with the feature fusion module, information from all aspects can be effectively leveraged.

(2) CNN in PDBFusNet: In this work, the CNN network, a branch of PDBFusNet, aims to address the limitations of Transformer in capturing short-term dependencies. Here, three convolutional blocks with two fully connected layers are considered as our CNN network as shown in Fig. 4(a)(2). Each convolutional block contains a convolutional layer activated by a Rectified Linear Unit (ReLU), a Batch Normalization (BN) layer, and a max-pooling layer. The convolutional layer performs convolution operations on the input features to capture the local features. The BN layer normalizes the extracted features to ensure that the input distribution is the same for each layer. The max-pooling layer down samples the input features to reduce the feature size. In detail, the size of the convolution kernels in the three convolutional layers is 16, 32 and 64, respectively. Finally, a Flattened layer is used after three convolutional blocks to reduce multi-dimensional features to a single dimension. Then the $Sigmoid$ activation function is used on the output of the fully connected layer, and a dropout layer is placed before the fully connected layer with a dropout rate of 0.5 to reduce overfitting.

The final feature representation $\mathbf{Z}$ is obtained by concatenating the output feature $\mathbf{Z}^{Trans}$ of the Transformer branch and the output feature $\mathbf{Z}^{CNN}$ of the CNN branch:

$$\mathbf{Z} = Concat\left(\mathbf{Z}^{Trans}, \mathbf{Z}^{CNN}\right) \tag{8}$$

Finally, the feature $\mathbf{Z}$ is delivered to the fully connected layer to obtain the results of the prediction. By minimizing the cross-entropy loss, PDBFusNet can be optimized.

4. Experiments

In this section, first, we present the details of the experimental settings and the evaluation metrics. Second, to validate the effectiveness of our proposal, compared with the baseline model the experimental results on the CHB-MIT dataset are presented. Finally, to verify the effectiveness of the feature fusion module and the dual branching, ablation experiments and parametric experiments are performed, analyzed and discussed.

4.1. Experiment settings and evaluation metrics

In the task of epileptic seizure prediction, dataset partitioning directly affects the accuracy and reliability of the final model. It is essential to ensure that the testing data, which is used to evaluate the model's performance, is independent of the training data and not influenced by any prior knowledge of the model. With careful consideration, we adopt a strict setting (LOOCV) to obtain reliable prediction results [27]. Specifically, for a patient with E epileptic seizure, each seizure is referred to a separate test dataset, while the remaining data are utilized as the training set. The model is trained and tested for the patient E times and the average of E results is computed to yield the patient's experimental results on the model.

Table 3
Performance comparison of the overall performance of the three models 3T-Transformer, CNN, and PDBFusNet.

Model	Acc. (%)	Sen. (%)	Spe. (%)	Pre. (%)
3T-Transformer	87.22	85.96	88.42	89.52
CNN	91.06	90.19	91.94	91.77
PDBFusNet	**95.76**	**95.81**	**95.71**	**95.71**

The best results are in boldface.

We implement our experiments on a Linux server with two NVIDIA GTX 1080Ti GPUs. Our experiments are completed in the Python 3.8 environment and Pytorch 2.0. As stated earlier in the data description, the number of channels N is 23. For the parameter settings of the PDBFusNet, the number of filter banks C is 12, and the time scale T is 41. The output size of the Transformer branch is 768, and the output size of the CNN branch is 256. Finally, the output concatenation size of the two branches is 1024. The experiments are performed using an Adagrad optimizer with a learning rate of 0.001.

We mainly evaluate the performance of PDBFusNet using four evaluation metrics. Accuracy is defined as the ratio of correct predictions to the total number of predictions made. Sensitivity measures the proportion of seizures that are correctly predicted out of the total number of seizures. Specificity calculates the ratio of correctly predicted negative samples to the total number of negative samples, while precision determines the proportion of seizures that are correctly predicted out of all predicted seizures. Mathematically, the four evaluation metrics are defined as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{9}$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{10}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{11}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{12}$$

where True Positive (TP) and True Negative (TN) refer to EEG segments correctly classified as preictal and interictal, respectively. False Positive (FP) and False Negative (FN) denote EEG segments incorrectly predicted as preictal and interictal, respectively. The short names Acc., Sen., Spe., and Pre. are used in the following to represent accuracy, sensitivity, specificity, and precision.

4.2. Baseline comparison

To validate the effectiveness of our proposal, we compared the results with the baseline model (three-tower Transformer model [10] short for 3T-Transformer, CNN model [9]). Both the CNN model and the 3T-Transformer model use time–frequency features extracted by STFT. The CNN model consists of three convolutional blocks. The 3T-Transformer model consists of a three-tower Transformer based on a Transformer encoder layer and a gating mechanism. Inspired by the 3T-Transformer, we propose a feature fusion module that assigns corresponding weights to different dimensions and obtains global features after full fusion. Furthermore, a CNN branch is introduced to extract local features. For a fair comparison, both models are executed on the same 15 patients from the CHB-MIT dataset under the LOOCV setting.

