Channel-Weighted Squeeze-and-Excitation Networks For Epileptic Seizure Detection

Nan Ke[1], Tong Lin[1,2*], Zhouchen Lin[1,3]
[1]*The Key Laboratory of Machine Perception (MOE), School of EECS, Peking University*
[2]*Peng Cheng Laboratory, Shenzhen 518052, China*
[3]*Pazhou Laboratory, Guangzhou 510330, China*
{kenan12860, lintong, zlin}@pku.edu.cn

***Abstract*—Epilepsy is a chronic neurological disorder that affects many people in the world. Automatic epileptic detection based on multi-channel electroencephalogram (EEG) signals is of great significance and has been widely studied. Recent deep learning models fail to consider the weights of different EEG channels since a few channels can play more important roles than other ones. In this paper, we propose an end-to-end epilepsy detection model, CW-SRNet, to solve this problem. We design a novel channel-weighted block (CW-Block) to capture the different importance of EEG channels automatically and dynamically. We combine our novel CW-Block with the squeeze-and-excitation residual network to improve epilepsy detection performance. Experiments on two public EEG datasets show that our model achieves state-of-the-art performance. Particularly on the CHB-MIT dataset, our model achieves an average sensitivity of 96.84% and an average specificity of 99.68%, outperforming other methods with clear margins.**

***Index Terms*—EEG, epilepsy seizure detection, deep learning, multi-channel weights, squeeze-and-excitation, residual networks**

I. INTRODUCTION

Epilepsy is a chronic neurological disorder caused by abnormal discharges of brain nerve cells [1]. According to the World Health Organization Report, around 50 million people worldwide suffer from epilepsy, making it one of the most common neurological diseases globally [2]. Electroencephalography (EEG, as shown in Figure 1) can record epileptic discharges directly and has been widely used in the diagnosis and treatment of epilepsy [3]. The analysis of EEG recordings requires highly professionally trained neurologists, which may take much time and is subject to the experience and proficiency of the neurologists. Therefore, automatic seizure detection is of great significance for clinical application and research [4].

Various machine learning techniques have been proposed in order to realize automatic epileptic seizure detection. The local mean decomposition (LMD) was firstly introduced by Smith [5]. Support vector machine (SVM) has also been used for classification of EEG signals [6]. However, most of these methods require complicated feature engineering based on mathematical or medical knowledge. The dominant frequency, power ratio of the EEG waves and width of the dominant peak are the common hand-crafted features extracted from temporal windows of EEG signals [7] [8]. Those feature engineering practices are often complicated and most algorithms also require precise seizure labels which are time-consuming for doctors to create.

The deep learning models have recently achieved some significant successes in the analysis of EEG signals. The convolutional neural networks (CNNs) have been widely used in the medical image analysis [9]. Compared to traditional artificial features, feature learning by deep CNNs obviously improves the performance of seizure detection [10]. However, many of the proposed CNN models do not explicitly consider the multi-channel characteristics of EEG data, and there is still much room for improvement in epileptic seizure detection.

In this paper, we propose a novel epileptic seizure detection model based on the deep residual networks and squeeze-and-excitation networks. The well-known residual networks using the shortcut can solve the problem of reduced accuracy in deeper layer to a certain extent [11]. The squeeze-and-excitation networks using SE blocks can capture the relationship between channels by explicitly modeling the interdependencies between the channels of its convolutional features [12]. Meanwhile, we design a novel channel-weighted block (CW-Block) to weight the multi-channel EEG data. Finally, we achieve state-of-the-art results on the Bonn dataset and the CHB-MIT dataset. The contributions of this paper are listed as follow:

- We propose a novel channel-weighted block (CW-Block) to automatically weight multi-channel EEG signals. Our experiments prove that this block can effectively capture EEG channel weights and improve the ability of the model to detect epilepsy.
- We propose an end-to-end seizure detection model, CW-SRNet. This model, designed for the multi-channel characteristics of EEG, uses the CW-blocks and SE blocks to improve the seizure detection performance.
- Experiments on two public EEG datasets show that our method achieves state-of-the-art results. Particularly on the CHB-MIT dataset, our model achieves an average sensitivity of 96.84%, which is three percentage points higher than previous SOTA models.

This work was supported by NSFC Tianyuan Fund for Mathematics (No. 12026606), and National Key R&D Program of China (No. 2018AAA0100300). * Tong Lin is the corresponding author.

