Synthetic Epileptic Brain Activities Using Generative Adversarial Networks

Damian Pascual¹, Amir Aminifar²
David Atienza², Philippe Ryvlin³, Roger Wattenhofer¹
¹Swiss Federal Institute of Technology Zurich (ETHZ), Switzerland.
²Swiss Federal Institute of Technology Lausanne (EPFL), Switzerland.
³Lausanne University Hospital (CHUV), Switzerland.
dpascual@ethz.ch

Abstract

Epilepsy is a chronic neurological disorder affecting more than 65 million people worldwide, defined by recurrent unprovoked seizures. The unpredictability of seizures not only degrades the quality of life of the patients, but it can also be life-threatening. Modern systems monitoring electroencephalography (EEG) signals are being currently developed with the view to detect epileptic seizures in order to alert caregivers and reduce the impact of seizures on patients’ quality of life. Such seizure detection systems employ state-of-the-art machine learning algorithms that require a considerably large amount of labeled personal data for training. However, acquiring EEG signals of epileptic seizures is a costly and time-consuming process for medical experts and patients, currently requiring in-hospital recordings in specialized units. In this work, we generate synthetic seizure-like brain electrical activities, i.e., EEG signals, that can be used to train seizure detection algorithms, alleviating the need for recorded data. First, we train a Generative Adversarial Network (GAN) with data from 30 epilepsy patients. Then, we generate synthetic personalized training sets for new, unseen patients, which overall yield higher detection performance than the real-data training sets. We demonstrate our results using the datasets from the EPILEPSIAE Project, one of the world’s largest public databases for seizure detection.

Introduction

Epilepsy is the fourth most common chronic neurological disorder worldwide (Hirtz et al. 2007), affecting over 65 million people. Epilepsy manifests itself by recurrent unprovoked seizures due to abnormal activity in the brain. The length of the seizures can range from few seconds to several minutes with a large variety of symptoms, including sensory auras, loss of awareness, behavioral arrest, automatic movements, and full body convulsions (Blumenfeld 2012). These symptoms not only degrade the quality of life of the patients, but they are also associated with a mortality rate 5 times higher among patients with recurrent seizures (Sperling et al. 1999) than in the corresponding group of the general population. One third of epilepsy patients suffer from drug-resistant uncontrolled seizures, which time of occurrence is usually unpredictable. A promising solution to reduce mortality and to improve the living standard and independence of epilepsy patients is continuous real-time monitoring using wearable technologies that collect and process EEG signals from the patient in real time and, upon occurrence of epileptic seizures, raise alerts to caregivers or family members (Goverdovsky et al. 2017; Sopic, Aminifar, and Atienza 2018; Debener et al. 2015). However, a fundamental barrier in developing reliable epileptic seizure detection systems is a lack of sufficient volume of training data. Indeed, modern detection systems are driven by machine-learning-based algorithms (Alotaiby et al. 2014; Acharya et al. 2018) that require a considerable amount of recorded seizures in order to reliably detect future seizures. Collecting and labeling EEG data from epilepsy patients is a costly process that currently requires in-hospital recording in specialized units. Such recordings are performed in clinical practice in a minority of patients and over short periods of time, typically a week, enabling to only record a few seizures per patient. This is a major limitation considering the privacy concerns that exist around sharing medical data, and the current trend towards personalized medicine. As a consequence, it is necessary to acquire significant amounts of new data for each patient.

The problem of scarce reliable training data is common in the field of artificial intelligence and it is particularly severe in the specific case of epilepsy monitoring. The most comprehensive solution to this problem consists of generating synthetic data that can be used to train the detection algorithms. However, generating high quality medical data is a challenging, and only recently substantial progress has been made, thanks to advances in deep generative models.

