Available at www.sciencedirect.com
ScienceDirect
journal homepage: www.elsevier.com/locate/bbe

Original Research Article

Generative adversarial network and convolutional neural network-based EEG imbalanced classification model for seizure detection

Bin Gao, Jiazheng Zhou, Yuying Yang, Jinxin Chi, Qi Yuan *

Shandong Province Key Laboratory of Medical Physics and Image Processing Technology, School of Physics and Electronics, Shandong Normal University, Jinan, China

ARTICLE INFO

Article history:
Received 19 June 2021
Received in revised form
21 September 2021
Accepted 16 November 2021
Available online 27 November 2021

Keywords:
Scalp EEG
Seizure detection
Generative adversarial network
1DCNN
Imbalanced classification

ABSTRACT

Automatic seizure detection technology is of great significance to reduce workloads of neurologists for epilepsy diagnosis and treatments. Imbalanced classification is a challenge in seizure detection from long-term continuous EEG recordings, as the durations of the seizure events are much shorter than the non-seizure periods. An imbalanced deep learning model is proposed in this paper to improve the performance of seizure detection. To modify imbalanced EEG data distribution, a generative adversarial network (GAN) that is a strong candidate for data enhancement is built to produce the seizure-period EEG data used for forming a more balanced training set. Next, a pyramidal one-dimensional convolutional neural network (1DCNN) is designed to deal with 1D EEG signals and trained on the augmented training set that consists of both original and generated EEG data. Compared to the conventional 2DCNNs, the deep architecture of the 1DCNN reduces the training parameters so as to greatly increase the training speed. The proposed method is evaluated on three publicly available EEG databases. After data augmentation by the GAN, the designed 1DCNN shows much better classification for seizure detection, achieving competitive results over the three EEG databases, which demonstrates the generalizability of this method across different databases. Comparison with other published methods indicates its enhanced detection performance for imbalanced EEG data.

* Corresponding author at: School of Physics and Electronics, Shandong Normal University, University Science and Technology Park Road on the 1st, Changqing District, Ji'nan, Shandong 250358, China.
E-mail address: yuanqi@sdnu.edu.cn (Q. Yuan).
https://doi.org/10.1016/j.bbe.2021.11.002

1. Introduction

Epilepsy is a chronic and recurrent neurological disorder caused by the sudden and excessive discharge of neurons in the brain [1,2]. There are more than 50 million people with epilepsy worldwide, 25 percent of whom have no effective surgical or drug treatment [3,4]. Epilepsy can cause not only loss of consciousness, but also disruptions in feelings, emotions, or mental functions [5,6]. The electroencephalogram (EEG) is an important clinical measurement tool that can record the electrical activities of brain neurons; it can be obtained by scalp electrodes on the surface of the head [7–10]. The EEG has been widely used in the clinical diagnosis and treatment of epilepsy. Traditional seizure detection requires visual analysis of EEG signals by a professionally trained neurologist [11–13]. However, with the increase in the volume of EEG data, visual examination has become a time-consuming and laborious task with an accuracy of detection that is heavily dependent on the experience and proficiency of the neurologist [14–16]. Therefore, a reliable automatic epileptic seizure detection technology would be of great significance for clinical application and research [17,18].

Automatic seizure detection technology has aroused the interest of numerous researchers and professionals, and many automatic seizure detection technologies have been developed. The earliest epileptic seizure detection technique was developed by Gotman et al. [19], who used intuitional characteristics such as peak amplitude, duration, slope, and sharpness to detect seizures. In recent years, with the rapid increase in processing speed of electronic devices and development of artificial intelligence, deep learning algorithms have attracted extensive attention from researchers all over the world. Deep learning algorithms are easy to execute, have excellent performance in data processing, and show great potential in the field of data processing and pattern recognition, so they have also been used in the field of seizure detection. A deep network model for multi-channel EEG signals designed by Park et al. extracted features from the temporal and spatial levels and achieved 90.50% and 85.60% accuracy rates on the two data sets [20]. Emami et al. applied CNN to long-term EEG including epileptic seizures, making the median rate of epileptic seizures detected by CNN in practical application 100% by minute, higher than 73.3% of BESA and 81.7% of Persyst [21]. Liang et al. proposed a Long-term Recurrent Convolutional Networks (LRCN) model which obtained a sensitivity of 84%, a specificity of 99% and a geometric mean (G-mean) of 91.19% [22]. The method proposed by Daoud H et al., which extracted important spatial features of different locations by convolutional neural network to predict the occurrence of epileptic seizures in advance, achieved the highest accuracy rate of 99.6%, the lowest false alarm rate of 0.004 h^{-1} and the very early prediction time of epileptic seizures of 1 h [23]. Ahmedt-Aristizabal et al. used the Long-Short Term Memory (LSTM) network to directly extract feature information from the original EEG data and accomplish seizure detection, achieving an accuracy rate of 95.54% [24]. Xu YK et al. proposed a CNN end-to-end deep learning solution, and the accuracy of sensI is 93.5%, the false prediction rate is 0.063/h, and the false prediction rate is 0.981 on two data sets of Kaggle intracranial and CHB-MIT scalp EEG [25]. Truong et al. used GAN to train in the CHB-MIT database in an unsupervised manner that ignored the seizure information and achieved an out-of-sample test area (AUC) of 77.68% under the operating characteristic curve [26]. In addition, the deep learning was combined with other signal processing techniques, such as wavelet decomposition [27], merger of the increasing and decreasing sequences (MIDS) [28], and feature scaling [29], to improve the performance of seizure detection. Takahashi et al. combined CNN processing EEG images with patient-specific EEG signal autoencoder (AE), and the results showed that the median false positive rate of AE-CNN was reduced to 0.034 h^{-1}, which was one fifth of the original CNN (0.17 h^{-1}) [30]. In this work, we designed a pyramidal one-dimensional convolutional neural network (1DCNN) which not only is suitable for dealing with 1D EEG signals but also can reduce the training parameters and thereby increase the training speed compared with the conventional 2DCNNs.

Epileptic EEG signals are a typical imbalanced data set because the duration of seizures is much shorter than that of non-seizures. Ward et al. proposed the DBBE-CLSTM algorithm as a high-performance algorithm for binary unbalanced time series classification tasks [31]. Yuan et al. proposed a weighted ELM algorithm for epileptic seizure detection with unbalanced EEG data distribution [12]. Birjandtalab et al. proposed an unbalanced learning method to improve the accuracy of highly unbalanced captured datasets [32]. Masum et al. studied the classification performance using a variety of ML methods under the influence of different balance ratios [33]. Alzaid compared the performance of ML model with the classical feature engineering method and RNN in seizure detection of unbalanced data sets [34].

In this work, a generative adversarial network (GAN) was built to resolve this imbalance problem. Generative adversarial networks (GANs) are strong candidates for data enhancement [35]. The potential of GANs in learning the underlying distribution of data allows the production of an unlimited number of real samples [36]. In addition to its compelling performance in synthetic image generation, recent studies have attempted to synthesize sequences or advanced feature samples for a variety of applications in the medical field [37–39]. In the work of Qin et al., a conditional generative adversarial network (CGAN) was used to generate high-quality virtual samples and rich data sets, which mitigated the over-fitting phenomenon and improved the performance of the classification model [40]. Zhou et al. proposed a high-performance automatic ECG classification system based on GANs that showed superior overall performance [41]. Liu et al. employed the Deep Convolution Generative Adversarial Networks (DCGAN) model to the CelebFaces Attributes Dataset (CelebA) under constraint and unconstraint conditions [42]. In the work of Chen et al., the Discriminant Metric-based Generative Adversarial Networks (DMGANS) were established to generate similar real samples from the perspective of depth metric learning [43]. Zhong et al. proposed Decoder-Encoder Generative Adversarial Networks (DEGANS) that could quickly learn and generate high quality images and were applied to handwritten numbers, faces, and object generation tasks [44].

In this study, a novel model that combines a GAN and 1DCNN is proposed for seizure detection. A GAN is built to generate the seizure-period EEG data used for training, thus effectively reducing the adverse effect of the imbalanced class distribution of EEG data. Then a 1DCNN is designed and trained by the original and generated EEG training samples to detect seizure periods in long-term EEG data. This model has yielded competitive results on the three publicly available EEG databases, and compared with other detection methods, this model improved detection performance by increasing sensitivity with unchanged specificity.

