

Cross-Patient Automatic Epileptic Seizure Detection Using Patient-Adversarial Neural Networks with Spatio-Temporal EEG Augmentation

Zongpeng Zhang^{a,b,*}, Taoyun Ji^{c,*}, Mingqing Xiao^b, Wen Wang^c, Guojing Yu^c,
Tong Lin^{b,**}, Yuwu Jiang^{c,***}, Xiaohua Zhou^{a,d,e}, Zhouchen Lin^{b,f}

^a*Department of Biostatistics, School of Public Health, Peking University, China*

^b*National Key Lab. of General AI, School of Intelligence Science and Technology, Peking
University, China*

^c*Department of Pediatrics, Peking University First Hospital, Beijing, China*

^d*Beijing International Center for Mathematical Research, Peking University, China*

^e*Peking University Chongqing Institute for Big Data, Chongqing, China*

^f*Institute for Artificial Intelligence, Peking University, China*

Abstract

Cross-patient automatic epileptic seizure detection through electroencephalogram (EEG) is significant for clinical application and research. However, most automatic seizure detection methods are patient-specific and have poor generalization ability to unseen patients. In this paper, we consider two aspects for solving the cross-patient generalization problem in deep learning approaches. Firstly, we propose a new data augmentation method to largely improve generalization. We analyze the statistical distribution of seizure EEG signals and propose spatio-temporal EEG augmentation (STEA) to generate synthetic training seizure data with spatio-temporal dependencies in EEG. Secondly, we propose patient-adversarial neural network (PANN) to learn a patient-invariant representation for better generalization. We add a discriminator to distinguish the identity of patients and perform adversarial optimization between the feature extractor and identity discriminator, so that only shared seizure features are maintained. We conduct experiments on both public and clinical EEG datasets

*Equal contributions

**Corresponding author

Email addresses: zongpeng.zhang98@gmail.com (Zongpeng Zhang), lintong@pku.edu.cn
(Tong Lin)

with various settings of different window lengths and divisions of training and testing sets. Our model significantly improves the performance of cross-patient detection compared with other methods and achieves state-of-the-art performance. Particularly, our proposed method achieves up to 95% sensitivity on CHB-MIT under 5s window length segments, and achieves about 85% AUC and sensitivity even on the clinical dataset. Our designed methods can enable detection models extended to unseen patients, with performance achieving the clinical application standards.

Keywords: automatic epileptic seizure detection, cross-patient, deep learning, EEG data augmentation, patient-adversarial neural networks

1. Introduction

Epilepsy is one of the most common chronic noncommunicable diseases of the brain which influences people of all ages (Mormann et al., 2007). According to the World Health Organization Report, epilepsy affects 50 million people around the world (World Health Organization, 2016). As a worldwide disease, epilepsy brings pain and inconvenience to patients and their families, which can even threaten their lives. Thus, the research on epilepsy diagnosis and treatment techniques is valuable and vital.

One of the indicators for epileptic seizure is the abnormal electroencephalography (EEG) pattern (Ahmad et al., 2016). As shown in Fig. 1, the brain action in epileptic EEG waves consists of the preictal stage, ictal stage, postictal stage, and interictal stage. These four stages represent the time nearing the seizure, the on-time period of the seizure, the interval succeeding a seizure, and the left-out period, respectively.

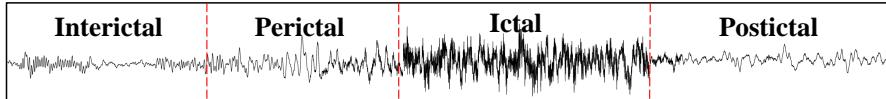


Figure 1: Illustration of four stages of the epileptic EEG.

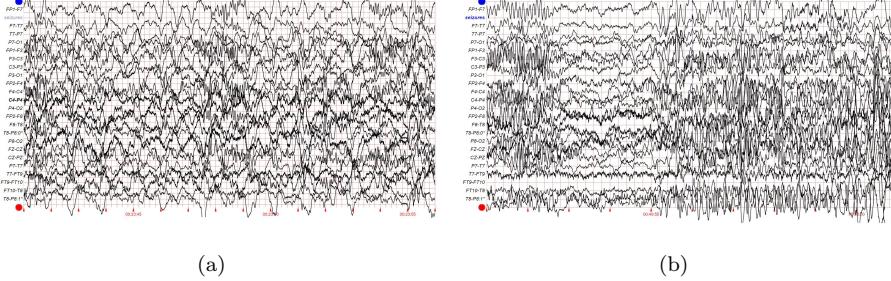
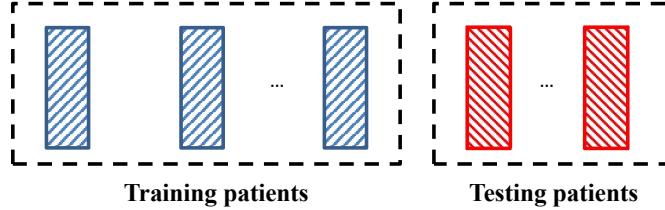


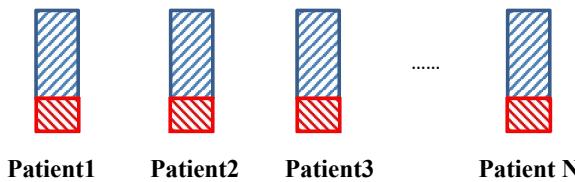
Figure 2: The waveform graph sampled from the CHB-MIT dataset (Shoeb & Guttag, 2010). Figure (a) is the normal period and Figure (b) is the epileptic seizure period. In the figures, x axis represents the time dimension, and y axis represents 23 electrode channels.

Seizure detection aims to distinguish whether an EEG wave is in the ictal stage. However, this is really difficult since the raw EEG signal has much noise as shown in Fig. 2, hence it invariably takes much time for manual diagnoses. Therefore, the automated methods to deal with EEG signals are important to ease and speed up the detection, which is the main focus of most research (Ahmad et al., 2016; Ke et al., 2021; Wei et al., 2019; Abdelhameed et al., 2018; Abiyev et al., 2020; O’Shea et al., 2020; Peng et al., 2022; Jiang et al., 2017; Li et al., 2021; Jiang et al., 2023). These methods formulate the EEG detection problem as a classification task by segmenting EEG signals in short time intervals and classifying whether each part is in the ictal stage, and the mainstream works apply deep neural networks to solve this problem (Ke et al., 2021; Wei et al., 2019; Abdelhameed et al., 2018; Abiyev et al., 2020; O’Shea et al., 2020; Daoud & Bayoumi, 2019; Li et al., 2020; Shen et al., 2023; He et al., 2022).

However, most previous works do not consider the cross-patient generalization problem, which limits their practical utility. Specifically, the traditional setting of previous work only considers the training and testing of the model on the same group of patients and fails to generalize to unseen ones. Differently, the cross-patient setting separates patient identities in the training and testing sets, which aims to train models with several individuals’ data and generalize to unseen patients. The illustration of the differences between these two settings is



(a) Cross-patient epileptic seizure detection, where the training patients and testing patients are different.



(b) Traditional setting for epileptic seizure detection, where the detection models are trained and tested on the same patients.

Figure 3: Two kinds of epileptic seizure detection settings. The blue block represents the training set and the red block represents the testing set.

35 shown in Fig. 3.

In reality, the cross-patient setting is much closer to the real-world setting and is essential for medical treatment, because the clinical applications usually prefer a universal model for different patients and require the generalization ability to adapt to fresh samples. With these advantages, cross-patient epilepsy 40 seizure detection methods are more convenient to be promoted and valuable for both reducing the efforts of well-experienced doctors and improving the diagnosis precision of junior doctors with inadequate training.

Nevertheless, the cross-patient setting is more difficult and there is still much room for improvement. This is because the seizure form of epilepsy varies a lot 45 among patients and is difficult to generalize. Additionally, since there are much more non-seizure periods than seizure periods for EEG signals, such a class-imbalance problem will cause performance degradation and poorer detection

generalization ability, which is more severe in the cross-patient setting. Most previous automatic detection methods perform poorly under the cross-patient setting. For example, the performance of the state-of-the-art model (Ke et al., 2021), which achieves 99% sensitivity in the traditional epileptic detection setting on the CHB-MIT dataset, drops to below 60% in the cross-patient setting. How to improve the generalization ability under the inter-patient EEG pattern variation is a key problem for practical applications of epileptic seizure detection.

In this paper, we improve the performance of cross-patient seizure detection in two aspects: data augmentation and adversarial optimization. Firstly, we propose the spatio-temporal EEG augmentation method to effectively increase the sample diversity, by incorporating more spatial information from different channels. In this way, the imbalance problem is alleviated and the generalization ability is improved. Secondly, we design a patient-adversarial neural network (PANN) that adds a discriminator for identity classification during training. With adversarial optimization between the epileptic feature extractor and identity discriminator, PANN is effective in eliminating patient-specific features and learning patient-invariant representations. Experiments show that our methods are successful in cross-patient seizure detection. Compared with other methods, we achieve state-of-the-art results on both the public CHB-MIT dataset and the clinical Peking University Hospital dataset. Our method demonstrates high performances of up to 95% sensitivity on CHB-MIT under 5s window length segments, and about 85% AUC and sensitivity even on the clinical dataset. Our contributions are listed below:

- We propose a new data augmentation method for EEG signals based on the newly discovered property of their distributions. The new data augmentation method can capture both temporal and spatial correlativity to better increase the sample diversity and significantly improve generalization.
- We propose a new adversarial optimization method with a patient-adversarial neural network to learn patient-invariant representations of EEG signals. It can eliminate patient-specific features to improve cross-patient general-

ization.

- 80 • We conduct various experiments on both public and clinical datasets under different settings. Compared with other methods, our model achieves state-of-the-art results in cross-patient seizure detection and provides great potential for clinical applications in practice.

85 We organize the paper as follows: Section 2 discusses the background and related work of epileptic seizure detection. In section 3, we introduce the problem setup of cross-patient seizure detection. Section 4 elaborates on our proposed methods. Section 5 shows experimental results and the model performance. Section 6 is the summary and discussion.

2. Related Work

2.1. Patient-specific Epileptic Seizure Detection

90 EEG-based epileptic detection has been a hot research topic recently. Some early works on seizure detection focus on feature engineering. For example, discrete Fourier transform (DFT) and discrete wavelet transform (DWT) have found popularity in seizure detection and prediction applications (Alotaiby et al., 2014). Similarly, empirical mode decomposition (EMD) (Martis et al., 2012), the local mean decomposition (LMD) (Smith, 2005), and local binary patterns(LBP) (Burrello et al., 2019) have also found a role in these applications.