In this work, we present the use of a Generative Adversarial Network (GAN) (Goodfellow et al. 2014) to produce high quality synthetic epileptic seizure signals. To the best of our knowledge, this is the first time that seizure EEG samples are generated and used to train epilepsy detection algorithms achieving state-of-the-art results. The contributions of our work are:

1. A GAN model capable of generating realistic seizure (ictal) EEG signals from non-seizure (inter-ictal) signals.
2. An evaluation framework to assess the quality of synthetic

1The code of both, the model and the evaluation experiments is available at https://github.com/dapascual/GAN_epilepsy

22 Jul 2019
EEG ictal samples.

3. A study of the generated ictal samples that shows that our synthetic data can train seizure detection algorithms and achieve (and even improve) state-of-the-art results. Our results imply that it is possible to train epilepsy monitoring systems using exclusively synthetic seizures, circumventing in this way the obstacles related to seizure recording.

Related Work

A large amount of research effort has been devoted in the last years to the generation of reliable synthetic data for medical applications. Several studies have used synthetic data in areas such as medical imaging (Frid-Adar et al. 2018; Shin et al. 2018) [Costa et al. 2018] and Intensive Care Unit (ICU) monitoring (Esteban, Hyland, and Rätsch 2017; Che et al. 2017) [Lipton, Kale, and Wetzel 2016] to augment existing training sets in order to improve detection accuracy. Although this data augmentation approach has proved effective, previous attempts to train only with synthetic data have reported such a strong degradation in performance (Esteban, Hyland, and Rätsch 2017; Shin et al. 2018) that it has not been possible so far to dispense with real training data. Therefore, the scenario where no real training data can be accessed and only a purely synthetic training set may be available remains unsolved. This is, however, a common scenario in several medical applications, including epilepsy, given the difficulties and privacy concerns associated with collecting and sharing medical data (Price and Cohen 2019).

In recent years, GANs have attained outstanding results in a wide variety of challenging areas such as computer vision (Choi et al. 2018; Karras et al. 2018), audio (Yang, Chou, and Yang 2017; Hsu et al. 2017) or natural language processing (Yang et al. 2017; Yu et al. 2017). However, the application of GANs to the generation of brain signals has obtained very limited success so far: Hartmann et al. (Hartmann, Schirrmeister, and Ball 2018) generated EEG-like signals, without demonstrating the quality of the synthetic data in any specific task or pathology detection. In addition, Aznan et al. (Aznan et al. 2019) generated synthetic EEG data to augment existing real training sets for Brain-Computer Interfaces, but they only obtained a small increase in accuracy and without targeting any evaluation on completely synthetic training sets. Corley et al. (Corley and Huang 2018) used a GAN to upsample the spatial resolution of EEG signals and, despite the improvement in visual quality, the resulting training set produced a degradation of 4–9% of accuracy in a mental imagery classification task in comparison to the original training set.

Furthermore, the current literature has not addressed the problem of generating ictal samples, which are rare events in the EEG recordings and generating synthetic samples is particularly important for such rare events. In this work, we tackle this problem and evaluate the quality and utility of the generated ictal samples on the task of seizure detection, demonstrating state-of-the-art performance.

Generative model

Generative Adversarial Networks are a class of deep generative models in which two neural networks are trained simultaneously, while competing in a two-player minimax game. One network is a discriminator that estimates whether a sample is real or synthetic. The other network is a generator whose task is to generate realistic synthetic samples that maximize the probability of the discriminator making a mistake. During training, the discriminator improves its ability to recognize synthetic samples and as a consequence, the generator learns to produce increasingly realistic samples to deceive the discriminator. In this adversarial setting, optimality is reached when the generator produces realistic samples such that the discriminator cannot tell whether they are real or synthetic.

Our model is a conditional GAN (Mirza and Osindero 2014) that, given inter-ictal EEG samples at the input, generates EEG samples of epileptic seizures. The rationale behind our design is that, while epileptic seizures are very costly to record, inter-ictal signals can be easily recorded. As a result, we condition the network on inter-ictal samples from the target patient in order to provide additional information to the generator that can be exploited to produce more realistic seizure samples. In this way, we can use an already existing database to train our GAN and then use the GAN to generate seizure samples for a new patient.

