

# Semi-supervised Seizure Prediction with Generative Adversarial Networks

Nhan Duy Truong\*, *Student Member, IEEE*, Luping Zhou\*, *Senior Member, IEEE*,  
Omid Kavehei\*<sup>†</sup>, *Senior Member, IEEE*

**Abstract**—Many outstanding studies have reported promising results in seizure prediction that is considered one of the most challenging predictive data analysis. This is mainly because electroencephalogram (EEG) bio-signal intensity is very small, in  $\mu\text{V}$  range, and there are significant sensing difficulties given physiological and non-physiological artifacts. In this article, we propose an approach that can make use of not only labeled EEG signals but also the unlabeled ones which are more accessible. We also suggest the use of data fusion to further improve the seizure prediction accuracy. Data fusion in our vision includes EEG signals, cardiogram signals, body temperature and time. We use the short-time Fourier transform on 28-s EEG windows as a pre-processing step. A generative adversarial network (GAN) is trained in an unsupervised manner where information of seizure onset is disregarded. The trained Discriminator of the GAN is then used as feature extractor. Features generated by the feature extractor are classified by two fully-connected layers (can be replaced by any classifier) for the labeled EEG signals. This semi-supervised seizure prediction method achieves area under the operating characteristic curve (AUC) of 77.68% and 75.47% for the CHBMIT scalp EEG dataset and the Freiburg Hospital intracranial EEG dataset, respectively. Unsupervised training without the need of labeling is important because not only it can be performed in real-time during EEG signal recording, but also it does not require feature engineering effort for each patient.

**Index Terms**—seizure prediction, adversarial networks, convolutional neural network, machine learning, iEEG, sEEG.

## I. INTRODUCTION

EPILEPSY affects almost 1% of the global population and considerably impacts the quality of life of those patients diagnosed with the disease [1]. Over the past two decades, a tremendous number of techniques on predicting seizure has been proposed with promising performance. An early approach based on similarity, correlation and energy of EEG signals achieve a modest sensitivity of 42% and a false prediction rate (FPR) less than 0.15/h tested with the Freiburg Hospital dataset [2]. The performance improved with the use phase coherence and synchronization information in EEG signals, resulting sensitivity 60% and FPR of 0.15/h in [3] and 95.4% and FPR of 0.36/h in [4]. Support vector machine with frequency bands of the spectral energy as inputs further boosted the performance to 98.3% and FPR of 0.29/h [5] tested with

the Freiburg Hospital dataset. In our recent work [1], we showed that convolutional neural networks (CNNs) can be used as an effective seizure prediction method. In this work, we exploit deep convolutional generative adversarial network (GAN) [6] as an unsupervised technique to extract features that can be used for seizure prediction task. The extracted features can be classified by any classifier (a neural network with two fully-connected layers in this work).

Structure of this article is as follows. We first introduce the datasets being used in this work. Next, we describe how EEG signals are pre-processed. Then we provide details on GAN and how it can be used as feature extractor for seizure prediction. Lastly, we evaluate our approach and discuss on the results on two different datasets.

TABLE I: Summary of the two datasets used in this paper.

Dataset	EEG type	No. of patients	No. of channels	No. of seizures*	Interictal hours
Freiburg	intracranial	13	6	59	311.4
CHB-MIT	scalp	13	22	64	209

\* We are considering leading seizures only. Seizures that are less than 30 min away from the previous one are considered as one seizure only, and the onset of leading seizure is used as the onset of the combined seizure.

## II. PROPOSED METHOD

### A. Dataset

Table I summarizes the two datasets being used in this work: the CHB-MIT dataset [7] and the Freiburg Hospital dataset [8]. The CHB-MIT dataset contains scalp EEG (sEEG) data of 23 pediatric patients with 844 hours of continuous sEEG recording and 163 seizures. Scalp EEG signals were captured using 22 electrodes at sampling rate of 256 Hz [7]. We define interictal periods that are at least 4 h away before seizure onset and after seizure ends. In this dataset, there are cases that multiple seizures occur close to each other. For the seizure prediction task, we are interested in predicting the leading seizures [1]. Besides, we only consider patients with less than 10 seizures a day for the prediction task because it is not very critical to perform the task for patients having a seizure every 2 hours on average. With the above definition and consideration, there are 13 patients with sufficient data (at least 3 leading seizures and 3 interictal hours).

