

Optimization Driven Variational Autoencoder GAN for Artifact Reduction in EEG Signals for Improved Neurological Disorder and Disability Assessment

Mohamed Yacin Sikkandar ^{1*}, S. Sabarunisha Begum ², Ahmad Alassaf¹, Ibrahim AlMohimeed¹, Khalid Alhussaini³, Adham Aleid ³, Abdulrahman Khalid Alhaidar⁴

¹Department of Medical Equipment Technology, College of Applied Medical Sciences, Majmaah University, Al Majmaah 11952, Saudi Arabia, m.sikkandar@mu.edu.sa, am.alassaf@mu.edu.sa, i.almohimeed@mu.edu.sa

²Department of Biotechnology, P.S.R. Engineering College, Sivakasi 626140, India, sabarunishabegum@psr.edu.in

³Department of Biomedical Technology, College of Applied Medical Sciences, King Saud University, Riyadh 12372, Saudi Arabia, kalhussaini@ksu.edu.sa, adaleid@ksu.edu.sa

⁴Support Services Department, Tumair General Hospital, Ministry of Health, Tumair 15316, Saudi Arabia, akalhaidar@moh.gov.sa

Abstract: This study introduces an innovative method for minimizing artifacts in electroencephalography (EEG) signals by integrating brainstorm optimization (BSO) with a variational autoencoder generative adversarial network (VAE-GAN), resulting in the BrOpt_VAGAN model. EEG signals are critical for the diagnosis of neurological disorders, for brain-computer interface (BCI) applications, and for the monitoring of neurological disabilities. However, EEG data often contains artifacts from physiological sources — such as electro-oculographic (EOG), electromyographic (EMG), and electrocardiographic (ECG) signals — which can distort the accuracy of brain activity readings. Our proposed BrOpt_VAGAN model combines BSO with a VAE-GAN framework to more effectively remove these artifacts, thus improving the clarity and accuracy of EEG signals. In this model, the VAE first reduces the raw EEG signals into a lower-dimensional representation that captures the essential signal patterns while filtering out the noise. The GAN component then refines this representation via adversarial training, effectively minimizing artifacts and improving the quality of the processed EEG data. BSO optimally adjusts the encoding and decoding parameters within the VAE-GAN structure, enabling the model to handle different noise levels and helps to find different neurological disorders. Preliminary results show that BrOpt_VAGAN performs significantly better with an accuracy of 98.5 % and an error rate of 11.23 %, enabling a clearer and more precise EEG signal reconstruction.

Keywords: EEG, ECG, artefacts, EMG, random noise, neural network, optimization, fractional calculus

1. INTRODUCTION

Numerous studies have shown that neurological problems are increasing at an alarming rate. The WHO states that one in four people worldwide will experience neurological problems at some point in their lives [1]. Neurological diseases are the second most common disease worldwide after ischemic heart disease. Neurological diseases affect both the brain and the nervous system of the human body [2]. Many neurological disorders are well-documented and rather prevalent, while many others are uncommon. Neurological difficulties include a variety of conditions, including epilepsy, learning disabilities, neuromuscular disorders, autism, Alzheimer's disease, attention deficit hyperactive disorder (ADHD), multiple sclerosis, Parkinson's disease,

sleep problems, and cerebral palsy. Mental illnesses are classified as "psychiatric diseases" and are primarily characterized by abnormalities in cognition, emotion, or behavior that lead to suffering or functional impairment. A variety of brain-imaging modalities are available to diagnose neurological disorders, including positron emission tomography (PET), near infrared spectroscopy (NIRS), magnetoencephalography (MEG), electroencephalography (EEG), and functional magnetic resonance imaging (fMRI) [3]. This paper emphasizes EEG analysis due to its cost-effectiveness, non-invasiveness, and portability, making it a widely used approach. An EEG systematically monitors and records the brain's electrical activity to assess the cerebral processes. Many studies use EEG data to detect neurological

diseases, neurodevelopmental problems, acute neurological events, and patient behavior [4]-[5]. Traumatic brain injury is the leading cause of disability and death in children worldwide. Over five million Americans are disabled as a result of a traumatic brain injury. Researchers believe that a computer-aided diagnosis (CAD) system trained on extensive patient data and physiologic signals and images using advanced signal processing and AI/ML techniques can help neurologists, neurosurgeons, radiologists, and other medical professionals improve clinical decision-making. Research in this area has increased significantly over the last ten years.

