Motor Imagery Classification Enhancement using Generative Adversarial Networks for EEG Spectrum Image Generation

Ahmed G. Habashi
Computer and Systems Engineering Department
Ain Shams University
Cairo, Egypt
a.habashy@ieee.org

Ahmed M. Azab
Biomedical Engineering Department
Technical Research Center
Cairo, Egypt
ahmed.m.azab@ieee.org

Seif Eldawlatly
Computer and Systems Engineering Department
Ain Shams University
Computer Science and Engineering Department
The American University in Cairo
Cairo, Egypt
seldawlatly@eng.asu.edu.eg

Gamal M. Aly
Computer and Systems Engineering Department
Ain Shams University
Cairo, Egypt
gamal.aly@eng.asu.edu.eg

2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS) | 979-8-3503-1224-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/CBMS58004.2023.00243

***Abstract*—The development of practical Brain-Computer Interface (BCI) systems has been hindered by significant issues related to data, specifically the lack of sufficient data needed for training. To address this challenge, generating synthetic data that mimics real recorded data has been proposed to augment the real data. One promising technique for data augmentation is through the use of Generative Adversarial Networks (GANs), which have been successfully applied in many other fields. This paper proposes a novel GAN-based approach for generating synthetic spectrum images of Motor Imagery (MI) Electroencephalogram (EEG). The proposed GAN is examined with two Convolutional Neural Network (CNN) architectures in the context of MI classification. Using the public dataset BCI competition IV, our findings reveal that the generated EEG spectrum images using GANs exhibit temporal, spectral, and spatial characteristics similar to the real ones. The average classification accuracy of right-hand versus left-hand MI using the proposed GAN/CNN models has improved to 76.71% with an enhancement of 2.5% in comparison to using the CNN applied to the real data only. These results suggest that using GANs could improve MI BCI systems with limited data.**

Keywords—EEG, GAN, Motor Imagery, Short Time Fourier Transform (STFT), BCI.

I. Introduction

Brain-Computer Interface (BCI) represents one promising technology that enables humans to communicate with computers using their brain activity [1]. Non-invasive BCI systems, which use electroencephalogram (EEG) signals, are considered more practical in comparison to invasive systems given their higher portability and simplicity [2]. The various advantages of using EEG drive conducting many studies with the aim of helping patients with disabilities to, for example, control a robotic arm, control a wheelchair, and use spelling devices [3].

BCI has two main categories: exogenous and endogenous BCIs [4]. An exogenous BCI is a stimulus-based BCI, with examples being steady-state visually evoked potentials and P300. On the other hand, an endogenous BCI is an imagination-based BCI, examples of which include motor imagery, visual imagery and imagined speech [1]. Among the endogenous BCI paradigms, motor imagery (MI) is one of the most widely studied due to its clear neuroscientific characteristics. MI is the process of imagining the movement when performing a certain task or moving a particular body part [5]. This process activates specific areas of the brain related to these movement. It is a key method in BCI studies, particularly for people with movement impairments such as paraplegia or stroke and does not require any external stimuli [6]. MI-based BCIs detect intended movement by recognizing changes in oscillatory activity, known as event-related desynchronization and synchronization (ERD and ERS), in certain EEG frequency bands [7].

Despite the advantages of MI-based BCIs, they suffer from a number of challenges. First, long training sessions are typically required to effectively use MI-based BCIs. MI is a skill that needs to be learned and practiced, that requires long training sessions (20-30 minutes) to properly set up MI-based BCI systems [8]. Second, EEG data suffers from limitations due to its low signal-to-noise ratio (SNR) [4]. Third, machine learning models, such as deep neural networks, which are being increasingly used in analyzing EEG signals require large training sets to achieve the accepted classification accuracy [9].