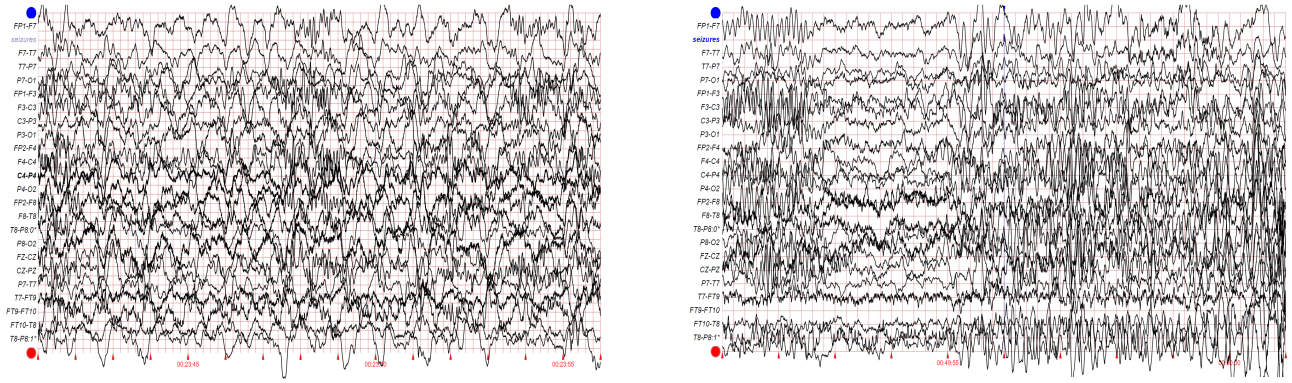

Fig. 1. The waveform graphs sampled from the CHB-MIT EEG dataset. the left figure is the normal period, and the right is the epileptic seizure period. There are 23 channels of this EEG signals.

The article is composed as follows. The Section II discusses the relevant work. The Section III elaborates our proposed detection model. We then describe the experimental setup and present the results in Section IV. We give further discussion in Section V before conclusion in Section VI.

II. RELATED WORK

In this section, we review some representative work for seizure detection.

A. Feature Engineering Methods

Early automated seizure detection relied on heuristic rules and threshold settings of human experts [13]. Following these heuristic methods, machine learning algorithms have been widely used in studies to detect seizures [14]. To convert the raw multi-channel EEG signals into the model input, many feature engineering methods based on time domain and frequency domain have been proposed. The short-time Fourier transform (STFT) was used by Kiymik [15] to decompose EEG signals into time-frequency representations. The wavelet transform (WT) also became an important tool for time-frequency decomposition of EEG signals [16]. Zhang and Parhi proposed the spectral power analysis method [17]. Ghosh-Dastidar combined the discrete wavelet transform (DWT) with nonlinear dynamics to detect seizures [18]. By decomposing a signal into a series of finite modulo functions, the empirical mode decomposition (EMD) has been successfully applied to medical signal analysis and is suitable for EEG analysis [19]. The local mean decomposition (LMD), firstly introduced by Smith [5], aims at decomposing an original EEG signal into a series products of frequency modulated signals and an envelope signal [20].

B. Deep Learning Methods

Recently, deep learning technologies have been applied to automatic feature learning for raw high-dimensional EEG data [21]. By constructing multi-layer neural networks, deep learning is able to capture relevant features during end-to-end training. Convolutional neural networks have been widely used in EEG signal processing [22] after making great achievements in visual tasks. Acharya [23] proposed a deep CNN consisting of 13 layers for automatic seizure detection. Wei [10] applied a 12 layers CNN with Wasserstein Generative Adversarial Nets (WGANs) for data enhancement. Abdelhameed [24] combined a one-dimensional CNN with a bi-directional long short-term memory (Bi-LSTM). Hu [25] used a Bi-LSTM with some time-frequence features. Tanveer [26] employed ensemble CNNs to enhance the predictive performance of the model.

III. METHODOLOGY

In this section, we introduce the structure of our proposed seizure detection model, CW-SRNet, which is mainly composed of the CW-Block and the SE-ResNet module.

A. CW-Block

During epileptic seizures, mostly only a few channels have obvious signal changes. So it is necessary to consider the importance of different channels. In order to adaptively capture the importance of channels, we design the channel-weighted block (CW-Block). After a simple pre-processing, the dimension of raw input EEG data is $T \times C$, where T is the total time of each channel and C is the number of EEG channels. Firstly, we use the 1D convolution to capture the global signal representation, which maps the input into a segment with size of $S \times C$. Then we put the segment into two-layer neural network and get the output after the softmax layer. Average pooling is used for every channel to get the final channel weights vector with the size of $1 \times C$. Finally, we multiply the data of each channel with its corresponding weights. This process is formulated as follows:

$$X_1 = \text{1D-Convolution}(Input), \tag{1}$$

$$X_2 = \text{Softmax}(FC_2(FC_1(X_1))), \tag{2}$$

$$X_3 = \text{Average Pooling}(X_2), \tag{3}$$

$$X_4 = X_3 \times Input, \tag{4}$$

where FC_i is a fully connected network and X_3 are channel weights. The procedure (1)-(4) forms a channel-weighted

