Physica A 584 (2021) 126376

Contents lists available at ScienceDirect

Physica A

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

Seizure prediction in scalp EEG based channel attention dual-input convolutional neural network☆

Biao Sun [a], Jia-Jun Lv [a], Lin-Ge Rui [a], Yu-Xuan Yang [a], Yun-Gang Chen [b], Chao Ma [a,*], Zhong-Ke Gao [a]

[a] School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China
[b] Jun sheng (Tianjin) Technology Development Co., Ltd, Tianjin 300072, China

ARTICLE INFO

Article history:
Received 19 April 2021
Received in revised form 19 August 2021
Available online 24 August 2021

Keywords:
Scalp EEG
Seizure prediction
Convolutional neural network
Deep learning
Time series analysis

ABSTRACT

Epilepsy is one of the most common brain diseases, and seizures usually occur randomly. Accurately predicting seizures enable doctors and patients to carry out medical prevention timely. In seizure prediction studies, single-domain information input (time domain or frequency domain, Etc.) neglects some parts' information from signals. In this paper, we propose a novel deep learning framework named channel attention dual-input convolutional neural network (CADCNN) to obtain the signal's useful information fully. The spatial–temporal features extracted by short-time Fourier transform (STFT) are fed to the CADCNN, and the raw EEG signals are fed for further feature extraction. With the fusion of two inputs from different domains and the combination of channel attention, CADCNN can learn faithful and distinguishable representations of EEG signals and boost the temporal, spectrum, and spatial information utilization capability. We evaluate the proposed method using the Boston Children's Hospital-MIT scalp EEG public datasets. Compared with other state-of-the-art methods, the sensitivity, false prediction rate, specificity, and AUC of our proposed method reach 97.1%, 0.029h, 95.6%, and 0.917, respectively, presenting better performance and higher prediction accuracy.

1. Introduction

Epilepsy is one of the most common severe brain diseases, seriously affecting 1% of the world's population [1,2]. From the study of epilepsy patients, when seizures occur, they are often accompanied by abnormal emotions and behaviors. In clinical practice, CT, MRI, and conventional EEG are the most common tools used by physicians to check the possibility of epilepsy [3,4].

In the epileptic EEG signal analysis, because seizures occur randomly [5], timely prediction of the upcoming seizures of patients with epilepsy can leave sufficient treatment time for patients and doctors. However, only professional trained medical doctors can judge whether the patient is sick or not [6,7]. In response to the time-consuming manual detection of seizures and the risk of human errors, the researchers researched new methods for predicting seizures in the field of artificial intelligence combined with epileptic EEG. With the rise of deep learning (DL) and machine learning (ML) in EEG signal analysis [8–12], some research on seizure prediction based on DL methods has been carried out. Zhang [13] proposed a method to extract the features of EEG signals using CSP. Then a shallow CNN was applied to classify preictal

☆ This work was supported in part by the National Natural Science Foundation of China under Grants Nos. 61873181, 61903270 and 61922062.
* Corresponding author.
E-mail address: chao.ma@tju.edu.cn (C. Ma).

https://doi.org/10.1016/j.physa.2021.126376
0378-4371/

and interictal in the CHB-MIT dataset, obtaining a sensitivity of 92.2% and a false prediction rate of 0.12/h. Ozcan [14] combined multiple features with 3D CNN and tested 16 patients to achieve a sensitivity of 85.7% and a false prediction rate of 0.096/h.

In the previous work, epilepsy events were divided into ictal state, preictal state, and interictal state [15]. Yang [16] revealed the synchronization between all channels by cross recurrence plots, with differences in recurrence rates between preictal and interictal states. All these research works laid the foundation for later studies. The previously proposed methods mainly extract the features of epileptic EEG signals from the time domain or frequency domain. Time-domain feature methods such as convergence and divergence short-term maximum Lyapunov index method mainly extract the temporal information of epileptic EEG signals [17,18]. Frequency domain feature methods such as Power Spectral Density (PSD) [19], Discrete Wavelet Transform (DWT) [20], and Short-Time Fourier Transform (STFT) [21] mainly extract the spectral information of epileptic EEG signals. STFT has a significant effect in extracting frequency domain information of epileptic EEG signals [22]. Truong [23] proposed a method to perform an STFT on the EEG signals to obtain the feature signals and performed convolution filtering on the feature signals, achieving a sensitivity of 81.2% and a false prediction rate of 0.16/h. Truong [24] also proposed a semi-supervised seizure prediction method combining STFT and GAN. In this paper, we use STFT to extract the spectral information of the EEG signals.

The CNN deep learning model's input form in the latest EEG research is primarily a single domain input, such as EEG signals or feature signals. While exciting results have been achieved, there are still many challenges to further improve CNN models' ability to extract EEG signal information. In these CNN models, since both the convolutional and pooling layers have the attributes of locally extracting features, they neglect some partial information of the EEG signals or feature signals. We introduce the attention mechanism, which can be explained intuitively using human visual mechanisms. The human visual system focuses on important information, concentrating on certain judgment-aiding elements and images while ignoring irrelevant information. Attention mechanisms in deep learning help the model focuses on rich information and extracts deep high-level features from the feature maps dynamically. Various forms of attention have been proposed in recent years [25–28]. Deep learning frameworks based on attention mechanisms are widely used in image processing [29], text recognition [30] and speech systems [31] and have also obtained excellent performance in EEG signal analysis [32]. Based on these studies, we introduce a channel attention mechanism based on Hu [33] and combine it with the two input forms of different domains, EEG signals, and feature signals. This paper's main contributions are listed as follows:

(1) We have developed a novel deep learning framework CADCNN to deal with EEG-based seizure prediction tasks. The experimental results show that our proposed method outperforms the existing methods considered in the comparison.
(2) CADCNN could elegantly utilize prior knowledge to enhance the · model's ability to capture spectral information. Moreover, few existing works employ another form (time domain or frequency domain, Etc.) of EEG signals to enhance the model's feature extraction ability.
(3) The framework fuses two inputs from different domains. It combines channel attention, which can learn faithful and distinguishable representations of EEG signals and boost the temporal, spectrum, and spatial information utilization capability.

2. Method

2.1. Channel attention dual-input convolutional neural network

In recent years, CNNs are widely used in the analysis of EEG and epileptic EEG signals [34–36], reflecting the widespread use of CNNs in the field of EEG analysis. We propose the channel attention dual-input convolutional neural network (CADCNN), in which the feature processing module is used to process feature signals and the signal processing module is used to process EEG signals. The structure of both the feature processing and signal processing modules consists of four Conv (convolutional) layers, channel attention, and a dense layer. Each Conv layer contains a ReLU activation function and a batch normalization layer. The first Conv layer in the feature processing module has 16 kernels with a kernel size of $1 \times 2 \times 1$ and a stride of $1 \times 1 \times 1$ for filtering in the temporal direction. The other three Conv layers with kernel numbers of 32, 64, and 64, kernel sizes of $18 \times 1 \times 1$, $1 \times 1 \times 7$, and $1 \times 1 \times 5$, and strides of $1 \times 1 \times 1$, $1 \times 1 \times 3$, and $1 \times 1 \times 3$, respectively, are used for channel feature fusion to extract spatial information and frequency direction feature fusion.