Due to the aforementioned challenges, recent research has suggested various data augmentation (DA) techniques to overcome such problems by generating synthetic EEG data that resembles the limited recorded signals [10]. Using Generative Adversarial Networks (GANs), as one of the promising DA techniques, has demonstrated a significant improvement in image, audio, and video data generation [11]. GANs can learn the possible distribution of actual data without making any assumptions and then produce synthetic samples via an adversarial training approach between a generator network and a discriminator network. One of the earliest studies which employed GANs for augmenting MI signals, is the study by Abdelfattah et al., in which a recurrent GAN (RGAN) model was introduced for generating synthetic EEG data to increase the dataset size [12]. The RGAN model demonstrated an improvement in the classification accuracy relative to autoencoders (AE) and variational autoencoders (VAE), respectively. Moreover, Fahimi et al. in [13], [14] used Deep Convolutional GANs (DCGANs) to generate synthetic MI data. They investigated the similarity between the generated and the real EEG data in time and frequency domains. Wasserstein GAN (WGAN) was also used in MI data analysis to handle training stability problems [11], [15], [16]. Instead of the classification of the MI data itself, the obtained spectrum images in [17], [18] were utilized for the

2372-9198/23/$31.00 ©2023 IEEE
DOI 10.1109/CBMS58004.2023.00071

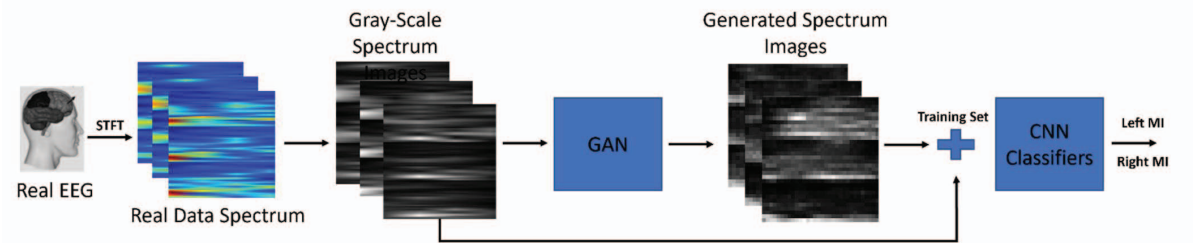


Fig. 1. The proposed approach for data augmentation using input EEG spectrum images obtained using STFT. The real data is used to train a GAN that generates synthetic real-like spectrum images. The real and generated images are used to train a CNN classifier to discriminate between right-hand and left-hand movements.

classification process. They used DCGAN and a combination of a conditional VAE network with GAN for the generation of MI-EEG brain signals. These studies demonstrated that the accuracy of recognizing movements from EEG signals when using the generated GAN data in addition to the recorded signals to train the MI BCI system was remarkably higher than the accuracy achieved when using the recorded signals only.

Here, we propose a new approach to improve the performance of deep learning-based EEG classification techniques by augmenting the original training set with EEG samples generated using GANs. Fig. 1 illustrates an overview of our approach. In this paper, we use gray-scale spectrum images of MI EEG signals with two different CNN models to examine the classification performance. We compare the performance of the CNN classifiers when training these classifiers with and without the augmented datasets. The key contribution of this paper is the proposal of a novel GAN/CNN model that can generate high-quality EEG data and augment its datasets with the synthesized data to overcome the challenge of limited data size.

II. Methods

A. Dataset Description

In this paper, we validated the effectiveness of the proposed methods by using the publicly available BCI competition IV dataset one [19]. The data consists of EEG signals recorded from seven subjects (A, B, C, D, E, F, and G). The data represents the recorded EEG signals for left-hand and right-hand motor imagery recorded with a 100 Hz sampling frequency. Each subject performed 200 trials as the training and test data. In our analysis, we used only EEG signals from three channels (C3, Cz, and C4) since these channels record motor cortex activity [20].