2. Materials and methods

2.1. EEG databases

2.1.1. CHB-MIT database

The EEG data used in this study came from the CHB-MIT Scalp EEG Database at Children's Hospital Boston, which contained EEG recordings of children with intractable seizures. The database can be downloaded from the PhysioNet website [45]. The subjects were monitored for several days after they were discontinued from antiepileptic drugs to determine the characteristics of their seizures and to evaluate them as candidates for surgical intervention. The database recorded a total of 23 cases from 22 patients (5 males, ages 3–22 years; 17 females, ages 1.5–19 years), in which chb21 was obtained from the same female subjects 1.5 years after chb01 [46]. The file subject information contained the sex and age of each subject except for case chb24, which was later added to this collection with the sex and age unknown.

The recordings were based on the international 10–20 EEG electrode location and nomenclature system, and all signals were sampled at 256 samples per second with a 16-bit resolution. As the number of channels available for different patients was not consistent, 23 channels were used in most of the cases, while 18 channels were used in a few. EEG data totaling about 433.54 h were used in this work, of which 222.27 h was used for training and 211.27 h was used for performance testing. More details of the database are shown in Table 1.

2.1.2. Bonn database

The Bonn database used for this work refers to the data described by Andrzejak et al [47]. The entire dataset consisted of five sets (denoted as Z, O, N, F and S, respectively), each set contained 100 single-channel EEG fragments with a duration of 23.6 s and a sampling rate of 173.6 Hz. All EEG signals were recorded using the same 128-channel amplifier system and the average common reference. The data was digitized at a rate of 173.61 samples per second with a resolution of 12 bits, and the spectral bandwidths of the acquisition system ranged from 0.5 Hz to 85 Hz.

2.1.3. New Delhi database

The data set was collected from Sir Ganga Ram Hospital under the supervision of an experienced neurophysiologist [48]. This set included seizure and seizure-free records obtained at 400 hz from six epileptic patients. Surface recording of both sets of data was performed using a 10–20 international electrode placement system.

Table 1 – The length of the training and test set.

Seizure (h)	Non-seizure (h)	Total (h)	
Training	1.57	220.7	222.27
Test	1.57	209.7	211.27
Total (h)	3.14	430.4	433.54

2.2. Data preprocessing

2.2.1. Segmentation

Since the CHB-MIT Scalp EEG data used in this study were long-term EEG signals, the data set was first preprocessed. In a single one-hour period, most patients had a dataset containing 23 channels, with each channel containing 921,600 sampling points. Segmentation was processed using a moving window with a length of 4 s, so one hour of data is divided into 900 segments with 1024 sampling points in each segment.

2.2.2. Training set and test set

In this study, it was first necessary to generate a training set and a test set for each epileptic patient. Half of all seizure period data were used for training and half for testing. Since the non-seizure datasets were very large in this database, about 30 h of non-seizure data were randomly selected for each patient, and 50% of the data were also used for training and 50% for testing. More details on the training set and test set are shown in Table 2.

2.3. Methods

A flow chart of the main process of this study is shown in Fig. 1. The seizure training data were first fed into the GAN to generate more seizure training data, then the generated and original seizure training data were combined to establish the new seizure training set. In the next step, the data of the new training set in the seizure period and the data of the non-seizure period were merged into the new training set. After that, the new training set was employed to train the designed 1DCNN for the classification of seizure and non-seizure EEG data.

2.3.1. Generative adversarial network

GAN was firstly proposed by Ian Goodfellow in 2014. The original purpose of creating GAN is to generate data that do not exist in the real world [41].

(a) Objective functions

In general, GAN usually consists of two parts: one is the generation model or generation network; the other is the discriminant model or discriminant network. These two parts oppose each other, with the generating model G trying to generate sample data that are not real, and the discriminating model D trying to determine whether the data are real or generated by the generating model. The real data and the

Table 2 – Specific information of training set and test set.

Case	Gender	Age	Seizure period training set	Seizure period test set	Non-Seizure period training set	Non-Seizure period test set
1	F	11	1277	1276	359,040	359,040
2	M	11	492	491	341,566	341,566
3	F	14	1153	1153	320,877	320,876
4	M	22	1060	1059	312,143	312,143
5	F	7	1603	1602	351,940	351,939
6	F	1.5	441	440	20,700	20,700
7	F	14.5	934	934	200,000	200,000
8	M	3.5	2643	2643	155,315	155,314
9	F	10	793	793	110,000	110,000
10	M	3	1211	1210	445,071	445,071
11	F	12	2318	2317	400,954	400,954
12	F	2	3838	3837	117,053	117,053
13	F	3	1438	1437	272,262	272,262
14	F	9	421	420	12,600	12,600
15	M	16	5589	5589	50,000	50,000
16	F	7	195	195	12,600	12,600
17	F	12	805	805	204,013	204,013
18	F	18	912	911	200,000	180,000
19	F	19	678	678	400,000	275,026
20	F	6	845	845	125,290	65,000
21	F	13	572	572	12,600	12,600
22	F	9	586	586	50,400	50,400
23	F	6	1220	1220	100,000	100,000
24	–	–	1469	1469	82,800	62,100

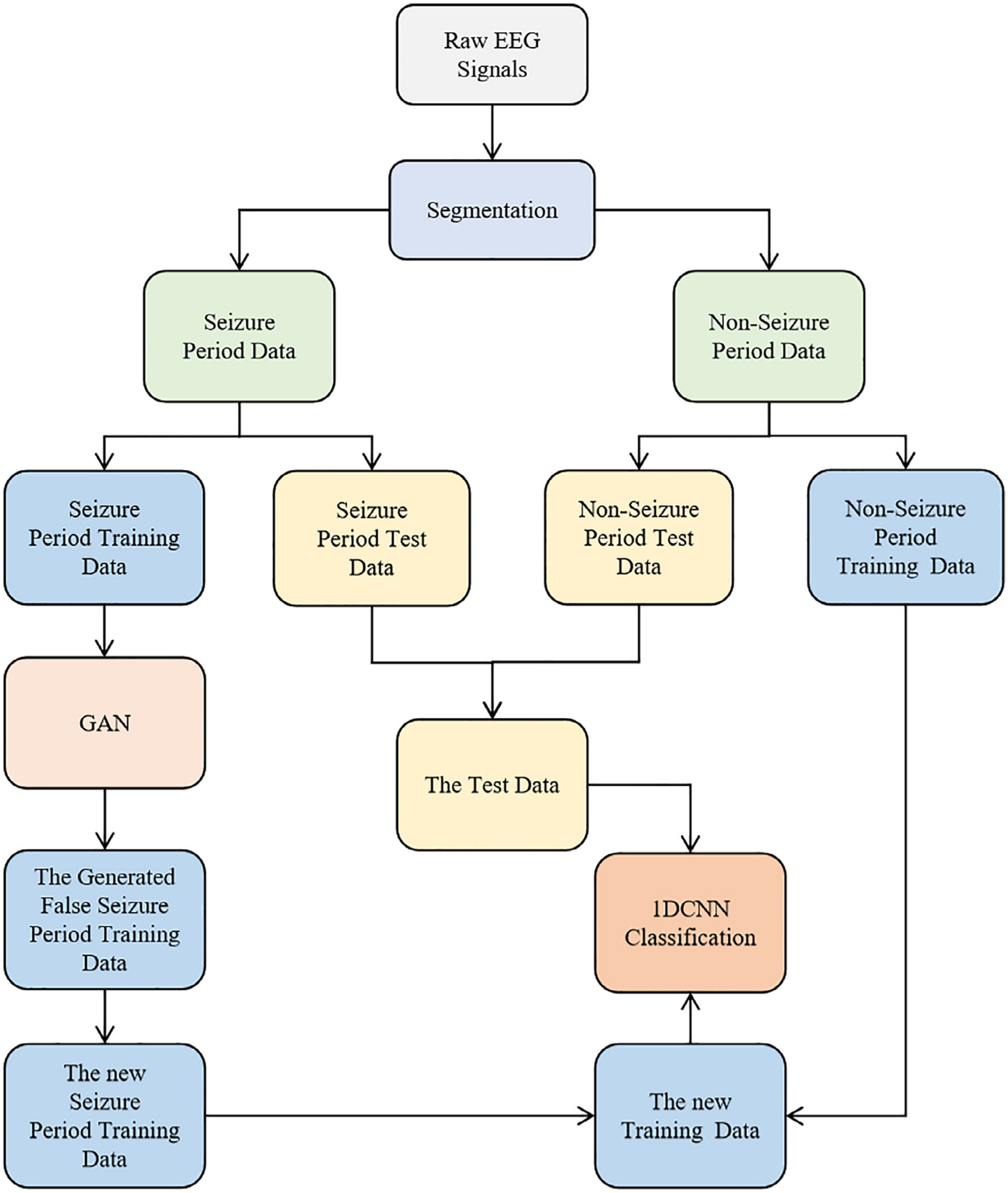


Fig. 1 – Flow chart of the proposed method.