The EEG signal contains rich Spatio-temporal information. The first Conv layer in the signal processing module performs channel feature fusion to extract the epileptic EEG signal's spatial information, which has 16 kernels with a kernel size of 18×1 and a stride of 1×1. The other three Conv layers perform temporal filtering to extract temporal features, where the number of kernels is 32, 64, and 64, and the kernel sizes are 1×7, 1×7, and 1×5, respectively, all in strides of 1×3. The features extracted from the two modules are flattened and connected to a fully connected layer with an output size of 256. We use a dropout rate of 0.4 and 0.3 before and after this fully connected layer, respectively. The output is concatenated together after the fully connected layers. After that, two fully connected layers with output sizes of 32 and 2 are added. The activation functions are the ReLU activation function and the softmax activation function, respectively.

The importance of individual feature channels in the feature map is likely to be different. Channel attention automatically obtains the importance of each channel by learning, and then can optimize network performance by selecting

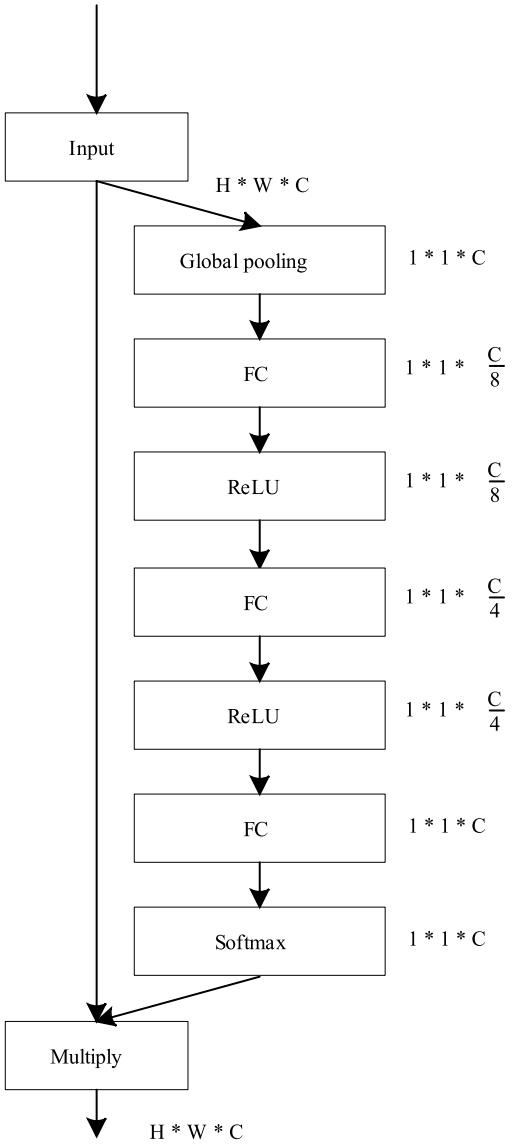


Fig. 1. Channel attention module.

valid information based on each channel's importance and suppressing irrelevant information. Through global average pooling, channel attention can capture the feature information of the signal from the global sensory field, allowing the global information to be utilized at lower levels of the network. The low-level network obtains the global receptive field, which is essential in the classification task of epileptic EEG signals. As shown in Fig. 1, the input data size $[H, W, C]$ (H is the feature length, W is the feature width, C is the number of feature channels) is squeezed into a real number $[1, 1, C]$ by global average pooling, called the squeeze operation. This real number characterizes the global receptive field on the feature channel. Eq. (1) shows that $x_k(i, j)$ denotes the input feature data on the kth channel at position (i, j), and the input data is $X = \{x_1, x_2, \ldots, x_k\}$, $k \in (1, 2, \ldots, C)$, where the dimensions of x_k and X are $x_k \in R^{H\times W}$ and $X \in R^{H\times W\times C}$. z_k is obtained from the kth channel feature data by global average pooling. z is obtained from the feature data of the C feature channels by Eq. (1), its dimensional information is $z \in R^C$.

$$z_k = \frac{1}{W \times H} \sum_{i=1}^{W} \sum_{j=1}^{H} x_k(i, j) \tag{1}$$

The data after the squeeze operation is downscaled and upscaled by a three-layer fully connected network to fuse the feature map information of each feature channel. As shown in Eq. (2), each feature channel's weights are learned by the

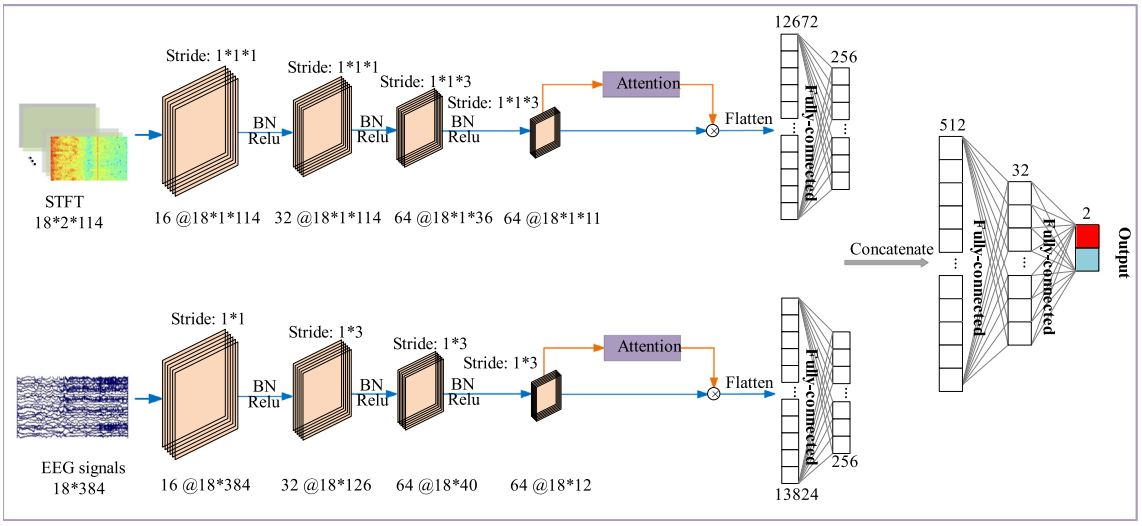


Fig. 2. Structure of channel attention dual-input convolutional neural network (CADCNN). e.g., 16@18 × 1 × 114 means 16 kernels, and the output size is 18 × 1 × 114. The batch normalization layer we note as BN. The kernel size of the feature processing module is 1 × 2 × 1, 18 × 1 × 1, 1 × 1 × 7, and 1 × 1 × 5. The kernel size of the signal processing module is 18 × 1, 1 × 7, 1 × 7, and 1 × 5.