The architecture of our GAN is modeled after the SEGAN from Pascual et al. (Pascual, Bonafonte, and Serra 2017). Our generator, which architecture is depicted in Figure 1, is a U-net (Ronneberger, Fischer, and Brox 2015) convolutional autoencoder network with weighted skip connections. Generally, an autoencoder consists of two symmetric parts, an encoder that processes the input sample and generates
a latent code, and a decoder that restores the original sample by decoding the latent code. However, in our case, the decoder does not restore the original inter-ictal sample but translates the latent code into an ictal sample. In order to introduce stochasticity into the model, Gaussian noise with mean 0 and standard deviation 1 is concatenated to the latent code. The skip connections multiply the feature maps at each layer of the encoder with a weight which is learnt during training, and then, the result of that operation is added to the corresponding feature map of the decoder. In this way, the weights of the skip connections regulate the amount of information that is fed from the encoder into the decoder. The discriminator of our GAN has the same structure as the encoder of the generator, but it includes an additional fully connected layer at the output. In this way, the discriminator outputs a single value between 1 and 0, where 1 represents the real class and 0 the synthetic class.

To obtain the network parameters of the model, an optimization problem is solved iteratively during the training stage in which the loss functions of the generator and the discriminator are alternatively minimized. In our model, these losses are based on the Least Squares GAN (LSGAN) (Mao et al. 2017). Consequently, the minimization objective of the discriminator is given by:

\[
\min_{\theta_D} \mathcal{L}_D(\theta_D) = \mathbb{E}_{x \sim p_g}[(D(x; \theta_D) - 1)^2] + \mathbb{E}_{x \sim p}(D(G(x); \theta_D))^2, \tag{1}
\]

where the function \( D \) corresponds to the discriminator and \( G \) to the generator and, as mentioned before, the output of the discriminator lies between 0 and 1. The input data \( x \) is sampled from the input data distribution \( p \), and \( \theta_D \) are the network parameters of the discriminator. The first term of the loss function pushes the discriminator to output 1 when the input is a real sample \( x \), whereas the second term is minimized when the discriminator outputs a 0 given a synthetic sample.

The generator’s loss includes a weighted \( L_1 \) regularization term that ensures that the generated signal is similar to the reference output signal \( y \). This constraint enforces the output to stay similar to EEG signals and makes the training more stable. The minimization objective of the generator, with network parameters \( \theta_G \), is thus:

\[
\min_{\theta_G} \mathcal{L}_G(\theta_G) = \mathbb{E}_{x \sim p_g}[(D(G(x; \theta_G) - 1)^2] + \lambda ||G(x; \theta_G) - y||_1, \tag{2}
\]

where \( \lambda \) is a hyperparameter that we fix to 100 in order to make both terms of the loss function of comparable magnitude, thus without letting the regularization term dominate the optimization problem. The first term of the loss encourages the generator to produce synthetic samples that are classified as 1, i.e., real, by the discriminator, which is adversarial with respect to the discriminator’s loss function. Hence, the competing interests of the generator and the discriminator during training drives the generator to produce more and more realistic samples.

**GAN Architectural Details**

The input to the generator are samples of length 2048 points (4 seconds of signal from 2 electrodes recorded at a frequency of 256 Hz). The encoder consists of eight blocks that alternate a convolutional layer with a max pooling layer with 2x2 filters and stride 2. The feature maps extracted at each block of the encoder yield the following shapes: 2048x1, 1024x64, 512x64, 256x128, 128x128, 64x256, 32x256, 16x512, 8x1024; where 2048x1 is the shape of the input and 8x1024 that of the latent encoding. Gaussian noise with mean 0 and standard deviation 1 and shape 8x1024 is concatenated to the latent code. The decoder is symmetric to the encoder, but it uses deconvolutions and dilations. Thus, the shapes of the feature maps at the decoder are 16x1024, 16x512, 32x256, 64x256, 128x128, 256x128, 512x64, 1024x64, 2048x1; where 2048x1 is the final output of the generator.