data generated by the generated model are fed into the discriminant model, and then the discriminant model learns to distinguish whether a given sample is real or not. In this kind of adversarial process, the two models compete with each other and constantly improve, both trying to get the best result. The value function is:

$$\min_G \max_D V(D,G) = E_{X\text{-}Pdata(x)}[\log D(x)] + E_{Z\text{-}PZ(z)}[\log(1 - D(G(z)))] \quad (1)$$

where x is the true EEG data, z is the noise vector input to G, $P_{data}(x)$ is the real data distribution, and $P_z(z)$ is the generated data distribution.

(b) Data generation based on GAN

The GAN structure used in this study is shown in Fig. 2.

Here, Z is the noise vector input for generating network G, G(Z) is the data generated by the generator, and X is the real data. Both of them are input at the same time to the identification network D, and finally the generated false data that are similar to the real data are output.

The random noise was sampled from a uniform distribution U[−1,1]. In the generator, the network contained four fully connected hidden layers, with 60, 80, 80, and 80 neurons, and one output layer with 1024 neurons, using tanh as an

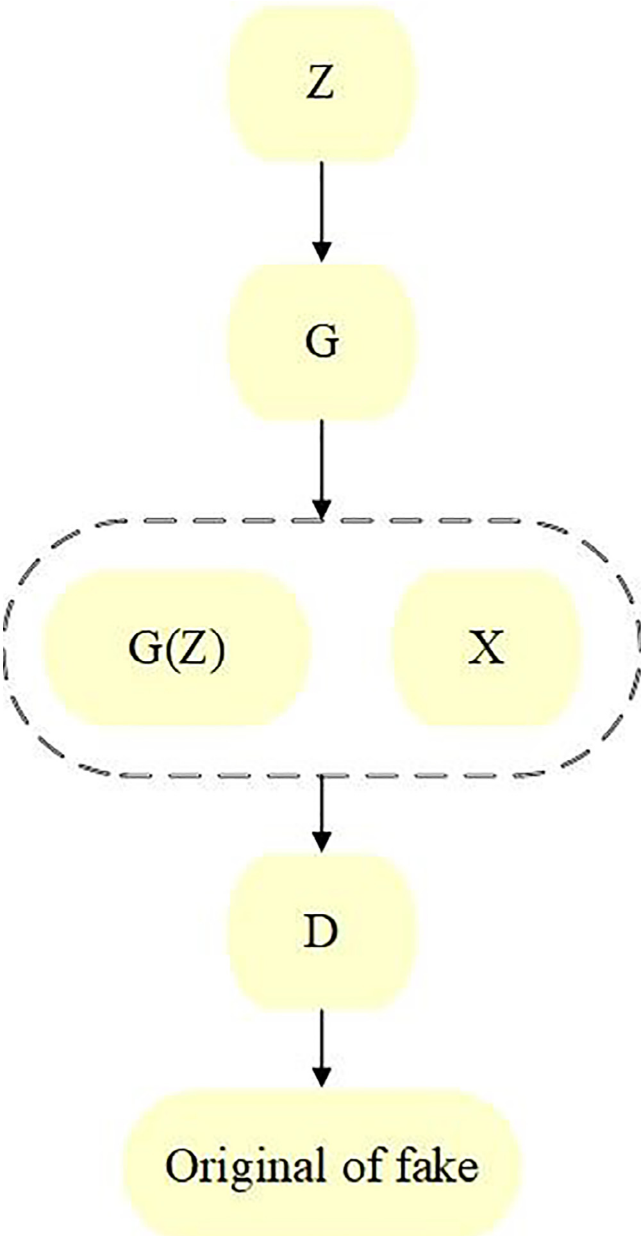


Fig. 2 – GAN structures.

activation function. The inputs to the discriminator were 1024 neurons, and the discriminator contained four fully connected hidden layers with 50, 50, 50, and 1 neurons, also using tanh as an activation function. The networks were optimized by an Adam optimizer. Each experiment iteration was 50 times. In each iteration, G was updated once while D was updated twice (one with real data and one with generated data). Batch size was set to 1024. The detailed operating environment and hyper-parameters are shown in Table 3.

Table 3 – Operating Environment And Hyper-parameters.

Language	Python (Anaconda 3.6.5)
Toolbox	TensorFlow 1.8.0
Numpy version	1.14.3
Dropout rate	0.5
Number of iterations	50
Batch size	1024
Optimizer	Adam

2.3.2. *EEG classification by deep learning*

(a) 1DCNN

The Convolutional Neural Network (CNN) was proposed by Yann LeCun of New York University in 1998. In essence, it is a multi-layer perceptron, and its success is due to its local connection and weight sharing method: On one hand, reducing the number of weights makes the network easy to optimize, and on the other hand, reducing the complexity of the model is able to decrease the risk of overfitting.

Generally, CNNs are mostly used for two-dimensional data, such as images. In order to use a CNN for processing the one-dimensional EEG signal data, we redesigned the convolutional kernels. The network structure of the improved CNN is shown in Fig. 3 [49].

The designed 1DCNN contains a total of 4 convolution layers, containing 32, 24, 16 and 8 cores. Its purpose is to extract a large number of microscopic structures at the lower level, and then synthesize higher-level features from the higher level. The number of features is small, but the difference is large. The pre-processed EEG signal data were input into the 1DCNN. Each convolution block contained a convolution layer (Conv), a batch normalization layer (BN) and a nonlinear activation layer (RELU). Conv1 contained a convolution kernel number of 32, a kernel size of 5*1, and the step size of 3. The number of convolution kernel of Conv2 was 24, the size of the convolution kernel was 5*1, and the step size was 3. The number of convolution kernel of Conv3 was 16, the size of the convolution kernel was 3*1, and the step size was 2. The number of convolution kernel of Conv4 was 8, the size of convolution kernel was 3*1, and the step size was 2. The output of Conv4 was input into the FC1 layer, which contained 20 nerve nodes. In order to avoid over-fitting, a Dropout layer was added between FC1 and FC2 to randomly discard information in neurons. The number of neurons in FC2 depended on the number of classes. The details of the network parameters are shown in Table 4.

(b) SoftMax classification

The SoftMax function was used as the activation function of the output layer for the final classification discrimination. SoftMax will conduct statistics on the probability of the type of sample data calculated by the full connection layer of the previous layer and get the final discrimination result. The algorithm principle of SoftMax activation function is:

$$SoftMax(z_i) = \frac{e^{zi}}{\sum_j e^{zj}} \quad (2)$$

where z_i represents the i-th element in vector z.

