SSRG International Journal of Computer Science and Engineering
ISSN: 2348 – 8387 / https://doi.org/10.14445/23488387/IJCSE-V9I10P101

Volume 9 Issue 10, 1-7, October 2022

Original Article

Seizure Prediction using Generative Adversarial Networks for EEG Data Synthesis

Garv Agarwal[1], Sai Sanjeet[2], Bibhu Datta Sahoo[3]

[1]*Lakeland High School, New York, USA*
[2] *Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur, India*

Received: 26 August 2022 Revised: 03 October 2022 Accepted: 15 October 2022 Published: 28 October 2022

Abstract - *Epilepsy is a common neurological disease characterized by seizures. Automatic prediction of these seizures can help clinicians prepare for and manage patient seizures due to prior knowledge of seizure onset. Automatic seizure prediction is done using electroencephalography (EEG) data containing brain activity representing seizures. Deep learning classifiers have been attempted in predicting seizure onset but are hindered due to a lack of high-quality preictal data in the dataset compared to the amount of interictal data. Solutions to the issues of data scarcity and data imbalance have been tried, such as under-sampling and various oversampling methods; however, these methods have not been successful in creating ample data. We propose a DCGAN that generates synthetic high-quality preictal data for the seizure prediction task. The synthetic data is compared to random oversampling of preictal data on the CHB-MIT Scalp EEG database using a CNN classifier with a 5-12% improvement.*

Keywords - *Electroencephalography, Preictal, Interictal, Generative Adversarial Networks, Seizure prediction.*

1. Introduction

Epilepsy is a neurological disease caused by sudden abnormal brain activity characterized by seizures. Sixty-five million people worldwide have epilepsy, making it the fourth most common chronic neurological disorder [1]. Due to the sudden and erratic nature of the seizures, predictive measures for epilepsy can help support preventative treatments and give patients time for precautionary measures.

To record brain activity for epileptic patients, electroencephalography (EEG) is used. EEG data for epileptic patients have been used for machine learning research, particularly through seizure detection and prediction algorithms. Such algorithms use extracted features of EEG from both the time and frequency domains to find the period where a seizure onset occurs.

Electroencephalography data can be recorded by placing the electrodes on the scalp of a patient, known as scalp EEG, or by implanting the electrodes within a patient's brain tissue, known as intracranial EEG (iEEG) [2]. Even though iEEG data has a better signal-to-noise ratio and is thus easier to extract features from [3], the invasiveness of the method makes it more difficult to collect.

Within EEG data containing epileptic seizures, there are four phases of the seizure visible in the data. Normal brain activity in the time between two seizures is known as the interictal state, abnormal brain activity around 60 to 90 minutes prior to seizure onset is known as the preictal state, the period from the onset of the seizure to the end of the seizure is known as the ictal state, and the period after a seizure where the patient returns to baseline condition is known as the postictal state [32]. These four states are seen in Fig. 1.

This EEG data can be used for deep learning-based seizure detection and prediction tasks. The seizure detection model focuses on detecting ongoing seizures and is primarily used to provide clinicians with seizure data that can be useful for epilepsy management [5]. Seizure detection primarily looks to differentiate between seizure and non-seizure data [6]. So classification algorithms for seizure detection assess only the preictal and ictal states to detect when exactly a seizure occurs. Predictive models for epilepsy can help clinicians by forecasting seizure onset, allowing clinicians to prepare medications. Such predictive models have developed fairly recently, as previous neuroscientists believed epileptic seizures were sudden until analysis of EEG recordings showed that seizures were predictable [7]. In order to have ample time to prevent seizures, interictal and preictal states are assessed as if the change in states between interictal and preictal states is detected; clinicians can use this knowledge to prevent seizures during the preictal state.