95 Recently, the mainstream method of seizure detection is using deep neural networks. Specifically, most methods leverage neural network models to integrate the feature extraction and classification procedure in an end-to-end way.
100 For example, Song et al. (2019) utilizes a neural mass model driven method with dynamic features for seizure detection, Wei et al. (2019) uses a 12-layer convolutional neural network (CNN), Abdelhameed et al. (2018) combines a one-dimensional CNN with a long short-term memory (LSTM), Ke et al. (2021) introduces convolution structure and explicitly models the multi-channel characteristics of convolutional features, and many other works (O’Shea et al., 2020;
105

Hu et al., 2018; Wei et al., 2019; Abiyev et al., 2020; Shen et al., 2023) also use CNN for seizure detection. Another common approach for seizure detection is recurrent neural networks (RNN). Abdelhameed et al. (2018) and Hu et al. (2020) adopt the deep bidirectional long short-term memory (Bi-LSTM) network. Most 110 recently, graph representations have also been applied to EEG signal mining. Wang et al. (2020) employs graph representation to characterize EEG signals and uses a novel deep model called Sequential Graph Convolutional Network (SGCN). Tang et al. (2021); He et al. (2022) represent the spatio-temporal dependencies in EEGs using a graph neural network and propose EEG graph structures that 115 capture the electrode geometry or dynamic brain correlativity. And Jia et al. (2020) propose an attention-based graph residual network.

However, all these methods do not consider the challenging cross-patient generalization problem and perform poorly in the cross-patient setting.

2.2. Cross-Patient Seizure Detection

120 In real medical treatment, the cross-patient setting is the actual mode. There are some existing cross-patient studies, but their performance still has much room for improvement.

Some works have used data augmentation methods to improve the accuracy 125 of cross-patient detection. Wei et al. (2019) first merge any two adjoining clutters or incomplete waves and then use Wasserstein Generative Adversarial Network (WGAN) for data augmentation. Gómez et al. (2020) utilize overlapping data augmentation strategies with time-shift for cross-patient detection. Peng et al. (2022) propose temporal information enhancement for feature augmentation to 130 improve cross-patient behavior. However, their augmentation methods mainly focus on the temporal dimension and do not consider the spatial information as in this work, thus their performance, especially sensitivity, has much room to improve.

Another effective cross-patient epilepsy detection method is to use feature 135 disentanglement to separate patient-specific features and common epilepsy features. Zhang et al. (2020) propose a so-called “adversarial” method to decompose

seizure and patient representation with two branches of neural networks for classification of seizure and patients as well as reconstruction. However, different from this work, they do not actually perform adversarial optimization to ensure that patient-specific features are eliminated in the seizure features
140 for seizure classification. Meanwhile, they require double models to extract decomposed features, while our proposed adversarial optimization of one model is more lightweight and efficient. Zhao et al. (2022) manually extract multi-view frequency domain and time-frequency domain features, and use two generators and two discriminators for feature generation and classification of seizure and
145 patients. Similar to (Zhang et al., 2020), they also do not perform real adversarial optimization with opposite objectives for feature elimination as in this work, and require more computation with two branch of models.

Some other methods applied meta learning for the cross-patient problem, such as MUPS (Meta Update Strategy) (Zhu et al., 2020) and MLCL (meta learning
150 on constrained transfer learning) (Duan et al., 2020), which divide training data into multiple meta tasks to construct effective cross-patient representations. This kind of methods involves alternating learning between two neural networks(meta nets and the main network), and both networks acquire knowledge from each other to improve themselves. The limitation of meta learning methods is that
155 the training process cannot guarantee effective representation, and the number of models is twice that of other methods, resulting in higher training difficulty.

Some works try to improve cross-patient settings mainly from the perspective of using effective network structures. For example, Dist-DCRNN (Tang et al.,
160 2021) applies a dist diffusion convolutional recurrent neural network which can model the spatiotemporal dependencies in EEGs, ConvLSTM(Yang et al., 2022) uses the convolutional long short-term memory network for cross-patient seizure detection, Dense CNN(Saab et al., 2020) exploits densely connected inception network trained by imperfect but plentiful archived annotations, neural memory networks (NMNs)(Ahmedt-Aristizabal et al., 2020) uses external memory
165 modules with trainable neural plasticity, etc. The above works are orthogonal to our methods, as those works are basically from the perspective of network

structure, while our methods are from the perspective of data and strategy. The above methods and our methods can be used together.

Some works have also used domain adaption (Nasiri & Clifford, 2021; He
170 & Wu, 2020; Xia et al., 2022) or domain generalization(Ayodele et al., 2020)
with multiple datasets to improve the model’s generalization ability. However,
domain generalization without the requirement of testing or additional data
is preferred for transfer learning, as utilizing both training and testing data
(domain adaption) is severely inconsistent with actual medical treatment.

175 **3. Problem Setup**

Seizure detection aims to distinguish whether a wave in a short time interval is
in the ictal stage and it is formulated as a binary classification task. After splitting
the whole EEG signals into short segments with same window size (t seconds),
we get the EEG segments and denote the available data as $(\mathbf{x}_i, \mathbf{y}_i), i = 1, \dots, N$,
180 where N is the number of segments, $\mathbf{y}_i \in \{0, 1\}$ is a class label, with $\mathbf{y}_i = 1$
corresponding to a seizure period, and $\mathbf{y}_i = 0$ corresponding to a non-seizure
period. The EEG signal of the i -th sample is denoted as $\mathbf{x}_i \in \mathbb{R}^{T \times C}$ where C
is the channel dimension and T is the product of time window size t and data
frequency f . The research goal is to design a classifier to correctly distinguish
185 whether the patient is on seizure according to t -second EEG signal duration.

Specifically, we focus on cross-patient seizure detection in this paper. Under
the basic problem of seizure detection, the cross-patient setting has a sample of
 M patients, where M_D patients are for model training and other M_T patients
are for model testing, with $M_D + M_T = M$. Cross-patient settings are the data
190 pattern of actual medical treatment. Besides, cross-patient seizure detection can
be reproducible, and generalizable to future patients.

The clinical application of cross-patient seizure detection is to provide assis-
tant judgments of seizure and give references of the start/end time of seizure to
help the doctors cure epilepsy. Specifically, we use the pretrained cross-patient
195 seizure model for orderly t -second EEG segments of one epilepsy patient. The

period of the first EEG segment detected to be ictal is marked as the start of the seizure, while the last detected EEG segment is marked as the end of the seizure, with the error range of t seconds.

4. Methodology

200 In this section, we elaborate on the two aspects of our method to solve the cross-patient problem: data augmentation and adversarial optimization. Data augmentation deals with the generalization problem mainly from the class-wise aspect, i.e. how to handle the class-imbalance problem; while adversarial optimization considers the patient-wise aspect, i.e. how to learn patient-invariant representations.
205

4.1. Spatio-Temporal EEG Augmentation (STEA)

As introduced in Introduction, the seizure data are few and the class-imbalance problem exists in the EEG signal. Compared to non-seizure periods, the number of seizure periods is fairly small. Such an imbalance problem will
210 cause performance degradation and poorer generalization, which is more severe for cross-patient generalization.

One key solution to this problem is data augmentation. However, existing few data augmentation methods of EEG signal do not work well in practice as they are mostly limited to simple linear transformation (Li et al., 2020). A possible
215 reason is that they only consider to add small perturbations but ignore the statistical characteristics, for example, the temporary and spatial correlativity over EEG signals from different time and EEG electrode channels.

To this end, we first analyze the temporal and spatial correlation of EEG signals. Specifically, given the data $\mathbf{x} \in \mathbb{R}^{T \times C}$, where T and C correspond to the
220 temporal and channel dimensions, respectively (different channels correspond to electronics at different spatial positions), we flatten the data of different channels into one vector with $T \times C$ elements and make augmentations with correlations among all dimensions. By analyzing the 3-D embedding mapping of the data

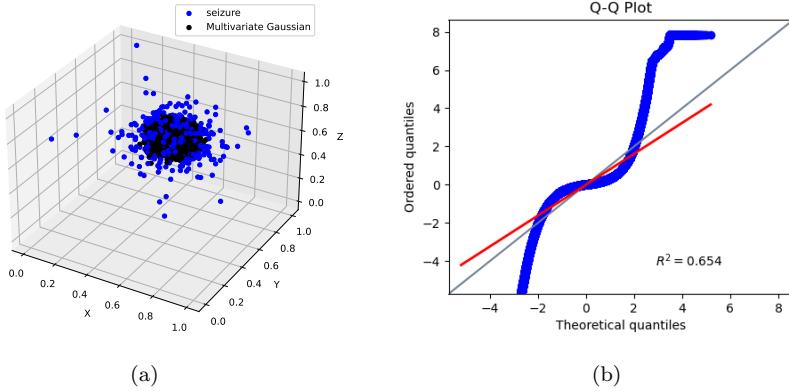


Figure 4: The visualization of the statistical characteristics of EEG signals.(a) The t-SNE visualization of EEG signals from positive seizure class (blue) and multi-Gaussian distribution samples (black) examples. The (X,Y,Z) labels are the corresponding coordinates in the 3D embedding space. (b) The quantile-quantile plot between seizure distribution (blue) and multivariate normal distribution (black). We consider each flattened 2-s seizure segment as a data point, where the flattened 2-s seizure segment is a column vector $\mathbf{x}^i \in \mathbb{R}^{TC}$.

with the t-distributed stochastic neighbor embedding (t-SNE) (Van der Maaten & Hinton, 2008), as shown in Fig. 4(a), we find that the EEG data may follow a distribution similar to the multidimensional Gaussian distribution. Then, we make the quantitative verification of the relation between seizure distribution and multidimensional Gaussian distribution. As shown in the quantile-quantile plot in Fig. 4(b), the two probability distributions' quantiles against each other are plotted in blue, and the best-fit linear regression for the data is plotted in red which is close to the ideal 45° line $y = x$. The coefficient of determination $\mathbf{R}^2 = 0.654$, which means a good fit when $\mathbf{R}^2 > 0.5$. \mathbf{R}^2 is the larger the better, and $0 \leq \mathbf{R}^2 \leq 1$. The \mathbf{R}^2 is calculated from all 2-second seizure segments (total number 3372) from 18 patients in the PKU1st dataset. Through the quantile-quantile plot, it can be seen that the population of EEG signals has close connectivity with multidimensional Gaussian distribution. This result demonstrates that the Gaussian distribution can account for a large part of the variety of seizure EEG signals, and therefore it can be an effective surrogate

distribution for data augmentation.

