technologies

MDPI

*Article*

# Pre-Ictal EEG Augmentation Based CDCGAN Model for Epileptic Seizure Prediction

**Xindi Huang [1], Hongying Meng [1,*] and Zhangyong Li [2]**

[1] Department of Electronic and Electrical Engineering, Brunel University of London, Uxbridge UB8 3PH, UK
[2] Research Centre of Biomedical Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
* Correspondence: hongying.meng@brunel.ac.uk

**Abstract**

Epilepsy is a common neurological disorder affecting over 50 million people worldwide, characterised by recurrent seizures accompanied by abnormal neuronal electrical activity. Electroencephalogram (EEG) is a technique for recording brain electrical signals, widely employed for epileptic seizure (ES) prediction due to its high temporal resolution, portability, and cost-effectiveness. However, reliable ES prediction based on EEG remains challenging, primarily owing to the limited duration of recorded pre-ictal states in publicly available datasets and the typically low signal-to-noise ratio (SNR) in non-invasive recordings. To mitigate these issues, we propose a Conditional Deep Convolutional Generative Adversarial Network (CDCGAN), which combines the representational power of Deep Convolutional Generative Adversarial Network (DCGAN) with the categorical conditioning mechanism of Conditional Generative Adversarial Network (CGAN) to generate class-specific EEG samples. By synthesising target samples, CDCGAN aims to alleviate class imbalance and enhance the quality of low-resolution spectral representations. To evaluate the practical utility of generated data, we trained a Convolutional Neural Network (CNN) on the augmented dataset and compared its performance against prior studies. Under the Leave-One-Seizure-Out cross-validation (LOSO-CV) protocol, our method achieved an average AUC of 0.876 at a 60% augmentation rate with 50 training epochs. The AUC improvement relative to corresponding control settings demonstrates that GAN-based data augmentation provides additional effective training samples for ES prediction while preserving task-relevant and discriminative pre-ictal EEG features.

**Keywords:** epileptic seizures prediction; electroencephalogram; biomedical signal processing; generative adversarial networks; deep learning



Academic Editor: Sheryl Berlin Brahnam

Received: 29 December 2025
Revised: 2 February 2026
Accepted: 9 February 2026
Published: 12 February 2026



## 1. Introduction

Epilepsy is a neurological disorder characterised by recurrent, unprovoked seizures and affects people across the lifespan. Although pharmacological therapy can control epileptic seizure (ES) for many patients, a substantial subset remains refractory, leaving them exposed to persistent risks such as injury and reduced quality of life [1]. In this context, seizure prediction aims to detect measurable changes in brain activity that precede an upcoming seizure and to issue an alert with sufficient lead time for intervention. Scalp electroencephalography (EEG), which captures electrical activity that is non-invasive, is widely used for this purpose because it provides a direct view of seizure-related neural dynamics [2].

ES prediction aims to issue an alarm before seizure onset so that preventive actions can be taken. In long-term EEG recordings, the signal can be divided into two broad categories: ictal segments, which correspond to the seizure event itself, and non-ictal segments, which correspond to all remaining periods when no seizure is occurring. The non-ictal part is further divided into inter-ictal and pre-ictal states. Inter-ictal refers to baseline activity far from any seizure event, whereas pre-ictal refers to the period preceding seizure onset in which seizure-related changes may emerge. To operationally define the pre-ictal interval for prediction, two time intervals relative to seizure onset are commonly introduced as follows: the seizure occurrence period (SOP) and the seizure prediction horizon (SPH). The SOP is defined as the pre-ictal interval immediately preceding seizure onset during which a seizure is expected to occur, and the SPH is a minimum lead time between an issued alarm and the SOP to allow sufficient time for intervention. Accordingly, an alarm is regarded as correct if (i) no seizure occurs during the SPH following the alarm and (ii) a seizure occurs within the subsequent SOP.

However, using EEG signals to predict ES is inherently challenging. The first is the limited amount of EEG data available, which limits the amount of data on which deep learning models can be trained. Also, public datasets now are unbalanced because when the EEG signal of a subject is recorded, most of the time the subject spends is between interictal rather than ES. This results in less data on the seizure period. However, such an unbalanced dataset allows the model to train on a smaller amount of data. If the original data is used for balancing, the data in the training set will be greatly reduced [2]. It will affect the performance of the model, especially for Convolutional Neural Network (CNN) models [3]. At the same time, non-invasive EEG recordings typically have a low signal-to-noise ratio (SNR), making it difficult to distinguish seizure-related activity from background noise.

Generative adversarial networks (GANs) [4] learn to synthesise data by training a generator and a discriminator. The generator produces candidate samples from latent noise, while the discriminator attempts to tell generated samples from real ones. This framework has been increasingly used for EEG data synthesis and augmentation. For instance, Hartmann et al. [5] introduced EEG-GAN based on a DCGAN-style architecture [6] and assessed real-synthetic similarity using widely used distribution-level metrics such as the Inception Score (IS), Frechet Inception Distance (FID), and sliced Wasserstein distance (SWD/WD).

Truong et al. [7] introduced GAN-based modelling for seizure prediction and evaluated it on several public EEG datasets, including CHB-MIT [8], Freiburg [9], and EPILEPSIAE [10]. Their pipeline converts inter-ictal and pre-ictal EEG segments into STFT spectrograms, trains a GAN on these representations, and leverages the discriminator as a learned feature extractor for downstream classification, reporting improvements over a CNN baseline [3]. Related efforts have also used GANs either for synthetic EEG generation (e.g., EpilepsyGAN for privacy-preserving data sharing [11]) or for explicit data augmentation, where STFT spectrogram windows are generated by a DCGAN and appended to the training set to improve prediction performance [12].

Xu et al. [13] proposed a GAN-based pre-ictal signal synthesis approach to generate multi-channel EEG samples. Specifically, temporal EEG segments were used to train four GAN variants with different architectures (DCGAN, DCWGAN, RGAN, and RWGAN). The generative performance was assessed by measuring the discrepancy between synthetic and real samples using frequency-domain RMSE (FDRMSE), FID, and WD, and DCWGAN was reported to achieve the best overall results. Xu et al. futher augmented the original dataset with DCWGAN-generated samples, which led to improvements in ES prediction accuracy and AUC-ROC. Yu et al. [14] employed a Conditional GAN (CGAN) for EEG data augmentation and introduced an additional refiner network to reduce the distribution gap between generated and real signals, targeting data scarcity and class imbalance. Yang

et al. [15] leveraged unsupervised learning to train the discriminator of a Wasserstein GAN as a feature extractor and subsequently applied a Bidirectional Long Short-Term Memory (Bi-LSTM) to perform ES prediction using the extracted EEG features. Subsequently, Feng et al. [16] proposed employing a Domain-Generative Adversarial Network (DGAN) to address model generalisation issues in ES detection. This network integrates a generative adversarial network with an adversarial domain adaptation module, designed to minimise variations in feature representations across patients and within individual patients. The method was evaluated on the publicly available CHB-MIT dataset and an additional proprietary dataset, yielding results that demonstrated superior performance compared to existing approaches.