Deep Learning (DL) classifiers have been a promising way to approach the seizure prediction task. Various methods

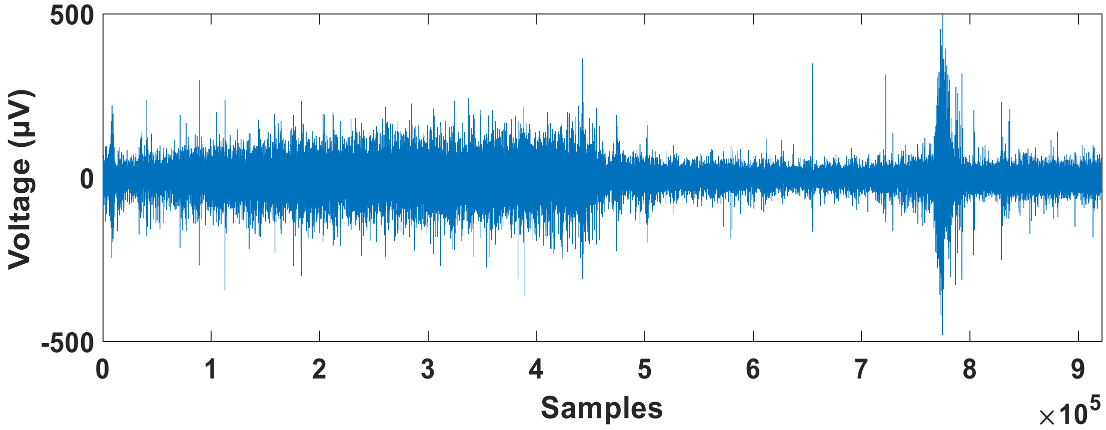

Fig. 1 One channel of EEG data depicting the interictal, preictal, ictal, and postictal states.

have been used to extract DL algorithms' features, including short-time Fourier Transform and wavelet Transform for converting EEG data from the time domain to the frequency domain.

There are also many variations between the deep learning classifiers themselves, as convolutional neural networks (CNN) and recurrent neural networks (RNN) have been used to approach classification between preictal and ictal states.

Though DL algorithms have been attempted to approach seizure prediction, the major challenge is that high amounts of data are required to obtain meaningful results. The current state-of-the-art machine learning-based approaches to epilepsy prediction require vast amounts of data [8]. However, this data is difficult to collect manually, as it requires patients to be clinically monitored in hospitals, leading to datasets containing few patients and a limited amount of EEG recording for each one [8]. Data imbalance is another major issue, as interictal data is much more abundant than preictal data within EEG datasets. Solutions to the data imbalance problem have been attempted through undersampling and oversampling methods like random oversampling [9] and SMOTE [10].

However, they have not successfully created ample data that seizure prediction classifiers can use.

Data scarcity and imbalance issues can be solved through synthetic data generation. Generative Adversarial Networks [11] have been widely used for data generation to create synthetic data of high quality. Synthetic data has been used for many machine-learning tasks, including medical imaging [33] [13] [14] and Intensive Care Unit (ICU) monitoring [15] [16] [17]. Data generated by GANs have even been used for epilepsy data [18], and this synthetic epilepsy data generated by GANs has been developed for the seizure prediction task in the past [19].

We propose a model that improves existing GAN-based seizure prediction models [19]. Our work contributes:

1. A Generative Adversarial Network that is used to generate synthetic data and resolve the issue of data scarcity of pre-ictal data on scalp-EEG data.
2. A CNN classification network for seizure prediction is trained and tested on real and synthetic data to determine the value of synthetic data for improving state-of-the-art results.

2. Epilepsy Prediction: A Review

The rest of this section will discuss various methods of epilepsy prediction. A summary of the various works discussed in this section will be in Table 1.

2.1. Feature Extraction

As EEG data containing epileptic seizures contains time domain data, it is intuitive to extract data from the time domain with Robust Generalized Synchrony [26]. However, extracting features from frequency domain data instead of time domain data is much more common. It is done in order to remove noise in the time domain and extract better features from frequency domain data. The Fast Fourier Transform (FFT) has been used to extract features from the frequency domain for both epilepsy detection and prediction by calculating the discrete Fourier transform with efficiency, particularly by Lee et al. 2017 [27] and Chu et al. 2017 [20]. Similarly, the Short Time Fourier Transform (STFT) implements various Fourier transforms over a sliding window. The exact calculation for the Fourier transform is shown below.