B. EEG Spectrum Image Pre-processing

EEG signals were filtered in the range 8–30 Hz by using a 6^{th} order Butterworth filter [21]. Short Time Fourier transform (STFT) was then used to convert the 4-sec EEG data (MI period) to 2D-EEG spectrum images [22]. The STFT divides the signal into small windows and get the Fourier transform of each one. The selected window size for the dataset is 128 samples. Next, the three channels spectrum images of each trial were vertically merged to get one image of size 32 × 32 for each trial. Finally, the resulting RGB image is converted to a gray-scale one to reduce the model complexity and accelerate both the training and the classification processes. The input data preprocessing was implemented in Python 3.9.7.

C. Generative Adversarial Networks (GANs)

In GANs, the Generator G and the Discriminator D are two antagonistic networks that work together to generate data that is as similar to the real data as possible [10]. The G network generates real-like data from a random noise input vector z while constantly attempting to capture the signal's distribution. In the meantime, the binary classifier D assesses the output of the generator and discriminates between fake $G(z)$ and real samples. Both networks are trained simultaneously, and as a result of their competition, high-quality data is ultimately generated. While the G network's training process strives to increase the likelihood that D classifies generated samples as real, the adversary training for the D network aims to increase the likelihood of identifying "real" data as opposed to "fake" data [5]. The loss function of this adversarial process V can be represented as [5]

$$min_G\, max_D\, V(D,G) = E_{x\in P_r}[log\, D(x)] + E_{z\in P}[log(1 - D(G(z)))] \quad (1)$$

where E is the expected value, P_r describes real data distribution, z represents the random noise vector from the latent space of the simple noise distribution P, $G(z)$ represents the data generated by G and $D(x)$ is the probability that x is a real data. In this optimization process, the Jensen-Shannon divergence is used as the function minimum [5].

D. Proposed GAN Model Architecture

As in the vanilla GANs [23], a generator and a discriminator are both components of our suggested GAN model for EEG spectrum images. Their architectures consist of layers with the parameters detailed in Table I and Table II. Similar to [18], we use a modified sequence of convolutional/deconvolution layers to enable CNNs to handle EEG data because of correlations across channels [24]. The networks details are given below.

a) The Discriminator Network

The Discriminator network predicts whether the input spectrum image is real or fake after taking raw image (a gray-scale image of size 32 × 32) from each subject as an input. The discriminator network consists of two convolutional layers with 128 filters each, 3 × 3 kernel sizes, strides of 2 × 2, “same” padding and a LeakyReLU activation function [25]. The convolution layers are followed by a flattening layer and 0.4 dropout layer. The Adam optimizer of stochastic gradient descent (SGD), with a learning rate of 0.0002 and a momentum of 0.5, was used to train the network. The model has no pooling layers and finally a single node in the output layer with a sigmoid activation function. Cross-entropy is

adopted as the loss function for this binary classification problem. Table I shows the detailed architecture and parameters of the discriminator model.

b) The Generator Network

The generator network is responsible for generating the gray spectrum MI image. This is accomplished by generating an image from an input point drawn from the latent space. The latent space is an arbitrary vector space of Gaussian-distributed values (100 dimensions). Eventually, the latent vector space represents a compressed representation of the output space which only the generator is capable of converting into plausible 32 × 32 pixels spectrum images. As shown in Table II, a dense layer is used as the first hidden layer with sufficient nodes to represent a low-resolution version of the output image, followed by two upsampling layers (deconvolution layer) with (4 × 4) kernel size and strides (2 × 2). Table II shows the detailed architecture and parameters of the discriminator model.

In the training procedure, the raw data was divided into training data and test data using 10-fold cross-validation. Randomly, 90% of the trials were selected as the training set, and the remaining 10% was used as the test set. The training data of each subject is fed to GAN separately to duplicate the number of the gray-scale spectrum images of MI data for each subject.