The comparison of the overall performance of the three models is shown in the Table 3. Generally, Table 3 shows an increase in accuracy of 4.7%, sensitivity of 5.62%, specificity of 3.77%, and precision of 3.94% when comparing the PDBFusNet and CNN. A comparison between PDBFusNet and 3T-Transformer in Table 3 reveals an increase in accuracy of 8.54%, sensitivity of 9.85%, specificity of 7.29%, and precision of 6.19% for PDBFusNet. To evaluate the significance of the results in Table 3, we use the Kruskal–Wallis test [60] as a non-parametric test statistic to analyze the results. The obtained p-values for accuracy, sensitivity, specificity, and precision were 0.01, 0.02, 0.03, and 0.04, respectively. The Kruskal–Wallis test results for all four metrics revealed p-values of less than 0.05, suggesting notable differences in the outcomes among the three models. This preliminary verification confirms the effectiveness of our proposal.

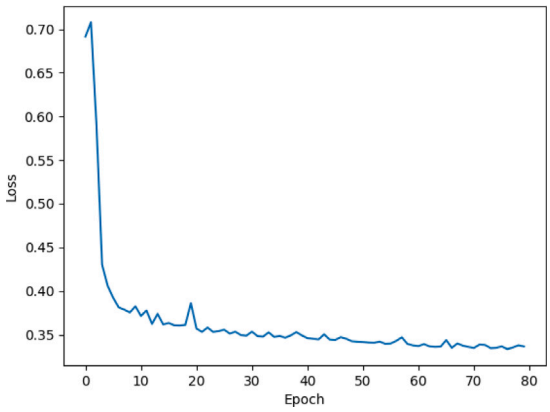


Fig. 6. Decreasing trend of loss values during training.

Table 4
Seizure prediction results of PDBFusNet on the CHB-MIT dataset.

Patient ID	Acc. (%)	Sen. (%)	Spe. (%)	Pre. (%)
chb01	99.98	100.0	99.97	99.97
chb02	89.48	95.38	83.58	87.12
chb03	97.74	98.08	97.40	97.46
chb05	97.11	96.51	97.71	97.66
chb06	88.89	83.14	94.64	92.45
chb07	83.97	82.70	85.24	84.90
chb08	99.33	98.67	100.0	100.0
chb10	93.37	89.76	96.97	94.70
chb11	99.65	99.92	99.37	99.38
chb14	99.75	99.79	99.71	99.71
chb19	98.83	98.87	98.79	98.79
chb20	99.95	99.96	99.94	99.94
chb21	95.55	96.88	94.21	94.37
chb22	93.10	97.51	88.68	89.66
chb23	99.75	100.0	99.50	99.50
Average	95.76	95.81	95.71	95.71

To validate the convergence of PDBFusNet, Fig. 6 shows the trend of PDBFusNet loss values decreasing with the training process. The experimental results depicted in Fig. 6 indicate that our model's performance experiences a sharp decline in the initial epochs followed by a stable phase throughout the remainder of the training process.

The experimental results in Table 4 show that the performance of the model varies among patients. This may be related to the data collected from the patients. PDBFusNet extracts local features through CNN, and the fusion module combines features of different dimensions. It solves the problem of using a Transformer to extract global features while ignoring local features and giving different dimension weights solves the problem of differences in the contribution of features to fusion results.

4.3. Parameters analysis

In this section, we perform experiments with the relevant parameters. We conduct three sets of experiments based on the standard range of output dimensions for CNN and transformers. In detail, the dimension size of the output is set to 128-256-512, 256-768-1024, and 512-1536-2048, respectively. Where, for example, 256-768-1028, 128 denotes the output size of the CNN branch, 768 denotes the output size of the Transformer branch, and 1024 denotes the output of the double-branch concatenation after the output size. The experimental results

Table 5
Performance comparison of dual-branch network output with different sizes.

Network output	Acc. (%)	Sen. (%)	Spe. (%)	Pre. (%)
128-384-512	93.88	94.97	92.80	93.77
256-768-1024	**95.76**	**95.81**	**95.71**	**95.71**
512-1536-2048	94.99	94.85	95.13	95.19

The best results are in boldface.