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot e^{-i\frac{2\pi nk}{N}} \qquad (1)$$

Table 1. Summary of Discussed Feature Extraction and Classification Methods for Seizure Prediction

Seizure Prediction Method	Feature Extraction	Classification Method	Sensitivity	Specificity	FPR/h
Chu et al. 2017 [20]	FFT	Threshold between epileptic states	86.7	86.7	0.367
Truong et al. 2018 [21]	STFT	CNN	81.2	84.0	0.367
Qin et al. 2019 [22]	STFT	CNN and ELM	95.85	-	0.045
Khan et al. 2017 [3]	DWT	CNN	87.8	85.8	0.142
Feudjio et al. 2021 [23]	DWT	Random Forest	99.1	98.0	-
Cho et al. 2017 [24]	EMD/Wavelet	SVM	80.54	80.50	-
Chaovalitwongse et al. 2007 [25]	Time-Series Data	KNN	81.29	72.86	-
Rasheed et al. 2021 [19]	STFT	GAN + CNN	96.0	-	0.05

STFT has been used to extract various features from frequency domain data by Truong et al. 2018 [21] and Qin et al. 2019. [22] The STFT can have various sliding windows, but 30-second windows are commonly used. While feature extraction methods using the Fourier transform have been attempted, the disadvantage of using the Fourier transform is that it tends to contain global frequency information, which is not necessarily important in seizure prediction due to the relatively fast oscillation of EEG data. Wavelet Transforms and Discrete Wavelet Transforms (DWT) are used to develop time-frequency representations of data by decomposing the function into wavelets [34] and have been used as a feature extraction method in epilepsy prediction by Khan et al. 2017 [3] and Feudjio et al. 2021 [23]. A less used but notable method for feature extraction includes Empirical Mode Decomposition (EMD), which attempts to extract features from data through the time-frequency-amplitude domains by decomposing the signal into intrinsic mode functions. It has been attempted by Cho et al. 2017 [24].

2.2. Classification Methods

After feature extraction has been performed, there are a myriad of classification methods that can be attempted to classify between interictal and preictal states for seizure prediction. Many feature extraction methods are traditional machine learning classifiers, such as k-nearest neighbor (KNN), which classifies test samples based on the samples' labels. Depending on the labels of the surrounding samples, the algorithm chooses to classify the test sample. The smaller the distance between the test sample and a "neighbor sample," the more similar the two samples are. This approach was attempted by Chaovalitwongse et al. 2007 [25], where an 81% sensitivity and a 73 % specificity were achieved using T-statistical distance. Naive Bayesian models have been attempted by Sharmila and Geethanjali 2016 [29] for the seizure detection problem, where the Naive Bayes algorithm uses Bayes' theorem for classification while assuming maximum independence between the different samples. They achieved accuracies between 90 % and -100 % for various patients. Support vector machines (SVM), such as the ones used by James and Gupta 2009 [30] for epilepsy prediction, attempt to develop a hyperplane that can clearly divide the samples to separate interictal and preictal data. A sensitivity of 86 % and a specificity of 80 % were achieved with this model. The random forest classifier uses a series of decision trees to make its predictions between interictal and preictal data. Feujdio et al. 2021 [23] used the random forest for prediction and obtained a 99.1 % mean sensitivity and a 98.0% specificity. Chu et al. 2017 [20] used a threshold for differentiating between interictal and preictal states, achieving an 86.67 sensitivity and specificity.

Though traditional machine learning models have had some success with the epilepsy prediction task, deep learning approaches to the epilepsy prediction task are much better, as they can engineer features automatically. The Convolutional Neural Network (CNN) is a feedforward deep learning algorithm that uses layers of neurons to extract valuable data that can be used for classification for seizure prediction. Truong et al. 2018 [21] achieved an 81.2% sensitivity and an 84% specificity with a 5-minute prediction time using a CNN and an STFT. Khan et al. 2018 [3] used the Discrete Wavelet Transform for feature extraction and achieved a sensitivity of 87.8% and a specificity of 85.8%. The Extreme Learning Machine is another feedfoward deep learning algorithm that does not need its neurons to be tuned. Qin et al. 2019 [22] used this network and an STFT to achieve a 95.85% sensitivity and a false prediction rate of 0.045 per hour.