The Freiburg Hospital dataset consists of intracranial EEG (iEEG) recordings of 21 patients with intractable

\*School of Electrical and Information Engineering, The University of Sydney, NSW 2006, Australia.

<sup>†</sup>Corresponding author: omid.kavehei@sydney.edu.au.

epilepsy. Due to the lack of availability of the dataset, we are only able to use data from 13 patients. A sampling rate of 256 Hz was used to record iEEG signals. In this dataset, there are 6 recording channels from 6 selected contacts where three of them are from epileptogenic regions, and the other three are from the remote regions. For each patient, there are at least 50 min preictal data and 24 h of interictal. More details about Freiburg dataset can be found in [2].

### B. Pre-processing

Since we will use a Generative Neural Network (GAN) architecture with three de-convolution layers, dimensions of GAN's input must be divisible by  $2^3$ , except the number of channels. Besides, we are interested in whether GAN can be effectively trained with non-patient specific data, all patients must have the same number of channels so that data from all patients can be combined. We follow the approach in [9] to select 16 channels for each patient in CHBMIT dataset. Subsequently, we use Short-Time Fourier Transform (STFT) to translate 28 seconds of time-series EEG signal into two-dimensional matrix comprised of frequency and time axes. For the STFT, we use cosine window of 1-second length and 50% overlap. The power line noise can be removed by excluding components at the frequency range of 47–53 Hz and 97–103 Hz if the power frequency is 50 Hz and components at the frequency range of 57–63 Hz and 117–123 Hz for the power line frequency of 60 Hz. The DC component (at 0 Hz) was also removed (see Fig. 1). We also trim components at the last two frequencies 127–128 Hz to have the final dimension of each pre-processed 28 s be  $n \times 56 \times 112$ , where  $n = 16$  for CHBMIT dataset and  $n = 6$  for Freiburg dataset.

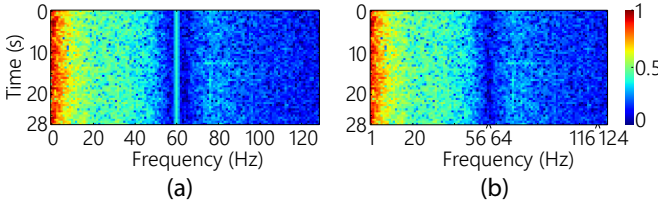


Fig. 1: (a) Example STFT of a 28 second window. (b) Same window after removing line noise.

### C. Adversarial Neural Network

In this paper, we use a Deep Convolutional Generative Adversarial Network (DCGAN) [6] as depicted in Fig. 2 as an unsupervised feature extraction technique. Note that here we explain for the CHBMIT dataset. The same explanation is applied for the other dataset with the change in input dimension as mentioned in Section II-B. The Generator takes a random sample of 100 data points from a uniform distribution  $\mathcal{U}(-1, 1)$  as input. The input is fully-connected with a hidden layer with output size of 6272 which is then reshaped to  $64 \times 7 \times 14$ . The hidden

layer is followed by three de-convolution layers with filter size  $5 \times 5$ , stride  $2 \times 2$ . Numbers of filters of the three de-convolution layers are 32, 16 and  $n$ , respectively. The Discriminator, on the other hand, is configured to discriminate the generated EEG signals from the original ones. The Discriminator consists of three convolution layers with filter size  $5 \times 5$ , stride  $2 \times 2$ . Numbers of filters of the three convolution layers are 16, 32 and 64, respectively. During training, the Generator tries to generate signals that "look" like the original ones while the Discriminator is optimized to detect those generated signals. As a result, the Discriminator learns how to extract unique features in the original EEG signals by adjusting its parameters in the three convolution layers. This training process is unsupervised because we do not provide labels (preictal or interictal) to the network.