E. CNN Models for Evaluation

To evaluate our proposed GAN-based approach for data augmentation, we investigated the performance of two CNN classifiers. Each CNN model is used as a classifier of our two-class classification problem (left-hand versus right-hand MI spectrum images). We examined the performance of each CNN model when separately trained using the original data of each subject (i.e. without augmentation) compared to their performance when trained using the original data combined with the synthetic spectrum images generated using the GAN. The classification accuracy is employed as a performance measure by calculating the ratio between correctly recognized test images to the total number of test images. A CNN is a neural network that consists of convolution layer(s), followed by pooling, and fully connected layers [26]. The convolution kernel connects neurons to the previous feature map. The convolution layer is responsible for extracting the input image features, while the pooling layer, the subsequent layer for each convolution layer, compresses the data and parameters and reduces overfitting. Finally, the fully connected layer converts the output matrix to an *n*-dimensional vector and that results in the distribution of the different classes prediction. The details of the proposed CNNs are illustrated as follows:

a) VGG-based CNN

The first CNN structure was proposed and implemented based on using VGG blocks as a known structure of neural network that achieved a remarkable performance for image classification [27]. We have used two VGG blocks of 32 and 16 filters with kernel size of 3 × 3. In addition, we employed the dropout technique and the SGD optimizer. The detailed structure of our proposed CNN model is shown in Table III. The SGD optimizer is used, with a learning rate of 0.0001, 0.99 momentum, and sparse categorical Cross-entropy loss function.

b) Audio-spectrum CNN

The second CNN model is adapted from the structure of the CNN reported in [28] originally developed for audio spectrum images with a modification for gray-scale input images. The architecture of this model adopts various blocks of Convolution with a kernel size of 3 × 3, Batch Normalization (BN), Max-pooling with a kernel size of 2 × 2, and again BN layers. The detailed structure is summarized in Table IV.

The VGG-based CNN was trained for 600 epochs, while the audio-spectrum one was trained for 100 epochs. This operation was repeated 10 times to obtain 10 models where the best achieved accuracies for each model were obtained, then the average accuracy per subject was calculated.

TABLE I. THE DISCRIMINATOR NETWORK DETAILED ARCHITECTURE

Layers	Filter Size	Output Dimension	Activation	Other Parameters
Input	--	(32, 32, 1)	--	--
Convolution	(3,3)	(32, 32, 128)	Leaky ReLU (alpha = 0.2)	Padding = 'same'
Convolution	(3,3)	(16, 16, 128)	Leaky ReLU (alpha = 0.2)	
Flatten	--	32,768	--	
Drop out	--	32,768	--	0.4
Output (Dense)	--	1	Sigmoid	--
optimizer = Adam (learning rate = 0.0002, beta_1=0.5) loss='binary_crossentropy'				

TABLE II. THE GENERATOR NETWORK DETAILED ARCHITECTURE

Layers	Filter Size	Output Dimension	Activation	Other Parameters
Input	--	(1, 1, 100)	--	Padding = 'same'
Dense	--	(8, 8, 4096)	--	
Deconvolution	(4,4)	(16, 16, 128)	Leaky ReLU (alpha = 0.2)	
Deconvolution	(4,4)	(32, 32, 128)	Leaky ReLU (alpha = 0.2)	
Output (Convolution)	(4,4)	(32, 32, 1)	Tanh	

TABLE III. VGG-BASED CNN STRUCTURE

Layers	Filter Size	Output Dimension	Activation	Other Parameters
Input	--	(32, 32, 1)	--	--
Drop out	--	(32, 32, 1)	--	0.2
Convolution	(3,3)	(32, 32, 16)	ReLU	Kernel initializer = 'he uniform' Kernel constraint = Max Norm (3) Padding = 'same'
Convolution	(3,3)	(32, 32, 16)	ReLU	
Max-pooling	(2,2)	(16, 16, 16)	--	
Drop out	--	(16, 16, 16)	--	0.2
Convolution	(3,3)	(16, 16, 32)	ReLU	As the previous Conv. layer
Convolution	(3,3)	(16, 16, 32)	ReLU	
Max-pooling	(2,2)	(8, 8, 32)	--	
Drop out	--	(8, 8, 32)	--	0.4
Convolution	(3,3)	(8, 8, 64)	ReLU	As the previous Conv. layer
Max-pooling	(2,2)	(4, 4, 64)	--	
Drop out	--	(4, 4, 64)	--	0.4
Flatten	--	1024	--	
Dense	--	128	ReLU	--
Drop out	--	128	--	0.4
Output (Dense)	--	2	Softmax	--