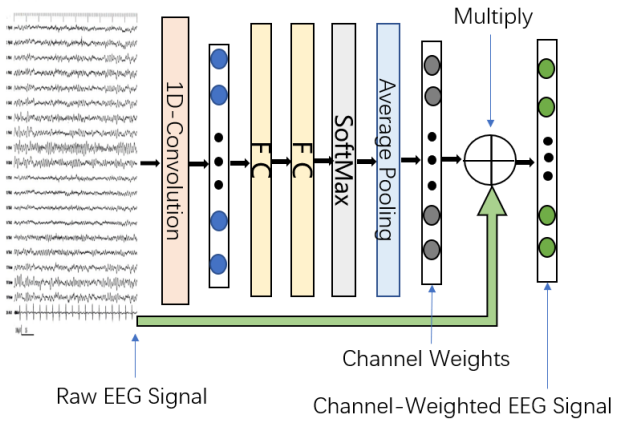

Fig. 2. The Structure of a CW-Block

block (CW-Block) and the structure of a CW-Block is shown in Figure 2. In Section IV and Section V we can see that this block can effectively assign the different channel weights and improve the performance of seizure detection models.

B. SE-ResNet module

This module is mainly based on deep residual networks [11] and squeeze-and-excitation networks [12]. The SE (squeeze-and-excitation) block is shown in Figure 3.

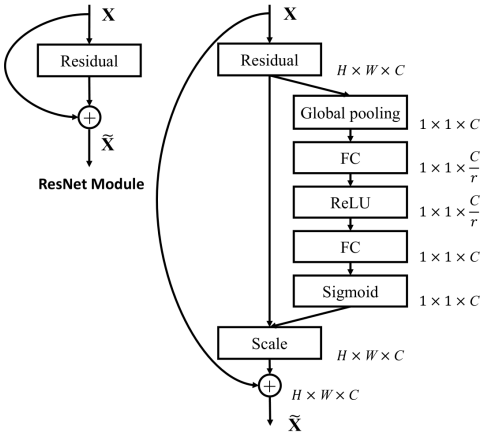

Fig. 3. The Structure of a SE Block

This SE block is designed to explicitly model the interdependencies between the channels of its convolutional features. We get SE-ResNet when we add SE blocks to the residual blocks in ResNet. To better deal with the multiple channels of EEG data , we replace part of the SE blocks with CW-Blocks. Different from the CW-Block as mentioned above, this 3D CW-Block operates on 3D data here. For feature maps with dimension $T \times C \times M$, where M is the number of feature maps, we let CW-Blocks only work on dimension C. Correspondingly, SE blocks only work on dimension M. In our method, 18-layer residual network is used as the backbone

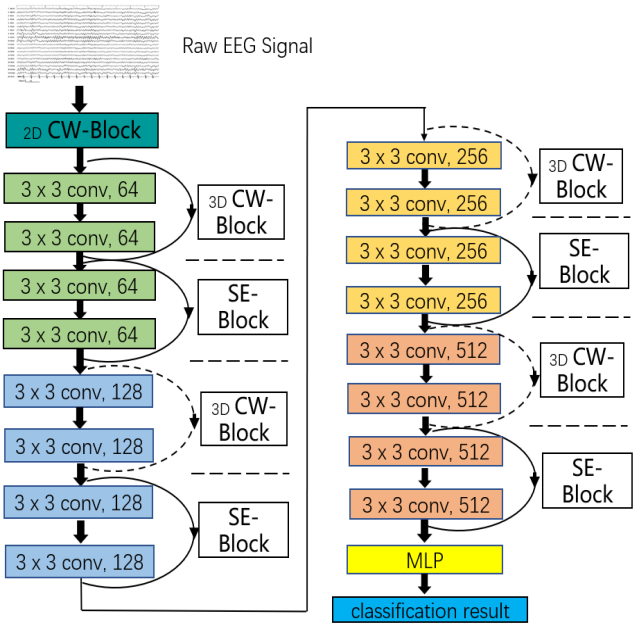

Fig. 4. The Architecture of Our CW-SRNet Model

model. Our proposed method alternately applies the CW-Block and the SE block to get channel weights in different dimensions. The layout of the CW-SRNet is given in Figure 4.

C. Focal Loss Optimization

The class imbalance problem in the epilepsy datasets is serious. The seizure EEG data is acquired through long duration recording, and the subject is in a non-seizure state for most of the time. So there are much more interictal segments than ictal segments. For majority of subjects, the ratio between seizure samples and non-seizure samples is even less than 0.1. This class imbalance will make the model prone to classify the segments as non-seizure. We use the focal loss to alleviate this problem.

The focal loss function was firstly proposed by [27], which is to address the problem of extreme imbalance between foreground and background classes during training in object detection scenarios. Focal loss is mainly used for object detection, but it is also applicable for classification problem with imbalanced EEG datasets. The focal loss is defined as follows:

$$\mathrm{FL}(p_t) = -\alpha_t(1-p_t)^{\gamma}\log(p_t). \tag{5}$$

where

$$p_t = \begin{cases} p & \text{if y = 1,} \\ 1-p & \text{otherwise.} \end{cases}$$

and

$$\alpha_t = \begin{cases} \alpha & \text{if y = 1,} \\ 1-\alpha & \text{otherwise.} \end{cases}$$

Here y denotes the ground truth of a sample.

