ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 978-1-7281-7605-5/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICASSP39728.2021.9414078

ENHANCING MULTI-CHANNEL EEG CLASSIFICATION WITH GRAMIAN TEMPORAL GENERATIVE ADVERSARIAL NETWORKS

Chi Nok Enoch Kan, Richard J Povinelli, Dong Hye Ye

Department of Electrical and Computer Engineering, Marquette University

ABSTRACT

Deep learning's requirements for large amounts of training data remains a challenge for researchers and developers. Generative Adversarial Network (GAN) is commonly used in medical image analysis to generate novel training images to help resolve this issue. While deep learning has many clinical applications in radiology, its applications in medical time series data such as electroencephalogram (EEG) are usually constrained to 1 dimension. Hence, there are few available GAN architectures that effectively synthesize single and multi-channel EEG. In this paper, we propose a novel method to synthesize multi-channel EEG in the form of Gramian Angular Field (GAF) images with a Gramian Temporal Generative Adversarial Network (GT-GAN). The proposed network is capable of generating realistic GAF images and enhances EEG anomaly detection accuracy in residual learning frameworks.

Index Terms— EEG, Generative Models, GAN, GAF

1. INTRODUCTION

In recent years, deep learning has been widely applied in the biomedical field to perform tasks such as medical image segmentation [1] and drug binding prediction [2]. Particularly in the neuroimaging domain, research efforts focus on applying deep learning to perform clinical tasks such as image-based stroke detection [3], Magenetic Resonance-Computed Tomography (MR-CT) modality transfer [4] and detection of neurodegenerative diseases [5].

Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks are commonly used to perform classification tasks involving medical time series such as electroencephalography (EEG) using the formation of a directed graph along temporal sequences in their node connections. Bresch et. al. [6] proposed a RNN architecture for sleep stage classification in single channel EEG. Zhang et. al. [7] proposed the use of a Long Short-Term Memory (LSTM) network to classify limb movements from EEG data. A novel approach proposed by Bashivan et. al. [8] involves preprocessing EEG data into 2D topological maps, thereby transforming time series classification tasks into conventional image classification tasks using a deep convolutional neural network (dCNN). In general, machine learning algorithms are able to perform well provided large of amounts of training data are available. This poses a significant challenge for the aforementioned algorithms to perform well in the medical domain since certain data, such as pediatric data, is hard to obtain due to risks associated with the collection procedures.

One effective way to generate more realistic training samples is the use of Generative Adversarial Networks. Generative Adversarial Networks (GANs) [9] have gained significant attention due to their ability to synthesize realistic images from noise. The original GAN architecture consists of two competing convolutional neural networks. The generator network is responsible for synthesising images from noise, whereas the discriminator network is responsible for discriminating real images from fake ones. GANs are extremely hard to train due to underlying issues such as mode collapse and non-convergence. New GAN architectures such as conditional GAN (cGAN) [10] and auxiliary classifier GAN (ACGAN) [11] are proposed to tackle these instability issues. Kan et. al. [12] modify an auxiliary classifier GAN (ACGAN) to conditionally generate pediatric CT image patches along with their pancreatic segmentation masks. As for generation of artificial medical time series, Hartmann et. al. [13] propose an EEG-GAN framework, which is a time series generation framework modified from gradient penalty Wasserstein GAN (WGAN-GP) for EEG generation. The proposed EEG-GAN framework, however, suffers from imablance between its discriminator and its generator. Moreover, EEG-GAN does not take into account of the temporal relationships in EEGs. Alternatively, a new loss function combining Wasserstein and temporal-spatial-frequency losses is proposed in Luo et. al.'s work [14] for EEG signal reconstruction. Nevertheless, none of the current EEG synthesis methods uses image representations of EEG that effectively encodes the temporal dependecies within EEG recordings. This is why our work is important in laying the foundation for image-based synthesis of medical time series data.