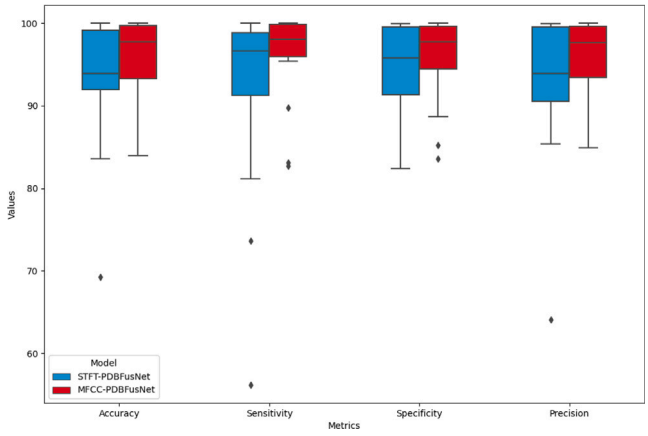


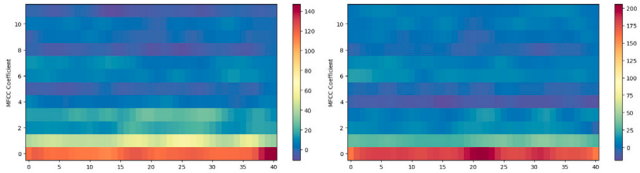
Fig. 7. Comparison of the performance of all patients on the four metrics of PDBFusNet based on STFT features and based on MFCC features.

are shown in Table 5 and the predictions are different when the output sizes are different. When the output size is 256-768-1024, the model performs best on the CHB-MIT dataset.

4.4. Ablation experiment

To further explore the effectiveness of parallel dual branches and the feature fusion module, we perform ablation experiments. The experimental results are shown in Table 6. The validation of major concerns in our proposal is analyzed: (1) remove the Transformer branch in PDBFusNet and the feature fusion module in this branch to study its impact on model performance. Comparing the fifth and fourth rows in Table 6, it can be observed that there are decreases in accuracy of 2.34%, sensitivity of 1.51%, specificity of 3.17%, and precision of 2.73%. Experimental results show the superiority of the Transformer in capturing the long-range dependence of EEG signals and the importance of global features in improving model performance [35]. (2) the CNN branch in PDBFusNet is removed to explore the impact of the CNN network on model performance. Table 6 shows a decrease in accuracy of 3.67%, sensitivity of 1.56%, specificity of 4.97%, and precision of 4.05% when comparing the fifth and third rows. The experimental results show that when the CNN branch is removed, the four metrics of the single-branch Transformer decrease significantly. The advantages of CNN in extracting local features and the effectiveness of parallel structure are verified. (3) We remove the feature fusion module from the Transformer branch. A comparison between the third and second rows in Table 6 reveals decreases in accuracy of 1.71%, sensitivity of 3.44%, specificity of 0.79%, and precision of 1.52%. It can be found that the importance of different dimensions is obtained by adding the feature fusion module, which gives weights to the output features on three different dimensions respectively. It can fully fuse the information on different dimensions of MFCC features to better obtain the ability to characterize the global information.

Since EEG-based feature extraction is critical in formulating a framework, we conducted an exploratory analysis to investigate the impact of alternative feature extraction methods on experimental results. STFT-based feature is also widely used [61,62]. MFCC captures more low-frequency information through a nonlinear scale, which makes it



(a) interictal (b) preictal

Fig. 8. MFCC feature maps for (a) interictal and (b) preictal epilepsy.

outstanding in characterizing non-stationary signals while being noise robust [53,63]. The performance of STFT-based features and MFCC-based features in PDBFusNet on each patient is shown in Fig. 7. The results show that the median boxplots based on the MFCC features are higher than those based on the STFT features on all four metrics, and the overall performance is better on all patients. For some outliers, it indicates that the performance of the model is correlated with the patient's EEG data. STFT extracts time–frequency features, while MFCC will lose part of the time domain information when extracting features. So the decrease of some metrics in a given patient is reasonable. The accuracy based on MFCC features increases by 2.84%, the sensitivity increases by 4.35%, the precision increases by 2.99%, and the specificity decreases by 1.33%. Overall, MFCC features have advantages over STFT features, proving the superiority of MFCC features in capturing low-frequency information and characterizing non-stationary signals.

4.5. Analytical experiments

Furthermore, we carry out several necessary experiments to pursue deeper insight into our PDBFusNe. First, the performance comparison of different combination ways of CNN and Transformer. Second, the advantage of MFCC over other features is verified.

Performance comparison of different combinations of CNN and Transformer. When networks are located differently, it affects the information and features that are received by the next network. We further explore the performance comparison of different combination methods of CNN and Transformer. The experimental results on the CHB-MIT dataset are shown in Table 7. The results show that the best performance is achieved when the CNN and the transformer are parallel structures. It is further proved that the parallel structure can extract local and global features better.

Performance comparison of different feature maps. To enhance clarity, we also showcase the variances in MFCC characteristics of EEG signals during the interictal and preictal stages of epileptic seizures. Fig. 8 depicts the variances in MFCC features during the interictal and preictal phases of patient chb03.

To better demonstrate the benefits of MFCC, we conducted additional preliminary experiments. Here, we conduct experiments to extract features from raw EEG signals using PSD, STFT, DWT, and MFCC, respectively. Five patients are randomly selected to calculate the average value of the four evaluation metrics, and the experimental results are shown in Table 8. Comparing the experimental results in the table, it can be found that the PDBFusNet model based on MFCC features has the best performance.