2.3.3. *Post-processing*

After classifier makes classification decision, channel fusion decision is made. In this paper, the EEG signal of this segment

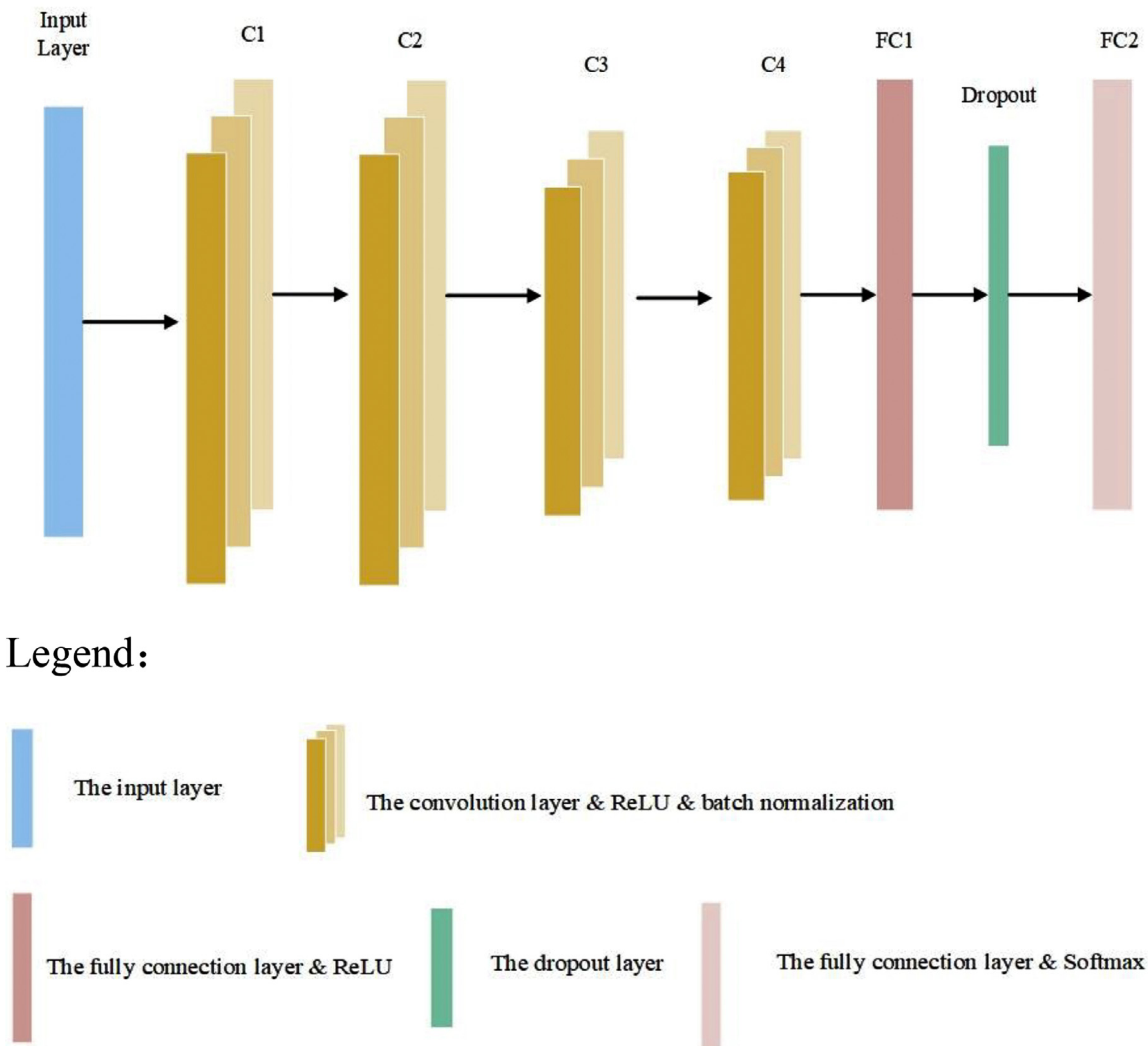


Fig. 3 – The network structure of the improved 1DCNN. Ci (i = 1, 2,3,4) represents a convolution layer; FCj (j = 1, 2) represents a fully connected layer.

is judged to be in the state of seizure only when the state of at least 12 channels was judged to be in the state of seizure by the decision maker. When the discriminant result of less than 12 channels in a certain segment of data was non-seizure, the state of the EEG signal of this segment was judged to be in the state of non-seizure.

2.3.4. *Evaluation indicators for detection model*

In this work, three statistical indicators were used for the performance evaluation of the proposed detection model.

Sensitivity (SEN) is the proportion of correctly classified positive samples to all positive samples. This measures the ability of the model to identify positive samples.

Table 4 – Details of the 1DCNN parameters.

Layer	Filter size	Filter number	Stride
Input	1024*1	–	–
Convolution + ReLU Batch Normalization	5*1	32	3
Convolution + ReLU Batch Normalization	5*1	24	3
Convolution + ReLU Batch Normalization	3*1	16	2
Convolution + ReLU Batch Normalization	3*1	8	2
Fully Connection + ReLU	20 outputs	–	–
Dropout(Dropout rate = 0.5)	–	–	–
Fully Connection + Softmax	2 outputs	–	–

$$SEN = \frac{TP}{TP + FN} \quad (3)$$

Specificity (SPE) is the proportion of correctly classified negative samples to all negative samples. This measures the ability of the model to identify negative samples.

$$SPE = \frac{TN}{TN + FP} \quad (4)$$

where TP is the True Positives (the number of the seizure segments that were correctly detected by the proposed method), TN is the True Negatives (the number of the non-seizure segments that were correctly classified by the proposed method), FP is the False Positives (the number of the non-seizure segments that were mistakenly detected by the proposed method), and FN is the False Negatives (the number of the seizure segments that were mistakenly classified by the proposed method).

G-mean: Since the durations of seizure events are much shorter than those of non-seizure events, seizure detection can be regarded as an imbalanced classification problem. G-mean is an informative indicator that is suitable for the evaluation of imbalanced classification in this work, and is defined as:

$$G - \text{mean} = \sqrt{SEN \times SPE} \quad (5)$$

3. Results and discussion

Since the number of seizures in the data set used in this experiment was too small, GAN was first used to generate more seizure data. The details of the original seizure period training data and the new seizure period training data (combined from the original seizure period training data and the generated false seizure period training data) are shown in Table 5.

The experimental results obtained by the proposed method are shown below. SEN, SPE and G-Mean are shown in Fig. 4. It can be seen that in the vast majority of cases SEN classification result was fine, and it is worth mentioning that the SEN of cases chb03, chb06, chb12, chb13, chb14, chb15, chb16, chb18, chb19, chb20, and chb22 all reached 100%. After calculation, the average SEN of GAN and 1DCNN classification in this database could reach 93.53%. The SPE indexes of the other cases were all around 99% except for chb15, which had an SPE index of 89.98%. In particular, the mean SPE index of the 24 cases reached 99.05% after the calculation. It is observed that the SPE indexes of this database were quite good when the classification of GAN and 1DCNN was combined. It is not difficult to see from the histogram that the G-mean index of all cases in the database was above 85% after the treatment of the method used in this experiment, and most of them were above 95%. After calculation, the average G-mean index of all cases reached 96.15%.

In addition, we conducted the experiment using 1DCNN to classify the original data set without data generation via the GAN, with the classification results shown in the histogram below. The SEN, SPE and G-Mean indexes are shown in Fig. 5, which shows that the result was not ideal. Moreover, the SEN indexes of cases chb11, chb15, and chb16 were all 0%. As can be seen from the figure, the SPE index of more than half of the cases reached 99%, and most of them were above 90%. It shows that the 1DCNN model used in this experiment had a good effect on the SPE index of the classification results of the CHB-MIT database. It is obvious that the G-

Table 5 – Details of the generated training set data.

Case	Gender	Age	Original seizure period training data (O)	New seizure period training data (N)	Ratio (O:N)
1	F	11	1277	2554	1:2
2	M	11	492	984	1:2
3	F	14	1153	2306	1:2
4	M	22	1060	3180	1:3
5	F	7	1603	3206	1:2
6	F	1.5	441	882	1:2
7	F	14.5	934	10,274	1:11
8	M	3.5	2643	29,073	1:11
9	F	10	793	4793	1:6
10	M	3	1211	2422	1:2
11	F	12	2318	6954	1:3
12	F	2	3838	7676	1:2
13	F	3	1438	2876	1:2
14	F	9	421	4210	1:10
15	M	16	5589	8000	1:1.4
16	F	7	195	2195	1:11
17	F	12	805	1610	1:2
18	F	18	912	1824	1:2
19	F	19	678	1356	1:2
20	F	6	845	5845	1:7
21	F	13	572	1572	1:2.8
22	F	9	586	3586	1:6
23	F	6	1220	2220	1:1.8
24	–	–	1469	2469	1:1.7