The focal loss function has two parameters α and γ. Lin et al. [27] used a 1:3 imbalanced dataset, and tested different α and γ. They found that the best α is 0.25 and the best γ is 2. However, most of the subjects that we used are extremely imbalanced than those in [27]. Therefore, the values used in [27] may not be applicable here. In this work, we propose to apply this focal loss function to the end of our proposed architecture and find the optimal parameters through experiments.

TABLE I
Data Information of the CHB-MIT Dataset

Patient	Age	Gender	Number of seizures	Total seizure time (s)	Total non-seizure time (hh:mm:ss)
1	F	11	7	474	37:31:42
2	M	11	2	172	34:16:34
3	F	14	7	402	33:03:13
4	M	22	2	276	59:27:32
5	F	7	5	558	35:57:44
6	F	1.5	10	104	58:45:12
7	F	14.5	3	325	62:01:26
8	M	3.5	5	919	17:57:44
9	F	10	3	197	64:14:10
10	M	3	7	447	45:02:22
11	F	12	3	806	31:36:11
20	F	6	8	229	24:05:08
21	F	13	4	199	31:01:02
22	F	9	3	204	31:00:11
23	F	6	7	244	24:35:14
Total	-	-	76	5556	590:35:25

IV. EXPERIMENTS

In this section, we compare our CW-SRNet model with other state-of-the-art methods for seizure detection.

A. EEG Datasets

a) Bonn EEG Dataset: This dataset was collected by Andrzejak [28] at Bonn University, Germany. It consists of two sets of surface EEG segments from healthy volunteers with and without eyes open (subsets A and B) and three sets of intracranial EEG recordings from epilepsy patients during the seizure-free period (subsets C and D) as well as during seizures (subset E). Each subset contains 100 single-channel EEG segments of 23.6 seconds duration with 4096 sampling indexes. The sampling rate and resolution are 173.61 Hz and 12 bits respectively [29].

b) CHB-MIT Dataset: The CHB-MIT EEG dataset was gathered at the Children's Hospital Boston [30]. This dataset consists of long-duration multi-channel EEG recordings from 23 pediatric patients with intractable seizures. The original scalp EEG recordings were obtained using 256 Hz sampling rate with a 16 bit resolution from electrodes. Those electrodes were placed according to the international 10–20 system of EEG electrode positions and nomenclature [31]. Because of the data disorder in some patients, we only use the data of fifteen patients. Channels 15 and 23 present the same bipolar combination, so channel 23 is discarded to avoid the redundant information. Details about the CHB-MIT dataset are described in Table I.

B. Setup

Some experimental paradigms in the Bonn dataset, such as single-channel EEG with tags assigned to long-term EEG fragments, are replaced by multi-channel recording and frame-by-frame detection [32]. In this study, we design two kinds of different cases as follows [33]. In Case I, subsets A, B, C, and D are grouped together as the normal class whereas the set E is considered as seizure class. In Case II, subsets A and B, subsets C and D, and subset E belong to healthy subjects, patients with interictal epilepsy, and patients with epilepsy during epilepsy, respectively.

For the CHB-MIT dataset, we only do simple preprocessing instead of complicated feature engineering. Firstly, we use downsampling to reduce the 256Hz EEG signal sampling frequency to 64Hz, which can reduce data redundancy and memory footprint. The downsampling can also speed up the training process. Then we split the continuous EEG into many two seconds segments. By means of the dataset seizure annotation, we can easily distinguish between interictal and ictal phase. We use a one-second sliding window to get seizure segments in the ictal phase. Because most of the time there is no seizure, there are much more non-seizure segments than seizure ones. So we have to randomly sample from interictal phase to get a certain amount of non-seizure sample. The ratio of negative and positive samples in our experiment is six.

All of the operations of these experiments were performed in the environment of PYTHON 3.6.1 and PYTORCH 1.5.0, running on an Intel core processor with a frequency of 3.40 GHz using an Nvidia TITAN RTX GPU.

C. Measurements

In our experiments, five statistical indicators are used for the performance evaluation of the proposed method. Some indicators are defined as follows:

$$\text{Sensitivity} = \frac{TP}{TP+FN}, \tag{6}$$

$$\text{Specificity} = \frac{TN}{TN+FP}, \tag{7}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}, \tag{8}$$

$$\text{Precision} = \frac{TP}{TP+FP}, \tag{9}$$

$$\text{F1-score} = \frac{2 \cdot \text{Sensitivity} \cdot \text{Precision}}{\text{Sensitivity} + \text{Precision}} \tag{10}$$

where TP (true positive) represents the number of cases correctly identified as seizure; FP (false positive) represents the number of cases incorrectly identified as seizure; TN (true negative) represents the number of cases correctly identified as non-seizure; and FN (false negative) represents the number of cases incorrectly identified as non-seizure.