2. METHODOLOGY

Our main goal is to encode EEG as images with as few transformation and feature selection steps as possible. For each EEG recording, 3 channels are chosen and transformed into a

3-channel Gramian Angular Field (GAF) image. We also propose and apply a novel GAN framework to synthesize artificial GAF images. Finally, we train a residual learning framework to detect anomalous EEGs based on their GAF images before and after the addition of synthesized GAF images.

2.1. Gramian Angular Field Transformation

In order to represent EEG data as 2D images, a robust transformation method is required to preserve both the spatial and temporal features of the original data. Wang and Oates [15] first proposed the use of Gramian Angular Summation/Difference Fields (GASF/ GADF) and Markov Transition Fields (MTF) to encode time series as images. Recall that Gram matrices are used in linear algebra to compute linear dependence of a given set of vectors $x_1 ... x_n$:

$$G(x_1 ... x_n) = \begin{pmatrix} < x_1, x_1 > & < x_1, x_2 > & ... & < x_1, x_n > \\ < x_2, x_1 > & < x_2, x_2 > & ... & < x_2, x_n > \\ \vdots & \vdots & \ddots & \vdots \\ < x_1, x_1 > & < x_1, x_2 > & ... & < x_1, x_n > \end{pmatrix}$$

where $< x_m, x_n >$ is defined as the inner product $\sum_{i=1}^{n} x_m \cdot x_n$ of two vectors x_m and x_n. Assuming both vectors follow L1-norm, we can simply express the inner product as follows:

$$< x_m, x_n >= ||x_m||_1 \cdot ||x_n||_1 \cdot \cos(\theta) = \cos(\theta)$$

where θ is the angular difference between the two vectors. Additionally, the vectors can be transformed into polar coordinates by taking the $\arccos$ of a given data point x_i in these vectors. As a result, a Gramian Angular Field (GAF) matrix can be constructed:

$$G(x_1 ... x_n) = \begin{pmatrix} \cos(\phi_1 + \phi_1) & \cos(\phi_1 + \phi_2) & ... & \cos(\phi_1 + \phi_n) \\ \cos(\phi_2 + \phi_1) & \cos(\phi_2 + \phi_2) & ... & \cos(\phi_2 + \phi_n) \\ \vdots & \vdots & \ddots & \vdots \\ \cos(\phi_n + \phi_1) & \cos(\phi_n + \phi_2) & ... & \cos(\phi_n + \phi_n) \end{pmatrix}$$

where ϕ_i is the polar coordinate angle of a given data point x_i. A major benefit of using Gramian matrices along with polar coordinates is that both the temporal and spatial dimensions are preserved in the transformation. A minor drawback of GAF transformation is the increasing size of the matrix given a vector of time series. Since the goal of the transformation is to create pairwise matrices, a GAF matrix of dimension d^2 will be created given a d-dimension time series.

2.2. Generative Adversarial Networks

Generative Adversarial Networks (GANs) are composed of two neural networks, namely a generator network G and a discriminator network D that compete against each other. The Generator attempts to transform random noise vector $\mathbf{z}$ into a generated image $G(\mathbf{z})$. On the other hand, the Discriminator attempts to maximize $log(D(\mathbf{z})) + log(1 - D(G(\mathbf{z}))$, the probability of assigning correct labels to both training images and images generated by G . The G network is trained to minimize the log of the inverted probability of D prediction of fake images $log(1 - D(G(\mathbf{z}))$. We seek to maximize the $D(G(\mathbf{z}))$ instead since minimization of the inverted probability is much harder in practice. In summary, GAN's objective function can be formulated as a minimax loss:

$$\min_{\mathbf{G}} \max_{\mathbf{D}} V(D, G) = \mathbb{E}_{\mathbf{x} \sim p_{data}(\mathbf{x})}[logD(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_{\mathbf{z}}(\mathbf{z})}[log(1 - D(G(\mathbf{z})))] \quad (1)$$

where $\mathbf{x} \sim p_{data}$ and $\mathbf{z} \sim p_{\mathbf{z}}$ represent the data distributions of the real and fake images respectively. This original minmax objective function aims to minimize the Jensen-Shannon divergence between the generated and real image data distributions.