The activation used is the leaky ReLU function (Maas, Hannun, and Ng 2013), except for the last block of the decoder, where we use the hyperbolic tangent function. All convolutions and deconvolutions are unbiased and spectral normalization (Miyato et al. 2018) is applied before each block in both, the generator and the discriminator. On top of that, in the discriminator we apply virtual batch normalization (Salimans et al. 2016).

To train the model we use the Adam (Kingma and Ba 2014) optimizer with 0 and 0.9 for the values of \( \beta_1 \) and \( \beta_2 \), respectively, and learning rates 0.0001 for the generator and 0.0004 for the discriminator. The size of the minibatches of data employed during training is 100 samples. All the hyperparameters employed are summarized in Table 1.

| Parameter | Value |
|-----------|-------|
| \( \beta_1 \) | 0 |
| \( \beta_2 \) | 0.9 |
| G learning rate | 0.0001 |
| D learning rate | 0.0004 |
| \( \lambda \) | 100 |
| Mini-batch size | 100 |

Table 1: Model hyperparameters.

**Training**

To train our model, we used data from the EPILEPSIAE project database (Hille et al. 2012), which is one of the world’s largest public databases for seizure detection. The dataset contains recordings from 30 different epilepsy patients with a total of 277 epileptic seizures that sum to a duration of 21,001 seconds altogether. The EEG data is collected at a sample frequency of 256 Hz and it is divided into recordings of one hour, each one corresponding to one recording session. The number of one-hour recordings varies for each patient in a range between 96 and 281 sessions.
In this work, we target the setup of real-world and stigma-free wearable monitoring devices (Hoppe et al. 2015) and thus we consider only the electrodes F7T3 and F8T4 in the standard 10–20 system (Klem et al. 1999), which can be easily hidden in glasses. We extract samples of four seconds of duration, since this length is effective to detect epileptic seizures. Given that the data was recorded at a frequency of 256 Hz, this results in samples of length 2048, i.e., 1024 per electrode. The four-second long ictal samples are collected with three seconds of overlap in order to augment the amount of training ictal samples, while the inter-ictal samples are collected with no overlap from recordings where no seizure occurred. Moreover, we do not apply any filtering or preprocessing step to the data and we work directly with the raw EEG signals. To construct the training set, we pair each ictal sample to an inter-ictal sample from the same patient. This means that each inter-ictal sample given at the input of the generator is associated to an ictal sample that is used as a reference to guide the training. In this manner, the generator learns to map inter-ictal samples to ictal samples for any given patient.

In order to train the GAN, the leave-one-out strategy is followed: for each target patient, the GAN is trained using the ictal and inter-ictal data coming from all other patients. The exact number of training samples depends on the number of seconds of seizure recording available in the database for all patients except for the left-out patient and, although it varies slightly, it is approximately 20,000 samples. Following this scheme, the GAN is trained independently for each patient and thus, we obtain one model per patient.

Evaluation

For each trained model, we generate between 2,000 and 6,000 ictal samples from inter-ictal EEG signals from the patient that was left out during training. The exact number of generated samples depends on the amount of inter-ictal data each patient has available. In Figure 2, some of the generated samples are shown in the time domain. The presence of the well-known delta–theta rhythm, i.e., rhythmic slow activity with a frequency of oscillation in 0.5–4 or 4–7 hertz, is a clear indication of the correct generation of the ictal discharge and epileptic seizure segment in the synthetically-generated EEG signals (Osorio et al. 2016).

However, in order to systematically evaluate the quality of the generated ictal samples beyond their visual appearance, we use them to train a state-of-the-art classifier based on the random forest algorithm (Díaz-Uriarte and De Andres 2006). The task of the classifier is to determine whether an incoming four-second sample is an ictal or an inter-ictal sample. Here, we follow the experiments performed in Sopic et al. (Sopic, Aminifar, and Atienza 2018), which are tailored to a stigma-free wearable device for epilepsy monitoring. Therefore, using that same classifier ensures that our synthetic samples can be used not only in a medical environment, but also in the more restrictive setting of continuous monitoring using wearable technologies.