parameter W, which is used to model the correlation between feature channels. σ denotes softmax, δ denotes ReLU, and the parameter dimension information is $W_1 \in R^{\frac{C}{8}\times C}$, $W_2 \in R^{\frac{C}{4}\times\frac{C}{8}}$ and $W_3 \in R^{C\times\frac{C}{4}}$.

$$s = \sigma\left(W_3\delta\left(W_2\delta\left(W_1 z\right)\right)\right) \tag{2}$$

The dimension of s is learned as $[1, 1, C]$, which describes each feature map's weight. As shown in Eq. (3), the learned weight s is the importance of each feature channel after feature selection, which is then applied to each of the original feature channels by multiplication to achieve recalibration of the original feature importance. $\widetilde{x_k}$ denotes the result after rescaling the original feature data's importance by the channel attention module on the kth channel.

$$\widetilde{x_k} = s_k \times x_k \tag{3}$$

In this study, through the combination of the EEG signals and the feature signals, the model learns more temporal and spatial information from the EEG signals and obtains the ability to extract spectrum information effectively. The model uses stochastic gradient descent (SGD) to update network parameters. SGD optimization algorithm is used with a learning rate of 0.01, a momentum of 0.9, and a weight decay of 1e-6. The addition of momentum to SGD can increase the stability of learning to a certain extent to learn faster and get rid of the local optimum. The structure of the CADCNN is shown in Fig. 2.

2.2. Seizure prediction method

Previous works defined two parameters for the seizure prediction problem: seizure prediction horizon (SPH) and seizure occurrence period (SOP). In this paper, we follow the definitions of the SOP and SPH proposed by Maiwald [37]. SOP is defined as the period of anticipated seizures, and SPH is defined as the time from the alarm to the start of the SOP. Among them, SPH is the buffer stage of seizure prediction, leaving enough time for patients with epilepsy to prepare mentally and take other measures. SOP should not be too long, which causes mental stress for the patient. From the clinical perspective, the SPH should not be too long to ensure predicting seizures' accuracy. Therefore, we define SOP as 30 min and SPH as 3 min in this paper. We set the alarm at the beginning of the SPH phase. When a seizure occurs, the alarm is raised and continues until the seizure ends in the SOP. The seizure prediction method based on SOP and SPH is shown in Fig. 3. We evaluate an accurate seizure prediction: when an alarm occurs, the seizures must come after the SPH and within the SOP.

The overall target of our proposed method is to recognize preictal from interictal states. When an alarm occurs, and there is no seizure in the SOP, it is called a false alarm. Therefore, an essential task in seizure prediction is to reduce the false alarm rate. We use the k of n method [23] to reduce seizure prediction's false alarm rate. An alarm is raised only when at least k of the n segments is classified as the preictal state. In this paper, we choose $k = 40$ and $n = 50$ based on [23] (this means that an alarm is raised out of 75 s of continuous data when at least 60 s of data are classified as the preictal state. Each of these segments is 1.5 s). We define 50 segments as a group, and if at least 40 segments in a group are classified as the preictal or interictal states, then the group is defined as the preictal or interictal state,

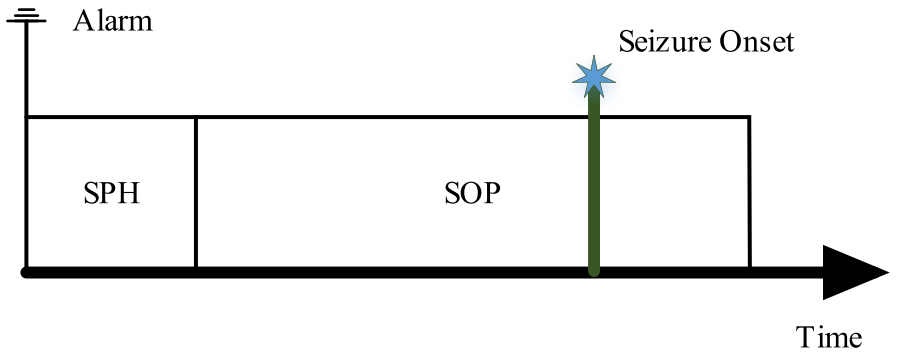


Fig. 3. Definition of seizure prediction horizon (SPH) and seizure occurrence period (SOP).

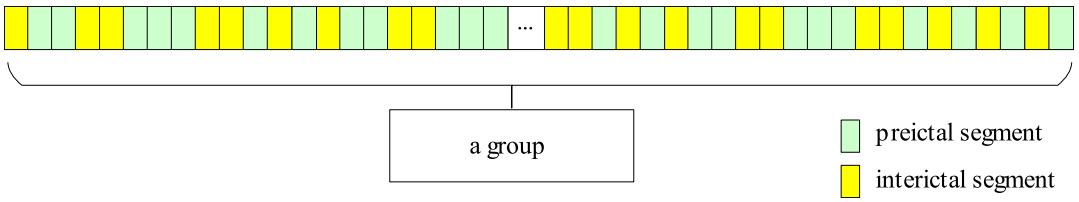


Fig. 4. A group is defined as a preictal state if at least 40 preictal segments are correctly classified out of 50 segments.

as shown in Fig. 4. Once a group is correctly classified as a preictal state, the alarm is raised, indicating that seizure will occur. When a group is incorrectly classified as a preictal state, the alarm is raised, which is called a false alarm. Therefore, we evaluate the model's effectiveness using metrics such as sensitivity, FPR, specificity, and AUC [38]. Sensitivity is used to evaluate seizure prediction accuracy and is defined as the percentage of correctly predicted seizures divided by the total number of seizures. False prediction rate (FPR) is an essential indicator for evaluating the false alarm rate of seizure prediction and is defined as the number of false alarms per hour. The false positives (FP) are defined by the number of segments that were falsely classified as preictal segments, false negatives (FN) are defined by the number of segments that were falsely classified as interictal segments, and true negatives (TN) are defined by the number of segments that were correctly classified as interictal segments. The specificity is calculated as $\frac{TN}{(TN+FP)}$. We use the AUC (Area Under Curve) as the metric for evaluating the model's performance. AUC is a thresholdless measurement, which can directly compare the performance of different methods.

3. Experiment and discussion

3.1. Dataset

The dataset used for this study is the Children's Hospital Boston, which consists of EEG recordings from pediatric subjects with intractable seizures [39,40]. The scalp EEG (sEEG) record contains 22 subjects (5 males, ages 3–22; 17 females, ages 1.5–19). The data sampling frequency is 256 Hz, and sEEG data from 22 electrodes are collected. The sEEG signals are collected using the international bipolar 10–20 system. sEEG recordings are stored in edf format, containing between 9 and 42 edf files per subject, and the data are preprocessed using MATLAB. The dataset contains recordings of sEEG, and a summary text with patient's information.

3.2. Experiment

Since there are multiple patients in the experiment using different electrodes, the use of different electrodes can lead to increased bias in data analysis, so we select the same 18 channels that all patients have: 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, and CZ-PZ.