TABLE IV. AUDIO-SPECTRUM CNN STRUCTURE

Layers	Filter Size	Output Dimension	Activation	Other Parameters
Input	--	(32, 32, 1)	--	--
Convolution	(3,3)	(16, 16, 32)	ReLU	Same Padding
BN	--	(16, 16, 32)	--	--
Max-pooling	(2,2)	(8, 8, 32)	--	
BN	--	(8, 8, 32)	--	--
Convolution	(3,3)	(8, 8, 64)	ReLU	Same Padding
BN	--	(8, 8, 64)	--	--
Max-pooling	(2,2)	(4, 4, 64)	--	
BN	--	(4, 4, 64)	--	--
Convolution	(3,3)	(4, 4, 128)	ReLU	Same Padding
BN	--	(4, 4, 128)	--	--
Max-pooling	(2,2)	(2, 2, 128)	--	
BN	--	(2, 2, 128)	--	--
Flatten	--	512	--	
Dense	--	256	ReLU	--
BN	--	256	--	--
Drop out	--	256	--	0.5
Output (Dense)	--	2	Softmax	--
optimizer = RMSprop loss = sparse categorical cross-entropy				

III. RESULTS

A. GAN Losses Over Training

We first examine the values of the loss function as the training of the GAN progresses. Fig. 2 shows the discriminator and generator losses over 300 epochs for the two classes of one sample subject (Subject F). The discriminator loss of real images is represented in blue while that of the fake images is represented in orange and the generator loss for the generated images is shown in green during training. At approximately 60 epochs, the model losses converge to some constant values. The other subjects' training data follow the same trend. This indicates that the training process results in a stable GAN model.

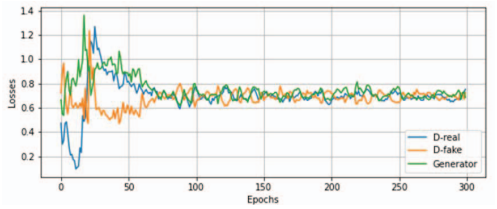
B. Generated EEG Images

We next investigate the time-frequency spectrums generated using the proposed GAN in comparison to the real ones. Fig. 3 depicts right-hand and left-hand motor imagery spectrums from the real EEG recordings versus the generated GAN images after epoch 300 for one sample subject. The figure demonstrates similarities between the real and generated spectrums. For instance, the generated spectrum of right-hand MI (Fig. 3b) obviously similar to the real one (Fig. 3a) as both exhibited a higher power density in the alpha frequency band (8–12.5 Hz) for channel C4 compared to the left-hand motor imagery (Fig. 3c and Fig. 3d). Likewise, the generated spectrum of the left-hand MI (Fig. 3d) reveals similarities to the real spectrum (Fig. 3c) with higher beta band power (18–26 Hz) in channels C3 and Cz compared to the right-hand MI spectrums. The figures demonstrate that the proposed GAN has successfully produced real-like images, which we hypothesize to have a positive impact on the training of the subsequent CNN classifiers.

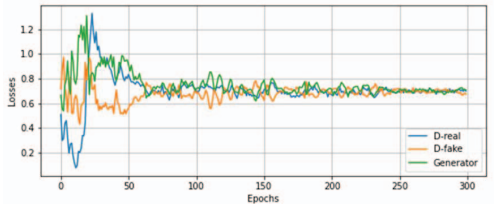
C. Classification Results

In order to assess the effectiveness of our approach, we trained two different CNN models with the real data in addition to the GAN-generated data and compared their performance when trained using the real data only. We employed the classification accuracy as a performance metric. Table V shows the classification accuracy achieved using each CNN model without augmentation and after using the generated spectrum images in training. Comparing the examined CNN architectures before using GANs, the results demonstrate that the proposed VGG-based CNN model is slightly better than audio-spectrum CNN with accuracy 74.64% ± 10.29, and 74.21% ± 9.98, respectively. However, using the second CNN was more efficient as it reached a comparable accuracy to that of VGG-based CNN in only 100 training epochs as opposed to 600 training epochs for VGG-based CNN.