Major issues with using machine learning and deep learning algorithms are the lack of data and the imbalance of data between the interictal and preictal states. To solve for data imbalance, undersampling and oversampling methods have been implemented. Random oversampling [9] and SMOTE [10] have been implemented for imbalances in the seizure detection task. Generative Adversarial Networks (GAN) are promising for solving data imbalance and scarcity. Generative Adversarial Networks are used to generate data to increase the accuracy of data-driven classifiers. It is done through a generator and a discriminator, where the generator generates random noise and makes adjustments to the network based on the information given to it by the discriminator, which differentiates between real data and data generated by the generator. Rasheed et al. 2021 [19] used a Deep Convolutional GAN to improve the accuracy of a CNN-based seizure prediction algorithm with a sensitivity of 88.21% and a false prediction rate of 0.14/h when training with synthetic data and testing with real data.

3. Materials and Methods

3.1. Dataset

The proposed framework is being run on the CHB-MIT Scalp EEG database [31]. The dataset consists of 23 cases from 22 patients, with each case containing between 9 to 42 continuous files of EEG data in the .edf data format. The scalp EEG data was mostly collected in 23 channels with 256 samples per second as a sampling rate. There are a collection of 182 seizures within the 23 cases. However, due to some consecutive seizures having a short time horizon between them, they were grouped together as if they were one seizure for this work, where the preictal period for the leading seizure was the preictal period for the consecutive seizures. It is intuitive, as predicting the leading seizure would lead clinicians to take precautions for any consecutive seizures proceeding the leading one.

3.2. Data Preprocessing

The task of seizure prediction involves distinguishing between two states: interictal and preictal. The interictal state, as mentioned earlier, is the state of normal activity between two seizures. For this work, a recording segment is considered interictal if it is at least 4 hours away from the nearest seizure. The preictal state is defined as the recording segment starting from 65 minutes before the seizure onset up to 5 minutes prior to the onset. The period of recording 30 minutes after the seizure offset is the postictal period. In case of seizures in quick succession, the part of recording defined as postictal is not considered to be preictal. The data corresponding to 3 patients, viz., chb01, chb02, and chb03 is preprocessed using the above-defined rules, resulting in 65 hours of interictal and 12 hours of preictal data.

3.3. Feature Extraction

For feature extraction, STFT was used to convert from the time domain to the frequency domain. We used a window length of one second (256 samples) with no overlap between them. After taking the magnitude of the conjugate values of the STFT, the D.C. noise was removed, and the 60Hz power line noise by removing the frequencies from 57-63Hz and 117-123Hz. As one-minute samples are being used as training data for both the DCGAN and the CNN, the continuous data is split into the shape (X, 23, 114, 60), where X is the number one minute samples, 23 is the number of channels, 114 is the relevant frequencies, and 60 is the number of seconds per sample.

3.4. Generation of Synthetic Data

A DCGAN is used to generate scalp EEG data for the CHB-MIT Scalp EEG Database. The generator takes 100 randomly generated samples of Gaussian noise as its input, which is then reshaped to 4096 through a dense hidden layer to 256 x 4 x 4. The rest of the network consists of deconvolutional layers. The first deconvolutional layer has 128 output channels, a kernel size of 7 x 5, and a stride of 2 x 2. This deconvolutional layer is followed by a rectified linear unit (ReLU), while normalization follows all the following deconvolutional layers in addition to ReLU. The second deconvolutional layer has 64 output channels with a 3 x 5 kernel size and a 2 x 2 stride, and the third deconvolutional layer has 32 output channels with a 5 x 5 kernel size and a 2 x 1 stride. The final deconvolutional layer has 23 output channels with a kernel size of 5 x 5 and a stride of 2 x 2. It led to the generation of samples with the size of (23 x 115 x 61), so the first 114 samples of the frequency dimension and the first 60 samples of the time dimension were kept to maintain the same shape as the real data.