Despite these advances, several challenges persist within the field. Firstly, improvements in distribution-level similarity metrics (such as FID/WD/FDRMSE) do not necessarily translate into enhanced performance for downstream ES prediction tasks, particularly under clinically relevant evaluation protocols. Secondly, generating pre-seizure samples of specific categories in a controlled manner remains challenging, thereby limiting the effectiveness of data augmentation in addressing class imbalance. Thirdly, the interaction between synthetic data quality and model generalisation capabilities is often insufficiently explored, whilst ambiguous data augmentation and segmentation strategies may yield unstable results or potential data leakage.

To address these limitations, we train a Conditional Deep Convolutional GAN (CDCGAN) on the CHB-MIT database to generate samples that capture pre-ictal characteristics in a class-conditional manner. The synthetic samples are then integrated with the original training set to construct an expanded dataset. Using an identical CNN-based modelling framework, we evaluate the expanded dataset against the original dataset under the Leave-One-Seizure-Out cross-validation (LOSO-CV) protocol. Figure 1 provides an overview of the proposed pipeline, including preprocessing, CDCGAN-based augmentation, detector training, and LOSO-CV evaluation.

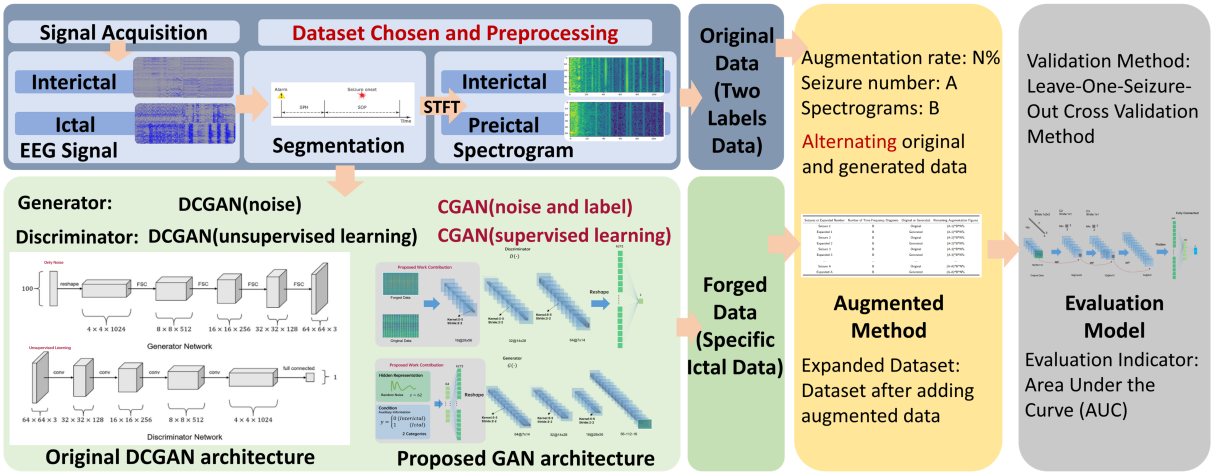


**Figure 1.** Overview of the proposed CDCGAN-based epileptic seizure (ES) prediction pipeline. Raw EEG signals are segmented into inter-ictal and pre-ictal intervals according to the SPH/SOP settings, and each 30 s segment is converted into a spectrogram using STFT. A conditional DCGAN (cDCGAN) is then trained to generate class-specific (pre-ictal) spectrograms, which are appended to the original training set according to a predefined augmentation rate (ratio of generated to original pre-ictal samples), forming an expanded dataset. The CNN detector is trained on the expanded dataset and evaluated using leave-one-seizure-out cross-validation (LOSO-CV); window-level ROC–AUC is computed for each held-out fold and averaged to obtain the patient-level AUC. Color blocks denote major stages of the workflow (blue: acquisition/preprocessing; green: GAN architecture and generation; yellow: data augmentation; gray: evaluation).

## 2. Method

### *2.1. Dataset*

The CHB-MIT [17,18] was collected at Boston Children's Hospital, in collaboration with the Massachusetts Institute of Technology, and contains long-term scalp EEG recordings from paediatric patients with intractable epilepsy. In total, 23 EEG recordings from 22 subjects are included. During data acquisition, patients were monitored continuously for several days, typically following withdrawal of anti-epileptic medication, to capture seizure events and to support clinical evaluation for potential surgical treatment. Overall, 182 seizure episodes with annotated onset and termination times are provided. Electrode placement and nomenclature follow the International 10–20 system [19].

In this study, we set SPH = 5 min and SOP = 30 min (Figure 2). For each annotated seizure with onset time $t_{\text{onset}}$, one pre-ictal SOP segment was extracted from $[t_{\text{onset}} - \text{SPH} - \text{SOP},\ t_{\text{onset}} - \text{SPH}]$. Only segments with the exact SOP duration were retained. With a sampling rate of 256 Hz, each SOP segment contains 460,800 samples.

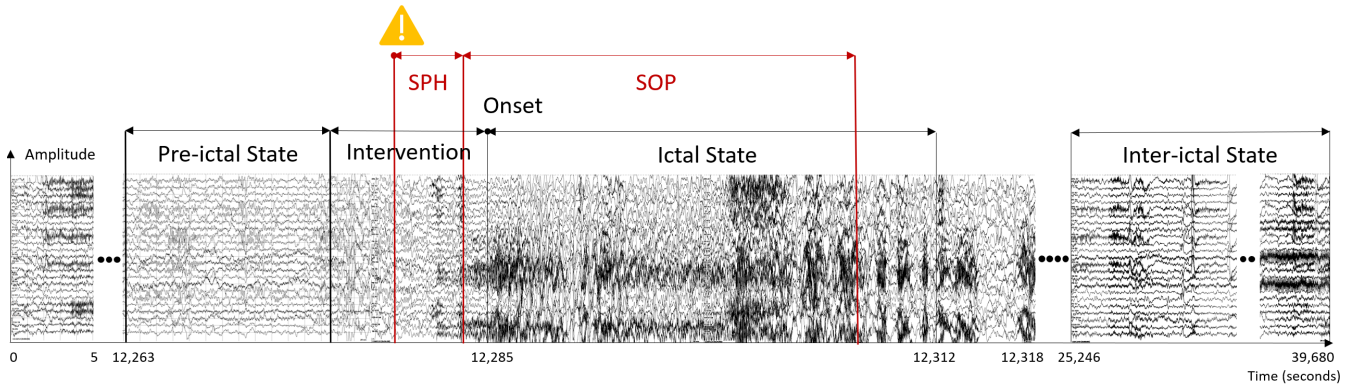


**Figure 2.** An EEG recording example from the CHB-MIT database [8] (e.g., the second seizure of Patient 1). The ictal stage is labelled from 12,285 to 12,312 s, and an inter-ictal interval is selected from 25,246 to 39,680 s. A correct prediction indicates that no seizure occurs during the SPH following an issued warning and that a seizure occurs within the subsequent SOP. The yellow triangle marker denotes the time at which the warning (alarm) is issued (start of SPH). Reproduced from [20].

Inter-ictal segments were obtained from the dataset segmentation file (segmentation.csv), which specifies inter-ictal intervals that are temporally separated from seizure events by design. Multiple inter-ictal intervals may correspond to one seizure. For each inter-ictal interval, we extracted the annotated EEG segment; when a full inter-ictal file is specified, the entire recording (typically 1 h, 921,600 samples) was used, otherwise only the annotated sub-interval was extracted.