To evaluate the effect of the synthesized EEG spectrums on the classification performance of the CNNs, we trained the CNNs with augmented the datasets, which comprised the raw

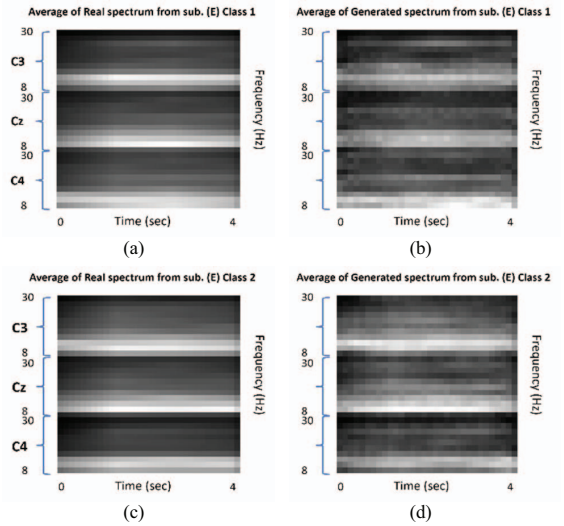


(a)



(b)

Fig. 2. GAN loss function versus epochs for (a) the 1st class (left-hand MI) and (b) the 2nd class (right-hand MI).



(a) (b) (c) (d)

Fig. 3. Average spectrum images of a sample subject (Subject E) of the real left/right hand MI-EEG signal (a, c) and the generated EEG spectrums using GANs (b, d). Class 1 represents left-hand motor imagery while Class 2 represents right-hand motor imagery.

TABLE V. CLASSIFICATION ACCURACY OBTAINED FOR DIFFERENT ARCHITECTURES

Subject	VGG-Based CNN	Audio-spec. CNN	GAN/ VGG-Based CNN	GAN/ Audio-spec. CNN	[29]
A	75.5	77	76	**80**	86.5
B	68.5	69.5	**73**	71	53.5
C	72.5	65.5	**69.5**	67	--
D	69.5	69	**75**	72	--
E	87.5	90	89	**90.5**	--
F	72.5	74	74	**77**	66.5
G	76.5	74.5	76	**79.5**	93.5
Mean	74.64 ± 10.29	74.21 ± 9.98	76.07 ± 10.03	**76.71 ± 10.06**	--

data combined with the same size of the generated data. Then, we use the same test dataset (selected before augmentation) for each subject separately to assess the classification accuracy. As shown in Table V, the artificially generated data improves the classification performance for all subjects. The average classification result with GAN data has improved by 2.5% on average reaching 76.71%. Comparing to one of the latest results which used the same dataset, our approach exceeds the accuracy reported by Miao et al. by 1.625% for the four subjects examined in that study [29]. These results indicate that the proposed GAN approach can produce high-quality synthetic EEG data, which may subsequently be used to improve the classification results of deep learning models used for MI-based BCI applications.

IV. CONCLUSION

We propose novel GAN and CNN architectures to synthesize MI-EEG samples for data augmentation. We adapted an approach based on the classification of spectrum images of MI data. The performance of this approach was assessed using the BCI IV-1 dataset. With more EEG data synthesized using our GAN model, the CNNs improved in classification performance. We compared the results of our proposed GAN/CNN approach with the latest state-of-the-art evaluated by the same dataset. We demonstrated that data augmentation is a viable strategy for improving the classification accuracy of deep neural network-based EEG processing algorithms, and that the proposed GAN/CNN architectures are capable of producing high-quality EEG data. Nonetheless, we aim to develop our work to successfully perform across several datasets for various classification tasks. Additionally, the use of different metrics to measure signal similarity and latest variations of the GAN framework to compare outcomes are both ideas for future research.