5. Discussions

In this section, we first compare with previous epilepsy prediction methods. Secondly, the limitations of the proposed method are discussed and an outlook for future work is given.

5.1. Comparison with state-of-the-art methods

The advancements in DL in recent years have led to significant progress in epileptic seizure prediction using DL methods. Here, we

Table 6
The comparison of experimental results from ablation experiments.

Transformer	Feature fusion module	CNN	Accuracy(%)	Sensitivity(%)	Specificity(%)	Precision(%)
✓	×	×	90.38	90.81	89.95	90.14
✓	✓	×	92.09	94.25	90.74	91.66
×	×	✓	93.42	94.30	92.54	92.98
✓	✓	✓	95.76	95.81	95.71	95.71

Among these, ✓indicates that this network or module has been added, and ×signifies that it has not been added.

Table 7
Performance comparison of different combinations of CNN and Transformer.

Combination method	Acc. (%)	Sen. (%)	Spe. (%)	Pre. (%)
CNN followed by Transformer	83.59	85.19	81.98	83.82
Transformer followed by CNN	88.01	92.55	83.46	86.43
CNN in parallel with Transformer	**95.76**	**95.81**	**95.71**	**95.71**

The best results are in boldface.

Table 8
Performance comparison of different features on PDBFusNet.

Model	Acc. (%)	Sen. (%)	Spe. (%)	Pre. (%)
DWT-PDBFusNet	89.89	89.89	89.81	89.63
STFT-PDBFusNet	93.03	91.44	94.62	93.41
PSD-PDBFusNet	84.20	82.41	86.00	85.58
MFCC-PDBFusNet	**96.12**	**97.51**	**94.73**	**95.46**

The best results are in boldface.

provide the currently and commonly used DL-based seizure prediction models in Table 9. It can be observed that there are discrepancies in the experimental setup of the current model. For instance, Yan et al. [10] and Bhattacharya et al. [34] achieve higher sensitivity on the CHB-MIT dataset in the hold-out experimental settings. Note that in the context of epileptic seizure prediction, LOOCV is a more rigorous setting than hold-out. The hold-out method refers to randomly dividing the experimental dataset into training and testing sets. It shuffles and mixes all EEG data so that some prior information in the test seizure is aware by the model during training in advance, destroying the independence between EEG data of different seizures. Since the model has learned the information about the test, it can achieve high performance [35]. Khan et al. [39], Cho et al. [64] and Büyükçakır et al. [65] use 10-fold Cross-Validation (10-fold CV). As a special case of K-fold CV, LOOCV divides the number of epileptic seizures in more detail and does not destroy the temporal continuity when testing. Models using K-fold CV train less available data than the models using LOOCV, which yields worse performance [68].

Under the same LOOCV setting, we further analyze from the perspective of data feature extraction. Zhao et al. [6] solely employ the time-domain features of original EEG data, overlooking the potential of utilizing the frequency information present in the EEG signal. Ozcan et al. [66] utilize a 3DCNN model for classification, resulting in increased computational costs without a substantial performance improvement. Truong et al. [9] and Zhang et al. [67] utilize STFT and Common Spatial Patterns (CSP), respectively, to extract time–frequency domain features from EEG signals, resulting in a commendable performance. Among them, the sensitivity obtained by using STFT to extract features is 81.2%, the sensitivity of using CSP to extract features is 92%, and the AUC is 0.90. However, using a pure CNN model can only model local features, ignoring the long-term dependencies of the EEG signal and limiting the performance of the model.

In summary, compared with the methods mentioned above, our model achieved an AUC of 0.955 and a sensitivity of 95.76% in the experiments using the LOOCV approach, both of which outperformed the method using the LOOCV approach, indicating that our proposed seizure prediction method can achieve better performance.

5.2. Limitations and future directions

Although the proposed seizure prediction model achieves high prediction performance, the current work still suffers from two limitations. Firstly, our proposed PDBFusNet model omits complex preprocessing (e.g., artifact removal) operations and the actual epileptiform discharges have artifacts and uncorrelated channels. Therefore, we will continue to explore the artifact removal algorithms [69] and channel selection mechanisms [70,71] to reduce the effects of noise and channel redundancy and further improve the performance of the model. Second, our approach was trained and tested on specific patients and did not take into account the effect of inter-patient variability on prediction. Therefore, in future work, supervised domain adaptation [72] or transfer learning strategies [73] are added to our model to reduce the impact of inter-individual patient variability.