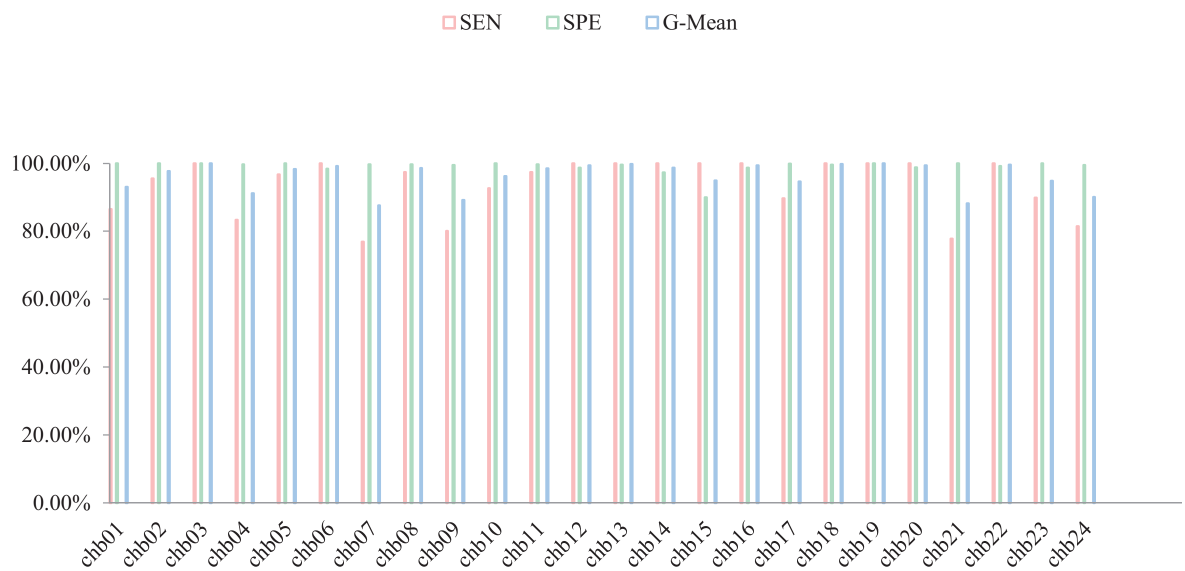


Fig. 4 – The SEN, SPE and G-Mean diagram of indicators of classification results.

Fig. 5 – The SEN, SPE and G-Mean index of classification results without the GAN.

mean value was also unsatisfactory. As can be seen from the figure, the G-mean of about one-third of the cases did not even reach 85%.

It is obvious that using 1DCNN alone without the use of GAN does not give an appropriate result to the classification of this database.

We compared the classification results of the CHB-MIT database by combining GAN with 1DCNN and using only 1DCNN in Fig. 6. In the bar chart, the blue part is the classification results using only 1DCNN, and the brown part is the classification results combined with GAN and 1DCNN. We used the mean values of SEN, SPE and G-mean of the classification results of all 24 cases in the CHB-MIT database in our experiment. As can be seen from the histogram, the classification method of GAN + 1DCNN could greatly improve the SEN index of classification results (increasing the SEN index from 69.95% to 93.53%) compared with the classification method using only 1DCNN. Because the 1DCNN used in this paper had a good effect on the SPE index of CHB-MIT database classification results, the GAN + 1DCNN classification method did not improve too much (only improving the SPE index from 98.64% to 99.05%). Finally, for the analysis of G-mean value, we increased it from 75.53% to 96.15%, which is a good improvement.

The above comparison between the GAN + 1DCNN and using only 1DCNN indicates that the augmentation of seizure training samples by the GAN can obviously improve the performance of seizure detection via increasing the SEN index. However, quite a few numbers of seizure patterns are known for one epileptic patient clinically. Fortunately, there is a plentiful supply of seizure patterns from other epileptic patients. Therefore, the cross-patient seizure detection method will be developed based on this model in our future work.

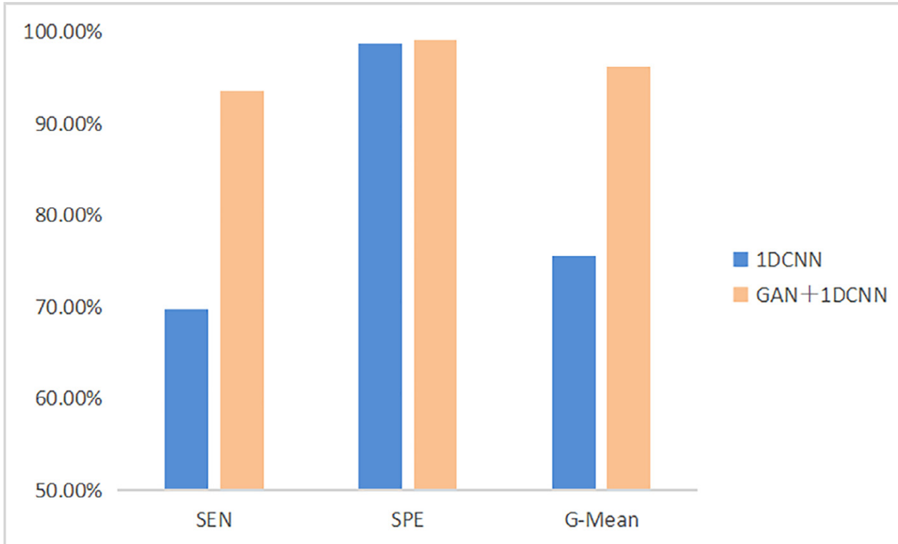
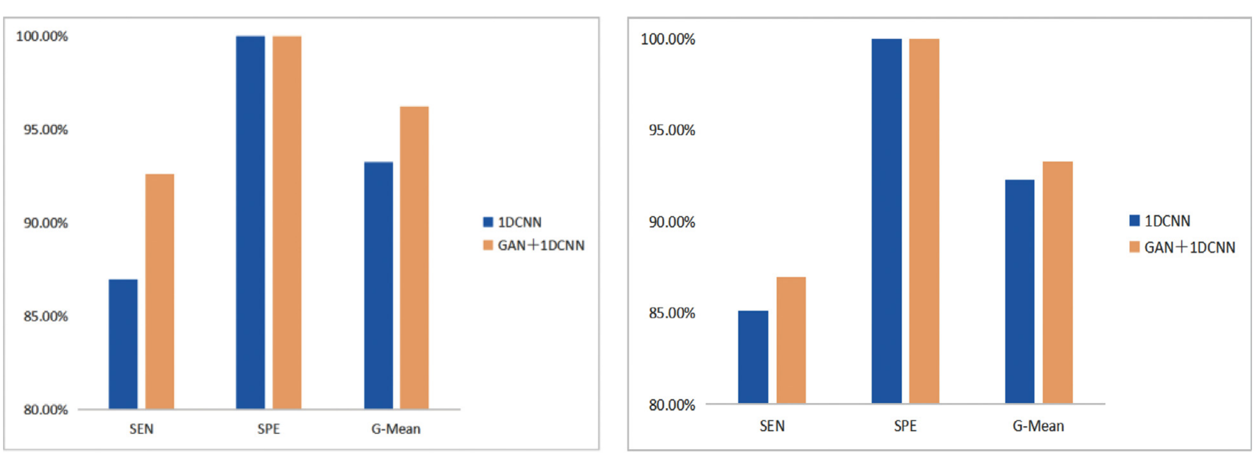


Fig. 6 – Comparison of the classification results of 1DCNN in combination with GAN and 1DCNN only.

In addition, we performed performance tests in Bonn (a) and New Delhi (b) databases. For Bonn dataset, the first 100 segments of data in S set and the first 200 segments of data in F set were taken as the training set, and the test set consisted of the last 100 segments of data in S set and the last 200 segments of data in F set. For New Delhi dataset, the first 10 recordings of seizure period and the first 20 recordings of non-seizure period were used as the training set, and the test set consisted of other 10 recordings of seizure period and other 20 recordings of non-seizure period. The classification results are shown in Fig. 7. It can be seen that the classification results obtained by this model on other databases are also excellent, and this method combining GAN and 1DCNN is indeed better than the 1DCNN model alone. At the same time, we also calculated the F1-score [50] of the two databases as 0.9667 and 0.9302 respectively.

In our initial ideal assumption, with the increase of epileptic period training data generated by GAN, it will be more and more helpful to the classification results of 1DCNN in the later stage. However, in the actual experiment, this did not turn out to be true. With the increase of generated data, the classification results may become better and better, because the generated seizure data effectively relieve the imbalanced distribution of seizure and non-seizure EEG signals. But the classification results may become worse, because the GAN generated too many data which belong to some of seizure patterns, which tends to lead to overfitting of the 1DCNN. A few cases are listed below to illustrate this point.