²⁴⁰ Owing to the above analyses, we propose the following spatio-temporal EEG augmentation (STEA) method to generate additional training data through the statistics of training data: It is important to note that the covariance

Algorithm 1 Spatio-Temporal EEG Augmentation (STEA)

Input: EEG seizure segments $\{\mathbf{X}_i\}$ in training set,

the channel number C , time window size t ,

data frequency f , time dimension $T = t \times f$

Procedure:

#Stage 1. reshape

Flatten each segment to 1-D vector $\mathbf{x}^i \in \mathbb{R}^{TC}$.

#Stage 2. calculate the statistical information

calculate the mean μ and covariance matrix Σ , where

$$\mu = \frac{1}{N} \sum_{i=1}^N \mathbf{x}^i, \quad \Sigma = (\Sigma_{pq})_{TC \times TC},$$

$$\Sigma_{pq} = \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_p^i - \mu_p)(\mathbf{x}_q^i - \mu_q).$$

#Stage 3. generate new data

generate the additional training data for seizure class by sampling $\hat{\mathbf{x}} \sim \mathcal{N}(\mu, \Sigma)$. Then reshape the vector to $\hat{\mathbf{X}} \in \mathbb{R}^{T \times C}$ for new data in the training set.

return $\hat{\mathbf{X}}$

matrix in Step 2 calculates the connectivity between any two channels in any two moments. This enables our STEA method to generate reasonable training samples considering both temporal and spatial correlation among existing training data. In practice, the augmentation is conducted after preprocessing of the raw data (the preprocessing details can be found in Section 5.1). An illustration of the data augmentation method is presented in Fig. 5. As the EEG signals are measured by electrodes placed in a manifold, they have non-Euclidean structures. ²⁴⁵ We can treat the EEG signal as a graph flow. Then the covariance matrix in our method can be viewed as calculating the correlation between two vertices in the graph. Some previous works try to model the spatial correlation of EEG signals

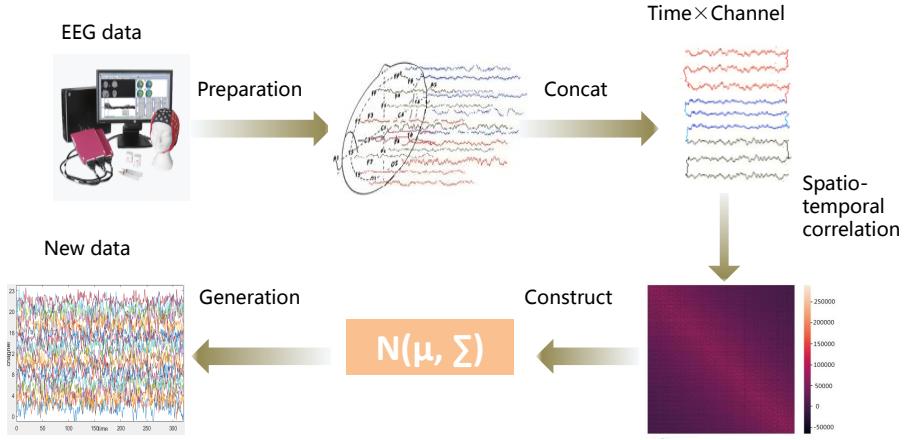


Figure 5: Illustration of the process of STEA. We first pre-compute the statistical information, i.e. the mean and covariance matrix, of the flattened training data. Then during training, we generate additional training data from the multivariate Gaussian distribution.

with graph neural networks (Wang et al., 2020; Tang et al., 2021; Jia et al., 2020), which calculates correlations only between adjacent vertices. In contrast, 255 our method considers more spatial correlations between nonadjacent vertices and also takes temporal dimensions into consideration. Additionally, our method can be viewed as introducing some randomness based on correlations, which can improve the robustness to uncertainty and the generalization ability of the model.

260 There are several advantages of our data augmentation method. STEA can not only alleviate the class-imbalance problem but also enlarge the sample diversity to cover more potential distributions so that the generalization ability to a different data distribution is improved. Therefore, it can largely improve performance under the cross-patient setting. Besides, our STEA method is 265 simple but highly effective, and can be combined with existing augmentation methods together.

As a comparison, we also consider some simple data augmentation (SDA) methods for EEG signals, e.g. add or subtract a constant value at every time point, orderly add and subtract computation, add random noise, scale the EEG

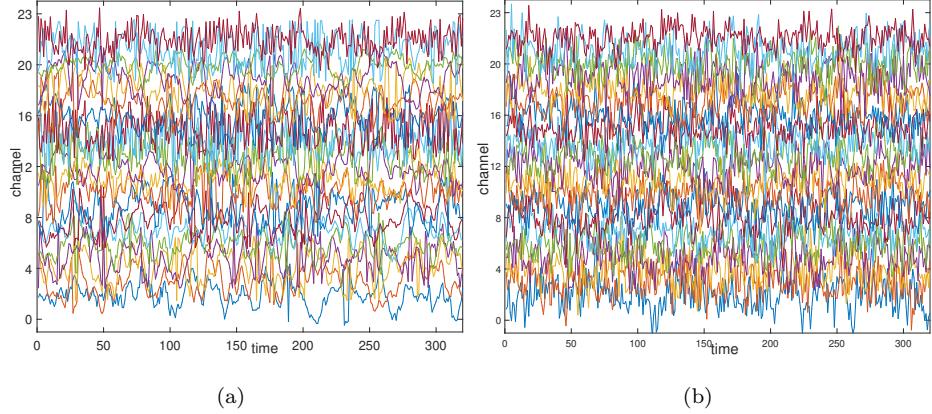


Figure 6: The visualization of STEA data augmentation. (a) the raw EEG signal; (b) the generated EEG signal.

²⁷⁰ signal, etc. And we also compare with the multivariate independent Gaussian generation (MIGG), which assumes that the signal from different channels is independent. Experiments in Section 5 will show the superiority of our method over them.

²⁷⁵ We illustrate the synthetic data of the seizure periods on CHB-MIT by STEA in Fig 6. It shows that the generated signal is similar to the real signal, which verifies the effectiveness of our data augmentation. Besides, the generated data can well learn the representation of seizure EEG signals.

4.2. Patient-Adversarial Neural Networks (PANN)

²⁸⁰ Another difficulty in cross-patient detection is the different distributions between the training set and testing set. However, the existing machine learning algorithms need the same distribution of training and testing. To solve this problem, we propose the adversarial optimization method combined with the newly designed patient-adversarial neural network (PANN) to learn patient-invariant representations. Compared with the common neural network model, PANN adds a discriminator to distinguish the identity of patients. With alternate adversarial training between the common model and the identity discriminator,

the identity-specific features can be largely restrained and the learned patient-invariant representation can improve the cross-patient generalization.

Specifically, the structure of PANN is shown in Fig. 7. PANN is composed of a feature extractor, a common classifier, and an identity classifier, whose parameters are represented by θ_f , θ_y and θ_I , respectively. Similar to common models, the feature extractor aims to learn representations of EEG signals, and the common classifier tries to distinguish the representations between seizure signals and non-seizure signals. The additional identity classifier plays the role of the discriminator to distinguish different patients.

Each EEG signal \mathbf{x} in the training data has two labels, i.e. the seizure label y and the identity label id . During training, the feature extractor maps the input \mathbf{x} to the embedding space. Then the output of the feature extractor is fed into two classifiers. One classifier is used for seizure detection with the prediction vector \hat{y} , and the other is to classify the identity with the prediction vector \hat{id} .

The objective function of PANN is composed of two competing losses: the identity classification loss \mathcal{L}_{id} and the adversarial identity-confusion loss \mathcal{L}_{adv} . The specific definitions are shown in Eqn. (1).

$$\begin{aligned}\mathcal{L}_{seiz}(y_i, \hat{y}_i) &= -(y_i \ln(\hat{y}_i) + (1 - y_i) \ln(1 - \hat{y}_i)), \\ \mathcal{L}_{id}(id_i, \hat{id}_i) &= -\sum_{j=1}^N id_{i,j} \cdot \ln(\hat{id}_{i,j}), \\ \mathcal{L}_{adv} &= \mathcal{L}_{seiz} - \lambda \mathcal{L}_{id}.\end{aligned}\quad (1)$$

The \mathcal{L}_{seiz} uses the binary cross entropy loss for the binary seizure classification task, and the \mathcal{L}_{id} uses the cross entropy loss for the identity classification task, where N represents the number of patients and $\hat{id}_{i,j}$ represents the prediction probability of i -th sample's id being j . The total identity-confusion loss \mathcal{L}_{adv} helps build domain-invariant features, where λ is the hyperparameter to balance seizure detection and identity confusion.

$$\theta_I = \arg \min \left(\frac{1}{n} \sum_{i=1}^n \mathcal{L}_{id}(id_i, \hat{id}_i) \right), \quad (2)$$

$$(\theta_f, \theta_y) = \arg \min \left(\frac{1}{n} \sum_{i=1}^n \mathcal{L}_{seiz}(y_i, \hat{y}_i) - \frac{\lambda}{n} \sum_{i=1}^n \mathcal{L}_{id}(id_i, \hat{id}_i) \right). \quad (3)$$