6. Conclusion

In this paper, we propose a Parallel Dual-Branch Fusion Network (PDBFusNet) to handle the EEG-based seizure prediction task. PDBFusNet harnesses the synergistic advantages of both CNN and Transformer architectures for simultaneously modeling local and global features. Specifically, the MFCC features are first extracted from the original EEG signal. PDBFusNet combines the synergistic advantages of CNN and Transformer. In addition, regarding the Transformer branch, we design a fusion module to integrate information from three dimensions of EEG signal time, frequency, and channel by assigning different weights. We conduct extensive experiments to validate the effectiveness of PDBFusNet. Experimental results show that PDBFusNet outperforms baseline, and further analytical and ablation studies demonstrate the necessity and importance of each proposed component. Compared with the state-of-the-art DL model, the results confirm our superiority in the strict experimental setting.

CRediT authorship contribution statement

Hongcheng Ma: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yajing Wu:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Methodology, Data curation. **Yongqiang Tang:** Writing – review & editing, Validation, Supervision, Methodology. **Rui Chen:** Writing – review & editing, Validation. **Tao Xu:** Investigation, Data curation. **Wensheng Zhang:** Supervision, Resources, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors are thankful for the financial support by the National Key Research and Development Program of China (No. 2021ZD0201600), the National Natural Science Foundation of China (Grant Nos. 62206292 and 62106266). This work was supported in part by Shandong Naoli Artificial Intelligence Tech. Co., Ltd, China.

Table 9
Performance comparison with previous methods on the CHB-MIT dataset.

Author	Features	Classifier	Validation strategy	No. of patients-seizures	Interictal-preictal intervals (min)	Average AUC-Sen (%)
Yan et al. [10]	STFT Spectrograms	Transformer	Hold-Out	21–111	240–30	NA-96.01
Bhattacharya et al. [34]	STFT Spectrograms	Transformer	Hold-Out	21–147	NA-30	NA-98.46
Khan et al. [39]	Wavelet Transform Coefficient	CNN	10-fold CV	15–18	10–10	86.6-87.8
Cho et al. [64]	Phase Locking Value	SVM	10-fold CV	21–65	30–5	NA-82.44
B̀uy′ukçakır et al. [65]	Hilbert Vibration Decomposition	MLP	10-fold CV	10–62	NA-30	NA-89.8
Zhao et al. [6]	Raw Data	AddNet-SCL	LOOCV	19–105	240-30, 240-60	0.929-93.0, 0.942-94.9
Ozcan et al. [66]	Spectral Power, Statistical Moments, Hjorth	3D CNN	LOOCV	16–77	60-60, 120-60, 240-60	NA-86.8, NA-87.0, NA-85.7
Truong et al. [9]	STFT Spectrograms	CNN	LOOCV	13–64	240–30	NA-81.2
Zhang et al. [67]	Common Spatial Pattern Statistics	CNN	LOOCV	23–156	30–30	0.900-92.0
Ours	MFCC	PDBFusNet	LOOCV	15–82	240–60	0.955-95.76

Where NA means not applicable in this work.

References

[1] M. Winterhalder, T. Maiwald, H.U. Voss, R. Aschenbrenner-Scheibe, J. Timmer, A. Schulze-Bonhage, The seizure prediction characteristic: a general framework to assess and compare seizure prediction methods, Epilepsy Behav. 4 (3) (2003) 318–325.

[2] World Health Organization, Epilepsy, 2023, https://www.who.int/news-room/fact-sheets/detail/epilepsy, (Accessed 9 February 2023).

[3] Y. Li, Y. Liu, W. Cui, Y. Guo, H. Huang, Z. Hu, Epileptic seizure detection in eeg signals using a unified temporal-spectral squeeze-and-excitation network, IEEE Trans. Neural Syst. Rehabil. Eng. 28 (4) (2020) 782–794.

[4] O. Faust, U.R. Acharya, H. Adeli, A. Adeli, Wavelet-based eeg processing for computer-aided seizure detection and epilepsy diagnosis, Seizure 26 (2015) 56–64.

[5] R. Jana, I. Mukherjee, Deep learning based efficient epileptic seizure prediction with eeg channel optimization, Biomed. Signal Process. Control. 68 (2021) 102767.

[6] Y. Zhao, C. Li, X. Liu, R. Qian, R. Song, X. Chen, Patient-specific seizure prediction via adder network and supervised contrastive learning, IEEE Trans. Neural Syst. Rehabil. Eng. 30 (2022) 1536–1547.

[7] B. Litt, K. Lehnertz, Seizure prediction and the preseizure period, Curr. Opin. Neurol. 15 (2) (2002) 173–177.

[8] H. Daoud, M.A. Bayoumi, Efficient epileptic seizure prediction based on deep learning, IEEE Trans. Biomed. Circuits Syst. 13 (5) (2019) 804–813.

[9] N.D. Truong, A.D. Nguyen, L. Kuhlmann, M.R. Bonyadi, J. Yang, S. Ippolito, O. Kavehei, Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram, Neural Netw. 105 (2018) 104–111.

[10] J. Yan, J. Li, H. Xu, Y. Yu, T. Xu, Seizure prediction based on transformer using scalp electroencephalogram, Appl. Sci. 12 (9) (2022) 4158.