The goal of the discriminator is to distinguish between real data and the synthetic data generated by the generator. The ability of the discriminator to classify data correctly is then fed into the generator to generate spectrograms that are indistinguishable from real spectrograms. The network consists of four convolutional layers, and each convolutional layer of the discriminator is followed by normalization and ReLU. The first convolutional layer of the network has 32 output channels with a kernel size of 5 x 5 and a stride of 2 x 2. The second convolutional layer has a similar kernel size and stride as the first layer but has 64 output channels. The third convolutional layer has 128 output channels with a kernel size of 4 x 4 and a stride of 2 x 2. The final convolutional layer has 256 output channels with a kernel size of 5 x 5 and a stride of 2 x 2. The output of the last convolutional layer is then flattened with a fully connected layer and a sigmoid activation function.

The discriminator and the generator use binary cross-entropy loss and the Adam optimizer with a learning rate of 10^{-3}. The DCGAN was trained using only preictal samples to generate purely preictal data, as the lack of preictal data causes data imbalance for each patient. The synthetic preictal data was then added to the training set of the classification algorithm.

3.5. Epileptic Seizure Classification

Due to the success of CNN for the seizure prediction task, a CNN was used for the seizure prediction classifier. This classifier is trained using pure real and synthetic data to augment the training set and remove data imbalance from a lack of preictal data. The exact architecture of the network is shown in Fig. 2.

The network consists of three convolutional layers, followed by a rectified linear unit (ReLU), a 2 x 2 pooling layer, a batch normalization layer, and a dropout of 0.5. The convolutional layers consist of a kernel size of 3 x 3, a stride of 1 x 1, and 0 paddings. The pooling layers consisted of a kernel size of 2 x 2 and a stride of 2 x 2. The network ends with a fully connected linear layer and a sigmoid activation function. Using a batch size of 128, 70% of the interictal and two-thirds of the preictal instances were used for training, while the rest was used for testing. The training was done with 20 epochs with binary cross-entropy loss, an Adam optimizer, and a learning rate of 10^{-4}. For training and testing the CNN on real data, random oversampling was used to solve for data imbalance.

4. Results and Discussion

We generated preictal samples for three patients from the CHB-MIT Scalp EEG Database and compared data augmentation using the DCGAN to random oversampling using the CNN classifier for epilepsy prediction. The results of the experiment are shown in Table 2.

For Patient 1, when using random oversampling of real preictal data, we achieved 85.5% sensitivity and 84.1% specificity on the testing set. When using preictal samples generated by the DCGAN, we achieved 99.5% sensitivity and 95.1% specificity, depicting a ≈12.3% increase in results. Similarly, for Patient 2, we achieved 85.0 % sensitivity and 80.0% specificity on the testing set when using random oversampling of real preictal data, compared to the 93.5% sensitivity and 83.7% specificity we achieved when using preictal samples generated by the DCGAN, which is around a 6% increase in results. For Patient 3, we found an ≈ 5% increase in results when using samples generated by the DCGAN, as we achieved 85.8 % sensitivity and 84.1% specificity on the testing set with random oversampling. In comparison, we achieved 99.5% sensitivity and 95.1% specificity using the samples generated by the DCGAN.

This work aimed to improve the performance of state-of-the-art deep-learning-based seizure prediction classification methods using samples generated by a DCGAN. With improvements in deep learning technology, the generation of high-quality data from a small number of training samples can lead to the possibility of more robust datasets. This work proved the viability of using synthetic data for the epilepsy prediction task, particularly with deep-learning classifiers. However, using DCGANs to augment the data of other classifiers is also possible. Our work shows that the results of current state-of-the-art seizure prediction classifiers can be improved using data augmentation with DCGANs. This work outperformed other traditional methods of data augmentation, particularly random oversampling. Notably, methods of undersampling were attempted as traditional data augmentation without synthetic data for this work. However, due to the low amounts of preictal data for each patient, undersampling led to the weak performance of the CNN classifier.

The use of generative methods for data augmentation is promising, as shown by this work, especially in the medical field, due to privacy concerns regarding patient data. It is difficult to obtain large amounts of data due to data sharing between hospitals and privacy concerns, so applying generative methods to create synthetic data improves seizure prediction methods' performance while maintaining patient data privacy.