Fig. 8 shows the classification results of cases chb07 (a) and chb08 (b) in the CHB-MIT database under different amount of data generated by GAN. The Original in the longitudinal axis is the data of the original seizure training set. We first



(a) (b)

Fig. 7 – Comparison of the classification results of 1DCNN in combination with GAN and 1DCNN only on the Bonn (a) and New Delhi (b) databases.

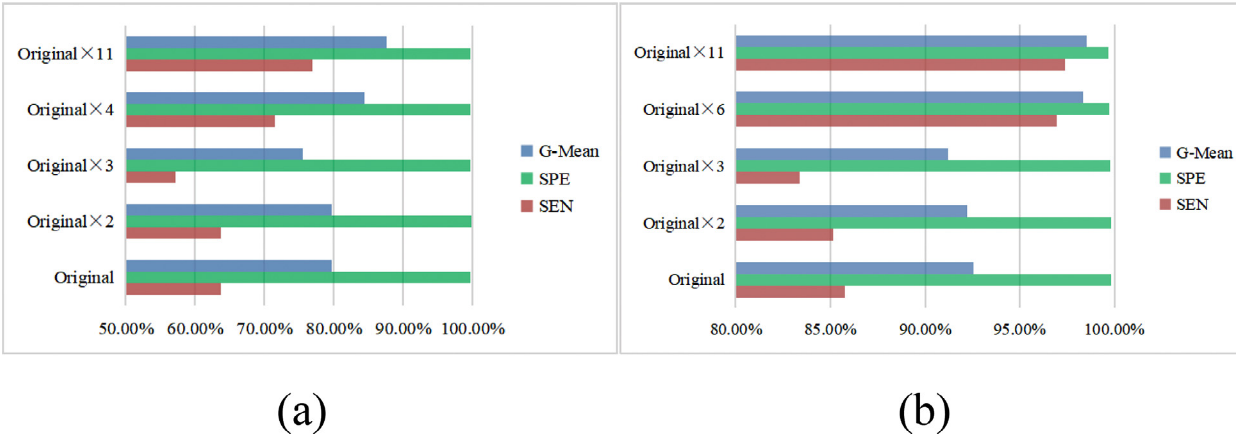


Fig. 8 – The classification results of chb07 (a) and chb08 (b) under different quantity data generated by GAN.

considered generating data several times of Original and input 1DCNN for classification experiment, namely as shown in the figure Original $\times$ 2, Original $\times$ 3, Original $\times$ 4, Original $\times$ 6 and Original $\times$ 11 (double, triple, quadruple, 6 times and 11 times of the original epileptic seizure training set data). It can be seen that the three indicators of SEN, SPE and G-mean in case chb07 did not gradually improve with the increase of generated data. Unexpectedly, the SEN index of this case was lower than that of the original data when triple the original data were generated. Finally, when we generated 11 times the data, our experimental results showed that all three indicators became ideal, so we chose the result of this time as our final experimental result. The classification results of case chb08 are similar to the problem described above. With the increase of GAN generation data, the SEN index and G-mean value of patients showed a trend of decreasing first and then increasing, while the SPE index basically did not change much. We also selected the optimal results after many experiments as our final experimental results.

The classification results of cases chb21 (a) and chb24 (b) under different amount of data generated by GAN are shown in Fig. 9. Unlike the previous two examples, the SEN index of case chb21 showed a trend of first rising, then falling, then rising again with the increase of generated data. After many experiments, the best result in the experiment process was saved as our final classification result. With the increase of GAN generated data, the SEN index of case chb24 showed a trend of first increasing and then decreasing, and then increasing and decreasing again. The results we ended up with were still the best of many experiments.

According to the 1DCNN network model used in this paper, we adopted the control variable method and change important network parameters: the optimizer and MiniBatchSize to compare the network performance.

The classification results of the 1DCNN model used in this study under the same MiniBatchSize and different optimizers are shown in Fig. 10. Half of the patients selected from the database used in this study were compared under three optimizers: Adam, RMSProp, and SGD. After conducting experi-

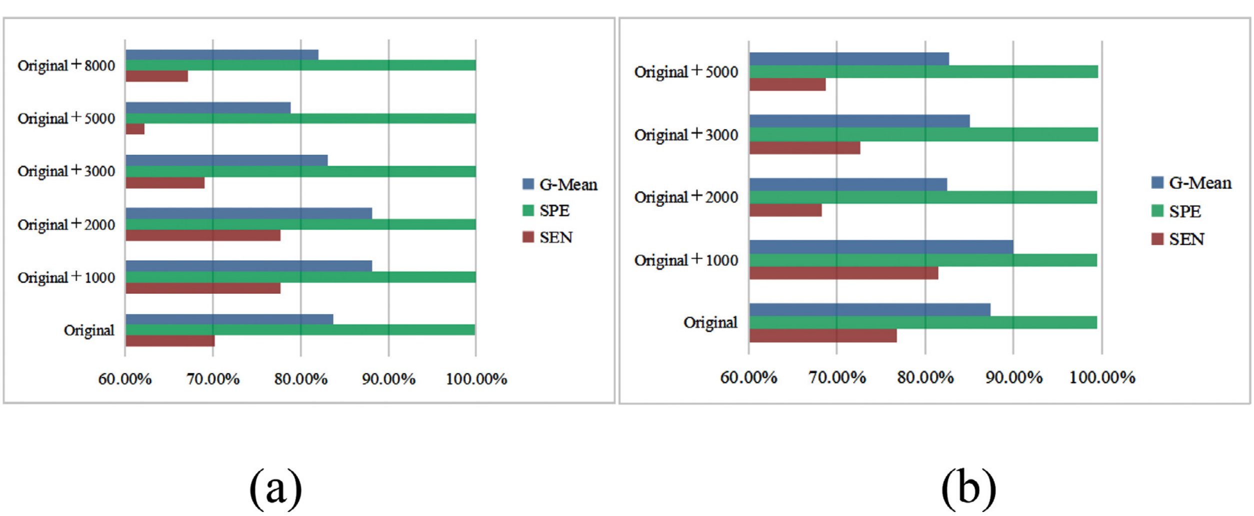


Fig. 9 – The classification results of chb21 (a) and chb24 (b) under different quantity data generated by GAN.

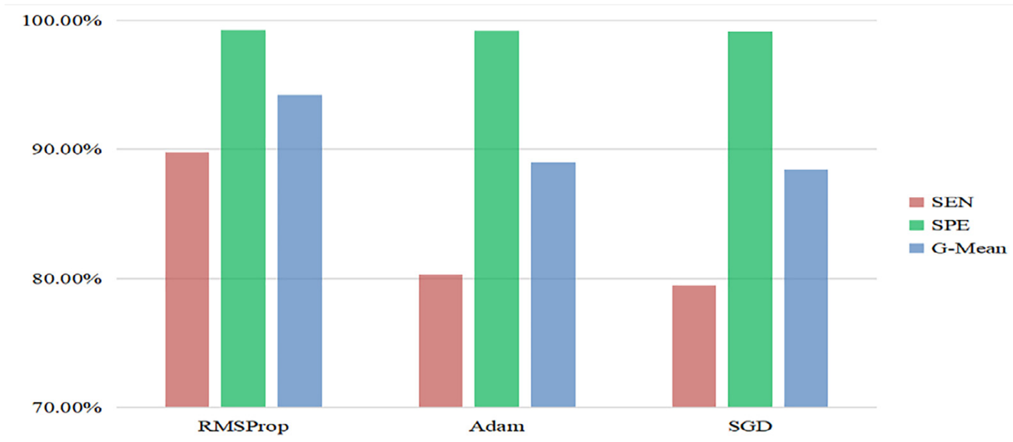


Fig. 10 – The classification results of 1DCNN under different optimizers.