This experiment defines the interictal state as at least 4 h before the seizure onset and at least 4 h after its end. We select continuous interictal data near seizure onset as the interictal sample set (IS Set) of this experiment, which lasts at least 4 h. The preictal data is defined as the preictal state data lasting 30 min before the seizure onset. The division of the data in the preictal and interictal states is shown in Fig. 5. If the selected patient has fewer than three seizures, using only one seizure to predict another seizure tends to trigger the model's overfitting to a particular patient, making it impossible to predict others' seizures completely. Therefore, at least two seizures must be selected in the training set, and another seizure is used for testing to reduce model overfitting. Considering that in some data segments, the time is

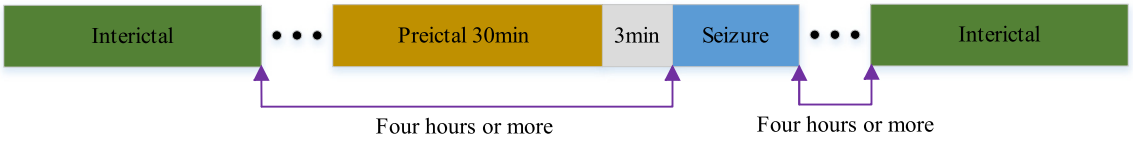


Fig. 5. Preictal and interictal state division.

Fig. 6. The EEG window is 1.5 s, and the sliding step is 0.75 s to expand the preictal data.

not continuous (e.g., chb19_27.edf and chb19_28.edf in chb19), the definition of preictal data selection is incompatible (e.g., chb13_47.edf and chb13_55.edf in chb13), the interictal data is not enough (e.g., chb08), so we select 17 patients in the entire CHB-MIT dataset for experiments.

STFT is widely used in signal processing, and it has been proved in many studies that STFT has advantages in seizure prediction [22,24]. We use an STFT to project the EEG sequences onto the time–frequency axis. STFT converts EEG signals into 2D representation. The STFT converts EEG signals into a 3D matrix used as an input to a CNN because of multi-dimensional epileptic EEG data. In this paper, we use an EEG window length of 1.5 s, and the 1.5s segment of the EEG signal is obtained. The STFT is used to transform the EEG signal segment into a three-dimensional matrix as the feature signal. EEG recordings of the CHB-MIT dataset are contaminated with power line noise at 60 Hz. Therefore, the frequency components in the range of 57–63 Hz and 117–123 Hz are excluded. The DC component (at 0 Hz) is also removed.

We use the leave-one-out method to train the model. If a subject has N seizures, N-1 seizures are used as the training set, and the remaining seizure is used as the testing set so that each seizure is tested as the testing set. When dividing the data samples, the amount of data in the preictal and interictal states is highly unbalanced, reaching a ratio of 1:15 between preictal and interictal data in multiple subjects. As shown in Fig. 6, we use an overlapping sample approach in this paper to balance the amount of data in the two categories, with each segment being 1.5 s, and we use a sliding step of 0.75 s to expand the preictal training samples. (e.g., chb18 contains four seizures, we perform 4-fold cross-validation with three preictal data (each preictal data length is 30 min) as training and one preictal data to test each fold of the experiment. We use the method of overlapping samples to expand the preictal data for training as 2398×3 segments). From our IS Set, one hour of data is randomly obtained as an interictal testing set. Simultaneously, we select the same number of interictal data segments as the preictal training data for training. We randomly select 20% of the data from each patient's preictal and interictal training sets as the validating set to prevent the model from overfitting.

3.3. Results

In this paper, our proposed method is implemented by the Keras and TensorFlow libraries in Python3. We use leave-one-out cross-validation to evaluate the performance of CADCNN and compare its performance with feature-only (the model only fed with the feature) according to sensitivity, FPR, specificity, and AUC. The structure and hyperparameters of the feature-only model are consistent with the feature processing module of CADCNN.

From Table 1, we compare the sensitivity, FPR, and specificity of the CADCNN and the feature-only. In terms of sensitivity, for all subjects, the CADCNN (97.1%) has a significant improvement over the feature-only (91.5%). It proves that the proposed CADCNN is superior to feature-only in the accuracy of seizure prediction. Evaluating the model's FPR, we get such a result that the CADCNN has a significant improvement over the feature-only, which is from 0.101/h to 0.029/h. This result indicates that the CADCNN is more advantageous in reducing the probability of false alarms. Also, we evaluate the probability of false alarms by the accuracy of interictal state classification. The performance of the CADCNN we developed has also been improved, where the specificity has increased from 93.8% to 95.6%. Moreover, its standard deviation has declined from 3.9% to 2.2%. It indicates that the CADCNN performs better and is more effective in reducing the false alarm rate.

As Fig. 7 shows that in this comparison, for each subject, the CADCNN has almost improved with its AUC value increased from 0.851 to 0.917 and standard deviation declined from 8.8% to 4.3%. Therefore, the proposed CADCNN has better classification performance. In this paper, the proposed CADCNN uses the dual-input structure to combine the signal sequences and the STFT features. Compared with multiple feature-only indicators, the results demonstrate that CADCNN can efficiently learn temporal and spatial information from the multi-channel EEG signals.

Table 1
The comparison between the performance of the CADCNN and the feature-only.

Patient	No.	Feature-only			CADCNN		
		Sp. (%)	Sn. (%)	FPR(/h)	Sp. (%)	Sn. (%)	FPR(/h)
Chb01	7	99.8	100	0	99.8	100	0
Chb02	3	92.7	66.7	0	95.5	100	0
Chb03	5	96.1	100	0	97.1	100	0
Chb04	3	93.3	33.3	0	98.5	100	0
Chb05	5	89.5	100	0	93.7	100	0
Chb06	9	82.6	55.6	0.89	77.9	66.7	0.33
Chb07	3	97.1	100	0	96.6	100	0
Chb09	3	98.7	100	0	99.8	100	0
Chb10	6	91.0	100	0.83	95.4	83.3	0.16
Chb13	4	90.0	100	0	99.4	100	0
Chb14	4	100	100	0	100	100	0
Chb16	5	86.2	100	0	86.5	100	0
Chb17	3	99.2	100	0	97.6	100	0
Chb18	4	100	100	0	100	100	0
Chb20	4	99.7	100	0	99.7	100	0
Chb21	4	89.9	100	0	93.0	100	0
Chb23	5	89.2	100	0	94.2	100	0
Total	**77**	**93.8**	**91.5**	**0.101**	**95.6**	**97.1**	**0.029**

[a] No.: The number of seizures. Sp.: Specificity. Sn.: Sensitivity.