2.3. Gramian Temporal GANs

We propose a novel GAN architecture which we call the Gramian Temporal Generative Adversarial Network (GT-GAN). Auxiliary Classifier Generative Adversarial Network (ACGAN) is used as our baseline architecture in which we task the discriminator to predict the diagnosis of a given GAF matrix (normal vs. abnormal). The last layer of GT-GAN's generator also enforces symmetry by taking the maximum values $max(M, M^T)$ between an input feature map M and its transposed matrix M^T. Given the success of self-normalizing activation functions in stabilizing training of deep networks, we replace piecewise linear activation functions (ReLU and leaky ReLU) in the generator with scaled exponential linear units (SELU).

To effectively synthesize both normal and abnormal EEG GAFs, our proposed GT-GAN uses a modified version of AC-GAN's objective function. In addition to classifying whether a given input image is real or fake, ACGAN's discriminator also predicts the class labels of the images. Mathematically speaking, the discriminator produces a probability distribution $P(C|X) = D(X)$ over the class labels of the images. The objective function of ACGAN can be defined as the log-likelihood of the correct source, L_S and the log-likelihood of the correct class, L_C, where X_r and X_f represent real and fake images respectively:

$$L_S = \mathbb{E}_{x_r, x_f}[\log P(S = r|X_r)] + \mathbb{E}_{x_r, x_f}[\log P(S = f|X_f)] \quad (2)$$

$$L_C = \mathbb{E}_{x_r, x_f}[\log P(C = c|X_r)] + \mathbb{E}_{x_r, x_f}[\log P(C = c|X_f)] \quad (3)$$

Given the success of relativistic GAN [16] in synthesizing high definition images with a slightly varied GAN objective function, we change L_s in our objective function to the log

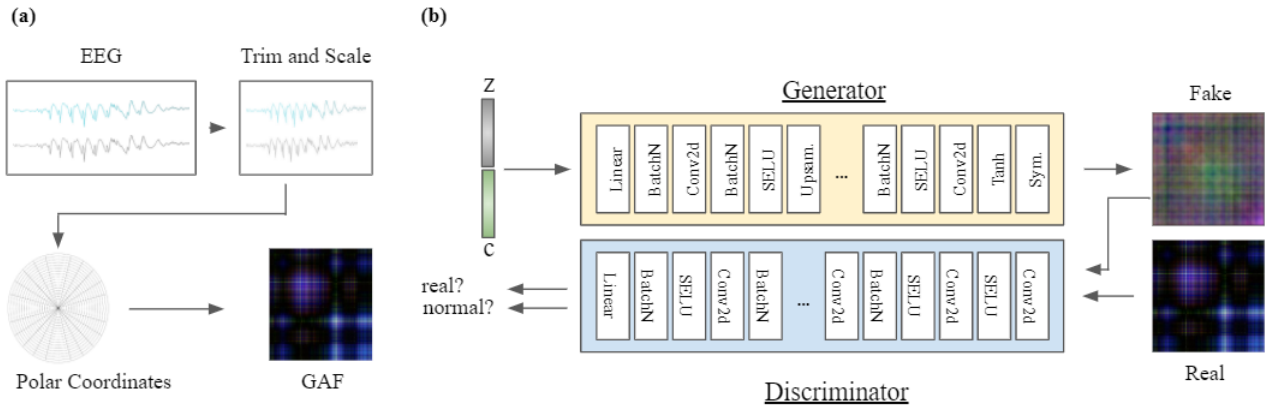

Fig. 1. **(a) Preprocessing Pipeline** An EEG series is first scaled to the range [-1, 1] using a min-max scaler, and its Gramian Angular Field (GAF) matrix is constructed by calculating pairwise polar coordinates of the scaled series. **(b) GT-GAN Model Architecture** GT-GAN generator contains multiple convolution layers and a symmetry layer to ensure symmetry in its generated images. Both the generator and the discriminator uses SELU, a self-normalizing activation function to ensure proper convergence.

probability of correct real image classification relative to incorrect fake image classification:

$$L'_S = \mathbb{E}_{x_r,x_f}\{\log P[(S=r|X_r)-(S=r|X_f)]\} \quad (4)$$

where GT-GAN's discriminator aims to maximize the sum of relativistic adversarial loss L'_S and classification loss L_C.