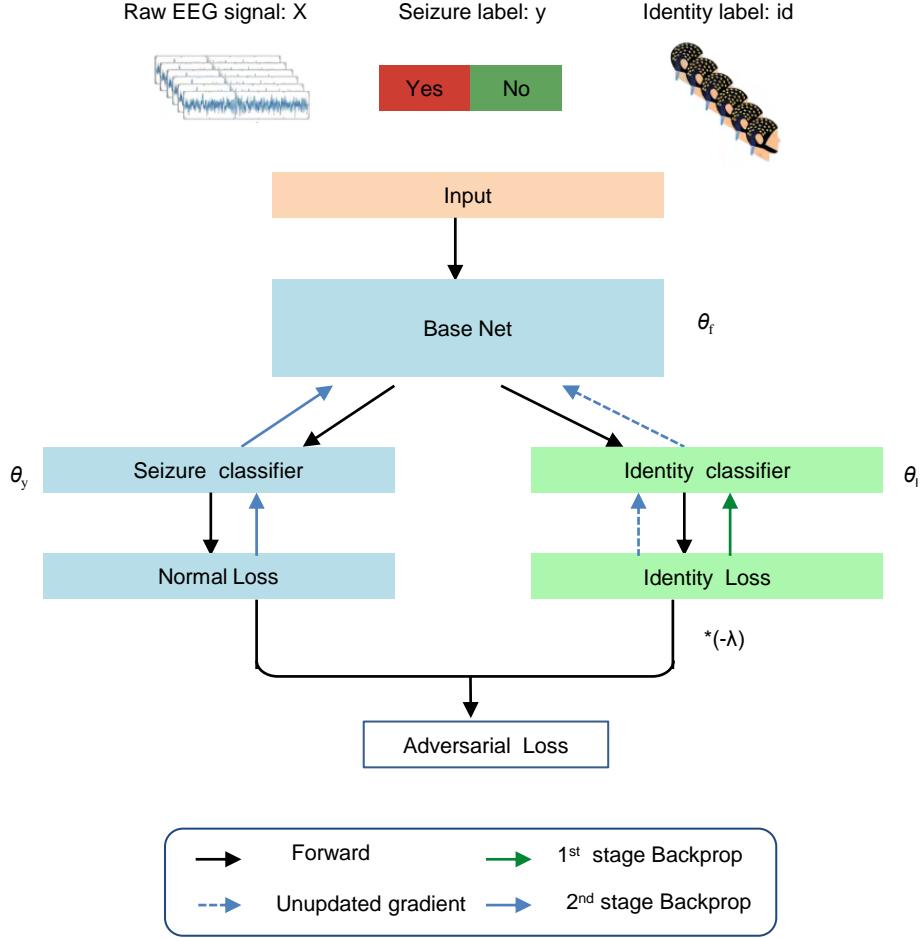


Figure 7: Illustration of the patient-adversarial neural network. The black lines represent the information propagation during the forward procedure. The blue and the green lines represent the alternate optimization procedure for the common classification model and the identity classifier, respectively, where the dashed lines mean propagating gradients without updating parameters in the path.

During training, we alternately minimize two losses \mathcal{L}_{id} and \mathcal{L}_{adv} with different directions for parameters, as illustrated in Algorithm 2. In the first stage of PANN training, we aim at minimizing \mathcal{L}_{id} , so that the identity classifier can clearly distinguish different patient identities, as Eqn. (2) shows. Then in the second training stage, the strategy minimizes \mathcal{L}_{adv} and updates the parameters

of the base net and seizure classifier as shown in Eqn. (3). The model has two tasks to accomplish: one is the normal seizure detection task, and the other is to eliminate the bias of individualized features.

$$\max_{\theta_f} \min_{\theta_I} \left[\frac{1}{n} \sum_{i=1}^n \mathcal{L}_{id}(id_i, \hat{id}_i) \right]. \quad (4)$$

Training the PANN network iteratively leads to the key min-max optimization problem in Eqn. (4). Our specially designed training strategy makes the feature extractor confused with different patient identities while the parameter of the identity classifier is updated in the opposite direction. Then the PANN model extracts patient-invariant features instead of individual features, so that the detection ability can be reproducible and generalizable to unseen patients. The newly designed patient individual classifier plays a role in the qualitative analysis of individual identification ability. Besides, Stage 1 of training makes sure that the result of the identity classifier is enough reliable and accurate. Stage 2 with the unupdated identity classifier makes sure that the confusion of identity classification is the contribution of the feature extractor instead of the identity classifier.

After training, the neural network model can extract the common features among different patients, i.e. the identity cannot be distinguished based on the feature. Such patient-invariant representation can better generalize to unseen different patients, and improve the cross-patient performance.

4.3. Implementation Details

We combine the STEA method with the PANN network together for cross-patient seizure detection. Firstly, we use the STEA method to augment enough seizure samples for training. Secondly, the feature extractor in PANN automatically extracts patient-invariant features through the training data input. Lastly, the seizure classifier in PANN uses the extracted feature to output the class-aware probability, and labels the input as the category with the highest probability. The overall framework can be found in the Appendix.

Algorithm 2 Training procedure of patient-adversarial neural network

Input: EEG signal \mathbf{x} , seizure label y , identity label id ,

iteration number K , patient adversarial neural
network parameters $\{\theta_f, \theta_y, \theta_I\}$

for iteration k from 0 **to** K **do**

#Stage 1. $\min \mathcal{L}_{id}$

Forward through individual branch

$$\hat{id} = g_I(f(\mathbf{x})).$$

$$\mathcal{L}_I = \mathcal{L}_{id}(id, \hat{id})$$

#Unupdate identity classifier params

$$\theta_I \leftarrow \theta_I - \eta \frac{\partial \mathcal{L}_I}{\partial \theta_I}.$$

#Stage 2. $\min (\mathcal{L}_{seiz} - \lambda \mathcal{L}_{id})$

Forward through two branches

$$\hat{y} = g_y(f(\mathbf{x})), \quad \hat{id} = g_I(f(\mathbf{x})).$$

$$\mathcal{L}_I = \mathcal{L}_{id}(id, \hat{id}), \quad \mathcal{L}_s = \mathcal{L}_{seiz}(y, \hat{y})$$

#Update feature extractor and seizure classifier

$$\theta_f \leftarrow \theta_f - \eta \left(\frac{\partial \mathcal{L}_n}{\partial \theta_f} - \lambda \frac{\partial \mathcal{L}_I}{\partial \theta_f} \right).$$

$$\theta_y \leftarrow \theta_y - \eta \frac{\partial \mathcal{L}_s}{\partial \theta_y}.$$

end for

325 As for the specific model structure, we choose successful deep neural network
models as the feature extractor and the seizure classifier. The Channel-Weighted
Squeeze-and-Excitation Network (CW-SRNet) that was proposed and discussed
in our previous research (Ke et al., 2021) is a flexible neural network architecture
that uses the attention mechanism to capture the different importance of EEG
330 channels automatically and dynamically. We apply this model as our feature
extractor backbone. A detailed illustration of the architecture is shown in the
appendix. The additional identity classifier of our PANN has two options. One
is a simple fully connected layer. The other is composed of one fully connected
layer with 100 neurons, one batch normalization layer, one ReLU activation
335 layer, and finally a fully connected classification layer.

We consider the common setting and the new “finetune” setting together in our experiments. The common cross-patient setting in our experiments is to train the model on training data from several individuals and test the performance on testing data from different patients, as illustrated in Fig. 3(a). Additionally,
340 we will consider a further setting where there are few-shot data from testing patients to slightly finetune the model for better generalization. This may also correspond to some real-world conditions that after the first arrival of new patients, we can collect their data and fine-tune the model so that the model can work better during their following diagnosis and treatment. This can also
345 validate the flexibility of our model to adapt to new patients. In the following experiments, we take both settings into consideration. The latter is marked as “finetune” in the results and other results belong to the former setting by default.

5. Experiments

In this section, we first evaluate the proposed methods in various settings
350 and we denote the combination of STEA and PANN as Augmented Adversarial Network (AAN). Besides, we also conduct a series of experiments for a comprehensive understanding of our proposed methods.

5.1. Dataset and Preprocessing

In our experiments, we use one public seizure dataset CHB-MIT (Shoeb & Guttag, 2010) and one clinical seizure dataset PKU1st as shown in Table 1. CHB-MIT dataset was collected by the Children’s Hospital Boston. There are 23 patients’ EEG signals which were sampled at 256 Hz. It is the most commonly used public dataset for EEG detection. All electroencephalograms were collected using the international 10–20 system of EEG electrode positions and nomenclature, and electroencephalograms were recorded using 23 channels.
360 Note that two seizure datasets are imbalanced, i.e., the amount of data for each class varies a lot.

The PKU1st dataset was the latest EEG data collected from the Department of Pediatrics of Peking University First Hospital and approved by the Ethics

Table 1: The information of seizure datasets.

Datasets	#Subjects	EEG channels	Place method	Sampling frequency	Seizure	Non seizure	Total record
CHB-MIT	23	21	international 10–20 system	256Hz	3.26h	964.59h	967.85h
PKU1st	19	19	international 10–20 system	500Hz	0.99h	72.18h	73.17h

Table 2: Comparison between the proposed method and conventional works for cross-patient seizure detection. “window” means time window size (s). “test N” means the number of test patients. The italics mean using additional few data for finetuning. And CW-SRNet (Ke et al., 2021) is our baseline method.

Dataset	Method	window	size	test number	Sensitivity	AUC	GMean	RAccuracy	Specificity
CHB-MIT	MIDS+WGAN (Wei et al., 2019)	5	1	74.08%	-	82.76%	83.27%	92.46%	
	CW-SRNet(Ke et al., 2021)	5	1	42.67%	96.18%	64.97%	70.08%	98.93%	
	AAN (ours)	5	1	99.74%	97.56%	93.54%	93.73%	87.72%	
	<i>AAN+finetune (ours)</i>	5	1	95.83%	94.15%	90.80%	90.94%	86.05%	
	First Seizures Model(Gómez et al., 2020)	4	1	78.48%	-	88.53%	89.18%	99.60%	
	SDG(Ayodele et al., 2020)	-	-	71.45%	-	73.69%	73.73%	76.0%	
	AAN(ours)	4	1	95.71%	97.98%	94.13%	94.15%	92.59%	
	<i>AAN+finetune(ours)</i>	4	1	99.37%	98.23%	87.32%	88.05%	76.73%	
	DCNN(Hossain et al., 2019)	2	1	93.42%	-	92.45%	92.47%	91.52%	
	AAN(ours)	2	1	95.20%	98.77%	95.53%	95.54%	95.88%	
	<i>AAN+finetune(ours)</i>	2	1	95.43%	92.36%	84.66%	85.27%	75.11%	
	CW-SRNet(Ke et al., 2021)	2	9	43.33%	79.55%	64.91%	70.29%	97.24%	
PKU1st	AAN (ours)	2	9	61.31%	88.23%	74.58%	76.02%	90.72%	
	<i>AAN+finetune(ours)</i>	2	9	90.38%	90.95%	81.04%	81.53%	72.67%	
	IBA(Zhao et al., 2022)	1	1	77.78%	93.61%	81.79%	83.31%	88.84%	
	AAN (ours)	1	1	92.43%	93.80%	85.71%	86.40%	80.37%	
	CW-SRNet(Ke et al., 2021)	5	1	2.91%	62.75%	16.58%	48.71%	94.51%	
	ConvLSTM(Yang et al., 2022)	5	1	3.9%	36.10%	16.58%	37.2%	70.5%	
	Dense CNN(Saab et al., 2020)	5	1	23.3%	79.7%	47.71%	60.5%	97.7%	
	Dist-DCRNN (Tang et al., 2021)	5	1	31.1%	72.4%	53.37%	61.35%	91.6%	
	Dist-DCRNN+AAN (ours)	5	1	67.0%	87.1%	77.18%	77.95%	88.9%	
	AAN (ours)	5	1	83.50%	86.59%	83.11%	83.11%	82.72%	
	<i>AAN+finetune (ours)</i>	5	1	81.93%	87.97%	82.43%	82.44%	82.94%	
	CW-SRNet(Ke et al., 2021)	2	9	30.62%	76.59%	54.52%	63.85%	97.08%	
	ConvLSTM(Yang et al., 2022)	2	9	22.97%	50.66%	42.59%	50.98%	78.98%	
	Dense CNN(Saab et al., 2020)	2	9	41.63%	64.98%	58.27%	61.6%	81.57%	
	Dist-DCRNN (Tang et al., 2021)	2	9	42.15%	56.65%	52.78%	54.12%	66.09%	
	Dist-DCRNN+AAN (ours)	2	9	47.8%	70.8%	62.61%	64.9%	82.0%	
	AAN (ours)	2	9	56.89%	72.91%	66.58%	67.40%	77.91%	
	<i>AAN+finetune (ours)</i>	2	9	82.24%	84.99%	76.06%	76.29%	70.60%	

365 Committee of the Peking University First Hospital (2021-225). It is composed of EEG recordings from 19 patients which were sampled at 500 Hz. And electroencephalograms were collected using 19 channels. The dataset details can be found in the appendix.