[11] A. Khorshidtalab, M.J.E. Salami, M. Hamedi, Robust classification of motor imagery eeg signals using statistical time–domain features, Physiol. Meas. 34 (11) (2013) 1563.

[12] C. Vidaurre, N. Krämer, B. Blankertz, A. Schlögl, Time domain parameters as a feature for eeg-based brain–computer interfaces, Neural Netw. 22 (9) (2009) 1313–1319.

[13] L. Ge, K.K. Parhi, Seizure detection using power spectral density via hyperdimensional computing, in: 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, 2021, pp. 7858–7862.

[14] K. Rasheed, J. Qadir, T.J. O'Brien, L. Kuhlmann, A. Razi, A generative model to synthesize eeg data for epileptic seizure prediction, IEEE Trans. Neural Syst. Rehabil. Eng. 29 (2021) 2322–2332.

[15] T. Tamanna, M.A. Rahman, S. Sultana, M.H. Haque, M.Z. Parvez, Predicting seizure onset based on time-frequency analysis of eeg signals, Chaos Solitons Fractals 145 (2021) 110796.

[16] H.T. Shiao, V. Cherkassky, J. Lee, B. Veber, E.E. Patterson, B.H. Brinkmann, G.A. Worrell, Svm-based system for prediction of epileptic seizures from ieeg signal, IEEE Trans. Biomed. Eng. 64 (5) (2016) 1011–1022.

[17] M.K. Hasan, M.A. Ahamed, M. Ahmad, M. Rashid, et al., Prediction of epileptic seizure by analysing time series eeg signal using-nn classifier, Appl. Bion. Biomech. (2017).

[18] Y. Wang, J. Cao, X. Lai, D. Hu, Epileptic state classification for seizure prediction with wavelet packet features and random forest, in: 2019 Chinese Control and Decision Conference, CCDC, 2019, pp. 3983–3987.

[19] X. Xu, M. Lin, T. Xu, Epilepsy seizures prediction based on nonlinear features of eeg signal and gradient boosting decision tree, Int. J. Environ. Res. Public Health 19 (18) (2022) 11326.

[20] T.N. Alotaiby, S.A. Alshebeili, F.M. Alotaibi, S.R. Alrshoud, et al., Epileptic seizure prediction using csp and lda for scalp eeg signals, Comput. Intell. Neurosci. (2017).

[21] A. Parvaiz, M.A. Khalid, R. Zafar, H. Ameer, M. Ali, M.M. Fraz, Vision transformers in medical computer vision—A contemplative retrospection, Eng. Appl. Artif. Intell. 122 (2023) 106126.

[22] X. Zhang, T. Wang, W. Luo, P. Huang, Multi-level fusion and attention-guided CNN for image dehazing, IEEE Trans. Circuits Syst. Video Technol. 31 (11) (2020) 4162–4173.

[23] B. Min, H. Ross, E. Sulem, A.P. Veyseh, T.H. Nguyen, O. Sainz, E. Agirre, I. Heintz, D. Roth, Recent advances in natural language processing via large pre-trained language models: A survey, ACM Comput. Surv. 56 (2) (2023) 1–40.

[24] G. Zeng, X. Xie, M. Chen, J. Weng, Adaptive population extremal optimization-based PID neural network for multivariable nonlinear control systems, Swarm Evol. Comput. 44 (2019) 320–334.

[25] H. Bao, Y. Zhu, Q. Li, Hybrid-scale contextual fusion network for medical image segmentation, Comput. Biol. Med. 152 (2023) 106439.

[26] Y. Chen, L. Feng, C. Zheng, T. Zhou, L. Liu, P. Liu, Pengfei, Y. Chen, Ldanet: Automatic lung parenchyma segmentation from CT images, Comput. Biol. Med. 155 (2023) 106659.

[27] C. Li, C. Shao, R. Song, G. Xu, X. Liu, R. Qian, X. Chen, Spatio-temporal mlp network for seizure prediction using eeg signals, Measurement 206 (2023) 112278.

[28] S. Zhang, D. Chen, R. Ranjan, H. Ke, Y. Tang, A.Y. Zomaya, A lightweight solution to epileptic seizure prediction based on eeg synchronization measurement, J. Supercomput. 77 (2021) 3914–3932.

[29] Y. Xu, J. Yang, S. Zhao, H. Wu, M. Sawan, An end-to-end deep learning approach for epileptic seizure prediction, in: 2020 2nd IEEE International Conference on Artificial Intelligence Circuits and Systems, AICAS, 2020, pp. 266–270.

[30] K.M. Tsiouris, V.C. Pezoulas, M. Zervakis, S. Konitsiotis, D.D. Koutsouris, D.I. 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.

[31] L. Tang, N. Xie, M. Zhao, X. Wu, Seizure prediction using multi-view features and improved convolutional gated recurrent network, IEEE Access 8 (2020) 172352-172361.