Because the number of available EEG channels varies across subjects in CHB-MIT, we adopted the channel selection and simplification procedure in [21] to obtain a consistent set of channels across patients (Table 1). Both pre-ictal and inter-ictal segments were partitioned into non-overlapping 30 s windows (7680 samples per window) and transformed into time–frequency representations using the same STFT-based preprocessing pipeline. Each 30 s window was represented as a spectrogram tensor of size $16 \times 56 \times 112$ (channels $\times$ time bins $\times$ frequency bins), which serves as the input to CDCGAN (for pre-ictal augmentation) and the downstream CNN predictor.

**Table 1.** Subject-wise primary EEG channels and seizure counts in the CHB-MIT dataset. For each patient, 16 primary channels were selected using the screening procedure described in [21] to obtain a consistent channel set across subjects. Channel indices correspond to the original channel numbering in the CHB-MIT recordings, and the last column reports the number of annotated seizure events for each patient.

| Patient | Primary Channels Number | Seizures Number |
|---|---|---|
| P1 | [8, 4, 6, 0, 5, 7, 16, 10, 20, 18, 3, 2, 12, 11, 15, 9] | 7 |
| P2 | [17, 10, 11, 16, 4, 5, 15, 0, 6, 7, 13, 12, 3, 21, 14, 2] | 3 |
| P3 | [17, 4, 10, 9, 8, 0, 16, 6, 7, 15, 3, 11, 21, 12, 14, 5] | 6 |
| P5 | [6, 7, 4, 12, 13, 15, 16, 5, 17, 3, 9, 8, 19, 18, 2, 1] | 5 |
| P9 | [9, 2, 17, 15, 19, 20, 18, 7, 1, 11, 8, 5, 13, 16, 4, 12] | 4 |
| P10 | [11, 7, 17, 10, 9, 6, 8, 4, 15, 3, 16, 20, 0, 19, 12, 13] | 6 |
| P13 | [6, 7, 15, 14, 3, 0, 10, 9, 11, 4, 5, 1, 8, 2, 16, 13] | 3 |
| P14 | [0, 16, 10, 4, 12, 9, 15, 1, 8, 11, 17, 7, 13, 6, 5, 20] | 5 |
| P18 | [7, 17, 3, 5, 16, 11, 15, 4, 10, 6, 9, 8, 18, 2, 0, 14] | 6 |
| P19 | [6, 16, 7, 21, 17, 5, 15, 3, 0, 14, 13, 11, 2, 18, 19, 10] | 3 |
| P20 | [14, 13, 0, 4, 21, 2, 18, 1, 6, 19, 5, 15, 10, 17, 16, 20] | 5 |
| P21 | [8, 9, 19, 4, 1, 18, 2, 12, 0, 16, 5, 20, 17, 21, 3, 13] | 4 |
| P23 | [19, 20, 21, 16, 11, 0, 4, 1, 7, 17, 10, 5, 6, 15, 3, 9] | 5 |

### *2.2. Preprocessing*

For preprocessing, the EEG recordings (16 channels, sampled at $f_s = 256$ Hz) were converted into time–frequency representations using the short-time Fourier transform (STFT). Each SOP segment was divided into 60 non-overlapping windows of 30 s (7680 samples per channel), and an STFT-based spectrogram was computed for every 30 s window. We adopt a time–frequency representation because pre-ictal dynamics are typically non-stationary and are often reflected by time-varying, frequency-dependent changes in EEG rhythms and spectral power; the spectrogram preserves both spectral content and its temporal evolution, which facilitates learning transient frequency-localised signatures preceding seizure onset [3,7].

We used the stft package to compute spectrograms with a frame length of 1 s (256 samples). Since the default setting uses an overlap coefficient of 2 (i.e., 50% overlap) and a cosine window, the hop size is 0.5 s (128 samples). With these settings, each 30 s window yields 59 time frames, and the one-sided spectrum contains 129 frequency bins (0–128 Hz with 1 Hz resolution). This configuration provides a practical time–frequency

resolution trade-off for scalp EEG: it is fine-grained enough to capture short-lived spectral modulations while remaining compact for stable GAN and CNN training.

The complex STFT output was converted to magnitude and stabilised by adding $10^{-6}$ before applying a base-10 logarithm. To suppress power-line interference in CHB-MIT, we removed frequency components within 57–63 Hz and 117–123 Hz (60 Hz mains frequency and its second harmonic), and excluded the DC component (0 Hz). After frequency removal, the spectrogram contains 114 frequency bins. Finally, for a fixed input size to the downstream models, we cropped the spectrogram to the first 56 time frames and 112 frequency bins, resulting in a tensor of size $16 \times 56 \times 112$ (channels $\times$ time bins $\times$ frequency bins) for each 30 s window.

### *2.3. Conditional Deep Convolutional Generative Adversarial Network (CDCGAN)*

The GAN adopted in this study is termed the CDCGAN, which integrates the convolutional architecture of DCGAN [6] with the label-conditioning mechanism of CGAN [22]. Figure 3 illustrates the overall CDCGAN architecture. In the discriminator, the input consists of either real spectrogram samples or synthetic samples produced by the generator. The discriminator employs a similar architecture to DCGAN, comprising three convolutional blocks containing 16, 32, and 64 feature maps respectively. Each block uses a $5 \times 5$ convolution kernel with a stride of 2 for progressive downsampling.