The idea of training a generative adversarial network is that the Discriminator and Generator compete each other and finally reach to an equilibrium [10]. However, when we first started training the DCGAN, we observed that the Discriminator converged too fast. This prevents the Generator from learning how to generate good STFT samples. As a result, classification between generated STFT samples and original ones becomes a trivial task. To overcome this, we update the Generator twice instead of once every mini-batch as suggested in [11] and configure a monitor of loss values the Discriminator ( $D_{\text{loss}}$ ) and Generator ( $G_{\text{loss}}$ ). The monitor stops the DCGAN training if  $D_{\text{loss}}$  keeps being larger than  $G_{\text{loss}}$  over  $k$  consecutive mini-batches. In this work, we used  $k = 20$  and Adam optimizer for gradient-based learning with a learning rate of  $1e^{-4}$ ,  $\beta_1 = 0.5$ ,  $\beta_2 = 0.999$ , and  $\epsilon = 1e^{-8}$ .

### D. Seizure prediction with features extracted by DCGAN

We use the trained convolution layers in the DCGAN's Discriminator as a feature extractor. Specifically, we feed STFT of 28-second EEG signals into the Discriminator and collect the flatten features at its last convolution layer's output. Those features can now be used with any classifier to perform the seizure prediction task. In this paper, we use a simple neural network consisting of two fully-connected layers with sigmoid activation and output sizes of 256 and 2, respectively. The former layer uses sigmoid activation function while the latter uses soft-max activation function. Both of the two layers have a drop-out rate of 0.5. We also apply a practice proposed in [1] to prevent over-fitting during the training of the neural network.

### E. System evaluation

Seizure prediction horizon (SPH) and seizure occurrence period (SOP) need to be defined before estimating the system's performance. In this paper, we follow the definition of SOP and SPH that was proposed in [2]. SOP is the interval where the seizure is expected to occur. The period between the alarm and beginning of SOP is called

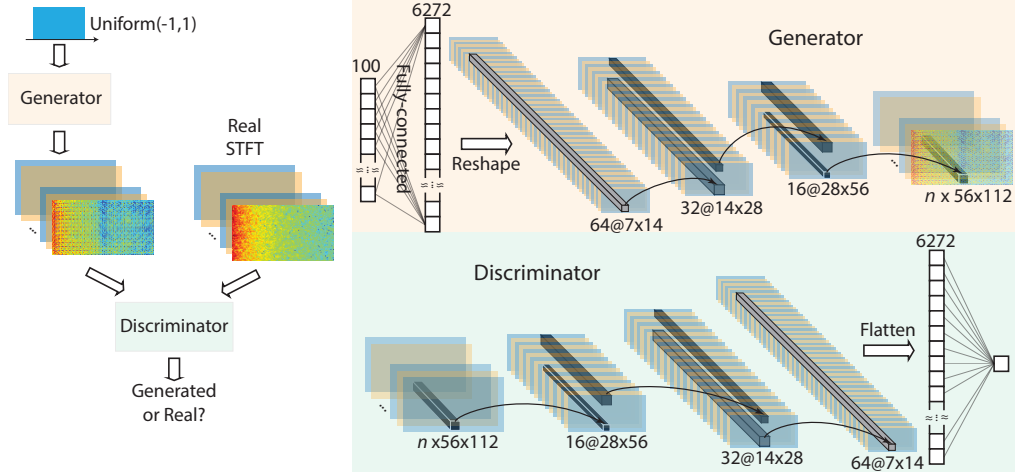


Fig. 2: The Generator takes a random sample of 100 data points from a uniform distribution  $\mathcal{U}(-1,1)$  as input. The input is fully-connected with a hidden layer with output size of 6272 which is then reshaped to  $64 \times 7 \times 14$ . The hidden layer is followed by three de-convolution layers with filter size  $5 \times 5$ , stride  $2 \times 2$ . The Discriminator consists of three convolution layers with filter size  $5 \times 5$ , stride  $2 \times 2$ .