3. EXPERIMENTS

3.1. Temple University Hospital Abnormal EEG Corpus

In order to test the efficacy of our proposed GT-GAN, we design an experiment to synthesize new GAF matrices given existing EEG data. The Temple University Hospital Abnormal EEG Corpus (TUAB) dataset [17], which contains 2785 patient EEGs in its training set and 280 patient EEGs in its evaluation set. EEG recordings vary in length, and they are classified by real clinicians as either 'normal' or 'abnormal' based on visual inspection. Most EEG records in this dataset contain patients of age $\geq$ 20 and $\leq$ 90, where the distribution of age roughly follows a normal distribution. The gender proportion in the dataset is also kept balanced.

The electrode placement on the head scalp of each patient follows the Standard 10-20 system [18], meaning that each adjacent electrode represents distances of either 10% or 20% of the total nasion-inion or right-left distance. All the EEG files in the TUAB dataset follow an unipolar AR montage, meaning that averages of certain numbers of electrodes are used as references.

3.2. Data Preprocessing

Both the training and evaluation datasets are transformed from multi-channel EEG to GAF matrices. We select three channels that are reported by previous literature to provide the most information regarding whether a given EEG is abnormal or not. These channels are P3-O1 (parietal), T5-O1 (temporal) and F7-T3 (frontal). For each channel, we select only the first 10 seconds of the EEG recording and preprocess the signal into a single 2500×2500 GAF matrix. Assuming a given EEG data of length t seconds with a sampling rate of r ms, the EEG time series will have a total of $d = \frac{1000t}{r}$ data points. The resulting GAF matrix will be of size $d \times d$. In order to run the GT-GAN efficiently given limited computing power, all the EEG recordings are downsampled by a factor of 10 to create 250×250 matrices. In real, clinical settings, clinicians are capable of distinguishing whether a given EEG is abnormal or not just by inspecting the first few seconds of the recording. Therefore, we hypothesize that using the first 10 seconds of the EEG recording is adequate for anomaly detection. In order to achieve a higher baseline classification accuracy, we use all normal and abnormal images from the training dataset to train our proposed GT-GAN. The rest of the images belong to the evaluation dataset and are used for our classification experiments with residual networks.

3.3. Evaluation

For each epoch, the generator is updated twice before the discriminator is updated. This training strategy is commonly used to stabilize the training process by maintaining balance between the generators and the discriminators. Both GT-

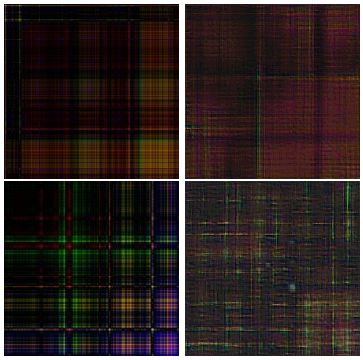

Fig. 2. Results of GT-GAN synthesis of GAF matrices. (Top left) GAF image of a ground truth normal EEG (top right) of a synthesized normal EEG (bottom left) of a ground truth abnormal EEG (bottom right) of a synthesized abnormal EEG

GAN's generator and disriminator converge around the 100th epoch for both normal and abnormal images using a batch size of 16 and an initial learning rate of 0.0002. Sample images are created by inputting random noise vectors and class labels into the trained GT-GAN model. As shown in **Figure 2**, the resulting images resemble the original training images. We identify the nearest neighbors of the synthesized images in the training dataset measured by L1 distance.