REFERENCES

[1] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain computer interfaces, a review," *Sensors*, vol. 12, no. 2. pp. 1211–1279, Feb. 2012. doi: 10.3390/s120201211.

[2] A. Biasiucci, B. Franceschiello, and M. M. Murray, "Electroencephalography," *Current Biology*, vol. 29, no. 3, pp. R80–R85, Feb. 2019, doi: 10.1016/J.CUB.2018.11.052.

[3] N. Padfield, K. Camilleri, T. Camilleri, S. Fabri, and M. Bugeja, "A Comprehensive Review of Endogenous EEG-Based BCIs for Dynamic Device Control," *Sensors 2022, Vol. 22, Page 5802*, vol. 22, no. 15, p. 5802, Aug. 2022, doi: 10.3390/S22155802.

[4] S. Kim, D. Lee, S. L.-I. Access, and undefined 2022, "Rethinking CNN Architecture for Enhancing Decoding Performance of Motor Imagery-Based EEG Signals," *ieeexplore.ieee.org*, Accessed: Jan. 22, 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9878333/

[5] J. Xie, S. Chen, Y. Zhang, D. Gao, and T. Liu, "Combining generative adversarial networks and multi-output CNN for motor imagery classification," *J Neural Eng*, vol. 18, no. 4, Aug. 2021, doi: 10.1088/1741-2552/abecc5.

[6] E. Sarasso, M. Gemma, F. Agosta, M. Filippi, and R. Gatti, "Action observation training to improve motor function recovery: a systematic review," *Arch Physiother*, vol. 5, no. 1, Dec. 2015, doi: 10.1186/S40945-015-0013-X.

[7] R. Abiri, S. Borhani, E. W. Sellers, Y. Jiang, and X. Zhao, "A comprehensive review of EEG-based brain–computer interface paradigms," *J Neural Eng*, vol. 16, no. 1, p. 011001, Jan. 2019, doi: 10.1088/1741-2552/AAF12E.

[8] A. M. Azab, J. Toth, L. S. Mihaylova, and M. Arvaneh, "A review on transfer learning approaches in brain-computer interface," *Signal Processing and Machine Learning for Brain-Machine Interfaces*, pp. 81–101, Jan. 2018, doi: 10.1049/PBCE114E_CH5/CITE/REFWORKS.

[9] G. Li, C. H. Lee, J. J. Jung, Y. C. Youn, and D. Camacho, "Deep learning for EEG data analytics: A survey," *Concurr Comput*, vol. 32, no. 18, p. e5199, Sep. 2020, doi: 10.1002/CPE.5199.

[10] Y. Abdelghaffar, A. Hashem, and S. Eldawlatly, "Generative Adversarial Networks for Augmenting EEG Data in P300-based Applications: A Comparative Study," *Proc IEEE Symp Comput Based Med Syst*, vol. 2022-July, pp. 177–182, 2022, doi: 10.1109/CBMS55023.2022.00038.

[11] I. Goodfellow *et al.*, "Generative adversarial networks," *Commun ACM*, vol. 63, no. 11, pp. 139–144, Oct. 2020, doi: 10.1145/3422622.

[12] S. M. Abdelfattah, G. M. Abdelrahman, and M. Wang, "Augmenting the Size of EEG datasets Using Generative Adversarial Networks," *Proceedings of the International Joint Conference on Neural Networks*, vol. 2018-July, Oct. 2018, doi: 10.1109/IJCNN.2018.8489727.

[13] F. Fahimi, Z. Zhang, W. B. Goh, K. K. Ang, and C. Guan, "Towards EEG generation using gans for bci applications," *2019 IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2019 - Proceedings*, May 2019, doi: 10.1109/BHI.2019.8834503.