In [6], a generalized EEG neural network (GENet) architecture is developed based on a convolutional neural network, which is able to identify various neurological disorders based on EEG data. This paradigm facilitates the execution of the essential functions for the categorization process. In [7], a deep neural networks (DNN)-based hybrid ensemble feature selection (HEFS) Framework for Parkinson's disease identification was proposed. Multi-level dimensionality reduction (MLDR) is applied to HEFS matrices. After normalizing the matrix scores, merging the scores, reconstructing a new dataset, and reducing the features using neighborhood component analysis (NCA), an accuracy rate of 97.08 % and an F1-score of 98.10 % were achieved. In [8], a unique expert system was introduced that utilizes just EEG information for the early diagnosis of schizophrenia. A deep learning network was developed to improve the accuracy of the image categorization outcomes.

In [9]-[11], the use of variable-frequency complex demodulation (VFCDM) and convolutional neural networks (CNN) to differentiate between healthy, interictal, and ictal states was investigated using EEG data. Time frequency spectrum (TFS) shows frequency changes across different states that correspond to fluctuations in brain activity. The LOSO CV method routinely achieves good performance, ranging from 90 % to 99 % across different combinations of healthy and epileptic states. In [12]-[14], the EEG temporal spatial network (ETSNet) is introduced, which includes a Squeeze-and-Excitation Block and several CNNs tailored for eyes-open and closed resting states. Several limitations are evident from the above studies:

1. The lack of standardized assessment measures and datasets makes comparison difficult.
2. The computational complexity of some deep learning models such as CNN with long short term memory (LSTM) may limit their practical use.

2. PROPOSED METHODOLOGY

An optimization-enhanced variational autoencoder generative adversarial network (OE-VAE-GAN) for artifact reduction in EEG signals could be a robust approach for cleaning EEG data, especially in clinical and research contexts where artifact presence (e.g., due to muscle movements, eye blinks, or ambient noise) compromises the quality of data analysis, as shown in Fig. 1.

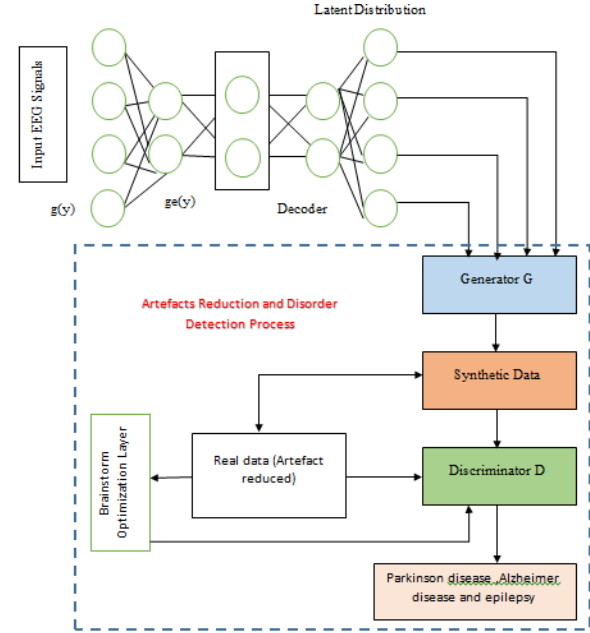


Fig. 1. BSO-VAE-GAN architecture for artifact reduction.

Probabilistic variational autoencoder (PVA) based filtering

First define the low-pass filter (LPF) and high-pass filter (HPF):

$$LPF[t] = \frac{S[t].w[t]}{\sum_{i=0}^{T-1} S[i].w[i]} \quad (1)$$

$$HPF[t] = \delta[t] - LPF[t] \quad (2)$$

where $S[t]$, $w[t]$, and $\delta[t]$ represent the sinc filter, the Hamming window, and the discrete unit impulse function. PVA consists of five main parts: feature extractor, encoder, sampler, feature generator, and signal reconstructor. The model's components work together to synthesize EEG signals. The feature extractor $g_x(\cdot)$ uses cascade filters to divide the input signal y into four amplitude-modulated subsets $x \in \{x_{HH}, x_{HL}, x_{LH}, x_{LL}\}$, which are the learning targets of the feature generator. The encoder $g_e(\cdot)$ learns the distribution parameters of the latent variable z and cutoff frequency θ . It makes two assumptions:

- $\theta_k \sim u(0, 1)$ for $k = 1, 2 \dots .6$ stands for six occurrences in the proposed model. The Bernoulli distributions approximates this distribution.
- $z_j \sim N(\mu_{z_j}, \sigma_{z_j}^2)$ for each dimension $j \in \{1, 2, \dots, J\}$, where j is a hyperparameter that controls the number of dimensions.

The samplers $g_z(\cdot)$ and $g_\theta(\cdot)$ provide a distinct data distribution estimate. The method uses reparameterization for backpropagation, sampling z_j and θ_k to allow the gradients to flow across the network node. Using the encoder settings, feature generator $g_{z'}(\cdot)$ creates four feature signals for the signal reconstructor. Signal reconstructor $g_y(\cdot)$ uses the generated feature subsets to reconstruct the signals while preserving their fundamentals. Monte Carlo (MC) can be used to estimate the predicted log likelihood of reconstructing raw signals y from reconstructed feature signals x .

$$-E_{q(x, z, \theta|y)}[\log p(y|x)] \approx -\frac{1}{L} \sum_{l=1}^L \log p(y|x^{(l)}) \quad (3)$$

where L is the number of samples. The mean square error (MSE) was reduced instead of the negative log likelihood from (4) to maintain convergence stability.

$$J(\varphi_y; y) = \frac{1}{N} \sum_{n=1}^N (y_n - g_n(x_n))^2 \quad (4)$$

where the sample index and the total number of training samples are n and N , respectively. The optimization of the latent variable Z reduces the Kullback-Leibler divergence (KLD) compared to the previous ones. Latent variable Z optimization under variational inference for $p(z)$ requires a reduction of the KLD relative to the prior.

$$E_{q(z|y)q(\theta|y)q(x|y, \theta)} \left[\log \frac{q(z|y)}{p(z)} \right] = KLD(q(z|y)p(z)) \quad (5)$$

The Gaussian distribution KLD and a VAE reparametrization method were used for this optimization.

Generative adversarial network with optimization process

A GAN consists of two CNNs, a generator and a discriminator, with opposing conditional arguments. In the discriminator, we used the Patch-GAN to classify each patch as true or generated. The discriminator should punish local signal patches to accurately mimic high-frequency components. The GAN training total loss function is:

$$L_{GAN}(G, D) = E_{x,y}[\log D(x, y)] + E_{x,y}[\log(1 - D(x, G(x, z)))] \quad (6)$$

The generator (G) minimizes the loss function $L_{GAN}(G, D)$, while the discriminator D maximizes it to discriminate between the generated samples $G(x, y)$ and the actual samples y . We provide the discriminator estimate error loss feedback to guarantee successful training. Therefore, the final goal function is:

$$G^* = ARG \min_G \max_D L_{GAN}(G, D) + \gamma L_{L_1(G)} \quad (7)$$

where $L_{L_1(G)}$ is an extra L_1 norm based on the generator function loss to approximate the ground truth output. γ is an adjustable parameter set to 100. The multi-modal feature-fusion of image and interaction information L_{GAN} causes the generative network to create the same pathological feature as the original signal:

$$L_{GAN} = \sum_{n=1}^N \log D \theta_x(G \theta_x(I^\alpha, t)) \quad (8)$$

where t is the interactive information and N is the number of samples. We use the visual perceptual loss to maintain the perceptual similarity. The loss function is:

$$L_{GAN/i,j} = \frac{1}{W_{i,j} H_{i,j}} \sum_{i=1}^{W_{i,j}} \sum_{j=1}^{H_{i,j}} (\varphi_{i,j}(I^f) - \varphi_{i,j}(G_{\theta_g}(I^{\alpha,t}))) \quad (9)$$