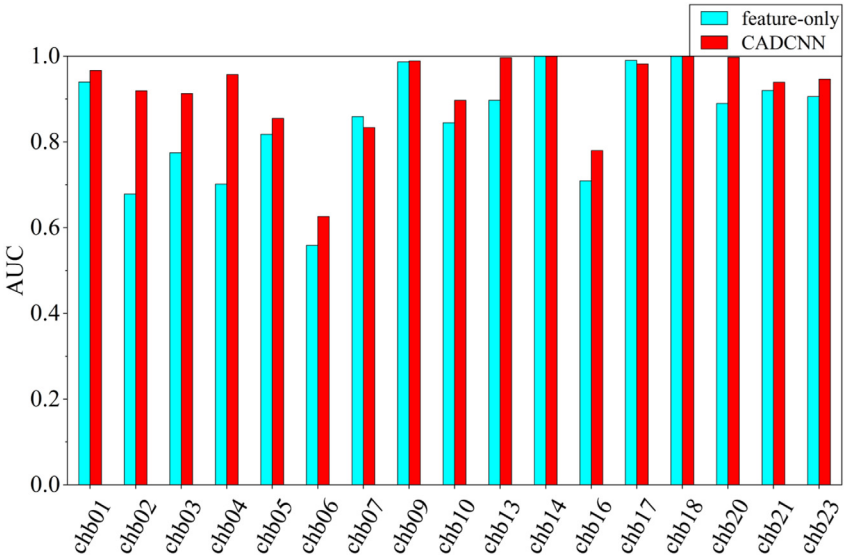


Fig. 7. The comparison between the AUC of the CADCNN and the feature-only.

3.4. Discussion

Our proposed dual-input structure achieves effective performance in seizure prediction and further validates our proposed method's reliability. We build some comparative models based on the CADCNN as follows:

CADCNN-a: it removed the channel attention.

CADCNN-b: the three-layer fully connected network of channel attention was replaced by a two-layer fully connected network (it removed the first fully connected layer).

CADCNN-c: an average pooling layer was added to both branches of the dual input structure.

CADCNN-d: a maxpooling layer was added to both branches of the dual input structure.

(Both CADCNN-c and CADCNN-d added pooling layers with kernel sizes of $1 \times 1 \times 2$ and 1×2, respectively, and strides of $1 \times 1 \times 2$ and 1×2, respectively.)

CADCNN-e: a convolutional layer was added after the convolutional layer of the dual-input structure with kernel sizes of $1 \times 1 \times 3$ and 1×3 for both branches and strides of $1 \times 1 \times 2$ and 1×3, respectively.

We compare two essential metrics of seizure prediction, sensitivity and FPR. Fig. 8(a) and Fig. 8(b) show that CADCNN-a does not introduce channel attention to obtain channels' importance and has reduced seizure prediction performance compared to CADCNN. Compared with CADCNN-b, channel attention in CADCNN learns the nonlinear relationships among

Table 2
Comparison of performance measures for our work with other reported methods.

Authors	Dataset	No.	No. s	Classifier	Sn. (%)	FPR(/h)	SPH
[37]	CHB-MIT	10	–	IndRNN	88.80	–	–
[42]	CHB-MIT	10	64	MLP	89.8	0.081	–
[43]	CHB-MIT	3	18	–	88.34	0.155	40/min
[44]	CHB-MIT	10	31	–	77	0.17	60/min
[46]	CHB-MIT	20	–	uMMD-AE	97.2	0.64	0
[23]	CHB-MIT	13	64	CNN	81.2	0.16	5/min
[47]	CHB-MIT	21	65	SVM	82.44	–	5/min
[14]	CHB-MIT	16	77	3D CNN	85.7	0.096	1/min
[45]	CHB-MIT	23	–	–	92.0	–	10/s
Our work	CHB-MIT	17	77	CADCNN	97.1	0.029	3/min

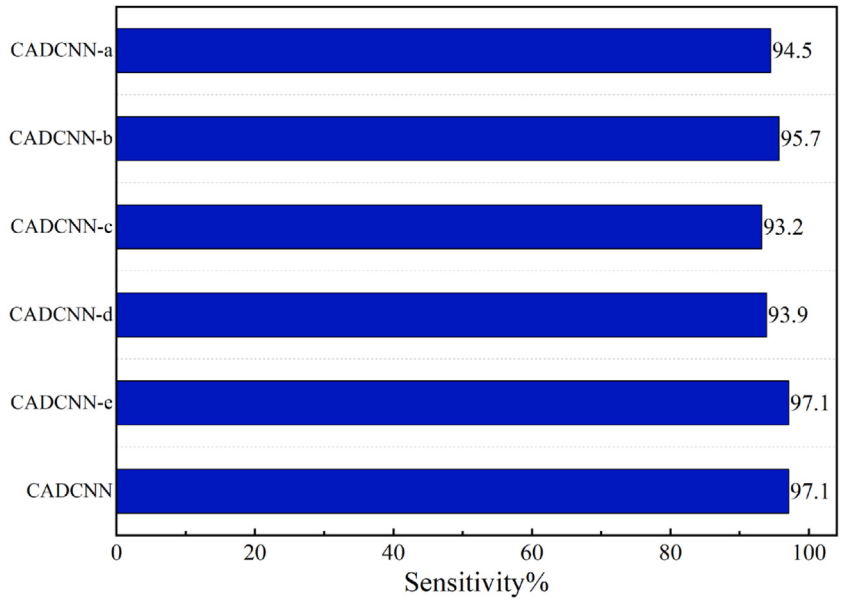
[a] No.: The number of used patients. No. s: The number of used seizures. Sn.: Sensitivity.
[b] IndRNN: independently recurrent neural network, MLP: Multilayer Perceptron, uMMD-AE: unified maximum mean discrepancy autoencoder.

channels at a deeper level, captures channel dependencies comprehensively, and achieves higher sensitivity. CADCNN-c and CADCNN-d perform pooling operations. Although it expands the perceptual field and reduces the number of parameters, it causes a partial loss of feature information and does not yield effective performance in this classification task. Compared with CADCNN, the dual-input structure in CADCNN-e performs deeper feature extraction by adding a convolutional layer. The higher FPR of CADCNN-e than CADCNN is mainly due to the overfitting of CADCNN-e during training with a deeper network structure. The model misclassifies more interictal segments as preictal segments in successive interictal sequences, which increases the FPR of CADCNN-e. Both frameworks achieve the same sensitivity, but our proposed CADCNN structure is the most streamlined and does not consume additional resources. Compared with all models, CADCNN achieves smaller FPR and reduces false alarms of seizures more effectively. Overall, the CADCNN performs better in the seizure prediction task, indicating that our proposed dual-input structure effectively captures the EEG feature information.