For both datasets, we preprocess the raw EEG data before training with the 370 following procedure. Firstly, the frequency of data was downsampled from 256 Hz to 64 Hz or from 500 Hz to 50 Hz, which can reduce data noise and memory

costs. Secondly, we split the EEG recording of patients into many short seconds segments, while we use a random sampling technique so that the ratio of positive samples to negative samples is about 1:5. The time window size in experiments
 375 is chosen to be 1 second, 2 seconds, 4 seconds or 5 seconds for a fair comparison with other methods

Then we split a portion of patients for model training, and the rest patients for testing. For training data, we conduct STEA augmentation for the seizure class to generate enough samples, so that the ratio of positive samples to negative samples in training data is about 1:1. There are two kinds of train test split
 380 in our experiments, fixed test set or leave one out mode. And the number of test patients is chosen to be 1 or 9. The specific setting keeps consistency with comparison methods.

5.2. Metrics

In our experiments, four statistical indicators are used for the performance evaluation, namely sensitivity, the area under the receiver operating curve (AUC), GMean, and RAccuracy. Both AUC and GMean are common metrics for the model performance on imbalanced datasets. RAccuracy is the revised accuracy that pays the same attention to different classes of imbalanced datasets. Given the True Positive (TP), False Positive (FP), True Negative (TN), False Negative (FN), $r = \frac{TN+FP}{TP+FN}$, some indicators are defined as follows:

$$Sensitivity = \frac{TP}{TP + FN}, \quad (5)$$

$$Specificity = \frac{TN}{TN + FP}, \quad (6)$$

$$RAccuracy = \frac{r \cdot TP + TN}{r \cdot (TP + FN) + TN + FP}, \quad (7)$$

$$GMean = \sqrt{Sensitivity \times Specificity}. \quad (8)$$

385 Besides, we also use ROC and AUC as our metrics. The receiver operating characteristic (ROC) curve can illustrate the diagnostic ability of a seizure detection model. AUC stands for the area under the ROC curve.

Sensitivity signifies the rate of correct identification of seizure events from an EEG recording, which is the most concerned indicator for cross-patient detection.
390 RAccuracy indicates the inherent ability to correctly distinguish whether it is on seizure or not. AUC is the average sensitivity of all possible specificity. The AUC metric can measure the overall learning ability between positive and negative samples.

5.3. Baseline Models

395 We use MIDS+WGAN(Wei et al., 2019), IBA(Zhao et al., 2022), First Seizures model(Gómez et al., 2020), SDG(Ayodele et al., 2020), DCNN(Hossain et al., 2019), CW-SRNet(Ke et al., 2021), Dist-DCRNN(Tang et al., 2021), ConvLSTM(Yang et al., 2022) and Dense CNN (Saab et al., 2020) as nine baseline models to compare their performances with ours. Apart from the nine
400 latest models on cross-patient seizure detection, we also choose one non-cross-patient model for comparison. The brief descriptions of the nine baseline models are as follows:

- **MIDS+WGAN**(Wei et al., 2019) starts with MIDS (the merger of the increasing and decreasing sequences) data prepossessing and WGAN (Wasserstein Generative Adversarial Nets) data augmentation, and then employs a 15-layer CNN architecture for cross-patient detection.
405
- **IBA**(Zhao et al., 2022) is a kind of multi-view learning method with feature disentanglement, and it uses two GAN models for cross-patient seizure detection.
- **First Seizures model**(Gómez et al., 2020) implements a fully convolutional network (FCN) with a time-shift between consecutive windows of 1/4 s for the seizure period of the dataset.
410
- **SDG**(Ayodele et al., 2020) uses the technique of supervised domain generalization with additional much more datasets for training. The backbone model is a CNN architecture for feature extraction, followed by an LSTM layer for seizure detection.
415

- **DCNN**(Hossain et al., 2019) applies a 7-layer deep CNN model with cropped training strategy after using a sliding window of 2 seconds to crop the EEG signals.
- 420 • **CW-SRNet**(Ke et al., 2021) exploits a custom CNN architecture composed of CW-Block with attention mechanism and SE-Block. CW-SRNet is the **non-cross-patient** (patient-specific) model with the state-of-the-art performance.
- 425 • **Dist-DCRNN**(Tang et al., 2021) is a dist diffusion convolutional recurrent neural network which can model the spatiotemporal dependencies in EEGs.
- 430 • **ConvLSTM**(Yang et al., 2022) uses the convolutional long short-term memory network for cross-patient seizure detection.
- 435 • **Dense CNN**(Saab et al., 2020) exploits densely connected inception network trained by imperfect but plentiful archived annotations.

The experiment settings of these methods are different in the aspects of time window size, test patient number, etc., and the details are shown in Table 2. As there is no standard evaluation in existing works and model performance depends on the specific setting being analyzed, we conduct all-round experiments under different settings for fair comparisons with the above methods.

In experiments, we implement our methods by Python3 and train our models by the Adam optimizer (Kingma & Ba, 2015) for 50 epochs. We choose the batch size as 64 and the learning rate is set as a constant value and will be specified for different settings. No dropout or weight decay is applied. The hyper-parameter λ is set as 10^{-3} , which is determined by grid search in the range of 10^{-5} to 10^{-2} at a log-uniformly interval on the validation set.

5.4. Comparison of Performance

We compare the performance of our AAN model against various state-of-the-art methods. We compare their performance in terms of sensitivity, AUC, RAccuracy, and GMean. And we evaluate our method on various datasets

⁴⁴⁵ including CHB-MIT and PKU1st. We also provide the result combined with the pretrain-finetune strategy which selects 20% data of the test patient for finetuning.

Table 2 shows the comparisons over MIDS+WGAN(Wei et al., 2019), IBA(Zhao et al., 2022), First Seizures model(Gómez et al., 2020), SDG(Ayodele et al., 2020), DCNN(Hossain et al., 2019), CW-SRNet(Ke et al., 2021), Dist-DCRNN(Tang et al., 2021), ConvLSTM(Yang et al., 2022) and Dense CNN (Saab et al., 2020) with our AAN method. Compared with other models, our method achieves better performance with four higher measurements (sensitivity, AUC, GMean, and RAccuracy) on both public and clinical datasets, due to effective adversarial optimization and powerful data augmentation. Besides our method can achieve more than 80% sensitivity and AUC under 5s window length and cross-patient setting, which can satisfy the need of the clinical application. It can be also seen that non-cross-patient models such as CW-SRNet cannot handle the cross-patient setting well. Our method significantly improves the performance of cross-patient seizure detection.
⁴⁵⁰
⁴⁵⁵
⁴⁶⁰

We also provide the model performance on the finetune setting, which is not considered for the above comparisons. It can be seen that “finetune” can help improve the model performance, especially under much tougher conditions, e.g. 2-second segments with 9 test patients, while its effect is not obvious under easy settings.
⁴⁶⁵

Among different settings, smaller window lengths and more test patients mean much more difficulty, as it requires more accurate detection and broader generalization. Besides, compared with CHB-MIT, the clinical PKU1st dataset has newer EEG data and less recording time, which is harder for seizure detection.
⁴⁷⁰ Moreover, the common setting can be more difficult than the “finetune” setting as it has no knowledge on test data.

In addition, the leave-one-out result of cross-patient detection on the CHB-MIT dataset is shown in Fig. 8, where we select each patient alone as the testing set and other patients for training. Among all patients, we have an average of 88.04% sensitivity, 90.98% AUC, 84.51% GMean, and 85.44% RAccuracy under
⁴⁷⁵

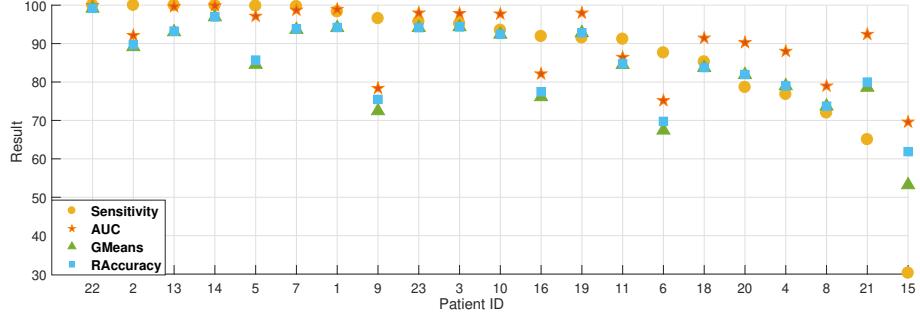


Figure 8: The leave-one-out (LOO) result of cross-patient detection on the CHB-MIT dataset. Patients in X-axis are sorted by sensitivity.