3. PERFORMANCE ANALYSIS

Dataset description

In this section, both study datasets are described. The CHB-MIT dataset [12] initially included 22 participants: 17 women aged 1.5-19 years and 5 men aged 3-22 years. The collection includes 198 seizures and 969 hours of EEG recordings. The number of seizures is lower than the number of seizure-free signals. The second dataset, KAU, was obtained at 256 Hz from two male scalp EEG patients aged 28 years. This dataset is similar to the CHB-MIT dataset. The subjects' ages were considered. The individuals in the CHB-MIT dataset are similar to these two cases. In both datasets, an age range of 1-28 years was chosen. This is significant as age considerably affects the clinical and electroencephalographic features of seizures [13]. Both individuals had 38-channel EEGs. They had two 495 s seizures and four 417 s seizures, respectively. The CHB-MIT dataset selects 18 out of 23 channels because they are similar to all recordings.

Experimental setup

The time-step value in this study varied from 0 to $m - t$, with t set to 50. The prior and downstream tasks were each trained for 30 epochs with a batch size of 64. The network was trained using Python, PyTorch, and an RTX 3060Ti GPU. The BrOpt_VAGAN model with its complex VAE and GAN components can be very computationally intensive. Real-time EEG analysis requires significant hardware resources (e.g., GPUs, TPUs) that may not be available in all medical settings, especially for portable or low-power devices. Here, the modeling findings of the proposed artifact elimination method using the BrOpt_VAGAN network are quickly investigated. The simulation is performed by comparing the results with a number of well-known methods, including HEFS+DNN [12], VFCDM+CNN [14]. These are done to reduce artifacts due to random noise. The metrics used for assessment in this study include MSE and signal to noise ratio (SNR).

The MSE measure describes the difference between the real reaction and the intended response:

$$MSE = \frac{1}{N} (O^n - D^n)^2 \quad (10)$$

Table 1 shows the accuracy and error calculation for the proposed BrOpt_VAGAN method in terms of pseudo-clean and noisy input.

Table 1. Accuracy performance of the proposed BrOpt_VAGAN model.

	Mixtures of artifact components	Accuracy [%]	Error [%]
Pseudo-clean	brain	98.5	12.41
	eye	96.2	11.53
	muscle	97.3	12.74
Noisy input	brain	98.6	11.84
	eye	95.9	11.90
	muscle	93.5	12.56

Fig. 2 shows that BrOpt_VAGAN consistently achieves the lowest *MSE* % with values between 11.2 % and 12.6 % and thus has superior accuracy. In contrast, HEFS+DNN shows stable but higher *MSE* values (20.1 % to 20.9 %), while VFCDM+CNN has the highest *MSE* % across all channels (21.3 % to 21.76 %). This shows that BrOpt_VAGAN is the most effective method for error reduction regardless of the number of channels, outperforming both HEFS+DNN and VFCDM+CNN.

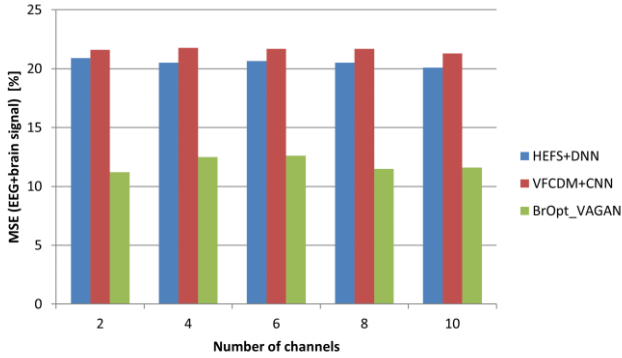


Fig. 2. Comparison of *MSE* with the EEG+brain signal artifact.

Fig. 3 shows the *MSE* of HEFS+DNN, VFCDM+CNN, and BrOpt_VAGAN for EEG signals with eye signal artifacts across channels (2 to 10). HEFS+DNN has stable *MSE* values between 19.4 % and 19.5 % and thus shows a constant performance. VFCDM+CNN has the highest *MSE* % (21.5 % to 21.9 %), indicating that it is less effective at handling artifacts. BrOpt_VAGAN achieves lower *MSE* values (15.3 % to 15.9 %) than VFCDM+CNN but slightly higher than HEFS+DNN, suggesting moderate effectiveness in reducing artifacts. Overall, HEFS+DNN is the most stable, while BrOpt_VAGAN performs better than VFCDM+CNN in error minimization.