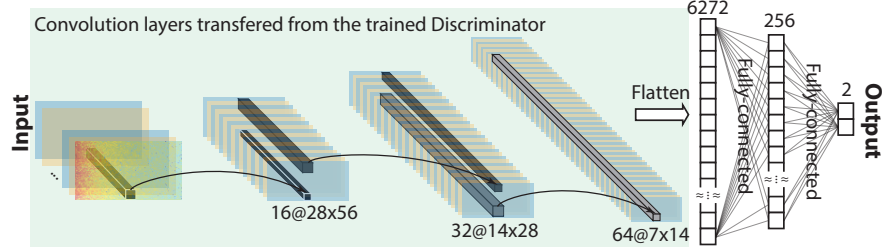


Fig. 3: Convolutional neural network architecture. This illustration is applied to Freiburg and CHB-MIT datasets. Input are STFT transforms of 28s windows of raw EEG signals. Features extracted by the three convolution blocks of the Discriminator are flattened and connected to a neural network consisting of 2 fully-connected layers.

SPH. For a correct prediction, a seizure onset must be after the SPH and within the SOP. Likewise, a false alarm rises when the prediction system returns a positive but there is no seizure occurring during SOP.

We use area under the receiver operating characteristics curve (AUC) with SPH of 5 min and SOP of 30 min. To have a robust evaluation, we follow a leave-one-out cross-validation approach as described in [1].

### III. RESULTS

In this section, we test our approach with two datasets: the CHB-MIT sEEG dataset and the Freiburg Hospital iEEG dataset. SOP = 30 min and SPH = 5 min were used in calculating all metrics in this paper. Each fold of leave-one-out cross-validation was executed twice, and average results with standard deviations were reported. Fig. 4 summarizes seizure prediction results with SOP of 30 min and SPH of 5 min. Our model training is performed on an NVIDIA P100 graphic card using Tensorflow 1.4.0 framework. The results are shown in Fig. 4.

Compared to the fully supervised CNN, semi-supervised GAN-CNN introduces  $\sim 6\%$  and  $\sim 12\%$  loss in AUC on average for the CHBMIT sEEG dataset and the Freiburg

Hospital iEEG dataset, respectively. Regarding the CHBMIT dataset, the semi-supervised approach can deliver good performance (with  $AUC \geq 80\%$ ) and comparable with that of the fully supervised training for 8 out of 13 patients. In respect of the Freiburg Hospital dataset, 6 out of 13 patients have high seizure prediction performance. The main contributors for AUC gap are Pat. 6 and 15 who have only 3 and 4 seizures, respectively.

### IV. DISCUSSION

We have shown that feature extraction for seizure prediction can be performed in an unsupervised way. Though the overall AUC degraded by  $\sim 6\text{--}12\%$  across the two datasets, our unsupervised feature extraction can help to minimize the EEG labeling task that is costly and time-consuming. In the field of computer vision, GAN can help to reduce the amount of labeled data without compromising the classification performance [12]. With GAN, we can use unlabeled EEG signals to train the GAN's Generator and Discriminator. The trained GAN's Discriminator then plays as a feature extractor. Extracted features from labeled EEG data (that can be much smaller than unlabeled one) can be fed to any classifier