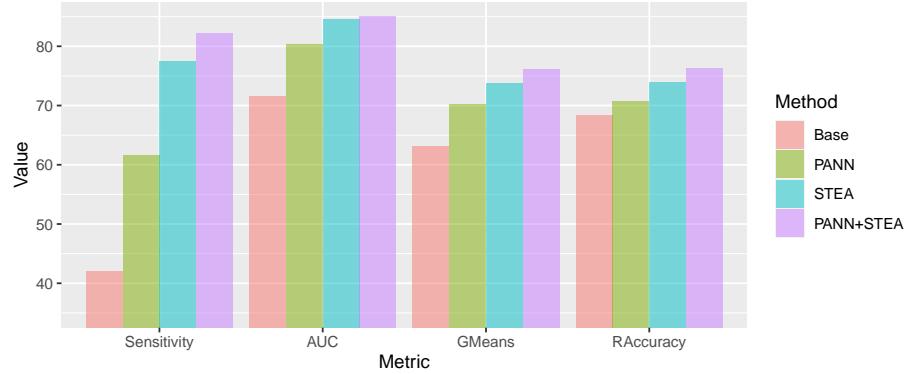


Figure 9: The ablation study of our method conducted on the PKU1st dataset.

4 seconds EEG segments. The single result of each patient is shown in Fig. 8 after sorting by sensitivity, which indicates the powerful generalization ability of our method. Besides, our method also has good leave-one-out performance on the PKU1st dataset, the detailed result can be found in Appendix D.

⁴⁸⁰ 5.5. Analysis Experiments

5.5.1. Ablation Study

For clarity of the effect of each part in our methods, we conduct the ablation study on two seconds segments of the PKU1st dataset with 9 patients test, as shown in Fig. 9. We derive four model variants as follows:

- 485 (1) Base: Baseline model CW-SRNet (Ke et al., 2021) without STEA or
PANN.
- (2) PANN: Baseline model with PANN module.
- (3) STEA: Baseline model with STEA method.
- (4) PANN+STEA: Baseline model with PANN and STEA together.
- 490 From the ablation study, for all four metrics, single PANN or STEA outperforms
the baseline model. Besides, the combination of STEA and PANN surpasses the
single one. So we can draw the following conclusions: PANN is beneficial for cross-
patient detection and STEA also contributes to the well-behaved cross-patient
performance. Moreover, our two methods can be used together for performance
495 improvements. In addition, STEA is more important than PANN as STEA
outperforms PANN. We also conduct the ablation study on the CHB-MIT
dataset, which can be found in the supplementary material.

5.5.2. *The Effectiveness Analysis of STEA*

We also compare our data augmentation method STEA with several data
500 augmentation strategies, including SDA, MIGG, and without data augmentation
(W/O DA). It can be seen from Table 3 that STEA achieves the state-of-the-art
performance. Besides, data augmentation can increase the performance on
imbalanced epilepsy datasets which implies that data augmentation is important
for cross-patient models to detect seizures. Moreover, the introduction of STEA
505 improves the model performance a lot, i.e., the sensitivity is increased from
37.68% to 75.9%, and STEA outperforms the other two augmentation methods.
STEA significantly improves the model’s understanding of EEG data.

5.5.3. *The Effectiveness Analysis of PANN*

In addition, we conduct experiments to verify the effectiveness of the PANN
510 structure. We compared the dynamics of the identity accuracy with and with-
out the adversarial optimization during training to show that our method can
effectively eliminate patient-specific features to enable better cross-patient gener-
alization. As shown in Fig. 10, the identity accuracy of adversarial optimization

Table 3: Comparisons of different data augmentation strategies on the PKU1st dataset under 2 seconds segments and 9 test patients without PANN module, where "W/O DA" represents without data augmentation.

Method	sensitivity	AUC	GMean	RAccuracy
W/O DA	37.68%	81.94%	59.49%	65.81%
SDA	67.25%	83.01%	69.54%	69.58%
MIGG	73.91%	85.34%	76.65%	76.70%
STEA	75.90%	86.61%	79.05%	79.12%

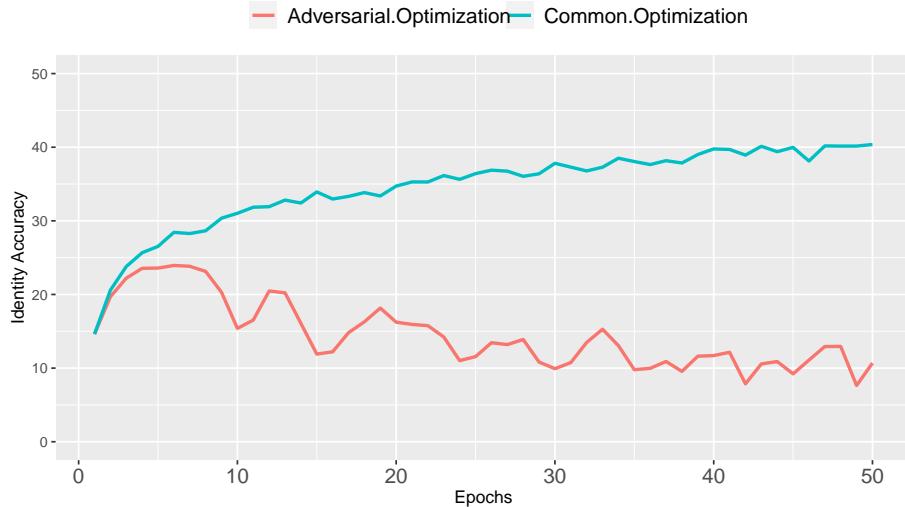


Figure 10: The comparison experiments between adversarial optimization and common optimization, where the identity accuracy is expected to be low so that features are independent to specific patients. Adversarial optimization can achieve decreasing identity accuracy which implies that our method can alleviate patient-specific features and improve cross-patient generalization.

(the red line) gradually decreases with fluctuation, indicating that our model learns to extract patient-invariant features without identity information. On the opposite, the identity accuracy of common optimization (the blue line) continually increases during training, showing that the model may overfit to patient-specific features which will harm the generalization. The contrast of the two optimization strategies shows that our PANN network can effectively learn

520 patient-invariant features to help the cross-patient generalization.

5.5.4. Online processing time

We record the processing time for the inference of an EEG segment by our trained model on one NVIDIA GeForce RTX 3090 GPU card. As shown in Table 4, the processing time is about 0.007 seconds per segment, and it can achieve fast online seizure detection.

Table 4: The online processing time of seizure datasets.

Seizure datasets	Time window size (second)	Processing time (second)
CHB-MIT	4	0.0072±0.00015
PKU1st	5	0.0073±0.000097

525

6. Conclusion

In this paper, we propose spatio-temporal EEG augmentation and patient-adversarial neural networks to significantly improve cross-patient automatic epileptic seizure detection. We first identify the property of the statistical 530 information of seizure EEG signals and propose a novel data augmentation method that uses both temporal and spatial information. Also, we propose a patient-adversarial neural network to eliminate patient-specific features and learn patient-invariant representations (i.e., shared epilepsy-related features) of EEG signals which can be generalized to unseen patients and improves model 535 generalization performance. Experiment results indicate that PANN and STEA are effective approaches for cross-patient detection, with significant improvement in terms of different evaluation metrics. As an automatic cross-patient approach, our method has promising possibilities for deployment in clinical practice. This work could relieve the workload of neurologists, help improve the diagnostic level 540 with inadequate experience, and have great potential significance in auxiliary diagnosis.

Acknowledgements

The work was supported by National High Level Hospital Clinical Research Funding (Multi-center Clinical Research Project of Peking University First Hospital)(2022CR60), National Key R&D Program of China (No. 2018AAA0100300 and 2021YFF0901400), the NSF China (No.s 62276004), NSFC Tianyuan Fund for Mathematics (No. 12026606). This research was partially funded by Novo Nordisk A/S.

Appendix A. Model details

Our baseline model is CW-SRNet (Ke et al., 2021) which is the state-of-the-art model for non-cross-patient seizure detection. The network structure can be seen in Fig. A.11. It is composed of the channel-weighted block (CW-Block) that captures the different importance of EEG channels, and the squeeze-and-excitation Block (SE-Block) that is designed to explicitly model the interdependencies among the channels of its convolutional features.

Appendix B. The PKU1st dataset details

The details of each patient in the PKU1st dataset are shown in Table B.5. The total seizure time is 3558 seconds (0.99h) and the total record time is 263416 seconds (73.17h).

Appendix C. Discussion of statistical characteristics of non-seizure EEG signals

It is worth noting that our STEA method is only applied to seizure signals in order to alleviate the class imbalance problem, because the number of seizure signals is much fewer than non-seizure ones. Therefore, we mainly illustrate seizure signals in Fig 4. We supplement the visualization of non-seizure signals as shown in Fig C.12 (total number 15500 from 18 patients in PKU1st dataset). From the quantile-quantile plot, it can be seen that the non-seizure data is

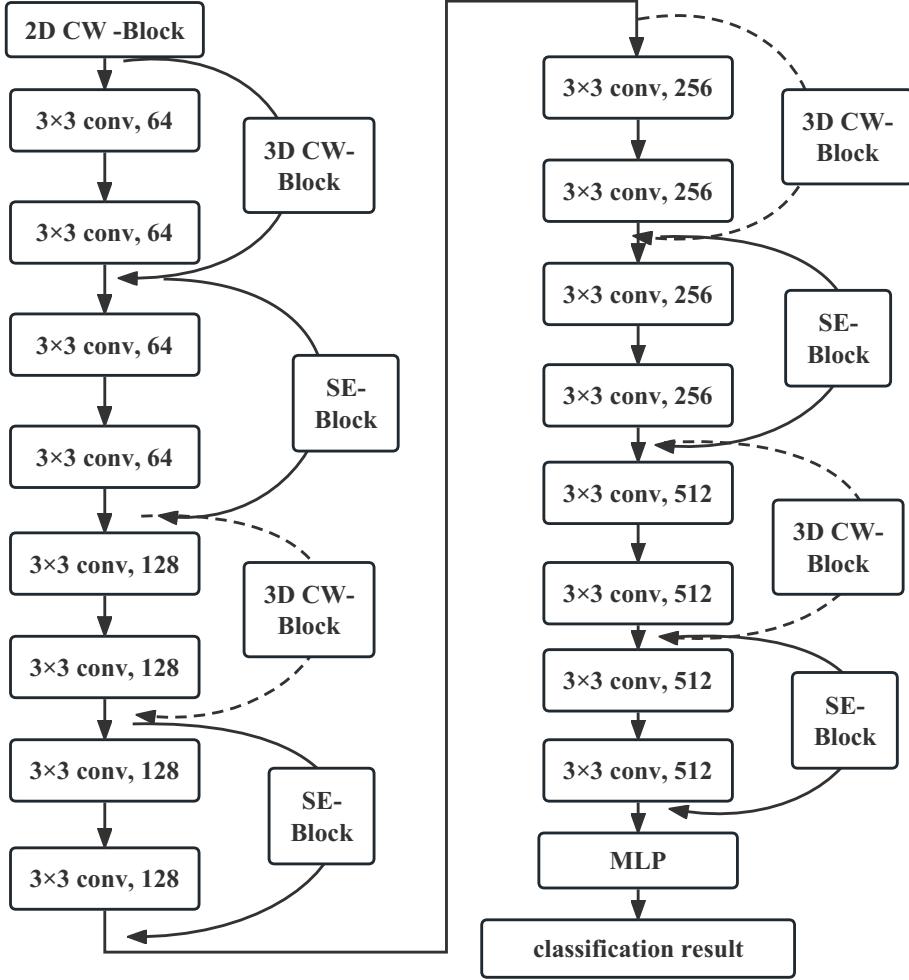


Figure A.11: The structure of CW-SR Net (Ke et al., 2021)

much farther from the multi-gaussian distribution than seizure data, as the corresponding blue data points are farther from the diagonal. We also use quantifiable metric \mathbf{R}^2 to calculate similarity with multi-gaussian distribution.
 570 It can be seen that \mathbf{R}^2 of the non-seizure data is 0.46, which is much lower than \mathbf{R}^2 of the seizure data. This indicates that seizure signals are more proper to use our method, supporting our real practice.