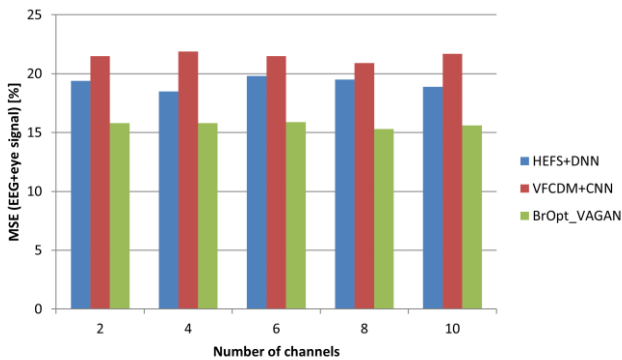


Fig. 3. Comparison of *MSE* with the EEG+eye signal artifact.

Fig. 4 compares the *MSE* of the aforementioned methods when processing EEG signals with muscle signal artifacts over different numbers of channels. HEFS+DNN consistently shows the highest *MSE* %, ranging from 27.4 % to 28.6 %, indicating a significant error in handling muscle artifacts. VFCDM+CNN achieves lower *MSE* values, ranging from 19.2 % to 19.5 %, indicating moderate effectiveness in artifact reduction. BrOpt_VAGAN consistently achieves the

lowest *MSE* %, with values between 12.3 % and 12.98 %, indicating superior performance in minimizing errors due to muscle artifacts. Overall, BrOpt_VAGAN is the most effective, followed by VFCDM+CNN, while HEFS+DNN is the least effective in this context.

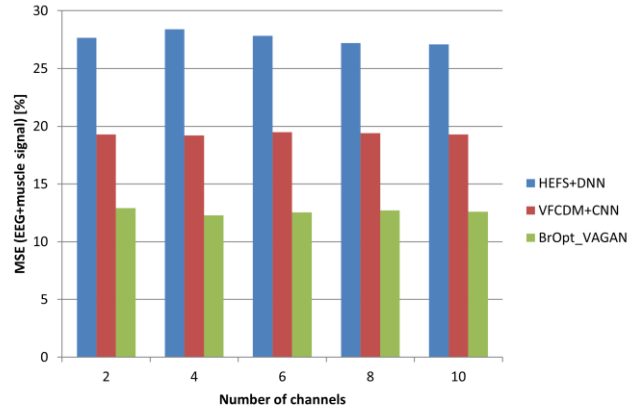


Fig. 4. Comparison of *MSE* with the EEG+muscle signal artifact.

4. CONCLUSION

The proposed study presents a BrOpt_VAGAN framework for automated classification of neurological disorders from raw EEG data. The experiments are performed with BrOpt_VAGAN, a publicly available benchmark dataset. The experiments were conducted under closed and open-eye conditions, as recommended in the CHB-MIT dataset publication and other research papers. The results show that the proposed method can improve the maximum performance by 98.7 % accuracy on the specified dataset using multiple channels. The performance improvement is shown for five-class classification, which confirms the effectiveness and efficiency of the BrOpt_VAGAN framework. Therefore, future work will focus on exploring other loss functions tailored to imbalanced data or incorporating ensemble methods that can also lead to improved accuracy. Furthermore, fine-tuning the self-supervised learning approach with larger and more diverse datasets could lead to better generalization and make the method more reliable in practical applications.

FUNDING

King Salman Center for Disability Research through Research Group Number KSRG-2023-339. The funders had no role in the study design, data collection and analysis, the decision to publish, or the preparation of the manuscript.

DATA AVAILABILITY STATEMENT

The dataset is available under the following reference, Little, Max A et al. "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease." IEEE transactions on bio-medical engineering vol. 56, 4 (2009): 1015.

ACKNOWLEDGMENTS

The authors extend their appreciation to the King Salman Center for Disability Research through Research Group Number KSRG-2023-339 for supporting this research.