One major challenge in GAN training is the lack of appropriate evaluation metric. Several evaluation metrics, such as the inception score [19] and multi-scale structural similarity (MS-SSIM) index [20] have been proposed to quantify the quality of images generated from GANs. However, it is hard to quantify the amount of classification accuracy gained from enriching training datasets with GT-GAN synthesized images based on these metrics. Moreover, baseline values are required for these metrics to be comparable and meaningful. Therefore, we choose to evaluate the quality of our synthesized images with their ability to improve current anomaly detection algorithms.

Recall that the original TUAB dataset is separated into training and evaluation datasets. Since the training dataset is used to create synthetic images with GT-GAN, we use the 80% of the evaluation dataset to train residual networks (ResNet) of varying depths (20-layer, 56-layer, 110-layer) to distinguish whether a given EEG is normal or not. Preprocessing steps of the evaluation dataset are identical to that of the training set, where EEG channels P3-O1, T5-O1 and F7-T3 are mapped into respective GAF matrices to form three-channel, colored images. We choose to use the ResNet [21] architecture proposed by He et. al. since it is one of the well-known image classification networks that achieve state-of-the-art results on benchmark datasets. Three ResNet models, ResNet-20, ResNet-56 and ResNet-110 are chosen to validate our synthesized images under varying network depths.

	Original	Enhanced
resnet-20	44.2	34.8
resnet-56	39.9	21.4
resnet-110	29.0	**13.5**
PCA-HMM	32.6	-
Epoch-based HMM-SdA	**22.1**	-
GMM-HMM	26.1	-

Table 1. Average error rates (%) before and after the addition of GT-GAN training images. Lowest error rate in each column is highlighted in bold. Resnet-20, resnet-56 and resnet-110 are trained for 50 epochs with 3-fold cross-validation. Baseline classification errors for the TUAB dataset with other models are also reported.

Upon training the ResNet models with 80% of the classification dataset, we use the remaining 20% to test the models' performances on EEG anomaly detection. Then we retrain the networks from scratch with the same data, but this time adding additional images generated from the trained GT-GAN models to double the amount of training images. Classification results are summarized in **Table 1**. It is evident that the additional training images generated by the GT-GAN improve classification accuracy by the ResNet models. We also compare our residual framework's classification errors to the ones reported by Vinit et. al. [17]. Classification error of resnet-110 on the enhanced dataset is lower than all of their proposed classification frameworks on the same dataset. Artifacts are common in GAN-generated images and our generated GAF images contain a negligible amounts of artifacts. Our classification experiments show that the classification performance gain outweighs the amount of artifacts present in the generated iamges.

4. CONCLUSION AND FUTURE WORK

In this study, we propose GT-GAN for synthesis of GAF matrices of EEG data. The proposed network utilizes temporal relationships embedded within the GAF images to synthesize multi-channel GAF images without the use of recurrent layers, which are computationally expensive and contain many parameters. Experimental results have shown that our proposed GT-GAN is capable of generating realistic GAF images. Anomaly classification accuracy of the TUAB dataset with ResNet-20, ResNet-56 and ResNet-110 is improved by the addition of images synthesized by our GT-GAN. Overall, GT-GAN coupled with ResNet achieves a higher classification accuracy ResNet alone. The fact that only 10 seconds of the EEG recordings are used for GAF construction illustrates the effectiveness of our method in preserving information from the original EEGs. We can extend our work to further reduce the computational complexity of GT-GAN with methods such as Piecewise Aggregation Approximation.

5. REFERENCES

[1] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," *CoRR*, vol. abs/1505.04597, 2015.

[2] Q. Zhao, F. Xiao, M. Yang, Y. Li, and J. Wang, "Attentiondta: prediction of drug–target binding affinity using attention model," in *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Nov 2019, pp. 64–69.

[3] C. Chin, B. Lin, G. Wu, T. Weng, C. Yang, R. Su, and Y. Pan, "An automated early ischemic stroke detection system using cnn deep learning algorithm," in *2017 IEEE 8th International Conference on Awareness Science and Technology (iCAST)*, Nov 2017, pp. 368–372.