We also use ROC and AUC for performance evaluation. The ROC Curve measures how accurately the model can distinguish between two classes, and AUC measures the entire area underneath the ROC curve.

TABLE II
Results on the Bonn Dataset.

Method	Case I				Case II			
	Accuracy(%)	F1-score(%)	Sen(%)	Spe(%)	Accuracy(%)	F1-score(%)	Sen(%)	Spe(%)
EMD+SVM [36]	0.9812	0.9658	-	-	0.9362	0.9273	-	-
CWT+SVM [37]	0.9910	0.9756	-	-	0.9522	0.9517	-	-
Deep ConvNet [4]	0.9874	0.9627	-	-	0.9285	0.9132	-	-
Statistic+DNN [34]	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Bi-LSTM+GAP [35]	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
CW-SRNet (Ours)	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

[1] Here Sen means Sensitivity and Spe means Specificity.

D. Results

a) Results on the Bonn Dataset: For the Bonn Dataset, results are shown in Table II. The empirical mode decomposition (EMD) based analysis and the continuous wavelet transform (CWT) are commonly used baseline methods. Deep ConvNet is a deep learning model designed for the EEG decoding. In the case I and case II, our proposed model achieves the best result with accuracy 1.0 and F1-score 1.0, the same as previous models in [34] and [35].

TABLE III
Results on the CHB-MIT Dataset for the Single Patient

patient	Sensitivity(%)	Specificity(%)	Accuracy(%)
1	98.95	99.58	99.53
2	93.55	99.80	99.61
3	96.39	99.79	99.53
4	97.14	99.80	99.62
5	100	99.38	99.45
6	87.50	100	99.71
7	96.00	99.90	99.62
8	92.86	98.98	98.09
9	94.23	99.90	99.61
10	100	99.90	99.91
11	99.38	99.40	99.39
20	98.04	99.59	99.51
21	100	99.80	99.80
22	100	99.80	99.80
23	98.70	99.59	99.53
Average	**96.84**	**99.68**	**99.51**

b) Results on the CHB-MIT Dataset: For the CHB-MIT Dataset, our results are shown in Table III. Almost for all patients, the specificity and accuracy are higher than 99%. This is partly because of the imbalanced data, which makes the model more inclined to predict negative samples. Our model achieves an average sensitivity of 96.84%. Table IV shows the result of other methods. We can see that our model achieves state-of-the-art performance.

V. DISCUSSIONS

A. Effectiveness of CW-Block

To further prove the effectiveness of our novel CW-Block, we do more experiments on the CHB-MIT Dataset. Different from the experiment for single patient presented above, we mix the data of all patients together. Because different patients may have different brain wave activity patterns and seizure patterns, this mixed experiments are more difficult. The results are shown in Table V. We choose two layers CNN (CNN2), four layers CNN (CNN4) and ResNet18 as references. As shown in Table V, CNN2+CW achieves a sensitivity of 83.19% compared with 78.43% of CNN2. CNN4+CW achieves a sensitivity of 85.71% compared with 81.38% of CNN4. ResNet18+CW achieves a sensitivity of 96.08% compared with 95.20% of ResNet18. And our proposed model gets the best result with a sensitivity of 97.90%, a specificity of 98.59%, an accuracy of 98.40% and AUC of 99.60%. This experiment successfully proves the effectiveness of the CW-Block.

B. Effectiveness of the Focal Loss

In the definition of the focal loss function, α and γ are two key parameters. According to the origin focal loss [27], the optimal parameter values are $\alpha = 0.5$ and $\gamma = 2$. We did further experiments to find the optimal parameters. The results are shown in Table VI. We can see that the model performance fluctuates with changes in parameters. When $\alpha = 0.75$ and $\gamma = 1$ we can get the best result with a sensitivity of 97.90%, a specificity of 98.59%, an accuracy of 98.40% and AUC of 99.60%, higher than the binary cross entropy loss (BCE Loss) result.

C. Visualization of Channel Weights

In order to see the role of different EEG channels more intuitively, we visualized the EEG signals and channel weights. We selected the third patient from the CHB-MIT dataset as an illustration. Figure 5 is the normal EEG and Figure 6 shows the epileptic seizure waveform after high-pass filtering with a frequency of 10. Notice that the middle section of the waveform changes drastically because of seizures in Figure 6 and only parts of channels have significant signal changes. On the right side of the Figure 6 we show the corresponding channel weights obtained by our model. We can see that channels more related to epileptic seizures are given higher weights, and irrelevant channels have lower weights. We can understand the role of CW-Block in this way, that is, by lowering the weight of channels that are not relevant to epilepsy, we can eliminate interference signals and get more accurate detection results.