Table B.5: The PKU1st dataset details.

Patient id	Seizure time(s)	Record Time(s)	Seizure number
1	548	11088	2
2	167	16041	1
3	219	11701	2
4	566	11605	2
5	55	29333	2
6	429	11721	3
7	95	12007	3
8	164	10193	2
9	104	8745	1
10	47	10488	2
11	73	16106	2
12	214	11680	3
13	26	10859	2
14	83	10986	2
15	176	32853	4
16	205	10800	2
17	46	11071	2
18	233	10732	2
19	108	15407	1
Total	3558	263416	

Appendix D. More detailed results

575 As shown in Fig.D.13, the overall framework is composed of several steps: after preprocessing the raw EEG data, our method uses the STEA technique to augment enough training positive samples. Then during training, the deep learning model PANN learns to extract patient-invariant features and utilizes extracted features for seizure classification.

580 The detailed data of our method’s ablation study on the PKU1st dataset are shown in Table D.6. The experiment is conducted under 2-second segments and 9 test patients with the pretrain-finetune strategy.

The ablation study is also conducted on the CHB-MIT dataset under 4-second segments and 1 test patient with the pretrain-finetune strategy. The details

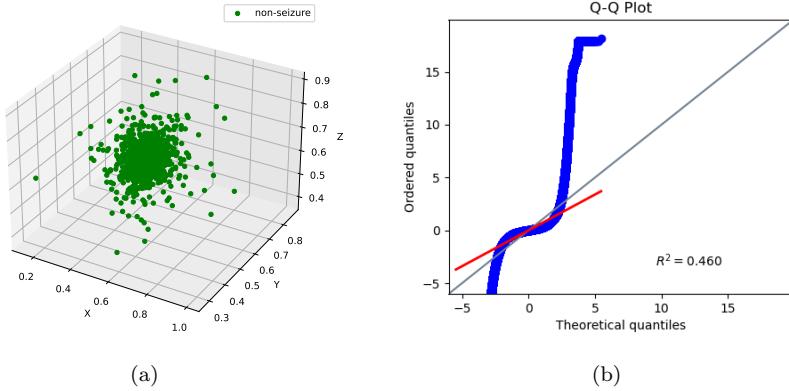


Figure C.12: The visualization of the statistical characteristics of EEG signals. (a) The t-SNE visualization of EEG signals from negative non-seizure class (green). (b) The quantile-quantile plot between non-seizure distribution (blue) and multivariate normal distribution (black).

Table D.6: The ablation study of our method on the PKU1st dataset.

STEA	PANN	sensitivity	AUC	GMean	RAccuracy
✗	✗	42.10%	71.49%	63.08%	68.31%
✗	✓	61.62%	80.27%	70.14%	70.73%
✓	✗	77.45%	84.55%	73.77%	73.86%
✓	✓	82.24%	84.99%	76.06%	76.29%

585 can be found in Table D.7 and Fig. D.14. From the ablation study, for all four
 metrics, single PANN or STEA outperforms the baseline model. Besides, the
 combination of STEA and PANN surpasses the single one. The performance
 result is consistent with the ablation study in the PKU1st dataset.

590 The detailed leave-one-out result on the CHB-MIT dataset with 4-second
 segments and 1 test patient is presented in Table D.8. The concrete metrics of
 each patient for testing are shown in rows.

595 Under the leave-one-out setting, we also test our method's performance on the
 PKU1st dataset, and our method can achieve an average of 77.54% sensitivity,
 87.14% AUC, 78.64% GMean and 79.77% RAccuracy under 5 seconds EEG
 segments. The detailed result can be found in Fig.D.15 and Table D.9. It shows

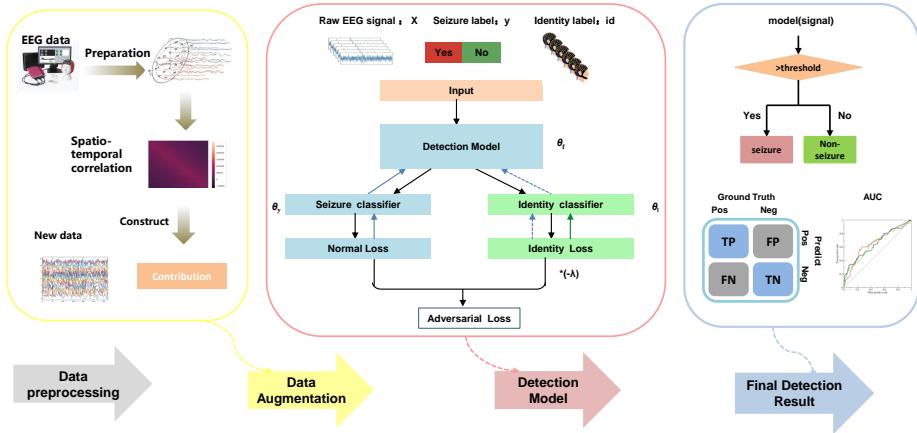


Figure D.13: The overall framework of our methods

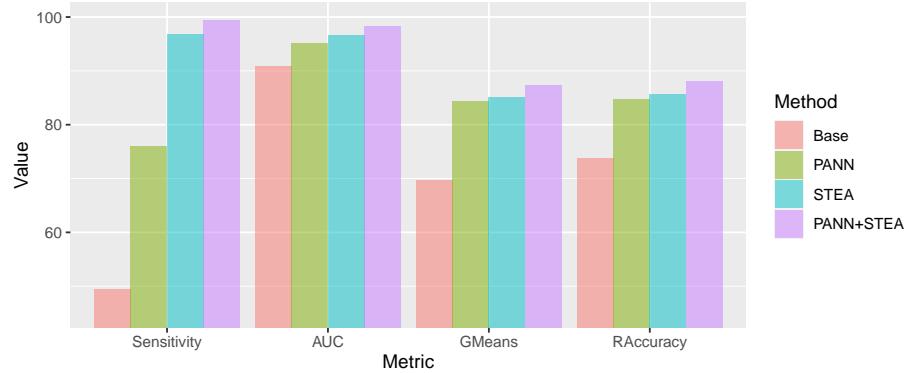


Figure D.14: The ablation study on CHB-MIT dataset

that our method can perform well in the leave-one-out setting.

The effectiveness analysis experiment of PANN is conducted on the CHB-MIT dataset under 4-second segments and 1 test patient without the pretrain-finetune strategy.

600 References

- Abdelhameed, A. M., Daoud, H. G., & Bayoumi, M. (2018). Epileptic seizure detection using deep convolutional autoencoder. In *2018 IEEE International Workshop on Signal Processing Systems (SiPS)* (pp. 223–228). IEEE.

Table D.7: The ablation study of our method on the CHB-MIT dataset.

STE A	PANN	sensitivity	AUC	GMean	R Accuracy
×	×	49.53%	90.8%	69.64%	73.73%
×	✓	76.03%	95.23%	84.30%	84.75%
✓	×	96.85%	96.66%	87.32%	88.05%
✓	✓	99.37%	99.87%	97.80%	97.83%

Table D.8: The leave-one-out result on the CHB-MIT dataset.

Patient	sensitivity	AUC	GMean	R Accuracy
Chb01	98.31	98.93	94.11	94.20
Chb02	100.00	92.12	89.14	89.73
Chb03	95.19	97.84	94.33	94.34
Chb04	76.80	88.00	78.97	79.01
Chb05	99.81	97.14	84.54	85.71
Chb06	87.61	75.17	67.42	69.75
Chb07	99.68	98.72	93.63	93.82
Chb08	71.97	78.97	73.69	73.71
Chb09	96.54	78.36	72.47	75.47
Chb10	93.5	97.72	92.39	92.40
Chb11	91.18	86.44	84.49	84.74
Chb13	99.97	99.65	93.02	93.26
Chb14	99.95	99.85	96.91	96.96
Chb15	30.28	69.58	53.20	61.88
Chb16	91.89	82.12	76.12	77.47
Chb18	85.26	91.45	83.74	83.76
Chb19	91.52	97.96	92.81	92.82
Chb20	78.63	90.27	81.88	81.9
Chb21	65.03	92.42	78.57	79.98
Chb22	100.00	99.92	99.17	99.17
Chb23	95.71	97.98	94.13	94.1
Average	88.04	90.98	84.51	85.44

Abiyev, R., Arslan, M., Bush Idoko, J., Sekeroglu, B., & Ilhan, A. (2020).

605 Identification of epileptic EEG signals using convolutional neural networks.
Applied Sciences, 10, 4089.

Ahmad, M., Saeed, M., Saleem, S., & Kamboh, A. M. (2016). Seizure detection

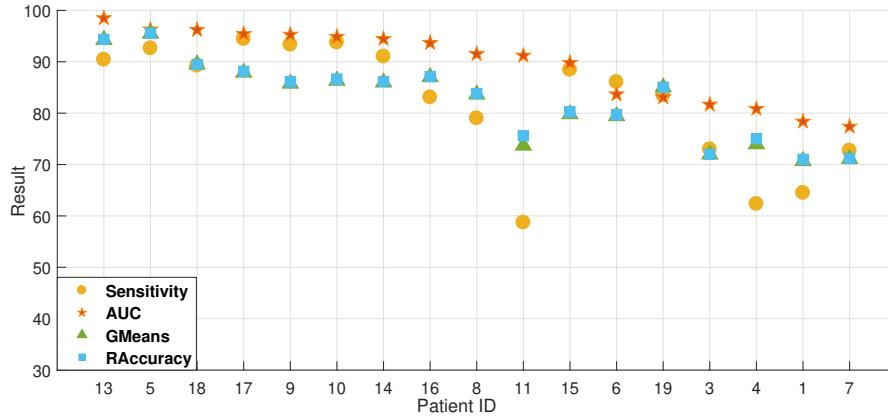


Figure D.15: The leave-one-out (LOO) result of cross-patient detection on the PKU1st dataset. Patients in X-axis are sorted by AUC.

using EEG: A survey of different techniques. In *2016 International Conference on Emerging Technologies (ICET)* (pp. 1–6). IEEE.