[4] Xiao Han, "Mr-based synthetic ct generation using a deep convolutional neural network method," *Medical Physics*, vol. 44, 02 2017.

[5] Taeho Jo, Kwangsik Nho, and Andrew J. Saykin, "Deep learning in alzheimer's disease: Diagnostic classification and prognostic prediction using neuroimaging data," in *Front. Aging Neurosci.*, 2019.

[6] Erik Bresch, Ulf Großekathöfer, and Gary Garcia-Molina, "Recurrent Deep Neural Networks for Real-Time Sleep Stage Classification From Single Channel EEG," *Frontiers in computational neuroscience*, vol. 12, pp. 85–85, Oct. 2018, Publisher: Frontiers Media S.A.

[7] Guangyi Zhang, Vandad Davoodnia, Alireza Sepas-Moghaddam, Yaoxue Zhang, and Ali Etemad, "Classification of hand movements from eeg using a deep attention-based lstm network," *IEEE Sensors Journal*, vol. 20, pp. 3113–3122, 2019.

[8] Pouya Bashivan, Irina Rish, Mohammed Yeasin, and Noel Codella, "Learning representations from eeg with deep recurrent-convolutional neural networks," *arXiv preprint arXiv:1511.06448*, 2015.

[9] 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 - Volume 2*, Cambridge, MA, USA, 2014, NIPS'14, pp. 2672–2680, MIT Press.

[10] Mehdi Mirza and Simon Osindero, "Conditional generative adversarial nets," *CoRR*, vol. abs/1411.1784, 2014.

[11] Augustus Odena, Christopher Olah, and Jonathon Shlens, "Conditional image synthesis with auxiliary classifier GANs," International Convention Centre, Sydney, Australia, 06–11 Aug 2017, vol. 70 of *Proceedings of Machine Learning Research*, pp. 2642–2651, PMLR.

[12] Chi Nok Enoch Kan, Najibakram Maheenaboobacker, and Dong Hye Ye, "Age-conditioned synthesis of pediatric computed tomography with auxiliary classifier generative adversarial networks," *IEEE International Symposium on on Biomedical Imaging (ISBI)*, 2020.

[13] Kay Gregor Hartmann, Robin Tibor Schirrmeister, and Tonio Ball, "EEG-GAN: Generative adversarial networks for electroencephalograhic (EEG) brain signals," *arXiv e-prints*, p. arXiv:1806.01875, June 2018.

[14] Tian-jian Luo, Yachao Fan, Lifei Chen, Gongde Guo, and Changle Zhou, "Eeg signal reconstruction using a generative adversarial network with wasserstein distance and temporal-spatial-frequency loss," *Frontiers in Neuroinformatics*, vol. 14, pp. 15, 2020.

[15] Zhiguang Wang and Tim Oates, "Imaging time-series to improve classification and imputation," *CoRR*, vol. abs/1506.00327, 2015.

[16] Alexia Jolicoeur-Martineau, "The relativistic discriminator: a key element missing from standard GAN," *CoRR*, vol. abs/1807.00734, 2018.

[17] Vinit Shah, Eva von Weltin, Silvia Lopez, James McHugh, Lily Veloso, Meysam Golmohammadi, Iyad Obeid, and Joseph Picone, "The temple university hospital seizure detection corpus," *Frontiers in Neuroinformatics*, vol. 12, 01 2018.

[18] Richard W. Homan, "The 10-20 electrode system and cerebral location," *American Journal of EEG Technology*, vol. 28, no. 4, pp. 269–279, 1988.

[19] Shane T. Barratt and Rishi Sharma, "A note on the inception score," *ArXiv*, vol. abs/1801.01973, 2018.

[20] Jake Snell, Karl Ridgeway, Renjie Liao, Brett D. Roads, Michael C. Mozer, and Richard S. Zemel, "Learning to generate images with perceptual similarity metrics," *2017 IEEE International Conference on Image Processing (ICIP)*, pp. 4277–4281, 2015.

[21] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.