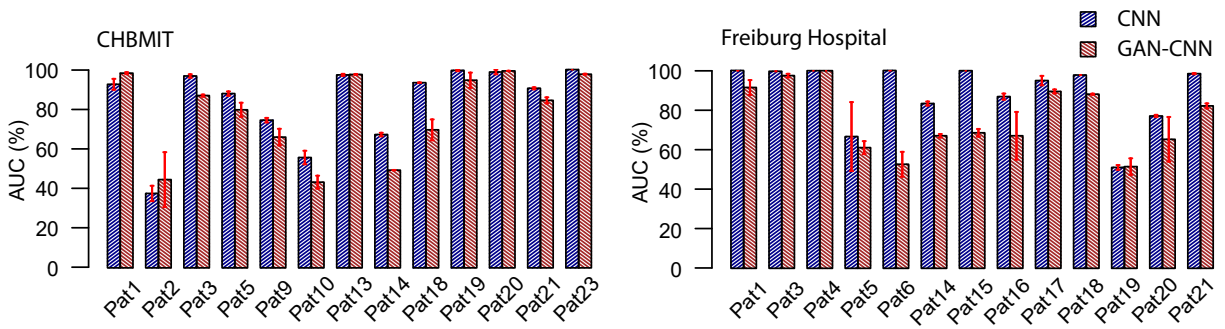


Fig. 4: Seizure prediction performance for the CHBMIT dataset (upper) and the Freiburg Hospital dataset (lower). Four methods are evaluated: (1) CNN: convolutional neural network, (2) GAN-CNN: unsupervised feature extraction using GAN and classification performed by a CNN.

(two fully-connected layers in our work) for the seizure prediction task.

## V. CONCLUSION

Seizure prediction capability has been studied and improved over the last four decades. A perfect prediction is yet available but with current prediction performance, it is useful to provide the patients with warning message so they can take some precautions for their safety. We have shown that feature extraction for seizure prediction can be done using unsupervised deep learning or GAN particularly. Seizure prediction can be implemented efficiently on low-power hardware. Though our working prototype that uses off-the-shelf devices does not provide impressive power consumption, it is obviously that power consumption can be greatly reduced with customized devices. This will help patients with epilepsy to have a more manageable life with a seizure prediction device.

## VI. ACKNOWLEDGMENT

This research is funded by a Sydney Research Accelerator (SOAR) Fellowship and an Early Career Research grant through the University of Sydney. N.D.T. acknowledges partial support provided via John Makepeace Bennett Gift Scholarship funded by the Australian Institute for Nanoscale Science and Technology (AINST) and administered by the University of Sydney Nano Institute. The authors also acknowledge the support by Sydney Informatics Hub via the University of Sydney's Core Research Facilities.

## REFERENCES

- [1] N. D. Truong *et al.*, "Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram," *Neural Networks*, vol. 105, pp. 104–111, 2018.
- [2] T. Maiwald *et al.*, "Comparison of three nonlinear seizure prediction methods by means of the seizure prediction characteristic," *Physica D: Nonlinear Phenomena*, vol. 194, no. 3-4, pp. 357–368, 2004.
- [3] M. Winterhalter *et al.*, "Spatio-temporal patient-individual assessment of synchronization changes for epileptic seizure prediction," *Clinical Neurophysiology*, vol. 117, no. 11, pp. 2399–2413, 2006.
- [4] M. Z. Parvez *et al.*, "Seizure prediction using undulated global and local features," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 1, pp. 208–217, 2017.
- [5] Y. Park *et al.*, "Seizure prediction with spectral power of EEG using cost-sensitive support vector machines," *Epilepsia*, vol. 52, no. 10, pp. 1761–1770, 2011.
- [6] A. Radford *et al.*, "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv preprint arXiv:1511.06434*, 2015.
- [7] A. H. Shueb, "Application of machine learning to epileptic seizure onset detection and treatment," Ph.D. dissertation, Massachusetts Institute of Technology, 2009.
- [8] University of Freiburg, "EEG Database at the Epilepsy Center of the University Hospital of Freiburg, Germany," 2003. [Online]. Available: <http://epilepsy.uni-freiburg.de>
- [9] N. D. Truong *et al.*, "Supervised learning in automatic channel selection for epileptic seizure detection," *Expert Systems with Applications*, vol. 86, pp. 199–207, 2017.
- [10] I. Goodfellow *et al.*, "Generative Adversarial Nets," *Advances in Neural Information Processing Systems*, pp. 2672–2680, 2014.
- [11] S.-H. Sun, "Deep Convolutional Generative Adversarial Networks in Tensorflow," 2017. [Online]. Available: <https://github.com/shaohua0116/DCGAN-Tensorflow>
- [12] D. P. Kingma *et al.*, "Semi-supervised Learning with Deep Generative Models," in *Advances in Neural Information Processing Systems 27*, Z. Ghahramani *et al.*, Eds. Curran Associates, Inc., 2014, pp. 3581–3589.