VI. CONCLUSION

In this paper, we propose a seizure detection model, CW-SRNet. To effectively use multi-channel signals, we design

TABLE IV
Results of Methods in Comparison on the CHB-MIT dataset

Method	Subjects	Duration (h)	Sensitivity	Specificity	Accuracy
Dyadic WT+LDA [38]	18	152.80	92.60	99.90	-
Discrete WT [39]	18	76.92	91.71	92.89	92.30
CNN+MIDS [10]	23	-	74.08	92.46	-
Bi-LSTM [25]	24	877.39	93.61	91.85	-
CE-STNet [33]	21	-	92.41	96.05	95.96
CW-SRNet (Ours)	15	590.58	**96.84**	**99.68**	**99.51**

TABLE V
Experiment Results of CW-Block on the Mixed CHB-MIT Dateset

Method	Sensitivity(%)	Specificity(%)	Accuracy(%)	AUC(%)
CNN2	78.43	92.19	88.37	93.04
CNN2+CW	**83.19**	**92.98**	**89.70**	**94.47**
CNN4	81.38	91.27	88.57	93.49
CNN4+CW	**85.71**	**91.15**	**89.41**	**94.98**
ResNet18	95.20	98.14	97.07	99.45
ResNet18+CW	**96.08**	**98.32**	**97.26**	**99.50**
CW-SRNet	**97.90**	**98.59**	**98.40**	**99.60**

[1] Here CW refers to the CW-Block we proposed .

TABLE VI
Influence of the Values of (α, γ) in the Focal Loss

(α,γ)	Sensitivity(%)	Specificity(%)	Accuracy(%)	AUC(%)
(0.05, 1)	86.13	99.23	95.59	99.36
(0.05, 2)	88.26	98.77	95.93	99.38
(0.05, 5)	90.95	98.67	96.58	99.49
(0.1, 1)	92.58	98.36	96.75	99.54
(0.1, 2)	90.23	98.27	96.10	99.07
(0.1, 5)	92.56	98.94	97.21	99.49
(0.25, 1)	86.39	99.40	95.78	99.53
(0.25, 2)	93.55	98.37	97.07	99.48
(0.25, 5)	93.55	98.64	97.26	99.53
(0.5, 1)	93.54	98.19	96.90	99.46
(0.5, 2)	**96.86**	**97.38**	**97.24**	**99.55**
(0.5, 5)	93.10	98.54	97.07	99.47
(0.75, 1)	**97.90**	**98.59**	**98.40**	**99.60**
(0.75, 2)	**96.68**	**96.28**	**96.39**	**99.42**
(0.75, 5)	0.9346	98.01	96.78	99.32
BCE Loss	**96.01**	**97.63**	**97.19**	**99.48**

a novel channel-weighted block (CW-Block) to automatically capture the EEG channel importance. The multi-channel raw EEG signals are sent to the first 2D CW-Block to get the weights for each EEG channel. Then we alternately add SE Blocks and 3D CW-Blocks to the 18-layer residual network. Finally the model outputs the epilepsy detection results. The classification performance has been evaluated on two public seizure datasets. Our proposed method achieves an accuracy of 100% on the Bonn dataset and achieves an average sensitivity of 96.84%, an average specificity of 99.68%, an average accuracy of 99.51% on the CHB-MIT dataset, which are the state-of-the-art results. These experiments convincingly prove the effectiveness of our proposed epilepsy detection model.

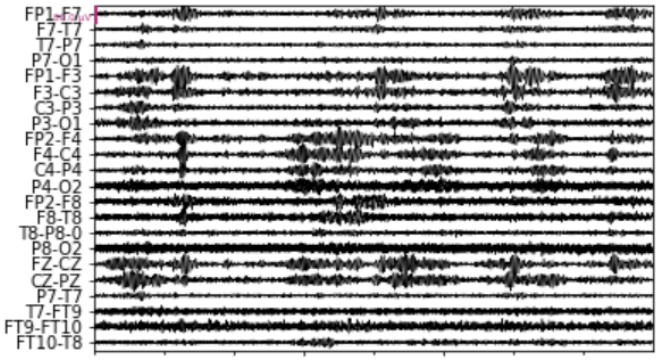

Fig. 5. An example of interictal EEG waveform from the CHB-MIT dataset after high-pass filtering with a frequency of 10.

REFERENCES

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

[2] World Health Organization, "A report about epilepsy," 2016, Available at: https://www.who.int/news-room/fact-sheets/detail/epilepsy.

[3] M. Ahmad, M. Saeed, S. Saleem, and A. M. Kamboh, "Seizure detection using EEG: A survey of different techniques," in *2016 International Conference on Emerging Technologies (ICET)*, pp. 1–6, 2016.

[4] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Human Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.

[5] J. S. Smith, "The local mean decomposition and its application to EEG perception data," *Journal of the Royal Society Interface*, vol. 2, no. 5, pp. 443–454, 2005.