[32] C. Cheng, B. You, Y. Liu, Y. Dai, Patient-specific method of sleep electroencephalography using wavelet packet transform and bi-lstm for epileptic seizure prediction, Biomed. Signal Process. Control 70 (2021) 102963.

[33] S.R. Shahamiri, V. Lal, D. Shah, Dysarthric speech transformer: A sequence-to-sequence dysarthric speech recognition system, IEEE Trans. Neural Syst. Rehabil. Eng. (2023).

[34] A. Bhattacharya, T. Baweja, S. Karri, Epileptic seizure prediction using deep transformer model, Int. J. Neural Syst. 32 (2) (2022) 2150058.

[35] C. Li, X. Huang, R. Song, R. Qian, X. Liu, X. Chen, Eeg-based seizure prediction via transformer guided cnn, Measurement 203 (2022) 111948.

[36] M.Z. Parvez, M. Paul, Seizure prediction using undulated global and local features, IEEE Trans. Biomed. Eng. 64 (1) (2016) 208–217.

[37] T. Dissanayake, T. Fernando, S. Denman, S. Sridharan, C. Fookes, Deep learning for patient-independent epileptic seizure prediction using scalp eeg signals, IEEE Sens. J. 21 (7) (2021) 9377–9388.

[38] H. Wang, W. Shi, C.S. Choy, Hardware design of real time epileptic seizure detection based on stft and svm, IEEE Access 6 (2018) 67277–67290.
[39] H. Khan, L. Marcuse, M. Fields, K. Swann, B. Yener, Focal onset seizure prediction using convolutional networks, IEEE Trans. Biomed. Eng. 65 (9) (2017) 2109–2118.
[40] M. Geng, W. Zhou, G. Liu, C. Li, Y. Zhang, Epileptic seizure detection based on stockwell transform and bidirectional long short-term memory, IEEE Trans. Neural Syst. Rehabil. Eng. 28 (3) (2020) 573–580.
[41] H. Ocak, Automatic detection of epileptic seizures in eeg using discrete wavelet transform and approximate entropy, Expert Syst. Appl. 36 (2) (2009) 2027–2036.
[42] A. Subasi, M.I. Gursoy, Eeg signal classification using pca, ica, lda and support vector machines, Expert Syst. Appl. 37 (12) (2010) 8659–8666.
[43] S. Patidar, T. Panigrahi, Detection of epileptic seizure using kraskov entropy applied on tunable-q wavelet transform of eeg signals, Biomed. Signal Process. Control 34 (2017) 74–80.
[44] S. Chen, X. Zhang, L. Chen, Z. Yang, Automatic diagnosis of epileptic seizure in electroencephalography signals using nonlinear dynamics features, IEEE Access 7 (2019) 61046–61056.
[45] K.D. Tzimourta, A.T. Tzallas, N. Giannakeas, L.G. Astrakas, D.G. Tsalikakis, P. Angelidis, M.G. Tsipouras, A robust methodology for classification of epileptic seizures in eeg signals, Health Technol. 9 (2019) 135–142.
[46] Q. Yuan, D. Wei, A seizure prediction method based on efficient features and blda, in: 2015 IEEE International Conference on Digital Signal Processing, DSP, 2015, pp. 177–181.
[47] N.D. Truong, A.D. Nguyen, L. Kuhlmann, M.R. Bonyadi, J. Yang, O. Kavehei, A generalised seizure prediction with convolutional neural networks for intracranial and scalp electroencephalogram data analysis, 2017, arXiv preprint arXiv:1707.01976.
[48] S.M. Usman, S. Khalid, S. Bashir, A deep learning based ensemble learning method for epileptic seizure prediction, Comput. Biol. Med. 136 (2021) 104710.
[49] R.V. Godoy, T.J. Reis, P.H. Polegato, G.J. Lahr, R.L. Saute, F.N. Nakano, H.R. Machado, A.C. Sakamoto, M. Becker, G.A. Caurin, Eeg-based epileptic seizure prediction using temporal multi-channel transformers, 2022, arXiv preprint arXiv:2209.11172.
[50] Y. Gao, A. Liu, X. Cui, R. Qian, X. Chen, A general sample-weighted framework for epileptic seizure prediction, Comput. Biol. Med. 150 (2022) 106169.
[51] A.L. Goldberger, L.A. 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.
[52] T. Dissanayake, T. Fernando, S. Denman, S. Sridharan, H. Ghaemmaghami, C. Fookes, A robust interpretable deep learning classifier for heart anomaly detection without segmentation, IEEE J. Biomed. Health Inform. 25 (6) (2020) 2162–2171.
[53] S. Guo, F. Zhang, A spcnn model for patient-independent prediction of epilepsy using mfcc features, in: 2022 12th International Conference on Information Science and Technology, ICIST, 2022, pp. 68–73.
