Enhancing Motor Imagery EEG Generation with Single Step Diffusion Models

Anonymous ICME submission

***Abstract*—Motor imagery-based brain-computer interfaces (MI-BCI) have attracted increasing attention in recent years. However, the development of reliable MI-BCI relies on a large amount of high-quality electroencephalography (EEG) data. Recent advances in diffusion probabilistic models (DPMs) have shown great potential in generating high-quality synthetic data across various domains. However, effective generating electroencephalogram (EEG) signals that align with neuroscientific characteristics remains a challenge. In this work, we propose DiffEEGLossNet, a novel deep generative framework based on DPMs designed to synthesize multi-channel EEG signals. Our approach incorporates motor imagery-specific loss constraints to enhance the neuroscientific characteristics of the generated EEG signals. We validate our model on the BCI Competition IV datasets 2a and 2b, demonstrating that the synthesized EEG signals exhibit high fidelity and significantly improve the performance of the EEG classification models when used for data augmentation. These findings highlight the potential of the proposed model in the advancement of MI-based EEG and other neuroengineering applications.**

***Index Terms*—motor imagery, diffusion models, data augmentation, EEG signal synthesis**

I. Introduction

Motor imagery (MI) can be defined as the user sending a command to a system through the imagination of a kinesthetic movement of his/her limbs [1]. MI is one of the most widely used cognitive tasks, which has advantages for the brain-computer interface (BCI) in both synchronous and asynchronous mode. MI-BCI has shown great potential to transform human-computer interaction, with applications ranging from rehabilitation [2], [3] to neuroprosthetics [4], providing greater convenience for individuals with disabilities [5]. Electroencephalography (EEG) represents the macroscopic activity on the surface of the human brain, making it an ideal choice for brain-computer interface (BCI) tasks [6].

The advent of deep learning (DL) has significantly enhanced the performance of MI-BCI. DL models can decode brain signals with high accuracy. However, decoding of MI-EEG signals is a challenging work because of low signal-to-noise ratio and high variability [7]. Most DL models are typically very complex, requiring a large number of free parameters (or degrees of freedom) to fit the data distribution. This necessitates large and high-quality EEG datasets. However, the size and quality of publicly available datasets are limited, and the acquisition of high-quality EEG data remains a labor intensive and costly endeavor. Therefore, data generation has become an effective way to address this problem [8].

The diffusion probabilistic model (DPM) [9] and generative adversarial networks (GAN) [10] represent state-of-the-art techniques in data generation, or as we might call it, data augmentation (DA) [11]. Although some studies have used GAN to generate EEG signals [12], [13], the challenges in training GAN, along with issues such as mode collapse in the output distribution [14], have not been fully resolved. Recent advances show the performance and effectiveness of DPM in both image generation and audio generation [15]. To our knowledge, there have been quite a few studies exploring the capabilities of DPM in generating multi-channel EEG signals [16]–[18].

The main contributions of this paper can be summarized as follows.

- We propose DiffEEGLossNet, a novel framework that leverages diffusion models enhanced with domain-specific knowledge from neuroscience to improve the generation of EEG signals.
- We incorporate ERD/ERS patterns as a loss function in the diffusion process, ensuring that the generated EEG signals retain essential physiological characteristics.
- Our experimental results indicate that the EEG signals generated by DiffEEGLossNet demonstrate higher quality and fidelity across multiple evaluation metrics. Furthermore, these generated samples effectively improve the performance of EEG classification models through robust data augmentation.
- We accelerate model inference through progressive distillation, improving the generation efficiency of the diffusion model, allowing DiffEEGLossNet to generate high-quality EEG signals with a single-step diffusion.

II. Related Works

Data augmentation (DA) has been widely used to improve the robustness and accuracy of DL by artificially increasing the number of training data. In the field of EEG, there are several traditional methods for DA [19].

In addition to traditional augmentation methods, deep generative models such as GAN [13] and Variational Autoencoder (VAE) [20] also achieved good performance for the EEG data augmentation task. Luo et al. enhanced EEG data for emotion recognition using VAEs [20] for feature generation and augmentation. Zhang et al. developed a novel augmentation framework, GANSER [13], which employed transformation operations to mask parts of EEG signals and forced the generator to synthesize potential EEG signals based on the unmasked parts to produce a wide variety of samples.

Beyond GANs and VAEs, Torma et al. proposed a framework based on DPMs named EEGWave [16], which generated high quality event-related potentials (ERP) and motor imagery (MI) signals. Soingern et al. [17] explored the use of a DPM based on WaveGrad [21] for motor imagery EEG data augmentation. Chen et al. proposed a Transformer-based DPM for generating multi-paradigm EEG signals and performing DA, hereinafter referred to as TransDiff [18].

However, these works using DPM to generate EEG signals rely solely on its powerful inference capabilities, and the generated signals cannot fully replicate the characteristics of brain cognition in certain scenarios. Therefore, in this work, we introduce prior knowledge from cognitive neuroscience to guide the diffusion process, enabling the diffusion model to generate high-quality EEG signals that better align with neuroscience.