REFERENCES

- [1] Raghavendra, U., Acharya, U. R., Adeli, H. (2020). Artificial intelligence techniques for automated diagnosis of neurological disorders. *European Neurology*, 82 (1-3), 41-64. <https://doi.org/10.1159/000504292>
- [2] Jayanthi, J., Kavitha, M., Jayasankar, T., Britto, A. S. F., Prakash, N. B. (2021). Segmentation of brain tumor magnetic resonance images using a teaching-learning optimization algorithm. *Computers, Materials & Continua*, 68 (3), 4191-4203. <https://doi.org/10.32604/cmc.2021.012252>
- [3] Stephe, S., Jayasankar, T., Kumar, K. V. (2022). Motor imagery EEG recognition using deep generative adversarial network with EMD for BCI applications. *Technical Gazette*, 29 (1), 92-100. <https://doi.org/10.17559/TV-20210121112228>
- [4] Tawhid, M. N. A., Siuly, S., Wang, K., Wang, H. (2024). GENet: A generic neural network for detecting various neurological disorders from EEG. *IEEE Transactions on Cognitive and Developmental Systems*, 16 (5), 1829-1842. <https://doi.org/10.1109/TCDS.2024.3386364>
- [5] Ouhmida, A., Saleh, S., Ammar, A., Raihani, A., Cherradi, B. (2024). HEFS-MLDR: A novel hybrid ensemble feature selection framework for improved deep neural network architecture in the diagnosis of Parkinson's disease. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-024-20276-x>
- [6] Al Fahoum, A., Zyout, A. (2024). Wavelet transform, reconstructed phase space, and deep learning neural networks for EEG-based schizophrenia detection. *International Journal of Neural Systems*, 34 (9), 2450046. <https://doi.org/10.1142/S0129065724500461>
- [7] Veeranki, Y. R., McNaboe, R., Posada-Quintero, H. F. (2023). EEG-based seizure detection using variable-frequency complex demodulation and convolutional neural networks. *Signals*, 4 (4), 816-835. <https://doi.org/10.3390/signals4040045>
- [8] Shah, S. J. H., Albishri, A., Kang, S. S., Lee, Y., Sponheim, S. R., Shim, M. (2023). ETSNet: A deep neural network for EEG-based temporal-spatial pattern recognition in psychiatric disorder and emotional distress classification. *Computers in Biology and Medicine*, 158, 106857. <https://doi.org/10.1016/j.compbiomed.2023.106857>
- [9] Fouad, I. A., Labib, F. E.-Z. M. (2023). Identification of Alzheimer's disease from central lobe EEG signals utilizing machine learning and residual neural network. *Biomedical Signal Processing and Control*, 86 (B), 105266. <https://doi.org/10.1016/j.bspc.2023.105266>
- [10] Holmes, G. L. (2012). Consequences of epilepsy through the ages: When is the die cast? *Epilepsy Currents*, 12 (4_suppl), 4-6. <https://doi.org/10.5698/1535-7511-12.4s.4>
- [11] Kaur, C., Singh, P., Sahni, S. (2021). EEG artifact removal system for depression using a hybrid denoising approach. *Basic and Clinical Neuroscience*, 12 (4), 465-476. <https://doi.org/10.32598/bcn.2021.1388.2>
- [12] Zhao, X., Liu, D., Ma, L., Ai, Q., Liu, Q., Xie, S. (2021). EEG signals de-noising with wavelet by optimizing threshold based on fruit fly optimization. In *ICNCC '20: Proceedings of the 2020 9th International Conference on Networks, Communication and Computing*. New York, US: Association for Computing Machinery, 71-77. <https://doi.org/10.1145/3447654.3447665>
- [13] Chen, X., Xu, X., Liu, A., McKeown, M. J., Wang, Z. J. (2018). The use of multivariate EMD and CCA for denoising muscle artefacts from few-channel EEG recordings. *IEEE Transactions on Instrumentation and Measurement*, 67 (2), 359-370. <https://doi.org/10.1109/TIM.2017.2759398>
- [14] Gorjan, D., Gramann, K., Pauw, K. D., Marusic, U. (2022). Removal of movement-induced EEG artefacts: Current state of the art and guidelines. *Journal of Neural Engineering*, 19 (1), 011004. <https://doi.org/10.1088/1741-2552/ac542c>

Received June 4, 2024
Accepted January 16, 2025