After generating the synthetic samples with the GAN, we have three sets of data for each patient: real inter-ictal samples, real ictal samples and synthetic ictal samples. Then, for each patient, we first construct the test set. The test set consists of all the ictal samples of the target patient without overlap and twice as many inter-ictal samples; we build an unbalanced test set with twice as many inter-ictal as ictal samples in order to better reproduce the real-world setting where the ictal samples are largely under-represented in the inference phase. The size of the test set changes slightly between patients, given that a different number of seizures are recorded for each patient.

In our evaluation, we target each patient independently assuming the scenario where the only data available for train-
ing are real inter-ictal samples and synthetic ictal samples from the target patient. Therefore, for each patient, we train the classifier with a balanced training set consisting of 2,000 synthetic seizures taken from all the synthetic samples generated for the target patient, which ranges between 2,000 and 6,000 samples, and 2,000 real inter-ictal samples. As a baseline for comparison, we consider the case where real ictal samples from all other patients and inter-ictal samples from the target patient are available. Thus, the baseline training set consists of 2,000 samples of real seizures randomly selected from all the patients in the database except for the target patient, as well as the same 2,000 inter-ictal samples used in the synthetic training set. In this way, in this evaluation framework, the synthetic ictal samples are strictly the only aspect that differs between the evaluation synthetic training set and the baseline.

Once the data is split in training and test sets, a feature extraction step is performed on the data. In the feature extraction stage we follow Sopic et al. (Sopic, Aminifar, and Atienza 2018) and extract 54 features of power and non-linearity per electrode, and, since we consider two electrodes, the total number of features is 108. These features are subsequently extracted for all the samples in both the training and the test sets. To calculate the non-linear features, the signal is decomposed using the discrete wavelet transform down to level seven. The nonlinear features extracted are: sixth and seventh level sample entropy (Richman and Moorman 2000) for \( k = 0.2 \) and \( k = 0.35 \); third, fourth, fifth, sixth and seventh level permutation entropy (Bandt and Pompe 2002) for \( n = 3, n = 5 \) and \( n = 7 \); third, fourth, fifth, sixth and seventh level, as well as raw signal, Shannon, Renyi and Tsallis entropies. The power features are: total power, total and relative band power in the bands delta [0.5,4] Hz, theta [4,8] Hz, alpha [8,12] Hz, beta [13,30] Hz, gamma [30,45] Hz as well as in the bands [0,0.1] Hz, [0.1,0.5] Hz, [12,13] Hz.

After the features are extracted, the baseline training set is used to train the random forest classifier with 500 trees and the resulting classifier is evaluated against the test set. The same procedure is then repeated for the synthetic training set. We repeat these experiments 15 times, shuffling the data each time, in order to make our results robust against different splits of data, as well as against different configurations of the random forest.

## Results

The detailed results of our experiments for each patient are reported in Table 2. The reported values are the geometric mean of sensitivity and specificity (Fleming and Wallace 1986). We observe that Patient 22 performs extremely poorly for both the baseline and the synthetic training sets and, therefore, it is not a relevant indicator of the classification quality. Consequently, it has been removed from the calculation of the total difference in performance. This total difference is calculated as the difference between the geometric mean of all patients in the synthetic case and the geometric mean of all patients in the baseline case. These results pass the Wilcoxon statistical significance test with a \( p \)-value of 0.0098 when patient 22 is already excluded, which indicates that the difference between the results obtained for the baseline and synthetic training sets is statistically significant.

Our experiments show that training with synthetic samples not only does it not degrade the performance, but yields a 1.2% improvement overall compared to training only with real samples from a generic database. On top of that, as detailed in Figure 3, 20 out of 29 patients, i.e., 69%, improve by more than 1%, while only for four out of 29 patients the performance decreases by more than 1%. An explanation for the performance improvement when using synthetic data is that, since our GAN generates ictal samples given inter-ictal samples from the same patient, the model generates synthetic seizures that retain a number of personal features.