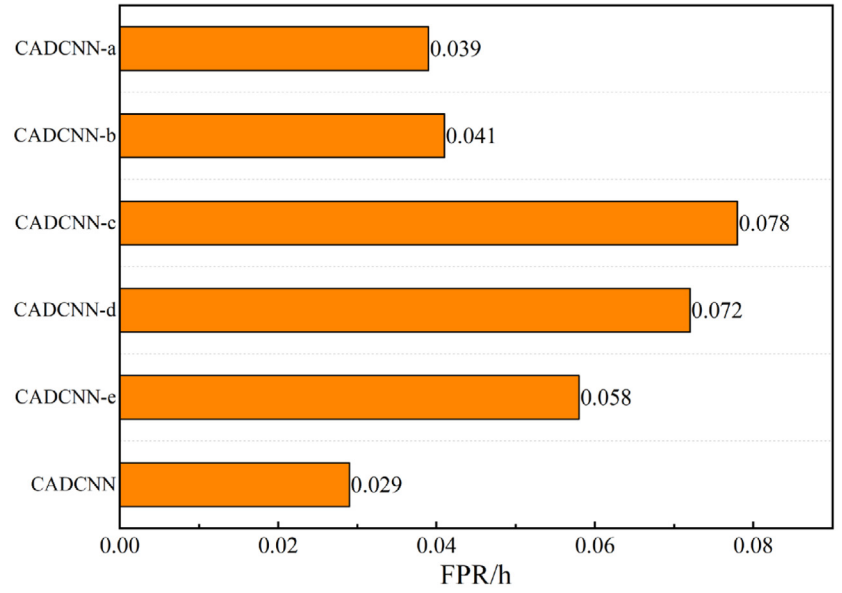
Several machine learning and deep learning methods have been introduced for predicting seizures using EEG signals in recent years. In Table 2, we compare the other methods on the CHB-MIT dataset. There is currently no unified seizure prediction definition, and there is no consistent division of preictal state time and interictal state time between different methods. These studies only consider some specific dataset selection, which leads to lousy generalization performance of model. i.e., the model which presents good effectiveness on one dataset has a terrible performance on another one.

From a clinical perspective, SPH is the buffer stage of seizure prediction, leaving enough time for epilepsy patients to prepare mentally and take other measures. Therefore, SPH should be considered in seizure prediction. Yao [41] proposed an independent recurrent neural network (IndRNN) with a dense structure and attention mechanism to exploit temporal and spatial discriminative features. An end-to-end learning strategy obtained a sensitivity of 88.80% and specificity of 88.60%. Büyükçakır [42] used the Hilbert Vibration Decomposition (HVD) method to decompose the EEG signal into seven subcomponents, achieving a sensitivity of 89.8% and a false prediction rate of 0.081/h. However, [41] and [42] did not consider the effect of SPH on the prediction effect. In addition, the SPH time should not be too long to reduce the accuracy of seizure prediction. Zandi [43] proposed a method based on positive zero-crossing interval analysis of the scalp EEG and achieved a high sensitivity of 88.34% and a false prediction rate of 0.155/h. The SPH used was 40 min. Myers [44] defined the SPH as 60 min and achieved a sensitivity of 77% and FPR of 0.17FP/h by calculating the phase/amplitude locking value (PLV/ALV). Likewise, the SPH should not be too short. Slimen [45] defined SPH as 10 s and proposed spike rate as a predictor of the seizure EEG signal. Actual clinical factors were not adequately included in the assessment in this work. When SPH is 0, the actual medical factors are ignored, which may cause a higher accuracy of seizure prediction. Tang [46] defined SPH as 0 and used principal component analysis with co-spatial patterns to extract information on the spectral and temporal characteristics of the EEG signal. The effectiveness of seizure prediction was mainly based on the screening of features.

In actual clinical practice, choosing an appropriate time for SOP and SPH can help improve the accuracy of seizure prediction and reduce patients' psychological pressure [14,23,45,47]. Cho [47] proposed a method to predict seizures based on noise-assisted multivariate empirical mode decomposition (NA-MEMD), phase-locked values (PLV), and SVM, obtaining a sensitivity of 82.44%. Then, Ozcan [14] combined multiple EEG features with a 3D CNN and obtained a sensitivity of 85.7% and a false prediction rate of 0.096/h. In contrast to these methods, our proposed CADCNN model uses two branches to fully combine the raw information and time–frequency feature information of the EEG signal. The channel attention in the model further focuses on the rich channel feature information in the EEG signal, which can effectively extract temporal, spectral, and spatial information. Moreover, we evaluate the AUC value of different methods on the CHB-MIT dataset. As shown in Fig. 9, we chose the data jointly owned in Truong [23], Zhang [13] for comparison with our method. We directly obtain a comparison of the classification performance of different methods. From the overall performance comparison perspective, the CADCNN we proposed is a more effective method to predict seizures.



(a)



(b)

Fig. 8. (a) and (b) are the sensitivity and FPR comparisons between all models, respectively.

4. Conclusion

Predicting seizures accurately and timely leaves patients with enough time to intervene. In this paper, the fusion of two different information domains is beneficial for the model to improve the representation capabilities of the useful

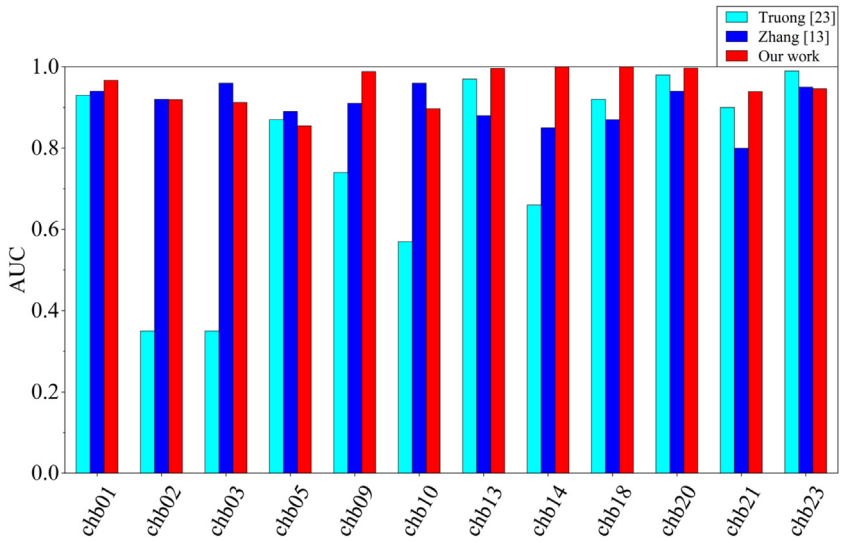


Fig. 9. Comparison between Truong [23], Zhang [13] and our work in AUC on the same subjects.

information of the EEG signals. Our proposed method achieves a sensitivity of 97.1% and a false prediction rate of 0.029/h and outperforms other state-of-the-art methods under consideration. We expect further development of this approach in other seizure prediction studies and apply it to additional datasets to achieve high accuracy seizure prediction.

CRediT authorship contribution statement

Biao Sun: Conceptualization, Investigation, Data curation, Software. **Jia-Jun Lv:** Methodology, Software, Validation, Data curation, Writing – original draft. **Lin-Ge Rui:** Visualization. **Yu-Xuan Yang:** Supervision. **Yun-Gang Chen:** Supervision. **Chao Ma:** Formal analysis, Writing – review & editing. **Zhong-Ke Gao:** Formal analysis, Supervision.

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.