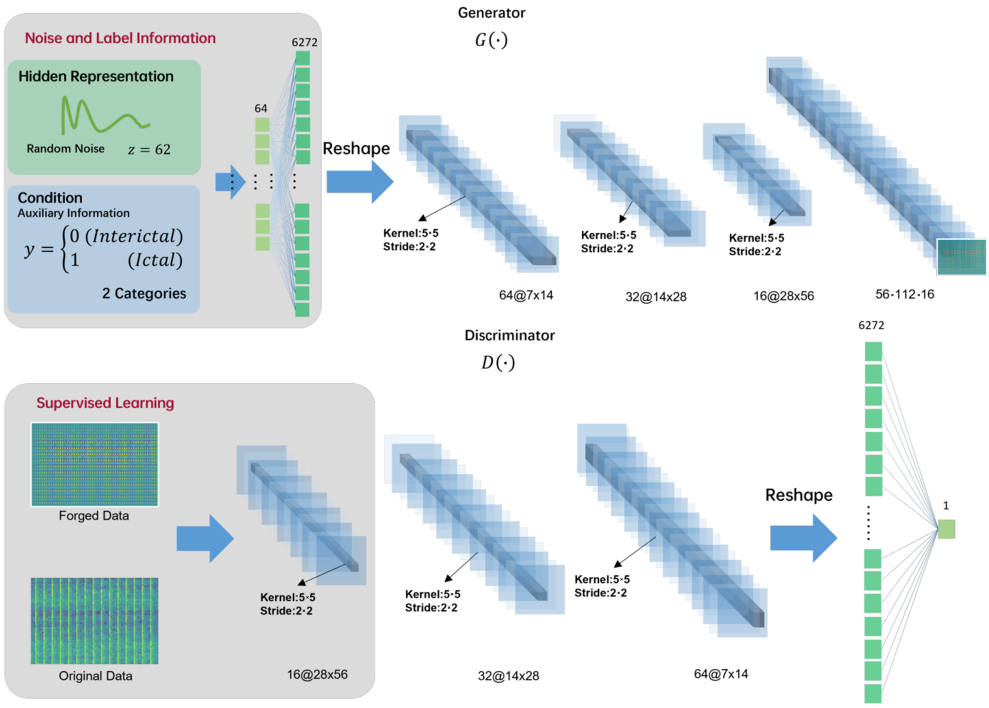


**Figure 3.** Architecture of the proposed Conditional Deep Convolutional GAN (CDCGAN). The generator $G(\cdot)$ takes a latent noise vector $z \in \mathbb{R}^{62}$ and a class condition $y \in \{0, 1\}$ (0: inter-ictal, 1: pre-ictal). The concatenated input is embedded into a 6272-dimensional hidden representation and reshaped into feature maps of size $64 \times 7 \times 14$. A sequence of transposed-convolution blocks (kernel $5 \times 5$, stride $2 \times 2$) progressively upsamples the representation to $32 \times 14 \times 28$, $16 \times 28 \times 56$, and finally generates a synthetic EEG time–frequency sample with size $16 \times 56 \times 112$. The discriminator $D(\cdot)$ receives either real (original) or forged (generated) samples and applies convolutional downsampling blocks (kernel $5 \times 5$, stride $2 \times 2$) to obtain $16 \times 28 \times 56$, $32 \times 14 \times 28$, and $64 \times 7 \times 14$ feature maps, followed by flattening (6272 units) and a sigmoid output that estimates the probability that the input is real. During adversarial training, $G$ is optimised to synthesise class-specific EEG samples, while $D$ is trained to distinguish real from generated data under the same conditioning.

Employing CNN within the GAN framework offers notable advantages for EEG modelling. Compared with early GAN variants based on fully connected networks, the convolutional structure enables the discriminator to capture hierarchical spatial patterns in the time–frequency representation, thereby facilitating the learning of more informative features. This is particularly beneficial for non-invasive scalp EEG, which often exhibits SNR and complex spatiotemporal characteristics that are difficult to model effectively with simple architectures such as shallow artificial neural networks (ANNs).

We apply batch normalisation once before each convolutional layer (three in total) to stabilise and accelerate training by reducing internal covariate shift [23]. The last layer of the discriminator is a fully connected layer followed by a leaky ReLU activation. The discriminator outputs a scalar score used to distinguish whether an input spectrogram is drawn from the real dataset or generated by the generator.

The structure of the generator is inspired by the Conditional GAN. The public datasets now are unbalanced because when the EEG signal of a subject is recorded, most of the time the subject spends is between interictal rather than seizures. This results in less data on the seizure period. However, such an unbalanced dataset allows the model to train on a smaller amount of data. If the original data is used for balancing, the data in the training set will be greatly reduced. It will affect the performance of the model, especially for CNN models [3]. Thus, the advantage of Conditional GAN is that it can generate a specific label so that it can solve unbalanced issues. From the construction of Conditional GAN, it can be found that the input of Conditional GAN contains the labels and original data. Thus, the training part of it is supervised learning in order to generate the specific label data.

The input of the generator is the noise and labels. To keep the same dimension of the discriminator, the shape of the noise is set to 62. Then the noise would be inputted into the fully-connected layers. The number of convolutional kernels in the generator is 32, 16 and 16 with the size 5 by 5 and strides 2. The setting of batch normalisation layers and the activate function is the same as the discriminator. For the label information, it is transformed into one hot label with two categories added to the discriminator and the generator.

Compared with the other GAN model, CDCGAN is more suitable for solving the issues in the EEG classification or prediction field. By combining the DCGAN and Conditional GAN, the CDCGAN can forge high-resolution time–frequency diagrams and learn more deep features. The CDCGAN can generate specific labels to solve unbalanced issues. The architecture of the complete CDCGAN is shown below.

### *2.4. Synthetic Data Generation and Sampling Strategy*

After CDCGAN generates class-specific samples, the synthetic spectrograms are appended to the original training set according to an augmentation rate. We define the augmentation rate as $r = N/100$, where $N\%$ denotes the ratio of generated spectrograms to the number of original spectrograms extracted from the patient's pre-ictal period (SOP). Let $A$ be the number of seizures for a patient and $B$ be the number of spectrograms extracted from the SOP of each seizure (in our setting, $B = 60$ corresponding to a 30 min SOP segmented into 60 non-overlapping 30 s windows). The target number of generated spectrograms is then computed as

$$G_{\text{target}} = \lceil r \cdot A \cdot B \rceil. \quad (1)$$

To preserve a seizure-wise organisation of samples, we allocate generated spectrograms in blocks following each seizure in the training set. Specifically, for seizure $i$ ($i = 1, \ldots, A$), we append a generated block of size

$$g_i = \min\big(B,\ \max\big(0,\ G_{\text{target}} - (i-1)B\big)\big), \tag{2}$$

and stop once $G_{\text{target}}$ samples have been added. Note that the addition referred to here pertains to appending samples to the training set at the spectrogram level, rather than concatenating EEG signals along the timeline. In practice, the generator is configured to produce up to 1000 synthetic spectrograms per patient. For a given augmentation rate, we compute the required number of synthetic samples $G_{\text{target}}$ (Equation (1)) and then select the first $G_{\text{target}}$ spectrograms from the generated pool (i.e., without random sampling). The detailed procedure is summarised in Table 2.

**Table 2.** Seizure-wise augmentation procedure for integrating generated pre-ictal spectrograms into the original training set. $r = N/100$ is the augmentation rate, and the target number of generated spectrograms is $G_{\text{target}} = \lceil rAB \rceil$ (Equation (1)). For seizure $i$, a generated block of size $g_i = \min\big(B, \max\big(0, G_{\text{target}} - (i-1)B\big)\big)$ is appended. The last column reports the remaining number of generated spectrograms needed after each generated block.

| Step | Spectrograms | Type | Remaining to Reach $G_{\text{target}}$ |
|---|---|---|---|
| Seizure 1 | $B$ | Original | $G_{\text{target}}$ |
| Expanded 1 | $g_1$ | Generated | $\max\big(0,\ G_{\text{target}} - B\big)$ |
| Seizure 2 | $B$ | Original | $\max\big(0,\ G_{\text{target}} - B\big)$ |
| Expanded 2 | $g_2$ | Generated | $\max\big(0,\ G_{\text{target}} - 2B\big)$ |
| Seizure 3 | $B$ | Original | $\max\big(0,\ G_{\text{target}} - 2B\big)$ |
| Expanded 3 | $g_3$ | Generated | $\max\big(0,\ G_{\text{target}} - 3B\big)$ |
| ... | ... | ... | ... |
| Seizure $A$ | $B$ | Original | $\max\big(0,\ G_{\text{target}} - (A-1)B\big)$ |
| Expanded $A$ | $g_A$ | Generated | $\max\big(0,\ G_{\text{target}} - AB\big)$ |

### *2.5. Evaluation Model*

After the data augmentation part, the expanded dataset can be evaluated by the CNN model. Figure 4 shows the construction of the evaluation model [3].

The evaluation network consists of three convolutional blocks, denoted as C1, C2, and C3. Each block performs hierarchical feature extraction, with max pooling applied to progressively reduce the spatial resolution of feature maps and improve computational efficiency. The resulting representations are then flattened and passed through fully connected layers. Finally, a soft-max activation is used to produce class probabilities and generate the predicted labels.