[14] F. Fahimi, S. Dosen, K. K. Ang, N. Mrachacz-Kersting, and C. Guan, "Generative Adversarial Networks-Based Data Augmentation for Brain-Computer Interface," *IEEE Trans Neural Netw Learn Syst*, vol. 32, no. 9, pp. 4039–4051, Sep. 2021, doi: 10.1109/TNNLS.2020.3016666.

[15] Y. Luo, S. Y. Zhang, W. L. Zheng, and B. L. Lu, "WGAN Domain Adaptation for EEG-Based Emotion Recognition," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11305 LNCS, pp. 275–286, Dec. 2018, doi: 10.1007/978-3-030-04221-9_25.

[16] K. G. Hartmann, R. T. Schirrmeister, and T. Ball, "EEG-GAN: Generative adversarial networks for electroencephalograhic (EEG) brain signals," Jun. 2018. [Online]. Available: http://arxiv.org/abs/1806.01875

[17] J. Yang, H. Yu, T. Shen, Y. Song, and Z. Chen, "4-class mi-eeg signal generation and recognition with cvae-gan," *Applied Sciences (Switzerland)*, vol. 11, no. 4, pp. 1–14, Feb. 2021, doi: 10.3390/app11041798.

[18] K. Zhang *et al.*, "Data augmentation for motor imagery signal classification based on a hybrid neural network," *Sensors (Switzerland)*, vol. 20, no. 16, pp. 1–20, Aug. 2020, doi: 10.3390/s20164485.

[19] "BCI Competition IV." https://www.bbci.de/competition/iv/#dataset1 (accessed Mar. 18, 2023).

[20] B. Xu *et al.*, "Wavelet Transform Time-Frequency Image and Convolutional Network-Based Motor Imagery EEG Classification," *IEEE Access*, vol. 7, pp. 6084–6093, 2019, doi: 10.1109/ACCESS.2018.2889093.

[21] I. W. Selesnick and C. Sidney Burrus, "Generalized digital butterworth filter design," *IEEE Transactions on Signal Processing*, vol. 46, no. 6, pp. 1688–1694, 1998, doi: 10.1109/78.678493.

[22] D. W. Griffin and J. S. Lim, "Signal Estimation from Modified Short-Time Fourier Transform," *IEEE Trans Acoust*, vol. 32, no. 2, pp. 236–243, 1984, doi: 10.1109/TASSP.1984.1164317.

[23] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Sherjil Ozair, and Aaron Courville, "Generative Adversarial Nets," 2014. Accessed: Nov. 27, 2021. [Online]. Available: https://proceedings.neurips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf

[24] I. A. Corley and Y. Huang, "Deep EEG super-resolution: Upsampling EEG spatial resolution with Generative Adversarial Networks," in *2018 IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2018*, Apr. 2018, vol. 2018-January, pp. 100–103. doi: 10.1109/BHI.2018.8333379.

[25] R. Nandhini Abirami, P. M. Durai Raj Vincent, K. Srinivasan, U. Tariq, and C. Y. Chang, "Deep CNN and Deep GAN in Computational Visual Perception-Driven Image Analysis," *Complexity*, vol. 2021, 2021, doi: 10.1155/2021/5541134.

[26] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *Journal of Big Data 2021 8:1*, vol. 8, no. 1, pp. 1–74, Mar. 2021, doi: 10.1186/S40537-021-00444-8.

[27] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, Sep. 2014, doi: 10.48550/arxiv.1409.1556.

[28] "Classify MNIST Audio using Spectrograms/Keras CNN | Kaggle." https://www.kaggle.com/code/christianlillelund/classify-mnist-audio-using-spectrograms-keras-cnn (accessed Jan. 24, 2023).

[29] Y. Miao *et al.*, "Learning Common Time-Frequency-Spatial Patterns for Motor Imagery Classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 699–707, 2021, doi: 10.1109/TNSRE.2021.3071140.