References

[1] C.P. Antonopoulos, Cyberphysical systems for epilepsy and related brain disorders: Multi-parametric monitoring and analysis for diagnosis and optimal disease management, Cyberphys. Syst. Epilepsy Relat. Brain Disord. Multi-Parametr. Monit. Anal. Diagn. Optim. Dis. Manag. (2015) 1–279, http://dx.doi.org/10.1007/978-3-319-20049-1.

[2] H.M. Emara, M. Elwekeil, T.E. Taha, A.S. El-Fishawy, E.S.M. El-Rabaie, T. Alotaiby, S.A. Alshebeili, F.E. Abd El-Samie, Hilbert transform and statistical analysis for channel selection and epileptic seizure prediction, Wirel. Pers. Commun. 116 (2021) 3371–3395, http://dx.doi.org/10.1007/s11277-020-07857-3.

[3] E.C. Mader, X.M. Xiang, P.W. Olejniczak, D. Miller, Ictal hypersalivation and salivary gland enlargement in a patient with acquired frontal lobe epilepsy, Cureus 13 (2021) http://dx.doi.org/10.7759/cureus.15319.

[4] A. Shoeibi, M. Khodatars, N. Ghassemi, M. Jafari, P. Moridian, R. Alizadehsani, M. Panahiazar, F. Khozeimeh, A. Zare, H. Hosseini-Nejad, A. Khosravi, A.F. Atiya, D. Aminshahidi, S. Hussain, M. Rouhani, S. Nahavandi, U.R. Acharya, Epileptic seizures detection using deep learning techniques: A review, Int. J. Environ. Res. Publ. Health. 18 (2021) http://dx.doi.org/10.3390/ijerph18115780.

[5] C.N. Cooray, A. Carvalho, G.K. Cooray, Noise induced quiescence of epileptic spike generation in patients with epilepsy, J. Comput. Neurosci. 49 (2021) 57–67, http://dx.doi.org/10.1007/s10827-020-00772-3.

[6] P. Thodoroff, J. Pineau, A. Lim, Learning robust features using deep learning for automatic seizure detection, 2016, pp. 1–12, http://arxiv.org/abs/1608.00220.

[7] F. Fürbass, P. Ossenblok, M. Hartmann, H. Perko, A.M. Skupch, G. Lindinger, L. Elezi, E. Pataraia, A.J. Colon, C. Baumgartner, T. Kluge, Prospective multi-center study of an automatic online seizure detection system for epilepsy monitoring units, Clin. Neurophysiol. 126 (2015) 1124–1131, http://dx.doi.org/10.1016/j.clinph.2014.09.023.

[8] C. Liu, B. Tan, M. Fu, J. Li, J. Wang, F. Hou, A. Yang, Automatic sleep staging with a single-channel EEG based on ensemble empirical mode decomposition, Phys. A Stat. Mech. Appl. 567 (2021) 125685, http://dx.doi.org/10.1016/j.physa.2020.125685.

[9] V.V. Grubov, E. Sitnikova, A.N. Pavlov, A.A. Koronovskii, A.E. Hramov, Recognizing of stereotypic patterns in epileptic EEG using empirical modes and wavelets, Phys. A Stat. Mech. Appl. 486 (2017) 206–217, http://dx.doi.org/10.1016/j.physa.2017.05.091.

[10] Z. Gao, Y. Li, Y. Yang, N. Dong, X. Yang, C. Grebogi, A coincidence-filtering-based approach for CNNs in EEG-based recognition, IEEE Trans. Ind. Inf. 16 (2020) 7159–7167, http://dx.doi.org/10.1109/TII.2019.2955447.

[11] Z.K. Gao, Q. Cai, Y.X. Yang, N. Dong, S.S. Zhang, Visibility graph from adaptive optimal kernel time-frequency representation for classification of epileptiform EEG, Int. J. Neural Syst. 27 (2017) 1–12, http://dx.doi.org/10.1142/S0129065717500058.