The evaluation CNN follows a hybrid 3D–2D design (Figure 4). Each input is a 5D spectrogram tensor of shape $(\text{batch}, 1, C, 56, 112)$, where $C$ denotes the number of EEG channels (16 when significant-channel selection is applied and 22 otherwise). In C1, a 3D convolution with 16 filters and kernel size $(C, 5, 5)$ is applied with stride $(1, 2, 2)$, so that the convolution spans the full channel dimension while downsampling the time–frequency plane; this is followed by a ReLU activation and 3D max pooling with pool size $(1, 2, 2)$. The resulting feature maps are reshaped to a 2D representation and passed through two 2D convolutional blocks (C2 and C3) with 32 and 64 filters, respectively, using $3 \times 3$ kernels and $2 \times 2$ max pooling. Finally, the features are flattened and fed into fully connected layers

with dropout (0.5), and a 2-unit softmax output produces class probabilities for inter-ictal versus pre-ictal classification.

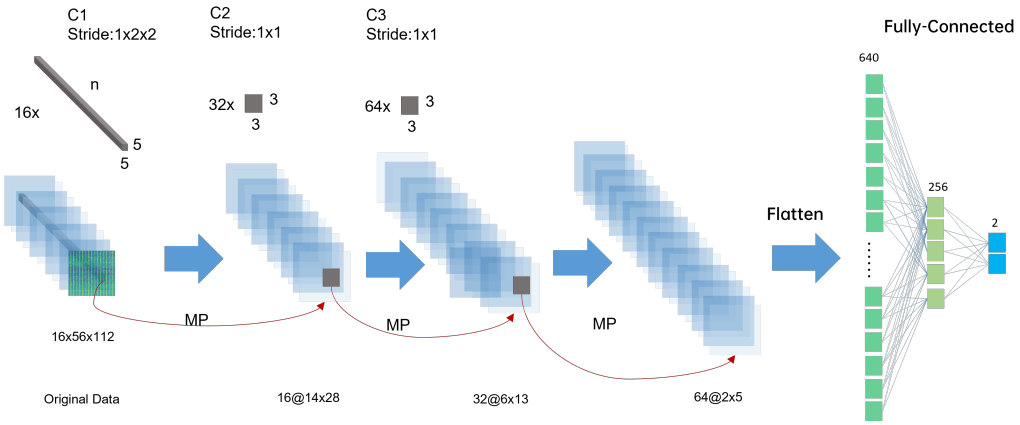


**Figure 4.** Architecture of the evaluation CNN used to assess CDCGAN-based data augmentation. Each input is a 5D spectrogram tensor of shape (batch, 1, $C$, 56, 112), where $C$ denotes the number of EEG channels ($C = 16$ when significant-channel selection is applied and $C = 22$ otherwise). C1 applies batch normalisation followed by a 3D convolution with 16 filters and kernel size $(C, 5, 5)$ (spanning the full channel dimension) using stride $(1, 2, 2)$, and 3D max pooling with pool size $(1, 2, 2)$. The output is reshaped and processed by two 2D convolutional blocks: C2 (32 filters) and C3 (64 filters), both using $3 \times 3$ kernels with stride $1 \times 1$ and 2D max pooling ($2 \times 2$). The feature maps are flattened and passed through fully connected layers with dropout (0.5): a 128-unit hidden layer (sigmoid) followed by a 2-unit softmax output for binary classification (inter-ictal vs. pre-ictal).

We evaluate the model using LOSO-CV [24]. For a patient with $N$ annotated seizures, in fold $i$, we hold out the SOP/pre-ictal windows extracted for seizure $i$ as the test set, while the remaining $(N - 1)$ seizures are used for training. This procedure is repeated $N$ times so that each seizure is tested exactly once.

Inter-ictal windows are assigned to folds according to the dataset segmentation file (segmentation.csv). Specifically, all inter-ictal intervals associated with the held-out seizure are included in the test set of that fold, whereas inter-ictal intervals associated with the remaining seizures are used for training. Within each fold, after forming the training set, we apply a deterministic sequential split for early stopping: the first 75% of training samples are used for optimisation and the remaining 25% are reserved as a validation set.

In our current implementation, CDCGAN is trained in a patient-specific manner using all available pre-ictal (SOP) samples from that patient prior to LOSO-CV. The generator is then used to produce additional samples that augment the training split in each fold.

## 3. Results

In training the generator and discriminator of CDCGAN, the training epoch is chosen as 50 and the learning rate of the generator and discriminator is the same $\alpha_1 = \alpha_2 = 1 \times 10^{-4}$. The leaky ReLU action function was used a slope of 0.2. The batch size is 64. The ratio of the number of training sets to the number of validation sets is 7 to 3. The generated samples for every patient are 1000. According to different augmentation rates, the generated samples will be added to the original data. In the evaluation model, the batch size is set to 32 and the training epoch is 100. Then, we used LOSO-CV to test the expanded dataset. The condition of early stop is that if the performance is not improved after 10 epochs, the training will be stopped. The dropout rate is set to 0.5. The optimiser was Adam with momentum $\beta_1 = 0.9$, $\beta_2 = 0.999$ and the learning rate $\alpha_3 = 1 \times 10^{-5}$.

### *3.1. Evaluation Metrics*

The primary evaluation metric is the area under the receiver operating characteristic curve (ROC-AUC). In this study, ROC-AUC is computed at the window level: for each 30 s spectrogram sample, the detector outputs a continuous score interpreted as the probability of the positive class (pre-ictal, $y = 1$). Within each LOSO-CV iteration, the ROC curve is obtained by sweeping the decision threshold applied to these scores and plotting the true positive rate (TPR, sensitivity) against the false positive rate (FPR). The AUC is then calculated as the area under the resulting ROC curve.

Under LOSO-CV, an AUC value is computed for the test fold in each iteration and then averaged across the $N$ folds (one fold per seizure) to obtain a patient-level AUC. Given a binary confusion matrix (Table 3), the rates used to construct the ROC curve are defined as $\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP+FN}}$, $\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP+TN}}$, $\mathrm{TNR} = \frac{\mathrm{TN}}{\mathrm{TN+FP}}$ (specificity), and $\mathrm{FNR} = \frac{\mathrm{FN}}{\mathrm{FN+TP}}$.

**Table 3.** Confusion-matrix definitions used to compute ROC-related quantities for ES prediction. Ground-truth labels (rows) are compared with predicted labels (columns), yielding true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN).

| Ground Truth \ Prediction | 1 (Pre-Ictal) | 0 (Inter-Ictal) | Total |
|---|---|---|---|
| 1 (pre-ictal) | TP | FN | TP+FN |
| 0 (inter-ictal) | FP | TN | FP+TN |
| Total | TP+FP | FN+TN | TP+FP+FN+TN |

### *3.2. Performance Under Different Augmentation Rate*

For each augmentation rate, we report a patient-level AUC obtained by averaging the window-level ROC-AUC across LOSO folds (one fold per seizure). We further report the mean AUC across patients for each augmentation setting, as shown in Figure 5. For each patient, the CDCGAN generator is configured to produce up to 1000 synthetic pre-ictal spectrograms; the required number of synthetic samples is then selected from this pool and appended to the training set according to the specified augmentation rate. In all experiments, we set a maximum epoch budget of 50 for the detector; however, early stopping may terminate training earlier based on a monitored objective (see Figure 6).