[6] J. Jin, X. Wang, and B. Wang, "Classification of direction perception EEG based on PCA-SVM," in *Third International Conference on Natural Computation (ICNC 2007)*, vol. 2, pp. 116–120, 2007.

[7] P. Celka and P. Colditz, "A computer-aided detection of EEG seizures in infants: a singular-spectrum approach and performance comparison," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 5, pp. 455–462, 2002.

[8] J. Gotman, D. Flanagan, J. Zhang, and B. Rosenblatt, "Automatic seizure detection in the newborn: methods and initial evaluation," *Electroencephalography and Clinical Neurophysiology*, vol. 103, no. 3, pp. 356–362, 1997.

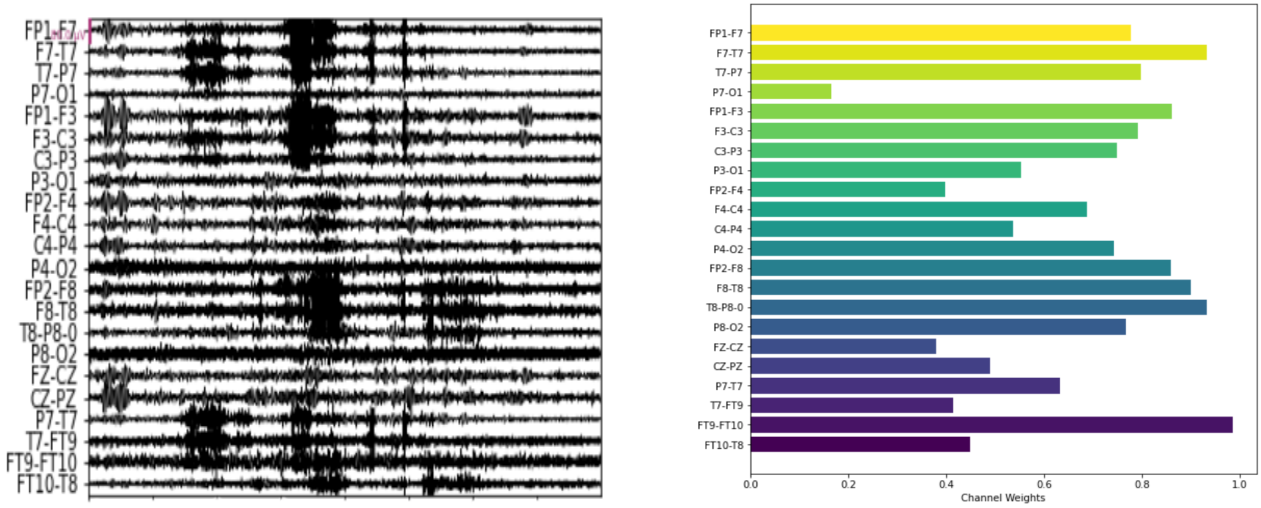

Fig. 6. The figure on the left is an example of epileptic seizure waveform from the CHB-MIT dataset after high-pass filtering with a frequency of 10, and the figure on the right is the corresponding channel weights obtained by our model.

[9] S. S. Yadav and S. M. Jadhav, "Deep convolutional neural network based medical image classification for disease diagnosis," *Journal of Big Data*, vol. 6, no. 1, pp. 1–18, 2019.

[10] Z. Wei, J. Zou, J. Zhang, and J. Xu, "Automatic epileptic EEG detection using convolutional neural network with improvements in time-domain," *Biomedical Signal Processing and Control*, vol. 53, p. 101551, 2019.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.

[12] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7132–7141, 2018.

[13] A. Liu, J. Hahn, G. Heldt, and R. Coen, "Detection of neonatal seizures through computerized EEG analysis," *Electroencephalography and Clinical Neurophysiology*, vol. 82, no. 1, pp. 30–37, 1992.

[14] B. Boashash and S. Ouelha, "Automatic signal abnormality detection using time-frequency features and machine learning: A newborn EEG seizure case study," *Knowledge-Based Systems*, vol. 106, pp. 38–50, 2016.

[15] M. Kıymık, İnan Güler, A. Dizibüyük, and M. Akın, "Comparison of stft and wavelet transform methods in determining epileptic seizure activity in EEG signals for real-time application," *Computers in Biology and Medicine*, vol. 35, no. 7, pp. 603–616, 2005.

[16] K. Gadhoumi, J.-M. Lina, and J. Gotman, "Discriminating preictal and interictal states in patients with temporal lobe epilepsy using wavelet analysis of intracerebral EEG," *Clinical Neurophysiology*, vol. 123, no. 10, pp. 1906–1916, 2012.

[17] Z. Zhang and K. K. Parhi, "Low-complexity seizure prediction from iEEG/sEEG using spectral power and ratios of spectral power," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 10, no. 3, pp. 693–706, 2015.