Regarding the patients for whom the performance degrades most significantly, i.e., Patients 2 and 21, their seizures are dominated by repetitive spiking. This pattern is relatively rare and is not well represented in the dataset (only 10.5% of the seizures). Therefore, our GAN model does not

### Table 2: Geometric mean of sensitivity and specificity per patient of our evaluation.

| Patient ID | Baseline (%) | Synthetic (%) | Difference (%) |
|------------|--------------|---------------|----------------|
| 1          | 73.49        | 80.06         | +6.57          |
| 2          | 79.36        | 70.30         | -9.06          |
| 3          | 77.59        | 82.84         | +5.25          |
| 4          | 76.34        | 78.33         | +1.99          |
| 5          | 64.86        | 68.19         | +3.33          |
| 6          | 74.10        | 74.74         | +0.64          |
| 7          | 68.11        | 68.59         | +0.48          |
| 8          | 81.41        | 86.14         | +4.73          |
| 9          | 76.74        | 80.67         | +3.93          |
| 10         | 66.84        | 65.87         | -0.97          |
| 11         | 81.03        | 83.66         | +2.63          |
| 12         | 63.00        | 66.26         | +3.26          |
| 13         | 77.20        | 78.54         | +1.34          |
| 14         | 74.32        | 76.51         | +2.19          |
| 15         | 74.25        | 74.07         | -0.18          |
| 16         | 78.11        | 80.64         | +2.53          |
| 17         | 65.27        | 67.84         | +2.57          |
| 18         | 66.20        | 71.62         | +5.42          |
| 19         | 76.95        | 78.13         | +1.18          |
| 20         | 73.42        | 68.67         | -4.75          |
| 21         | 79.18        | 71.61         | -7.57          |
| 22         | 26.88        | 12.62         | -14.26         |
| 23         | 77.05        | 78.62         | +1.57          |
| 24         | 78.28        | 77.87         | -0.41          |
| 25         | 77.02        | 75.65         | -1.37          |
| 26         | 74.36        | 76.15         | +1.79          |
| 27         | 76.00        | 78.00         | +2.0           |
| 28         | 81.97        | 83.07         | +1.10          |
| 29         | 75.73        | 78.41         | +2.68          |
| 30         | 79.80        | 82.29         | +2.49          |
| TOTAL      | 74.57        | 75.78         | +1.21          |
Figure 3: Performance difference in the classification task between synthetic and real training sets. The vertical axis represents the number of patients and the horizontal one the total difference, where larger is better. 69% of the patients improve by more than 1% and among those, seven improve by more than 3% and up to 6%, while only for four out of 29 patients the performance decreases by more than 1%.

capture this behavior with as much precision as it does capture other patterns such as theta or delta rhythms. In fact, Patient 22 also suffers from seizures with repetitive spiking and our experiments show that even the state-of-the-art techniques fail in detection of such seizures. Finally, Patient 20 has only 4 seizures in this dataset, which is the lowest number of seizures in the entire dataset and hinders robust evaluation of our model.

Discussion

In this work, we have presented a GAN model that generates synthetic EEG signals of epileptic seizures. To the best of our knowledge, for the first time in the medical domain, we have generated synthetic data sets that can train detection algorithms and achieve (and even improve) state-of-the-art results based on real data, which demonstrates the quality of our synthetic data. Our results underline that, in the most common scenario in which no recordings of epileptic seizures of a new patient are available, training using exclusively synthetic seizures achieves state-of-the-art performance. Hence, using an existing database, deep generative models can generate data to train a system to monitor any given new unseen patient. This solution circumvents both the costs related to seizure recording and labeling, as well as the privacy concerns derived from sharing personal and sensitive data.

Our work emphasizes that the application in medicine of deep generative models, such as GANs, can potentially solve many of the open challenges in the field and help bridging the gap on the adoption of continuous monitoring systems for patients suffering from chronic disorders. Further research into unpaired and conditional deep generative models may improve the quality and, therefore, the performance of synthetic training sets allowing for personalized medicine based on synthetic data.