[12] Q. Cai, Z. Gao, J. An, S. Gao, C. Grebogi, A graph-temporal fused dual-input convolutional neural network for detecting sleep stages from EEG signals, IEEE Trans. Circuits Syst. II Express Briefs 68 (2021) 777–781, http://dx.doi.org/10.1109/TCSII.2020.3014514.
[13] 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. Heal. Inf. 24 (2020) 465–474, http://dx.doi.org/10.1109/JBHI.2019.2933046.
[14] A.R. Özcan, S. Erturk, Seizure prediction in scalp EEG using 3D convolutional neural networks with an image-based approach, IEEE Trans. Neural Syst. Rehabil. Eng. 27 (2019) 2284–2293, http://dx.doi.org/10.1109/TNSRE.2019.2943707.
[15] B. Litt, K. Lehnertz, Seizure prediction and the preseizure period, Curr. Opin. Neurol. 15 (2002) 173–177, http://dx.doi.org/10.1097/00019052-200204000-00008.
[16] C. Yang, G. Luan, Z. Liu, Q. Wang, Dynamical analysis of epileptic characteristics based on recurrence quantification of SEEG recordings, Phys. A Stat. Mech. Appl. 523 (2019) 507–515, http://dx.doi.org/10.1016/j.physa.2019.02.017.
[17] L.D. Iasemidis, D.S. Shiau, W. Chaovalitwongse, J.C. Sackellares, P.M. Pardalos, J.C. Principe, P.R. Carney, A. Prasad, B. Veeramani, K. Tsakalis, Adaptive epileptic seizure prediction system, IEEE Trans. Biomed. Eng. 50 (2003) 616–627, http://dx.doi.org/10.1109/TBME.2003.810689.
[18] M. Winterhalder, B. Schelter, T. Maiwald, A. Brandt, A. Schad, A. Schulze-Bonhage, J. Timmer, Spatio-temporal patient-individual assessment of synchronization changes for epileptic seizure prediction, Clin. Neurophysiol. 117 (2006) 2399–2413, http://dx.doi.org/10.1016/j.clinph.2006.07.312.
[19] B. Krishnan, S. Tousseyn, C.S. Nayak, T. Aung, A. Kheder, Z.I. Wang, G. Wu, J. Gonzalez-Martinez, D. Nair, R. Burgess, L. Iasemidis, I. Najm, J. Bulacio, A.V. Alexopoulos, Neurovascular networks in epilepsy: Correlating ictal blood perfusion with intracranial electrophysiology, Neuroimage 231 (2021) 117838, http://dx.doi.org/10.1016/j.neuroimage.2021.117838.
[20] S. Syed Rafiammal, D. Najumnissa Jamal, S. Kaja Mohideen, Detection of epilepsy seizure in adults using discrete wavelet transform and cluster nearest neighborhood classifier, Iran. J. Sci. Technol. - Trans. Electr. Eng. 6 (2021) http://dx.doi.org/10.1007/s40998-021-00437-6.
[21] M.K. Kiymik, I. Güler, A. Dizibüyük, M. Akin, Comparison of STFT and wavelet transform methods in determining epileptic seizure activity in EEG signals for real-time application, Comput. Biol. Med. 35 (2005) 603–616, http://dx.doi.org/10.1016/j.compbiomed.2004.05.001.
[22] S. Muhammad Usman, S. Khalid, M.H. Aslam, Epileptic seizures prediction using deep learning techniques, IEEE Access 8 (2020) 39998–40007, http://dx.doi.org/10.1109/ACCESS.2020.2976866.
[23] 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, http://dx.doi.org/10.1016/j.neunet.2018.04.018.
[24] N.D. Truong, L. Zhou, O. Kavehei, Semi-supervised seizure prediction with generative adversarial networks, in: Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, 2019, pp. 2369–2372, http://dx.doi.org/10.1109/EMBC.2019.8857755.
[25] P. Shaw, J. Uszkoreit, A. Vaswani, Self-attention with relative position representations, in: NAACL HLT 2018-2018 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. - Proc. Conf., Vol. 2, 2018, pp. 464–468, http://dx.doi.org/10.18653/v1/n18-2074.
[26] Y. Cui, Z. Chen, S. Wei, S. Wang, T. Liu, G. Hu, Attention-over-attention neural networks for reading comprehension, in: ACL 2017-55th Annu. Meet. Assoc. Comput. Linguist. Proc. Conf. (Long Pap. 1), 2017, pp. 593–602, http://dx.doi.org/10.18653/v1/P17-1055.
[27] X. Lu, W. Wang, C. Ma, J. Shen, L. Shao, F. Porikli, See more know more: Unsupervised video object segmentation with co-attention siamese networks, in: Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2019-, 2019, pp. 3618–3627, http://dx.doi.org/10.1109/CVPR.2019.00374.
[28] W. Hu, J. Cao, X. Lai, J. Liu, Mean amplitude spectrum based epileptic state classification for seizure prediction using convolutional neural networks, J. Ambient Intell. Humaniz. Comput. (2019) http://dx.doi.org/10.1007/s12652-019-01220-6.
[29] Y. Zhu, L. Xie, J. Wang, L. Zheng, Attention-based pyramid aggregation network for visual place recognition, in: MM 2018 - Proc. 2018 ACM Multimed. Conf, 2018, pp. 99–107, http://dx.doi.org/10.1145/3240508.3240525.
[30] X. Wu, Z. Du, Y. Guo, H. Fujita, Hierarchical attention based long short-term memory for Chinese lyric generation, Appl. Intell. 49 (2019) 44–52, http://dx.doi.org/10.1007/s10489-018-1206-2.
[31] W. Chan, N. Jaitly, Q. Le, O. Vinyals, Listen, Listen, attend and spell: A neural network for large vocabulary conversational speech recognition, in: ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc. 2016-, 2016, pp. 4960–4964, http://dx.doi.org/10.1109/ICASSP.2016.7472621.
[32] Q. Yao, R. Wang, X. Fan, J. Liu, Y. Li, Multi-class Arrhythmia detection from 12-lead varied-length ECG using Attention-based Time-Incremental Convolutional Neural Network, Inf. Fusion. 53 (2020) 174–182, http://dx.doi.org/10.1016/j.inffus.2019.06.024.
[33] J. Hu, L. Shen, S. Albanie, G. Sun, E. Wu, Squeeze-and-excitation networks, IEEE Trans. Pattern Anal. Mach. Intell. 42 (2020) 2011–2023, http://dx.doi.org/10.1109/TPAMI.2019.2913372.
[34] G. Liu, W. Zhou, M. Geng, Automatic seizure detection based on S-transform and deep convolutional neural network, Int. J. Neural Syst. 30 (2020) 1–15, http://dx.doi.org/10.1142/S0129065719500242.
[35] Z. Yu, W. Nie, W. Zhou, F. Xu, S. Yuan, Y. Leng, Q. Yuan, Epileptic seizure prediction based on local mean decomposition and deep convolutional neural network, J. Supercomput. 76 (2020) 3462–3476, http://dx.doi.org/10.1007/s11227-018-2600-6.
[36] Ö. Yıldırım, U.B. Baloglu, U.R. Acharya, A deep convolutional neural network model for automated identification of abnormal EEG signals, Neural Comput. Appl. 32 (2020) 15857–15868, http://dx.doi.org/10.1007/s00521-018-3889-z.
[37] T. Maiwald, M. Winterhalder, R. Aschenbrenner-Scheibe, H.U. Voss, A. Schulze-Bonhage, J. Timmer, Comparison of three nonlinear seizure prediction methods by means of the seizure prediction characteristic, Phys. D Nonlinear Phenom. 194 (2004) 357–368, http://dx.doi.org/10.1016/j.physd.2004.02.013.
[38] Z. Yang, Q. Xu, S. Member, S. Bao, Learning with Multiclass AUC : Theory and Algorithms, Vol. 8828, 2021.
[39] A.H. Shoeb, Application of Machine Learning To Epileptic Seizure Onset Detection and Treatment, (Ph.D. thesis), 2009, doi:http://dx.doi.org/.
[40] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C. Peng, H.E. Stanley, PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals, Circulation 101 (2000) 15–20, http://dx.doi.org/10.1161/01.CIR.101.23.e215.
[41] X. Yao, Q. Cheng, G.-Q. Zhang, Automated classification of seizures against nonseizures: A deep learning approach, 2019, pp. 1–12, http://arxiv.org/abs/1906.02745.
[42] B. Büyükçakır, F. Elmaz, A.Y. Mutlu, Hilbert vibration decomposition-based epileptic seizure prediction with neural network, Comput. Biol. Med. 119 (2020) 103665, http://dx.doi.org/10.1016/j.compbiomed.2020.103665.
[43] A.S. Zandi, R. Tafreshi, M. Javidan, G.A. Dumont, Predicting epileptic seizures in scalp EEG based on a variational bayesian gaussian mixture model of zero-crossing intervals, IEEE Trans. Biomed. Eng. 60 (2013) 1401–1413, http://dx.doi.org/10.1109/TBME.2012.2237399.
[44] M.H. Myers, A. Padmanabha, G. Hossain, A.L. De Jongh Curry, C.D. Blaha, Seizure prediction and detection via phase and amplitude lock values, Front. Hum. Neurosci. 10 (2016) 1–9, http://dx.doi.org/10.3389/fnhum.2016.00080.
[45] I. Ben Slimen, L. Boubchir, H. Seddik, Epileptic seizure prediction based on EEG spikes detection of ictal-preictal states, J. Biomed. Res. 34 (2020) 162–169, http://dx.doi.org/10.7555/JBR.34.20190097.
[46] F.G. Tang, Y. Liu, Y. Li, Z.W. Peng, A unified multi-level spectral–temporal feature learning framework for patient-specific seizure onset detection in EEG signals, Knowl.-Based Syst. 205 (2020) 106152, http://dx.doi.org/10.1016/j.knosys.2020.106152.
[47] 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. Rehab. Eng. 25 (2017) 1309–1318, http://dx.doi.org/10.1109/TNSRE.2016.2618937.