[54] R. Chen, Y. Tang, W. Zhang, W. Feng, Deep multi-view semi-supervised clustering with sample pairwise constraints, Neurocomputing 500 (2022) 832–845.
[55] R. Chen, Y. Tang, W. Zhang, W. Feng, Adaptive-weighted deep multi-view clustering with uniform scale representation, Neural Netw. 171 (2024) 114–126.
[56] R. Chen, Y. Tang, Y. Xie, W. Feng, W. Zhang, Semisupervised progressive representation learning for deep multiview clustering, IEEE Trans. Neural Netw. Learn. Syst. (2023) http://dx.doi.org/10.1109/TNNLS.2023.3278379.
[57] R. Chen, Y. Tang, X. Cai, X. Yuan, W. Feng, W. Zhang, Graph structure aware contrastive multi-view clustering, IEEE Trans. Big Data (2023) http://dx.doi.org/10.1109/TBDATA.2023.3334674.
[58] P. Wu, Z. Wang, B. Zheng, H. Li, F. Alsaadi, N. Zeng, AGGN: Attention-based glioma grading network with multi-scale feature extraction and multi-modal information fusion, Comput. Biol. Med. 152 (2023) 106457.
[59] T. Zhen, L. Yan, Real-time control strategy of exoskeleton locomotion trajectory based on multi-modal fusion, J. Bionic Eng. 20 (6) (2023) 2670–2682.
[60] W.H. Kruskal, W.A. Wallis, Use of ranks in one-criterion variance analysis, J. Amer. Statist. Assoc. 47 (260) (1952) 583–621.
[61] X. Yang, J. Zhao, Q. Sun, J. Lu, X. Ma, An effective dual self-attention residual network for seizure prediction, IEEE Trans. Neural Syst. Rehabil. Eng. 29 (2021) 1604–1613.
[62] N.D. Truong, L. Kuhlmann, M.R. Bonyadi, D. Querlioz, L. Zhou, O. Kavehei, Epileptic seizure forecasting with generative adversarial networks, IEEE Access 7 (2019) 143999-144009.
[63] J. Tanlamai, A. Pattanateepapon, A. Thakkinstian, C. Limotai, Nonconvulsive seizure and status epilepticus detection with deep learning in high-risk adult critically ill, in: 2022 3rd International Conference on Big Data Analytics and Practices, IBDAP, 2022, pp. 37–42.
[64] D. Cho, B. Min, J. Kim, B. Lee, EEG-based prediction of epileptic seizures using phase synchronization elicited from noise-assisted multivariate empirical mode decomposition, IEEE Trans. Neural Syst. Rehabil. Eng. 25 (8) (2016) 1309–1318.
[65] B. Büyükçakır, F. Elmaz, A.Y. Mutlu, Hilbert vibration decomposition-based epileptic seizure prediction with neural network, Comput. Biol. Med. 119 (2020) 103665.
[66] A.R. Ozcan, S. Erturk, Seizure prediction in scalp eeg using 3d convolutional neural networks with an image-based approach, IEEE Trans. Neural Syst. Rehabil. Eng. 27 (11) (2019) 2284–2293.
[67] 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 Inform. 24 (2) (2019) 465–474.
[68] Y. Li, Y. Liu, Y. Guo, X. Liao, B. Hu, T. Yu, Spatio-temporal-spectral hierarchical graph convolutional network with semisupervised active learning for patient-specific seizure prediction, IEEE Trans. Cybern. 52 (11) (2021) 12189–12204.
[69] N. Ille, Y. Nakao, S. Yano, T. Taura, A. Ebert, H. Bornfleth, S. Asagi, K. Kozawa, I. Itabashi, T. Sato, R. Sakuraba, R. Tsuda, Y. Kakisaka, K. Jin, N. Nakasato, Ongoing EEG artifact correction using blind source separation, Clin. Neurophysiol. 158 (2024) 149–158.
[70] R. Jana, I. Mukherjee, Efficient seizure prediction and EEG channel selection based on multi-objective optimization, IEEE Access (2023).
[71] G. Ghorbanzadeh, Z. Nabizadeh, N. Karimi, P. Khadivi, A. Emami, S. Samavi, DGAFF: Deep genetic algorithm fitness formation for EEG bio-signal channel selection, Biomed. Signal Process. Control 79 (2023) 104119.
[72] D. Liang, A. Liu, Y. Gao, C. Li, R. Qian, X. Chen, Semi-supervised domain-adaptive seizure prediction via feature alignment and consistency regularization, IEEE Trans. Instrum. Meas. 72 (2023) 1–12.
[73] S. Hu, J. Liu, R. Yang, Y. Wang, A. Wang, K. Li, W. Liu, C. Yang, Exploring the applicability of transfer learning and feature engineering in epilepsy prediction using hybrid transformer model, IEEE Trans. Neural Syst. Rehabil. Eng. 31 (2023) 1321–1332.