We define the augmentation rate as the ratio of generated pre-ictal spectrograms to the number of original pre-ictal spectrograms for a given patient. Let $A$ denote the number of seizures for that patient. Since each seizure contributes one 30 min SOP that is converted into 60 spectrograms, the patient has $60A$ original pre-ictal spectrograms in total. An augmentation rate of $r$ therefore corresponds to adding approximately $\lceil r \cdot 60A \rceil$ generated pre-ictal spectrograms to the training set.

As shown in Figure 5, the AUC reaches its highest values at augmentation rates of approximately 60% and 100%. We adopt 60% as the default setting because it achieves performance comparable to 100% while injecting fewer synthetic samples, thereby reducing the risk of distribution shift induced by generated data and lowering computational cost. Compared with substantially higher oversampling ratios (e.g., 5×–10×) reported in [13], these results suggest that a moderate level of targeted pre-ictal augmentation is sufficient under the CHB-MIT setting and our training protocol.

To complement the AUC results, we also report representative learning curves to illustrate the training dynamics under different augmentation rates and the motivation for early stopping. Within each LOSO iteration, the training portion is split sequentially into a training subset (first 75%) and a monitoring subset (last 25%) for model selection. As shown in Figure 6, the training loss quickly approaches zero and the training accuracy

saturates near 1.0 across settings, while the monitoring loss exhibits a U-shaped trend and tends to increase after approximately 12–15 epochs, indicating the onset of overfitting. Early stopping is applied with a patience of 10 epochs by monitoring the objective $J = \mathcal{L}_{\text{val}} + \mathcal{L}_{\text{train}}$, and training is also stopped when $J < 10^{-3}$. We note that these curves are used for illustrating optimisation behaviour and model selection; final performance is reported as ROC-AUC on the held-out seizure fold under LOSO-CV.

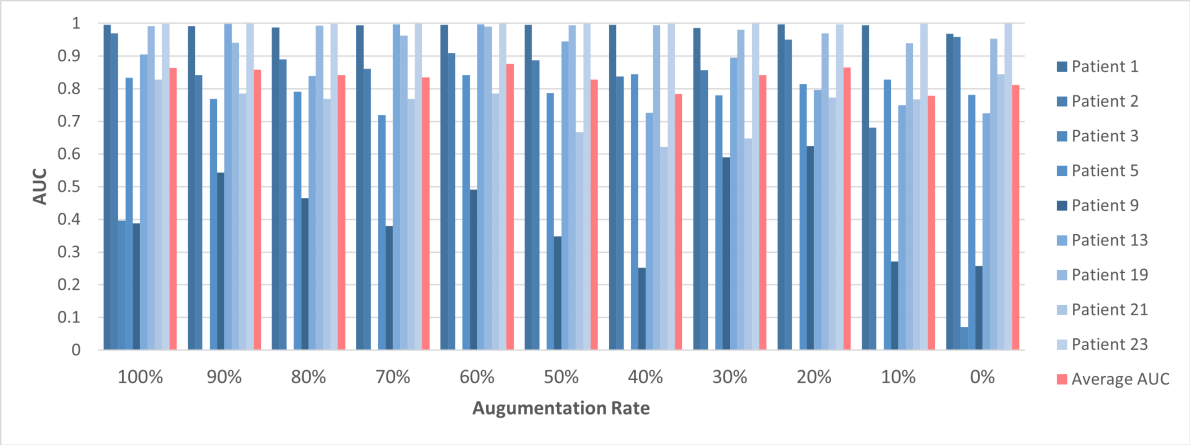


**Figure 5.** Effect of augmentation rate on ES prediction performance (window-level ROC-AUC). For each patient and each augmentation setting, we compute ROC-AUC on the test fold in LOSO-CV using the predicted probability of the positive class (pre-ictal, $y = 1$) for each 30 s spectrogram window. The patient-level AUC (blue bars) is obtained by averaging the AUC values over the $N$ LOSO folds (one fold per seizure). The red bars report the mean AUC averaged across patients for each augmentation rate.

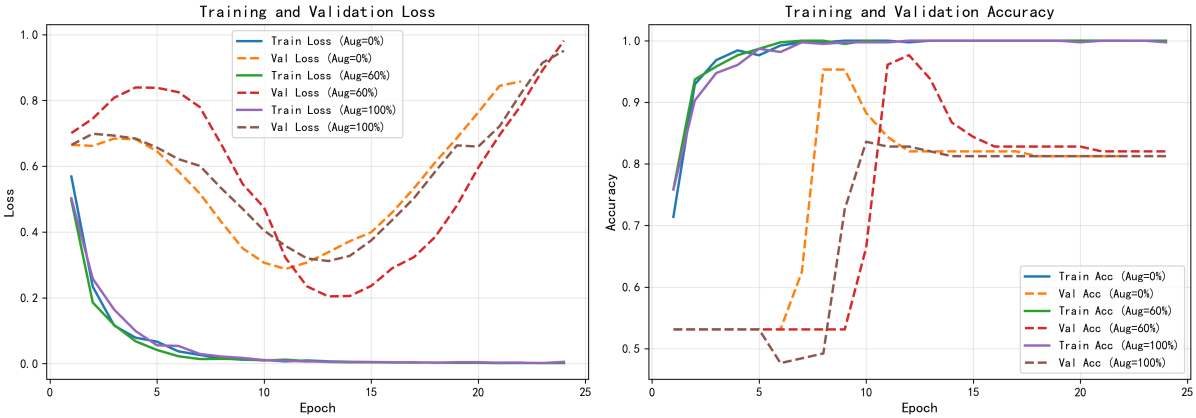


**Figure 6.** Representative learning curves of the evaluation CNN under different augmentation rates (Aug = 0%, 60%, and 100%) within a LOSO iteration. (**Left**): training and monitoring loss; (**right**): training and monitoring accuracy. The monitoring set is a held-in split of the training portion (first 75% for training and last 25% for monitoring) used for early stopping and model selection. Final evaluation is performed using ROC-AUC computed on the held-out seizure fold under LOSO-CV.

### *3.3. Performance Under Different Synthetic Specific Category Samples*

By analysing Table 4, it can be founded that in most of the patients, the AUC of the expanded dataset in the proposed work is larger than the previous work which means the forged data can improve the existing performance of the model. For the previous work [3] which is the paper related to the evaluation model, it can be found that the expanded data can improve the performance of the evaluation model which means the forged data by CDCGAN can contribute to the performance. For the previous work [7,13], they employed the GANs to whether to generate the time domain signal with multiple

channels or classification the different labels using the trained discriminator as the CNN model. The higher AUC can prove that the data generated by CDCGAN is superior to these two methods.