[18] S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, "Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 9, pp. 1545–1551, 2007.

[19] R. J. Martis, U. R. Acharya, J. H. Tan, A. Petznick, R. Yanti, C. K. Chua, E. K. Ng, and L. Tong, "Application of empirical mode decomposition (EMD) for automated detection of epilepsy using EEG signals," *International Journal of Neural Systems*, vol. 22, no. 06, p. 1250027, 2012.

[20] Z. Yu, W. Nie, W. Zhou, F. Xu, S. Yuan, Y. Leng, and Q. Yuan, "Epileptic seizure prediction based on local mean decomposition and deep convolutional neural network," *The Journal of Supercomputing*, vol. 76, no. 5, pp. 3462–3476, 2020.

[21] A. Antoniades, L. Spyrou, C. C. Took, and S. Sanei, "Deep learning for epileptic intracranial EEG data," in *2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)*, pp. 1–6, 2016.

[22] P. Thodoroff, J. Pineau, and A. Lim, "Learning robust features using deep learning for automatic seizure detection," in *Machine Learning for Healthcare Conference*, pp. 178–190, 2016.

[23] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals," *Computers in Biology and Medicine*, vol. 100, pp. 270–278, 2018.

[24] A. M. Abdelhameed, H. G. Daoud, and M. Bayoumi, "Epileptic seizure detection using deep convolutional autoencoder," in *2018 IEEE International Workshop on Signal Processing Systems (SiPS)*, pp. 223–228, 2018.

[25] X. Hu, S. Yuan, F. Xu, Y. Leng, K. Yuan, and Q. Yuan, "Scalp EEG classification using deep Bi-LSTM network for seizure detection," *Computers in Biology and Medicine*, vol. 124, p. 103919, 2020.

[26] M. A. Tanveer, M. J. Khan, H. Sajid, and N. Naseer, "Convolutional neural networks ensemble model for neonatal seizure detection," *Journal of Neuroscience Methods*, vol. 358, p. 109197, 2021.

[27] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," *IEEE International Conference on Computer Vision*, 2017, pp. 2980–2988.

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

[29] A. Shoeibi, N. Ghassemi, R. Alizadehsani, M. Rouhani, H. Hosseini-Nejad, A. Khosravi, M. Panahiazar, and S. Nahavandi, "A comprehensive comparison of handcrafted features and convolutional autoencoders for epileptic seizures detection in EEG signals," *Expert Systems with Applications*, vol. 163, pp. 113788, 2021.

[30] A. H. Shoeb, "Application of machine learning to epileptic seizure onset detection and treatment," Ph.D. dissertation, Massachusetts Institute of Technology, 2009.

[31] S. Janjarasjitt, "Epileptic seizure classifications of single-channel scalp EEG data using wavelet-based features and svm," *Medical & Biological Engineering & Computing*, vol. 55, no. 10, pp. 1743–1761, 2017.

[32] Y. Yuan, G. Xun, K. Jia, and A. Zhang, "A multi-view deep learning framework for EEG seizure detection," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 1, pp. 83–94, 2018.

[33] Y. Li, Y. Liu, W.-G. Cui, Y.-Z. Guo, H. Huang, and Z.-Y. Hu, "Epileptic seizure detection in EEG signals using a unified temporal-spectral

squeeze-and-excitation network," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 4, pp. 782–794, 2020.

[34] R. Sharma, R. B. Pachori, and P. Sircar, "Seizures classification based on higher order statistics and deep neural network," *Biomedical Signal Processing and Control*, vol. 59, p. 101921, 2020.

[35] D. Thara, B. Premasudha, R. S. Nayak, T. Murthy, G. A. Prabhu, and N. Hanoon, "Electroencephalogram for epileptic seizure detection using stacked bidirectional lstm_gap neural network," *Evolutionary Intelligence*, vol. 14, no. 2, pp. 823–833, 2021.

[36] F. Riaz, A. Hassan, S. Rehman, I. K. Niazi, and K. Dremstrup, "EMD-based temporal and spectral features for the classification of EEG signals using supervised learning," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 1, pp. 28–35, 2015.

[37] Y. Li, W. Cui, M. Luo, K. Li, and L. Wang, "Epileptic seizure detection based on time-frequency images of EEG signals using gaussian mixture model and gray level co-occurrence matrix features," *International Journal of Neural Systems*, vol. 28, no. 07, p. 1850003, 2018.

[38] L. Orosco, A. G. Correa, P. Diez, and E. Laciar, "Patient non-specific algorithm for seizures detection in scalp EEG," *Computers in Biology and Medicine*, vol. 71, pp. 128–134, 2016.

[39] D. Chen, S. Wan, J. Xiang, and F. S. Bao, "A high-performance seizure detection algorithm based on discrete wavelet transform (DWT) and EEG," *PloS One*, vol. 12, no. 3, 2017.