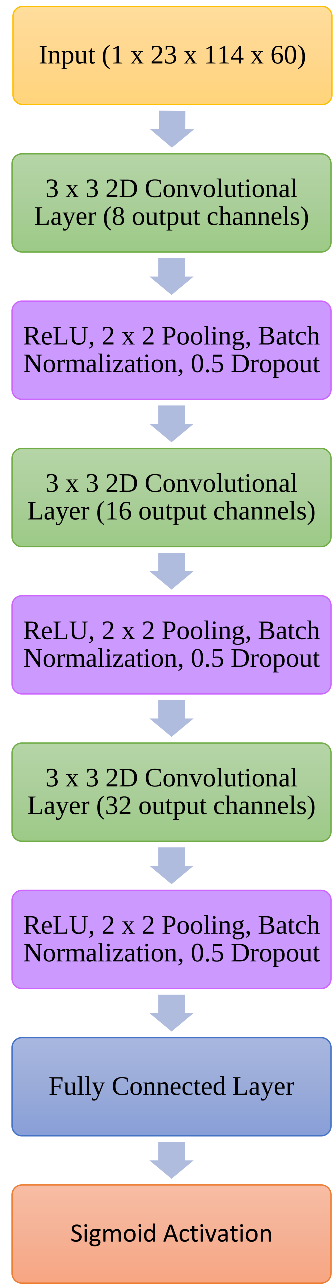

Fig. 2 Architecture for CNN-based Seizure Prediction Classifier

Table 2. Results of synthetic preictal samples of scalp EEG data using various combinations of synthetic and real data for training the CNN classifier. Synthesized data was generated from the DCGAN and was trained on the CHB-MIT scalp EEG dataset

Patient	Train and Test on Real Data		Train on Synthetic Data and Test on Real Data	
	Sensitivity	Specificity	Sensitivity	Specificity
Patient 1	85.8%	84.1%	99.5%	95.1%
Patient 2	85.0%	80.0%	93.2%	83.7%
Patient 3	95.3%	77.2%	97.7%	85.4%

5. Conclusion

Seizure prediction research attempts to develop a system that can alarm patients and clinicians before seizures occur. It has been attempted in various ways, but deep-learning-based approaches seem to be the most popular and successful in their performance. However, even with the success of these deep-learning-based approaches, there is a major lack of data, primarily preictal data, that can be used to train deep-learning networks to achieve satisfactory results. There are two explanations for this, the first being the difficulty in collecting data from seizure patients, as it requires patients to undergo seizures while being monitored, and the second being the relative abundance of interictal data, as that data is much more commonly collected during the monitoring of seizure patients. It ultimately leads to data scarcity and imbalance, which hinders the performance of deep-learning classifiers. We proposed a DCGAN to solve data scarcity and imbalance by generating synthetic samples of preictal data. Through the validation of the DCGAN using a CNN classifier for the seizure prediction task, we achieved up to a 12% increase in seizure performance compared to using random oversampling for data augmentation. Though our proposed framework does indeed increase the performance of state-of-the-art seizure prediction algorithms, to implement this prediction system in a real-world setting, some embedded devices will be required, which can be used to monitor patient brain activity and predict seizure onset in real time. This real-world implementation can be further explored in future work.

References

[1] D. Hirtz, D. J. Thurman, K. Gwinn-Hardy, M. Mohamed, A. R. Chaud-huri, and R. Zalutsky, "How Common are the "Common" Neurologic Disorders?," *Neurology,* vol. 68, no. 5, pp. 326–337, 2007.

[2] S. M. Usman, S. Khalid, R. Akhtar, Z. Bortolotto, Z. Bashir, and H. Qiu, "Using Scalp Eeg and Intracranial Eeg Signals for Predicting Epileptic Seizures: Review of Available Methodologies," *Seizure,* vol. 71, pp. 258–269, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1059131119302213.