610 Ahmedt-Aristizabal, D., Fernando, T., Denman, S., Petersson, L., Aburn, M. J., & Fookes, C. (2020). Neural memory networks for seizure type classification. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (pp. 569–575). IEEE.

Alotaiby, T. N., Alshebeili, S. A., Alshawi, T., Ahmad, I., El-Samie, A., & 615 Fathi, E. (2014). EEG seizure detection and prediction algorithms: a survey. *EURASIP Journal on Advances in Signal Processing*, 2014, 1–21.

Ayodele, K. P., Ikezogwo, W. O., Komolafe, M. A., & Ogunbona, P. (2020). Supervised domain generalization for integration of disparate scalp EEG datasets for automatic epileptic seizure detection. *Computers in Biology and 620 Medicine*, 120, 103757.

Burrello, A., Schindler, K., Benini, L., & Rahimi, A. (2019). Hyperdimensional computing with local binary patterns: One-shot learning of seizure onset and identification of ictogenic brain regions using short-time ieeg recordings. *IEEE Transactions on Biomedical Engineering*, 67, 601–613.

Table D.9: The leave-one-out result on the PKU1st dataset.

Patient	sensitivity	AUC	GMean	RAccuracy
1	64.49	78.37	70.70	71
2	58.37	54.25	49.47	50.15
3	72.97	81.67	71.98	71.99
4	62.34	80.85	73.96	75.04
5	92.63	96.30	95.51	95.56
6	86.06	83.67	79.43	79.69
7	72.73	77.39	71.13	71.15
8	79.01	91.52	83.65	83.79
9	93.33	95.29	85.76	86.07
10	93.75	94.87	86.34	86.63
11	58.73	91.21	73.70	75.61
12	18.75	79.17	42.95	58.58
13	90.41	98.49	94.26	94.34
14	91.03	94.45	86.00	86.14
15	88.44	89.78	79.81	80.23
16	83.08	93.68	87.03	87.12
17	94.44	95.40	87.91	88.14
18	89.24	96.22	89.46	89.47
19	83.50	83.11	85.03	85.05
Average	77.54	87.14	78.64	79.77

625 Daoud, H., & Bayoumi, M. A. (2019). Efficient epileptic seizure prediction based
on deep learning. *IEEE Transactions on Biomedical Circuits and Systems*,
13, 804–813.

Duan, T., Shaikh, M. A., Chauhan, M., Chu, J., Srihari, R. K., Pathak, A.,
& Srihari, S. N. (2020). Meta learn on constrained transfer learning for low
630 resource cross subject eeg classification. *IEEE Access*, 8, 224791–224802.

Gómez, C., Arbeláez, P., Navarrete, M., Alvarado-Rojas, C., Le Van Quyen, M.,
& Valderrama, M. (2020). Automatic seizure detection based on imaged-EEG
signals through fully convolutional networks. *Scientific Reports*, 10, 1–13.

He, H., & Wu, D. (2020). Different set domain adaptation for brain-computer

- 635 interfaces: A label alignment approach. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28, 1091–1108.
- He, J., Cui, J., Zhang, G., Xue, M., Chu, D., & Zhao, Y. (2022). Spatial-temporal seizure detection with graph attention network and bi-directional lstm architecture. *Biomedical Signal Processing and Control*, 78, 103908.
- 640 Hossain, M. S., Amin, S. U., Alsulaiman, M., & Muhammad, G. (2019). Applying deep learning for epilepsy seizure detection and brain mapping visualization. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 15, 1–17.
- Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (pp. 7132–7141).
- 645 Hu, X., Yuan, S., Xu, F., Leng, Y., Yuan, K., & Yuan, Q. (2020). Scalp EEG classification using deep bi-lstm network for seizure detection. *Computers in Biology and Medicine*, 124, 103919.
- Jia, S., Hou, Y., Shi, Y., & Li, Y. (2020). Attention-based graph resnet for motor intent detection from raw EEG signals. *arXiv preprint arXiv:2007.13484*, .
- Jiang, L., He, J., Pan, H., Wu, D., Jiang, T., & Liu, J. (2023). Seizure detection algorithm based on improved functional brain network structure feature extraction. *Biomedical Signal Processing and Control*, 79, 104053.
- 655 Jiang, Y., Wu, D., Deng, Z., Qian, P., Wang, J., Wang, G., Chung, F.-L., Choi, K.-S., & Wang, S. (2017). Seizure classification from eeg signals using transfer learning, semi-supervised learning and TSK fuzzy system. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25, 2270–2284.
- Ke, N., Lin, T., & Lin, Z. (2021). Channel-weighted squeeze-and-excitation networks for epileptic seizure detection. In *2021 IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI)* (pp. 666–673). IEEE.

- Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization.
In *ICLR (Poster)*.
- Li, Y., Liu, Y., Cui, W.-G., Guo, Y.-Z., Huang, H., & Hu, Z.-Y. (2020). Epileptic
665 seizure detection in EEG signals using a unified temporal-spectral squeeze-and-
excitation network. *IEEE Transactions on Neural Systems and Rehabilitation
Engineering*, *28*, 782–794.
- Li, Y., Liu, Y., Guo, Y.-Z., Liao, X.-F., Hu, B., & Yu, T. (2021). Spatio-temporal-
spectral hierarchical graph convolutional network with semisupervised active
670 learning for patient-specific seizure prediction. *IEEE Transactions on Cyber-
netics*, *52*, 12189–12204.
- Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal
of Machine Learning Research*, *9*, 2579–2605.
- Martis, R. J., Acharya, U. R., Tan, J. H., Petznick, A., Yanti, R., Chua, C. K.,
675 Ng, E. K., & Tong, L. (2012). Application of empirical mode decomposition
(emd) for automated detection of epilepsy using EEG signals. *International
Journal of Neural Systems*, *22*, 1250027.
- Mormann, F., Andrzejak, R. G., Elger, C. E., & Lehnertz, K. (2007). Seizure
prediction: the long and winding road. *Brain*, *130*, 314–333.
- 680 Nasiri, S., & Clifford, G. D. (2021). Generalizable seizure detection model using
generating transferable adversarial features. *IEEE Signal Processing Letters*,
28, 568–572.
- O’Shea, A., Lightbody, G., Boylan, G., & Temko, A. (2020). Neonatal seizure
detection from raw multi-channel EEG using a fully convolutional architecture.
685 *Neural Networks*, *123*, 12–25.
- Peng, R., Zhao, C., Jiang, J., Kuang, G., Cui, Y., Xu, Y., Du, H., Shao, J.,
& Wu, D. (2022). TIE-EEGNet: Temporal information enhanced eegnet for
seizure subtype classification. *IEEE Transactions on Neural Systems and
Rehabilitation Engineering*, *30*, 2567–2576.

- 690 Saab, K., Dunnmon, J., Ré, C., Rubin, D., & Lee-Messer, C. (2020). Weak supervision as an efficient approach for automated seizure detection in electroencephalography. *NPJ digital medicine*, *3*, 59.
- 695 Shen, M., Wen, P., Song, B., & Li, Y. (2023). Real-time epilepsy seizure detection based on eeg using tunable-q wavelet transform and convolutional neural network. *Biomedical Signal Processing and Control*, *82*, 104566.
- Shoeb, A. H., & Guttag, J. V. (2010). Application of machine learning to epileptic seizure detection. In *Proceedings of the 27th international conference on machine learning (ICML-10)* (pp. 975–982).
- 700 Smith, J. S. (2005). The local mean decomposition and its application to EEG perception data. *Journal of the Royal Society Interface*, *2*, 443–454.
- Song, J.-L., Li, Q., Zhang, B., Westover, M. B., & Zhang, R. (2019). A new neural mass model driven method and its application in early epileptic seizure detection. *IEEE Transactions on Biomedical Engineering*, *67*, 2194–2205.
- 705 Tang, S., Dunnmon, J., Saab, K. K., Zhang, X., Huang, Q., Dubost, F., Rubin, D., & Lee-Messer, C. (2021). Self-supervised graph neural networks for improved electroencephalographic seizure analysis. In *International Conference on Learning Representations*.
- 710 Wang, J., Liang, S., He, D., Wang, Y., Wu, Y., & Zhang, Y. (2020). A sequential graph convolutional network with frequency-domain complex network of EEG signals for epilepsy detection. In *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 785–792). IEEE.
- Wei, Z., Zou, J., Zhang, J., & Xu, J. (2019). Automatic epileptic EEG detection using convolutional neural network with improvements in time-domain. *Biomedical Signal Processing and Control*, *53*, 101551.
- 715 World Health Organization (2016). A report about epilepsy. <https://www.who.int/news-room/fact-sheets/detail/epilepsy>.

- Xia, K., Deng, L., Duch, W., & Wu, D. (2022). Privacy-preserving domain adaptation for motor imagery-based brain-computer interfaces. *IEEE Transactions on Biomedical Engineering*, 69, 3365–3376.
- 720 Yang, Y., Truong, N. D., Maher, C., Nikpour, A., & Kavehei, O. (2022). Continental generalization of a human-in-the-loop ai system for clinical seizure recognition. *Expert Systems with Applications*, 207, 118083.
- Zhang, X., Yao, L., Dong, M., Liu, Z., Zhang, Y., & Li, Y. (2020). Adversarial representation learning for robust patient-independent epileptic seizure detection. *IEEE Journal of Biomedical and Health Informatics*, 24, 2852–2859.
725
- Zhao, Y., Zhang, G., Zhang, Y., Xiao, T., Wang, Z., Xu, F., & Zheng, Y. (2022). Multi-view cross-subject seizure detection with information bottleneck attribution. *Journal of Neural Engineering*, 19, 046011.
- Zhu, Y., Saqib, M., Ham, E., Belhareth, S., Hoffman, R., & Wang, M. D. (2020). Mitigating patient-to-patient variation in EEG seizure detection using meta transfer learning. In *2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE)* (pp. 548–555). IEEE.
730