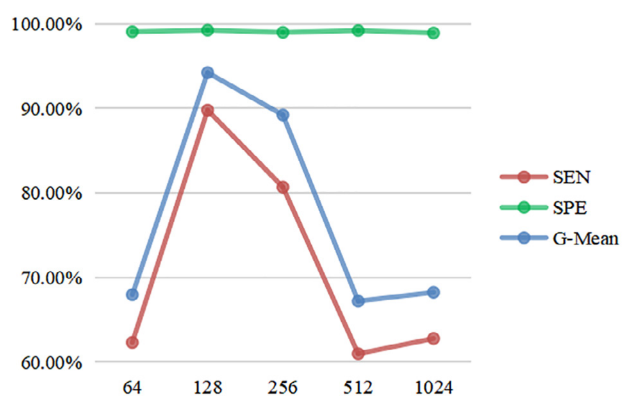


Fig. 11 – The classification results of 1DCNN under different MiniBatchSizes.

ments on each of them, SEN, SPE, and G-mean were calculated and the average values of the three indexes were finally obtained. It is not difficult to see from Fig. 10 that the classification results of the RMSProp optimizer were better than those of the other two optimizers. It is proved that the effect of the optimizer used in this experiment was optimal.

Fig. 11 shows the classification results of the 1DCNN network model used in this paper under different MiniBatchSize conditions of the RMSProp optimizer. The horizontal axis data in this picture are different MiniBatchSizes used in this comparison test: 64, 128, 256, 512, and 1024. It can be seen that basically the classification result was best when the MiniBatchSize was 128, so the 1DCNN network model with the MiniBatchSize of 128 was used in our experiment.

Table 6 shows a comparison of the classification results of the method used in this paper with other published methods on the CHB-MIT EEG database. Zabihi et al. proposed a multichannel EEG seizure detection method based on phase space trajectory dynamics and obtained the SEN score of 88.27% and SPE score of 93.21% on 25% training data of the CHB-MIT benchmark database [51]. A non-specific seizure detection strategy based on wavelet transform and linear discriminant analysis (LDA) proposed by Orosco et al. achieved a SEN index of 92.6% and an SPE index of 99.9% in the CHB-MIT dataset [52]. The feature extraction method proposed by Samiee et al., based on sparse rational decomposition and local Gabor binary mode, achieved 70.4% sensitivity and 99.1% specificity in 163 h of EEG recording in CHB-MIT scalp EEG database [53]. Chen et al. selected band features of 7 wavelets, and coif 3 achieved 92.30% accuracy in CHB-MIT database by using 7 features of 6 bands [54]. Liang et al. constructed an 18-layer long-term recursive convolutional network (LRCN) to evaluate 198 cases of seizure from 23 patients, achieving a SEN index of 84% and an SPE index of 99% [22]. The epileptic seizure prediction method based on public Space Mode (CSP) and convolutional neural network (CNN) proposed by Zhang et al. achieved a SEN index of 92.2% on the CHB-MIT dataset [55]. Kaziha et al. proposed a convolutional neural network model that automatically classifies raw EEG data from seizure and non-seizure signals without preprocessing, resulting in a SEN index of 82.35% and an SPE index of 100% [56]. The algorithm proposed by Ciurea et al. to extract time domain features in CHB-MIT data set obtained SEN index of 91.99% and SPE index of 93.38% [57]. Zanetti et al. proposed a robust seizure detection method suitable for wearable platforms and verified it in the CHB-MIT database with SEN index of 96.6% and SPE index of 92.5% [58]. Wang et al. proposed a stacked 1DCNN model combining random selection and data enhancement (RS-DA) strategy, which achieved a SEN index of 88.14% and an SPE index of 99.62% in the CHB-MIT dataset [59].

Among these comparison methods, Orosco et al. made the experimental G-mean value of 96.18%, only 0.03% higher than this work, and they used only 152.8 h of the CHB-MIT database for the performance test, which is about a third of the length of our experiment. The SPE index obtained by Kaziha et al. was as high as 100%, which was indeed better than the present study, but its SEN index was only 82.35%, much lower than the 93.35% in this experiment. In addition to the two methods mentioned above, the classification results of the other eight methods listed in Table 6 are all lower than the method proposed in this paper, indicating that the seizure detection method proposed in this work has effective classification performance for the CHB-MIT database.

Table 6 – The comparisons of results with other published methods using the CHB-MIT EEG database

Papers, Year	Duration (h)	Method	Performance metrics(%) SEN	SPE	G-Mean
Zabihi et al. (2016) [51]	–	25% training rate	88.27	93.21	90.70
		50% training rate	89.10	94.80	91.91
Orosco et al. (2016) [52]	152.8	Dyadic WT, LDA	92.6	99.9	96.18
Samiee et al. (2017) [53]	163	2D mapping and textual feature LGBP	70.4	99.1	83.53
Chen et al. (2017) [54]	76.92	patient-specific Dyadic WT (coif 3), SVM	91.71	92.89	92.30
Liang et al. (2018) [22]	–	LRCN(CNN+) LSTM	84	99	91.19
Zhang et al. (2019) [55]	–	CSP,CNN	92.2	–	–
Kaziha et al. (2020) [56]	10.11	CNN	82.35	100	96.74
Ciurea et al. (2020) [57]	–	Extracts time-domain Features	91.99	93.38	92.69
Zanetti et al. (2020) [58]	–	A wearable platform	96.6	92.5	94.5
Wang et al. (2021) [59]	–	RS-DA + 1D-CNN	88.14	99.62	93.70
This work GAN+	433.54	GAN+	93.53	99.05	96.15

4. Conclusions

In this paper, a novel epileptic seizure detection method combining GAN and 1DCNN is proposed. We firstly used the GAN to generate more seizure period training data, and then 1DCNN was used to classify the processed data. The experiments proved that this method had good classification results for the three EEG databases, which demonstrated the generalizability of this method across different databases.

Conflict of interest statement

We declared that we have no conflicts of interest to this work.

Acknowledgment

This work was supported by National Natural Science Foundation of China (No. 61501283).

REFERENCES

[1] Geier C, Lehnertz K. Which brain regions are important for seizure dynamics in epileptic networks? influence of link identification and EEG recording montage on node centralities. Int J Neural Syst 2017;27(1):1650033.
[2] Beghi E, Giussani G, Sander JW. The natural history and prognosis of epilepsy. Epileptic Disord 2015;17(3):243–53.