[3] H. Khan, L. Marcuse, M. Fields, K. Swann, and B. Yener, "Focal Onset Seizure Prediction using Convolutional Networks," *IEEE Transactions on Biomedical Engineering,* vol. 65, no. 9, pp. 2109–2118, 2018.

[4] Ms.Manjula B.M and Dr.Chirag Sharma, "FPGA Implementation of BCG Signal Filtering Scheme by using Weight Update Process," *SSRG International Journal of VLSI & Signal Processing*, vol. 3, no. 3, pp. 1-7, 2016. Crossref, https://doi.org/10.14445/23942584/IJVSP-V3I5P101

[5] S. Ramgopal, S. Thome-Souza, M. Jackson, N. E. Kadish, I. Sanchez, Fernandez, J. Klehm, W. Bosl, C. Reinsberger, S. Schachter, and T. Loddenkemper, "Seizure Detection, Seizure Prediction, and Closed-Loop Warning Systems in Epilepsy," *Epilepsy Behavior,* vol. 37, pp. 291–307, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1525505014002297.

[6] M. K. Siddiqui, R. Morales-Menendez, X. Huang, and N. Hussain, "A Review of Epileptic Seizure Detection using Machine Learning Classifiers," *Brain Informatics,* vol. 7, no. 1, 2020.

[7] B. Litt and J. Echauz, "Prediction of Epileptic Seizures," *The Lancet Neurology,* vol. 1, no. 1, pp. 22–30, 2002.

[8] T. N. Alotaiby, S. A. Alshebeili, T. Alshawi, I. Ahmad, and F. E. Abd El-Samie, "Eeg Seizure Detection and Prediction Algorithms: A Survey," *EURASIP Journal on Advances in Signal Processing,* vol. 2014, no. 1, 2014.

[9] M. Masum, "Integrated Machine Learning Approaches to Improve Classification Performance and Feature Extraction Process for EEG Dataset," *Analytics and Data Science Dissertations,* 2021.

[10] P. Kaur, V. Bharti, and S. Maji, "Enhanced Epileptic Seizure Detection using Imbalanced Classification," *International Journal of Recent Technology and Engineering,* vol. 9, no. 1, pp. 2412–2420, 2020.

[11] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Nets," in *Proceedings of the 27th International Conference on Neural Information Processing Systems,* ser. NIPS'14, Cambridge, MA, USA: MIT Press, vol. 2, pp. 2672–2680, 2014.

[12] Anu V. S.and Paul Thomas, "An Improved Method for Classification of Epileptic EEG Signals based on Spectral Features using k-NN," *SSRG International Journal of Electronics and Communication Engineering,* vol. 2, no. 7, pp. 22-25, 2015. Crossref, https://doi.org/10.14445/23488549/IJECE-V2I7P108.

[13] H.-C. Shin, N. A. Tenenholtz, J. K. Rogers, C. G. Schwarz, M. L. Senjem, J. L. Gunter, K. P. Andriole, and M. Michalski, "Medical Image Synthesis for Data Augmentation and Anonymization Using Generative Adversarial Networks," *Simulation and Synthesis in Medical Imaging,* pp. 1–11, 2018.

[14] P. Costa, A. Galdran, M. I. Meyer, M. Niemeijer, M. Abramoff,` A. M. Mendonc¸a, and A. Campilho, "End-to-End Adversarial Retinal Image Synthesis," *IEEE Transactions on Medical Imaging,* vol. 37, no. 3, pp. 781–791, 2018.

[15] C. Esteban, S. L. Hyland, and G. Ratsch,¨ "Real-Valued (Medical) Time Series Generation with Recurrent Conditional Gans," 2017.

[16] Z. Che, Y. Cheng, S. Zhai, Z. Sun, and Y. Liu, "Boosting Deep Learning Risk Prediction with Generative Adversarial Networks for Electronic Health Records," *2017 IEEE International Conference on Data Mining (ICDM)*, 2017.

[17] Z. C. Lipton, D. C. Kale, and R. Wetzel, "Modeling Missing Data in Clinical Time Series with RNNs," 2016.

[18] D. Pascual, A. Aminifar, D. Atienza, P. Ryvlin, and R. Wattenhofer, "Synthetic Epileptic Brain Activities using Generative Adversarial Networks," 2019.