**Table 4.** Patient-wise AUC comparison between the proposed approach and representative prior studies. Results for our method are reported under 50 training epochs and a 60% augmentation rate. We additionally report an ablation setting where label-conditioning is removed (unconditional DCGAN), while keeping the rest of the pipeline unchanged. The last two rows report the mean and sample standard deviation (SD) across patients; for Xu et al. [13], the statistics are computed over the patients for which AUC values are reported. Boldface indicates the highest AUC in each row (per patient and the mean).

| Patient Number | Truong et al. [3] | Truong et al. [7] | Xu et al. [13] | DCGAN | Our Method |
|---|---|---|---|---|---|
| Patient 1 | 0.924 | **0.998** | 0.912 | 0.986 | 0.996 |
| Patient 2 | 0.359 | 0.171 | \ | 0.639 | **0.910** |
| Patient 5 | **0.880** | 0.841 | \ | 0.837 | 0.841 |
| Patient 9 | **0.744** | 0.555 | 0.518 | 0.519 | 0.491 |
| Patient 13 | 0.973 | 0.984 | \ | 0.930 | **0.997** |
| Patient 19 | **0.993** | 0.971 | 0.494 | 0.990 | 0.991 |
| Patient 21 | **0.905** | 0.812 | \ | 0.840 | 0.785 |
| Patient 23 | 0.993 | 0.943 | 0.810 | 0.996 | **0.999** |
| Mean | 0.846 | 0.784 | 0.684 | 0.842 | **0.876** |
| SD | 0.213 | 0.287 | 0.209 | 0.177 | 0.176 |

### *3.4. Ablation Study*

We conducted an ablation study to quantify the contribution of the conditional mechanism in CDCGAN. Specifically, we removed the label-conditioning and trained an unconditional DCGAN under the same experimental protocol (same data splits, augmentation rate, detector architecture, and training epochs). Under this ablated setting, the patient-wise ROC-AUC values are 0.986 (P1), 0.639 (P2), 0.837 (P5), 0.519 (P9), 0.930 (P13), 0.990 (P19), 0.840 (P21), and 0.996 (P23), yielding a mean ROC-AUC of 0.842 (SD = 0.177) across patients. Compared with the full CDCGAN-based augmentation (mean ROC-AUC = 0.876), the performance drop indicates that label-conditioning is an important factor for generating class-consistent pre-ictal samples that benefit the downstream predictor.

## 4. Discussion

Table 5 provides a structured comparison between the proposed method and representative prior studies. The comparison covers the target task (ES prediction), whether GAN-based data augmentation is adopted, whether class-specific data generation is performed, whether both classes are used to train the GAN, the use of oversampling, the type of synthesised EEG signals (raw and/or long-term), the evaluation protocol (LOSC-CV), and whether an improvement in predictive performance is reported. Notably, among the compared works, only [25] and the proposed method generate class-specific samples, while the proposed method additionally avoids oversampling during GAN training and evaluates performance under LOSC-CV. As summarised in the final row, the proposed method reports improved performance relative to its baseline setting.

**Table 5.** Comparison of the proposed method with representative prior studies in terms of task setting, data augmentation (DA) strategy, GAN-based synthesis design, evaluation protocol, and reported performance improvement. Each column corresponds to a reference method and each row indicates whether a specific design choice (e.g., ES prediction task, use of GAN for DA, class-conditional generation, oversampling, raw or long-term EEG synthesis, LOSO-CV evaluation) is adopted.

| Researches | [7] | [11] | [26] | [27] | [12] | [25] | [13] | This Method |
|---|---|---|---|---|---|---|---|---|
| Epileptic seizure prediction task? | ✓ | × | ✓ | × | ✓ | × | ✓ | ✓ |
| GAN is used for DA? | × | ✓ | × | ✓ | × | ✓ | ✓ | ✓ |
| GAN is used for generating specific data? | × | × | × | × | × | ✓ | × | ✓ |
| Both categories' signals are trained in GAN? | ✓ | ✓ | × | × | ✓ | × | ✓ | ✓ |
| Oversampling is used in the training part? | ✓ | × | × | ✓ | × | × | × | × |
| Synthesise raw EEG signals? | × | ✓ | × | ✓ | × | ✓ | ✓ | ✓ |
| Synthesise long-term EEG signals? | × | × | × | ✓ | × | ✓ | ✓ | ✓ |
| LOSO-CV? | ✓ | × | × | × | × | × | ✓ | ✓ |
| Did performance improve or not? | × | × | × | × | ✓ | × | ✓ | ✓ |

### *4.1. Advantages*

First, unlike studies that apply generic oversampling or augment both classes indiscriminately, our framework uses a GAN-based model to selectively augment pre-ictal EEG segments. This targeted augmentation directly addresses the scarcity and class-imbalance problem in ES prediction, where pre-ictal samples are typically much fewer than inter-ictal samples. Second, we avoid oversampling during GAN training. This design reduces computational cost and, more importantly, mitigates the risk of repeatedly reusing noise-dominated patterns, which is a practical concern for non-invasive scalp EEG with relatively low SNR. Third, the conditional learning mechanism enables class-aware generation: by conditioning on labels, the generator learns class-specific characteristics and can synthesise pre-ictal signals while reducing feature leakage from the other class, thereby improving separability for downstream prediction.

### *4.2. Limitations*

Despite the observed performance improvements, several limitations warrant consideration. First, the proposed approach remains sensitive to the quality and diversity of the generated pre-ictal spectrograms. Mode collapse or insufficient coverage of heterogeneous pre-ictal patterns in the generator may lead the detector to overfit to a limited set of synthetic structures, potentially compromising robustness and generalisation to unseen recordings or patients.

Second, our current augmentation strategy is patient-specific: for each patient, the generator is trained using that patient's pre-ictal data to model an individual distribution, and the generated samples are subsequently used to augment detector training. While this setting aligns with a personalised deployment scenario, it may overestimate performance compared with more stringent protocols, such as fold-wise augmentation (where the generator is trained using training folds only) or cross-patient evaluation. Accordingly, the reported results should be interpreted as patient-specific performance on CHB-MIT under our current training protocol, and may not directly generalise to cross-patient settings.

Third, distribution-level similarity does not necessarily guarantee clinical validity. Although the generated samples can improve downstream ROC-AUC, GAN-generated

EEG spectrograms may not fully preserve clinically meaningful pre-ictal biomarkers. To assess physiological plausibility, we conducted quantitative similarity analyses between real and generated pre-ictal samples across all patients. Specifically, we compared the distributions of PSD-derived power and the mean spectrogram log-power ($\log_{10}$) for the full-band range and for each canonical sub-band ($\delta$, $\theta$, $\alpha$, $\beta$, $\gamma$). The results show partial distributional alignment between real and synthetic samples, with relatively closer agreement in higher-frequency bands (particularly $\beta/\gamma$) and more pronounced shifts in lower-frequency bands (notably $\delta/\theta$). This indicates that the current generator captures some high-frequency spectral characteristics more reliably than low-frequency content. Overall, these distribution-level comparisons suggest that the generated data do not perfectly match the full physiological distribution of real pre-ictal EEG; therefore, further validation is warranted, including expert review and methodological improvements to enhance low-frequency fidelity and overall distribution alignment.

## 5. Conclusions

We proposed a new structure of CDCGAN applied in order to solve the issues in the ES prediction. The results of the CDCGAN are better than the previous work with the value of AUC equal to 0.876 when the epoch equals 50 and the augmentation rate equals 60%. This method gives great ideas to solve the issues in the unbalanced datasets and low-resolution time–frequency diagrams in the ES prediction field.