[3] Mahmoodian N, Boese A, Friebe M, Haddadnia J. Epileptic seizure detection using cross-bispectrum of electroencephalogram signal. Seizure 2019;66:4–11.
[4] Keezer MR, Sisodiya SM, Sander JW. Comorbidities of epilepsy: current concepts and future perspectives. Lancet Neurol 2016;15(1):106–15.
[5] World Health Organization. seizure [cited April 8, 2016], http://www.who.int/mediacentre/factsheets/fs999/en/, 2016.
[6] Löscher W, Potschka H, Sisodiya SM, Vezzani A, Barker EL. Drug resistance in epilepsy: clinical impact, potential mechanisms, and new innovative treatment options. Pharmacol Rev 2020;72(3):606–38.
[7] Niedermeyer E, Silva FHLD. Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, 5th edn. (Wolters Kluwer, 2004), Lippincott Williams & Wilkins, Philadelphia, p. 526.
[8] Gandhi T, Panigrahi BK, Bhatia M, Anand S. Expert model for detection of epileptic activity in EEG signature. Expert Syst Appl 2010;37(4):3513–20.
[9] Sanei S, Chambers JA. EEG signal processing. UK: Cardiff University, lohn Wiley & Sons, Ltd; 2007.
[10] Symonds JD, Zuberi SM, Johnson MR. Advances in epilepsy gene discovery and implications for epilepsy diagnosis and treatment. Curr Opin Neurol 2017;30(2):193–9.
[11] McSharry PE, He T, Smith LA, Tarassenko L. Linear and nonlinear methods for automatic seizure detection in scalp electroencephalogram recordings. Med Biol Eng Comput 2002;40(4):447–61.
[12] Yuan Qi, Zhou W, Zhang L, Zhang F, Xu F, Leng Y, et al. Epileptic seizure detection based on imbalanced classification and wavelet packet transform. Seizure 2017;50:99–108.
[13] Husari KS, Dubey D. Autoimmune epilepsy. Neurotherapeutics 2019;16(3):685–702.
[14] Martinez-del-Rincon J, Santofimia MJ, del Toro X, Barba J, Romero F, Navas P, et al. Non-linear classifiers applied to EEG analysis for epilepsy seizure detection. Expert Systems Appl 2017;86:99–112.
[15] Kaleem M, Guergachi A, Krishnan S. Patient-specific seizure detection in long-term EEG using wavelet decomposition. Biomed Signal Process Control 2018;46:157–65.
[16] Vezzani A, Fujinami RS, White HS, Preux P-M, Blümcke I, Sander JW, et al. Infections, inflammation and epilepsy. Acta Neuropathol 2016;131(2):211–34.
[17] Janjarasjitt S. Epileptic seizure classifications of single-channel scalp EEG data using wavelet-based features and SVM. Med Biol Eng Comput 2017;55(10):1743–61.
[18] Balestrini S, Sisodiya SM. Pharmacogenomics in epilepsy. Neurosci Lett 2018;667:27–39.
[19] Gotman J. Automatic recognition of epileptic seizures in the EEG. Electroencephalogr Clin Neurophysiol 1982;54(5): 530–40.
[20] Park C, Choi G, Kim J, et al. Epileptic seizure detection for multi-channel EEG with deep convolutional neural network. In: IEEE International Conference on Electronics Information and Communication. p. 1–5.
[21] Emami A, Kunii N, Matsuo T, Shinozaki T, Kawai K, Takahashi H. Seizure detection by convolutional neural network-based analysis of scalp electroencephalography plot images. NeuroImage: Clinical 2019;22:101684.
[22] Liang W, Pei H, Cai Q, Wang Y. Scalp EEG epileptogenic zone recognition and localization based on long-term recurrent convolutional network. Neurocomputing 2020;396:569–76.
[23] Daoud H, Bayoumi MA. Efficient epileptic seizure prediction based on deep learning. Trans Biomed Circuits Systems 2019;13:804–13.
[24] Ahmedt-Aristizabal D, Fookes C, Nguyen K, et al. Deep classification of epileptic signals. In: 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. p. 332–5.
[25] Xu YK, Yang J, Zhao S, Wu H, Sawan M. An end-to-end deep learning approach for epileptic seizure prediction. In: 2020 2nd IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS). p. 266–70.
[26] Truong ND, Kuhlmann L, Bonyadi MR, Querlioz D, Zhou L, Kavehei O. Epileptic seizure forecasting with generative adversarial networks. IEEE Access 2019;7:143999–4009.
[27] Petrosian A, Prokhorov D, Homan R, Dasheiff R, Wunsch D. Recurrent neural network-based prediction of epileptic seizures in intra-and extracranial EEG. Neurocomputing 2000;30(1-4):201–18.
[28] Wei Z, Zou J, Zhang J, Xu J. Automatic epileptic EEG detection using convolutional neural network with improvements in time-domain. Biomed Signal Process Control 2019;53:101551.
[29] Thara DK, PremaSudha BG, Xiong F. Auto-detection of epileptic seizure events using deep neural network with different feature scaling techniques. Pattern Recogn Lett 2019;128(1):544–50.
[30] Takahashi H, Emami A, Shinozaki T, Kunii N, Matsuo T, Kawai K. Convolutional neural network with autoencoder-assisted multiclass labelling for seizure detection based on scalp electroencephalography. Comput Biol Med 2020;125:104016.
[31] Ward M, Malmsten K, Salamy H, Min CH. Data balanced bagging ensemble of convolutional-LSTM neural networks for time series data classification with an imbalanced dataset. In: IEEE International Symposium on Circuits and Systems (ISCAS). p. 1–5.
[32] Birjandtalab J, Jarmale VN, Nourani M, Harvey J. Imbalance learning using neural networks for seizure detection. In: IEEE Biomedical Circuits and Systems Conference (BioCAS). p. 1–4.
[33] Masum M et al. Analysis of sampling techniques towards epileptic seizure detection from imbalanced dataset. In: IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC). p. 684–92.
[34] Alzaid M. A comparative study between classical feature engineering and RNNs for seizure detection in imbalanced data. In: IEEE Signal Processing in Medicine and Biology Symposium (SPMB). p. 1–5.
[35] Antoniou A, Storkey A, and Edwards H. Data augmentation generative adversarial networks. arXiv:1711.04340, Mar. 2018.
[36] Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, WardeFarley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets, in Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2, ser. NIPS'14. Cambridge, MA, USA: MIT Press, 2014, pp.2672-2680.
[37] Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. Neurocomputing 2018;321:321–31.
[38] Hartmann KG, Schirrmeister RT, Ball T. EEGGAN: Generative adversarial networks for electroencephalographic (EEG) brain signals. arXiv:1806.01875, Jun. 2018.
[39] Esteban C, Hyland SL, Rätsch G. Real-valued (medical) time series generation with recurrent conditional GANs. arXiv:1706.02633, Dec. 2017.
[40] Qin X, Wang Z, Yao J, Zhou Q, Zhao P-F, Wang Z-Y, et al. Using a one-dimensional convolutional neural network with a conditional generative adversarial network to classify plant electrical signals. Comput Electron Agric 2020;174:105464.
[41] Zhou Z, Zhai X, Tin C. Fully automatic electrocardiogram classification system based on generative adversarial network with auxiliary classifier. Expert Syst Appl 2021;174:114809.
[42] Liu S, Yu M, Li M, Xu Q. The research of virtual face based on Deep Convolutional Generative Adversarial Networks using TensorFlow. Phys A 2019;521:667–80.

[43] Chen Z, Wang C, Wu H, Shang K, Wang J. DMGAN: Discriminative Metric-based Generative Adversarial Networks. Knowl-Based Syst 2020;192:105370.
[44] Zhong G, Gao W, Liu Y, Yang Y, Wang D-H, Huang K. Generative adversarial networks with decoder-encoder output noises. Neural Networks 2020;127:19–28.
[45] http://physionet.org/physiobank/database/chbmit/.
[46] Tsiouris , Pezoulas VC, Zervakis M, Konitsiotis S, Koutsouris DD, Fotiadis DI. A Long Short-Term Memory deep learning network for the prediction of epileptic seizures using EEG signals. Comput Biol Med 2018;99:24–37.
[47] Andrzejak R, Lehnertz K, Mormann F, Rieke C, David P, Elger C. Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. Phys Rev E 2001;64(6):061907-1-8.
[48] Jurcak V, Tsuzuki D, Dan I. 10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems. Neuroimage 2007;34(4):1600–11.
[49] Li Y, Zou Li, Jiang Li, Zhou X. Fault diagnosis of rotating machinery based on combination of deep belief network and one-dimensional convolutional neural network. IEEE Access 2019;7:165710–23.
[50] DeVries Z, Locke E, Hoda M, Moravek D, Phan K, Stratton A, et al. Using a national surgical database to predict complications following posterior lumbar surgery and comparing the area under the curve and F1-score for the assessment of prognostic capability. Spine J 2021;21 (7):1135–42.
[51] Zabihi M, Kiranyaz S, Rad A, et al. Analysis of high-dimensional phase space via Poincaré section for patient-specific seizure detection. IEEE Transactions on Neural Systems and Rehabilitation Engineering 2016;24:386–98.
[52] Orosco L, Correa AG, Diez P, Laciar E. Patient non-specific algorithm for seizures detection in scalp EEG. Comput Biol Med 2016;71:128–34.
[53] Samiee K, Kovács P, Gabbouj M. Epileptic seizure detection in long-term EEG records using sparse rational decomposition and local Gabor binary patterns feature extraction. Knowl-Based Syst 2017;118:228–40.
[54] Chen D, Wan S, Xiang J, Bao FS, Chung CK. A high-performance seizure detection algorithm based on Discrete Wavelet Transform (DWT) and EEG. PLoS ONE 2017;12(3): e0173138.
[55] Zhang Y, Guo Y, Yang Po, Chen W, Lo B. Epilepsy seizure prediction on EEG using common spatial pattern and convolutional neural network. Biomed Health Informatics 2020;24(2):465–74.
[56] Kaziha O, Bonny T. A convolutional neural network for seizure detection. Advances in Science and Engineering Technology International Conferences (ASET), 2020.
[57] Ciurea A, Manoila C, Tautan A, Ionescu B. Low latency automated epileptic seizure detection: individualized vs. global approaches. International Conference on e-Health and Bioengineering (EHB), 2020.
[58] Zanetti R, Aminifar A, Atienza D. Robust epileptic seizure detection on wearable systems with reduced false-alarm rate. In: 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). p. 4248–51.
[59] Wang X, Wang X, Liu W, Chang Z, Kärkkäinen T, Cong F. One dimensional convolutional neural networks for seizure onset detection using long-term scalp and intracranial EEG. Neurocomputing 2021;459:212–22.