[19] K. Rasheed, J. Qadir, T. J. O'Brien, L. Kuhlmann, and A. Razi, "A Generative Model to Synthesize EEG Data for Epileptic Seizure Prediction," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 2322–2332, 2021.

[20] H. Chu, C. K. Chung, W. Jeong, and K.-H. Cho, "Predicting Epileptic Seizures from Scalp EEG Based on Attractor State Analysis," *Computer Methods and Programs in Biomedicine*, vol. 143, pp. 75–87, May 2017.

[21] N. D. Truong, A. D. Nguyen, L. Kuhlmann, M. R. Bonyadi, J. Yang, S. Ippolito, and O. Kavehei, "Convolutional Neural Networks for Seizure Prediction using Intracranial and Scalp Electroencephalogram," *Neural Networks*, vol. 105, p. 104–111, May 2018.

[22] Y. Qin, H. Zheng, W. Chen, Q. Qin, C. Han, and Y. Che, "Patient-Specific Seizure Prediction with Scalp EEG using Convolutional Neural Network and Extreme Learning Machine," in *2020 39th Chinese Control Conference (CCC)*, 2020, pp. 7622–7625.

[23] C. Feudjio, V. Djimna Noyum, Y. Peuriekeu Mofenjou, R. Rockefeller, and E. Fokoue, "A Novel Use of Discrete Wavelet Transform Features in the Prediction of Epileptic Seizures from EEG Data," vol. 01, 2021.

[24] D. Cho, B. Min, J. Kim, and B. Lee, "EEG-Based Prediction of Epileptic Seizures using Phase Synchronization Elicited from Noise-Assisted Multivariate Empirical Mode Decomposition," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 8, pp. 1309–1318, 2017.

[25] W. A. Chaovalitwongse, Y.-J. Fan, and R. C. Sachdeo, "On the Time Series K-Nearest Neighbor Classification of Abnormal Brain Activity," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 37, no. 6, pp. 1005–1016, 2007.

[26] L. Shunan, L. Donghui, D. Bin, W. Xile, W. Jiang, and W. L. Chan, "A Novel Feature Extraction Method for Epilepsy EEG Signals Based on Robust Generalized Synchrony Analysis," *25th Chinese Control and Decision Conference (CCDC)*, 2013.

[27] J. Lee, J. Park, S. Yang, H. Kim, Y. S. Choi, H. J. Kim, H. W. Lee, and B.-U. Lee, "Early Seizure Detection by Applying Frequency-Based Algorithm Derived from the Principal Component Analysis," *Frontiers in Neuroinformatics*, vol. 11, 2017.

[28] Pannag Desai K N and Ravi N, "Magnetic Resonance Imaging of the Brain in Adults Presenting With New Onset Seizures," *SSRG International Journal of Medical Science*, vol. 2, no. 3, pp. 23-33, 2015. Crossref, https://doi.org/10.14445/23939117/IJMS-V2I3P105

[29] A. Sharmila and P. Geethanjali, "DWT based Detection of Epileptic Seizure from EEG Signals Using Naive Bayes and K-Nn Classifiers," *IEEE Access*, vol. 4, pp. 7716–7727, 2016.

[30] C. James and D. Gupta, "Seizure Prediction for Epilepsy using a Multi-Stage Phase Synchrony Based System," *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2009.

[31] A. H. Shoeb, "*Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment*," Ph.D. Dissertation, 2009.

[32] R. S. Fisher and J. J. Engel, "Definition of the Postictal State: When Does it Start and End?" *Eplilepsy Behavior*, vol. 19. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1525505010004592

[33] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "Gan-Based Synthetic Medical Image Augmentation for Increased CNN Performance in Liver Lesion Classification," *Neurocomputing*, vol. 321, pp. 321–331, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231218310749

[34] A. S. Al-Fahoum and A. A. Al-Fraihat, "Methods of EEG Signal Features Extraction using Linear Analysis in Frequency and Time-Frequency Domains," *ISRN Neuroscience*, vol. 2014, pp. 1–7, 2014.