**Author Contributions:** Conceptualization, methodology, X.H., H.M. and Z.L.; investigation, formal analysis, and writing—original draft preparation, X.H. and H.M.; writing—review and editing, X.H., H.M. and Z.L.; funding acquisition, H.M. and Z.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Royal Society (IEC\NSFC\223285) and National Natural Science Foundation of China (General Program) No. 62171073.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The CHB-MIT Scalp EEG Database used in this study is publicly available at https://physionet.org/content/chbmit/1.0.0/ (accessed on 13 December 2025).

**Conflicts of Interest:** The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. Czuczwar, S.J. (Ed.) *Epilepsy*; Exon Publications: Brisbane City, Australia, 2022.
2. Huang, X.; Meng, H.; Li, Z. Deep learning for epileptic seizure prediction from EEG signals: A review. *Biomed. Signal Process. Control* **2026**, *117*, 109518. [CrossRef]
3. Truong, N.D.; Nguyen, A.D.; Kuhlmann, L.; Bonyadi, M.R.; Yang, J.; Ippolito, S.; Kavehei, O. Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram. *Neural Netw.* **2018**, *105*, 104–111. [CrossRef] [PubMed]
4. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **2014**, *27*.
5. Hartmann, K.G.; Schirrmeister, R.T.; Ball, T. EEG-GAN: Generative adversarial networks for electroencephalograhic (EEG) brain signals. *arXiv* **2018**, arXiv:1806.01875. [CrossRef]
6. Radford, A.; Metz, L.; Chintala, S. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv* **2015**, arXiv:1511.06434.
7. Truong, N.D.; Kuhlmann, L.; Bonyadi, M.R.; Querlioz, D.; Zhou, L.; Kavehei, O. Epileptic seizure forecasting with generative adversarial networks. *IEEE Access* **2019**, *7*, 143999–144009. [CrossRef]

8. Shoeb, A.H. Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2009.
9. Haddad, T.; Ben-Hamida, N.; Talbi, L.; Lakhssassi, A.; Aouini, S. Temporal epilepsy seizures monitoring and prediction using cross-correlation and chaos theory. *Healthc. Technol. Lett.* **2014**, *1*, 45–50. [CrossRef] [PubMed]
10. Kuhlmann, L.; Karoly, P.; Freestone, D.R.; Brinkmann, B.H.; Temko, A.; Barachant, A.; Li, F.; Titericz, G., Jr.; Lang, B.W.; Lavery, D.; et al. Epilepsyecosystem. org: Crowd-sourcing reproducible seizure prediction with long-term human intracranial EEG. *Brain* **2018**, *141*, 2619–2630. [PubMed]
11. Pascual, D.; Amirshahi, A.; Aminifar, A.; Atienza, D.; Ryvlin, P.; Wattenhofer, R. Epilepsygan: Synthetic epileptic brain activities with privacy preservation. *IEEE Trans. Biomed. Eng.* **2020**, *68*, 2435–2446. [CrossRef] [PubMed]
12. Rasheed, K.; Qadir, J.; O'Brien, T.J.; Kuhlmann, L.; Razi, A. A generative model to synthesize EEG data for epileptic seizure prediction. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2021**, *29*, 2322–2332. [CrossRef] [PubMed]
13. Xu, Y.; Yang, J.; Sawan, M. Multichannel synthetic preictal EEG signals to enhance the prediction of epileptic seizures. *IEEE Trans. Biomed. Eng.* **2022**, *69*, 3516–3525. [CrossRef] [PubMed]
14. Yu, T.; Cui, B.; Xu, Y.; Liu, X. Refine EEG Spectrogram Synthesized by Generative Adversarial Network for Improving The Prediction of Epileptic Seizures. In *Proceedings of the 2023 11th International IEEE/EMBS Conference on Neural Engineering (NER), Baltimore, MD, USA, 24–27 April 2023*; IEEE: Piscataway Township, NJ, USA, 2023; pp. 1–4.
15. Yang, X.; Liu, L.; Li, Z.; Xia, Y.; Fan, Z.; Zhou, J. Semi-Supervised Seizure Prediction Model Combining Generative Adversarial Networks and Long Short-Term Memory Networks. *Appl. Sci.* **2023**, *13*, 11631. [CrossRef]
16. Feng, Y.; Zhou, T.; Tian, Y.; Wu, C.; Wang, J.; Feng, J.; Li, J. An EEG-based patient-independent epileptic seizure detection method based on domain generative adversarial network. *Health Inf. Sci. Syst.* **2025**, *13*, 50. [CrossRef] [PubMed]
17. Guttag, J. *CHB-MIT Scalp EEG Database*, version 1.0.0; PhysioNet: London, UK, 2010. [CrossRef]
18. Goldberger, A.L.; Amaral, L.A.N.; Glass, L.; Hausdorff, J.M.; Ivanov, P.C.; Mark, R.G.; Mietus, J.E.; Moody, G.B.; Peng, C.K.; Stanley, H.E. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* **2000**, *101*, e215–e220. [CrossRef] [PubMed]
19. Homan, R.W.; Herman, J.; Purdy, P. Cerebral location of international 10–20 system electrode placement. *Electroencephalogr. Clin. Neurophysiol.* **1987**, *66*, 376–382. [CrossRef] [PubMed]
20. Huang, X.; Meng, H.; Li, Z. Epileptic Seizure Detection Using Hyperdimensional Computing and Binary Naive Bayes Classifier. *Bioengineering* **2025**, *12*, 1327. [CrossRef] [PubMed]
21. Truong, N.D.; Kuhlmann, L.; Bonyadi, M.R.; Yang, J.; Faulks, A.; Kavehei, O. Supervised learning in automatic channel selection for epileptic seizure detection. *Expert Syst. Appl.* **2017**, *86*, 199–207. [CrossRef]
22. Mirza, M.; Osindero, S. Conditional generative adversarial nets. *arXiv* **2014**, arXiv:1411.1784. [CrossRef]
23. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *Proceedings of the International Conference on Machine Learning, Lille, France, 6–11 July 2015*; PMLR: Cambridge, MA, USA, 2015; pp. 448–456.
24. Georgis-Yap, Z.; Popovic, M.R.; Khan, S.S. Preictal-Interictal Classification for Seizure Prediction. In Proceedings of the Canadian AI, Toronto, ON, Canada, 30 May–3 June 2022.
25. Gao, B.; Zhou, J.; Yang, Y.; Chi, J.; Yuan, Q. Generative adversarial network and convolutional neural network-based EEG imbalanced classification model for seizure detection. *Biocybern. Biomed. Eng.* **2022**, *42*, 1–15. [CrossRef]
26. You, S.; Cho, B.H.; Yook, S.; Kim, J.Y.; Shon, Y.M.; Seo, D.W.; Kim, I.Y. Unsupervised automatic seizure detection for focal-onset seizures recorded with behind-the-ear EEG using an anomaly-detecting generative adversarial network. *Comput. Methods Programs Biomed.* **2020**, *193*, 105472. [CrossRef] [PubMed]
27. Guan, Y.; Koerner, J.; Valiante, T.A.; Genov, R.; O'Leary, G. Generative adversarial network-based synthetic seizure dataset augmentation. In *Proceedings of the 2021 10th International IEEE/EMBS Conference on Neural Engineering (NER), Virtual, 4–6 May 2021*; IEEE: Piscataway Township, NJ, USA, 2021; pp. 797–800.