Acknowledgements

This work has been partially supported by the Human Brain Project (HBP) SGA2 (GA No. 785907), the ML-edge Swiss National Science Foundation Research project (GA No. 200020_182009/1), and the MyPreHealth research project (Hasler Foundation project No. 16073).

References

[Acharya et al. 2018] Acharya, U. R.; Oh, S. L.; Hagiwara, Y.; Tan, J. H.; and Adeli, H. 2018. Deep convolutional neural network for the automated detection and diagnosis of seizure using eeg signals. Computers in biology and medicine 100:270–278.

[Alotaiby et al. 2014] Alotaiby, T. N.; Alshebeili, S. A.; Alshawi, T.; Ahmad, I.; and El-Samie, F. E. A. 2014. Eeg seizure detection and prediction algorithms: a survey. EURASIP Journal on Advances in Signal Processing 2014(1):183.

[Aznan et al. 2019] Aznan, N. K. N.; Atapour-Abarghouei, A.; Bonner, S.; Connolly, J.; Moubayed, N. A.; and Breckon, T. 2019. Simulating brain signals: Creating synthetic eeg data via neural-based generative models for improved ssvep classification. arXiv preprint arXiv:1901.07429.

[Blundell and Pompe 2002] Bandt, C., and Pompe, B. 2002. Permutation entropy: a natural complexity measure for time series. Physical review letters 88(17):174102.

[Blumenfeld 2012] Blumenfeld, H. 2012. Impaired consciousness in epilepsy. The Lancet Neurology 11(9):814–826.
[Che et al. 2017] Che, Z.; Cheng, Y.; Zhai, S.; Sun, Z.; and Liu, Y. 2017. Boosting deep learning risk prediction with generative adversarial networks for electronic health records. In 2017 IEEE International Conference on Data Mining (ICDM), 787–792. IEEE.

[Choi et al. 2018] Choi, Y.; Choi, M.; Kim, M.; Ha, J.-W.; Kim, S.; and Choo, J. 2018. Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[Corley and Huang 2018] Corley, I. A., and Huang, Y. 2018. Deep eeg super-resolution: Upsampling eeg spatial resolution with generative adversarial networks. In 2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), 100–103. IEEE.

[Costa et al. 2018] Costa, P.; Galdran, A.; Meyer, M. I.; Hartmann, K. G.; Schirrmeister, R. T.; and Ball, T. 2018. Eeg-gan: Generative adversarial networks for electroencephalographic (eeg) brain signals. arXiv preprint arXiv:1806.01875.

[Hirtz et al. 2007] Hirtz, D.; Thurman, D.; Gwinn-Hardy, K.; Mohamed, M.; Chaudhuri, A.; and Zalutsky, R. 2007. How common are the common neurologic disorders? Neurology 68(5):326–337.

[Hoppe et al. 2015] Hoppe, C.; Feldmann, M.; Blachut, B.; Surges, R.; Elger, C. E.; and Helmstaedter, C. 2015. Novel techniques for automated seizure registration: patients’ wants and needs. Epilepsy & Behavior 52:1–7.

[Hsu et al. 2017] Hsu, C.-C.; Wang, H.-T.; Wu, Y.-C.; Tsao, Y.; and Wang, H.-M. 2017. Voice conversion from unaligned corpora using variational autoencoding wasserstein generative adversarial networks. In INTERSPEECH.

[Ihle et al. 2012] Ihle, M.; Feldwisch-Drentrup, H.; Teixeira, C. A.; Witon, A.; Schelter, B.; Timmer, J.; and Schulze-Bonhage, A. 2012. Epilepsia—a european epilepsy database. Computer methods and programs in biomedicine 106(3):127–138.

[Karras et al. 2018] Karras, T.; Aila, T.; Laine, S.; and Lehtinen, J. 2018. Progressive growing of GANs for improved quality, stability, and variation. In International Conference on Learning Representations.

[Kingma and Ba 2014] Kingma, D. P., and Ba, J. 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

[Klem et al. 1999] Klem, G. H.; Lüders, H. O.; Jasper, H.; Elger, C.; et al. 1999. The ten-twenty electrode system of the international federation. Electroencephalogr Clin Neurophysiol 52(3):3–6.

[Lipton, Kale, and Wetzel 2016] Lipton, Z. C.; Kale, D. C.; and Wetzel, R. 2016. Modeling missing data in clinical time series with mns. Machine Learning for Healthcare.

[Maas, Hannun, and Ng 2013] Maas, A. L.; Hannun, A. Y.; and Ng, A. Y. 2013. Rectifier nonlinearities improve neural network acoustic models. In Proc. icml, volume 30, 3.

[Mao et al. 2017] Mao, X.; Li, Q.; Xie, H.; Lau, R. Y.; Wang, Z.; and Paul Smolley, S. 2017. Least squares generative adversarial networks. In The IEEE International Conference on Computer Vision (ICCV).

[Mirza and Osindero 2014] Mirza, M., and Osindero, S. 2014. Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784.

[Miyato et al. 2018] Miyato, T.; Kataoka, T.; Koyama, M.; and Yoshida, Y. 2018. Spectral normalization for generative adversarial networks. In 6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings.

[Osorio et al. 2016] Osorio, I.; Zaveri, H.; Frei, M.; and Arthur, S. 2016. Epilepsy: The Intersection of Neurosciences, Biology, Mathematics, Engineering, and Physics. CRC Press.

[Pascual, Bonafonte, and Serrà 2017] Pascual, S.; Bonafonte, A.; and Serrà, J. 2017. Segan: Speech enhancement generative adversarial network. In INTERSPEECH.
[Price and Cohen 2019] Price, W. N., and Cohen, I. G. 2019. Privacy in the age of medical big data. *Nature medicine* 25(1):37.

[Richman and Moorman 2000] Richman, J. S., and Moorman, J. R. 2000. Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology* 278(6):H2039–H2049.

[Ronneberger, Fischer, and Brox 2015] Ronneberger, O.; Fischer, P.; and Brox, T. 2015. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, 234–241. Springer.

[Salimans et al. 2016] Salimans, T.; Goodfellow, I.; Zaremba, W.; Cheung, V.; Radford, A.; and Chen, X. 2016. Improved techniques for training gans. In *Advances in neural information processing systems*, 2234–2242.

[Shin et al. 2018] Shin, H.-C.; Tenenholtz, N. A.; Rogers, J. K.; Schwarz, C. G.; Senjem, M. L.; Gunter, J. L.; Andriole, K. P.; and Michalski, M. 2018. Medical image synthesis for data augmentation and anonymization using generative adversarial networks. In *International Workshop on Simulation and Synthesis in Medical Imaging*, 1–11. Springer.

[Sopic, Aminifar, and Atienza 2018] Sopic, D.; Aminifar, A.; and Atienza, D. 2018. e-Glass: A wearable system for real-time detection of epileptic seizures. In *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, 1–5. IEEE.

[Sperling et al. 1999] Sperling, M. R.; Feldman, H.; Kinman, J.; Liporace, J. D.; and O’Connor, M. J. 1999. Seizure control and mortality in epilepsy. *Annals of neurology* 46(1):45–50.

[Yang et al. 2017] Yang, Z.; Chen, W.; Wang, F.; and Xu, B. 2017. Improving neural machine translation with conditional sequence generative adversarial nets. *arXiv preprint arXiv:1703.04887*.

[Yang, Chou, and Yang 2017] Yang, L.-C.; Chou, S.-Y.; and Yang, Y.-H. 2017. Midinet: A convolutional generative adversarial network for symbolic-domain music generation. In *ISMIR*.

[Yu et al. 2017] Yu, L.; Zhang, W.; Wang, J.; and Yu, Y. 2017. Seggan: Sequence generative adversarial nets with policy gradient. In *Thirty-First AAAI Conference on Artificial Intelligence*. 