III. Methodology

A. Preliminaries

Motor imagery (MI) tasks induce specific patterns in EEG signals known as event-related desynchronization (ERD) and event-related synchronization (ERS). These patterns typically occur in the alpha (8-12 Hz) and beta (13-30 Hz) frequency bands during and after motor imagery [22].

ERD often occurs contralaterally in the sensorimotor cortex, indicating active motor imagery processing, while ERS reflects a post-imagery synchronization rebound, marking the end of the cognitive process [23]. The degree of ERD/ERS is quantified by comparing the power during the motor imagery task to a baseline period using the following formula:

$$\text{ERD/ERS\%} = \left(\frac{X - M}{M}\right) \times 100 \tag{1}$$

where X represents the EEG signal power during the task, and M is the mean power during the baseline period. This quantification provides a robust metric for assessing the presence and intensity of motor imagery, which is crucial for applications in brain-computer interfaces (BCI) and neurofeedback training [24].

B. DiffEEGLossNet

We propose a novel diffusion model-based deep generative framework for generating high-quality EEG signals. Our architecture consists of three main components: the EEG signal preprocessing module, the diffusion inference module, and the motor imagery loss integration module. The structural illustration is provided in Fig.1.

1) EEG signal preprocessing module: The raw signals are band-pass-filtered using a zero-phase filter to remove noise and artifacts. After resampling, the original data labels are uniformly mapped and the signals are normalized to ensure consistency between samples. Furthermore, the class label y_{class} and the subject label $y_{subject}$ are extracted as embedding information for the residual features in the diffusion inference module, with y_{class} being equally important for the loss integration module. The signal data x_{input} obtained from the preprocessing module serves as input to the subsequent framework.

2) Diffusion inference module: The diffusion inference module aims to synthesize high-quality multi-channel EEG signals through a powerful diffusion model. The continuous-time diffusion model [25] better fits the temporal characteristics of the EEG samples.

After preprocessing, the distribution of the training data set is represented as $p(x)$, where $x \in \mathbb{R}^{E\times L}$ denotes the EEG data with electrodes E and the sequence length L. In the continuous-time diffusion framework, a time variable $t \in [0, 1]$ is introduced and the latent variable z_t has the same shape as the training data samples. The forward diffusion process can be described as:

$$q(z_t|x) = \mathcal{N}(z_t; \alpha_t x, \sigma_t^2 \mathbf{I}) \tag{2}$$

where α_t and σ_t^2 are smooth, differentiable, positive scalar-valued functions. The logarithmic signal-to-noise ratio $\lambda_t = \log(\alpha_t^2/\sigma_t^2)$ decreases monotonically as $t \to 1$. For any $0 \leq s \leq t \leq 1$, the reverse process is modeled as:

$$q(z_s|z_t, x) = \mathcal{N}(z_s; \tilde{\mu}_{s|t}(z_t, x), \tilde{\sigma}_{s|t}^2 \mathbf{I}) \tag{3}$$

where $\tilde{\mu}_{s|t}(z_t, x) = e^{\lambda_t - \lambda_s} \frac{\alpha_s}{\alpha_t} z_t + (1 - e^{\lambda_t - \lambda_s})\alpha_s x$.

During inference, the neural network $\tilde{x}_\theta(z_t, \lambda_t)$ predicts the data x, and the reverse process iteratively denoises the latent variable z_t using the formula for z_s from Salimans et al. [26], until $t = 0$. Training involves minimizing the following objective function:

$$L_{\text{diff}}(\theta) = \min_\theta L(\theta) = \mathbb{E}_{\epsilon,t}\left[\omega(\lambda_t)\|x - \tilde{x}_\theta(z_t, \lambda_t)\|_2^2\right] \tag{4}$$

where $\omega(\lambda_t)$ is chosen based on prior work [26]. Furthermore, our approach integrates the ERD/ERS patterns into the loss function to align the generated EEG signals with motor imagery research. The details of this integration are detailed in the motor imagery loss module.

In terms of implementation, the diffusion inference module is based on the work of Torma et al. in EEGWave [16], and can be expressed as $\tilde{x}_\theta : \mathbb{R}^{E\times L} \times \mathbb{R} \to \mathbb{R}^{E\times L}$. It employs non-causal bidirectional dilated convolutions, which help maintain global context and consistency throughout the epochs. We use 2D convolution to build it on the residual layers of DiffWave [15], and λ_t is embedded in the residual features through two global linear layers and one local residual layer. The class and subject information extracted from the preprocessing module is also embedded in the same manner. The network parameters and specific implementation details are provided in Section I-A of the **supplemental material**.

3) Motor imagery loss integration module: The motor imagery loss integration module is a critical part of DiffEEGLossNet. During the training process of the diffusion inference module, we divide the design of the loss function into two parts. The first part is the original variational lower bound (ELBO) of the diffusion model, denoted L_{diff} in Equation (4). The second part is a specially designed loss for motor imagery tasks. Based on the studies mentioned in the preliminaries section, we quantified the contralateral activation ERD/ERS

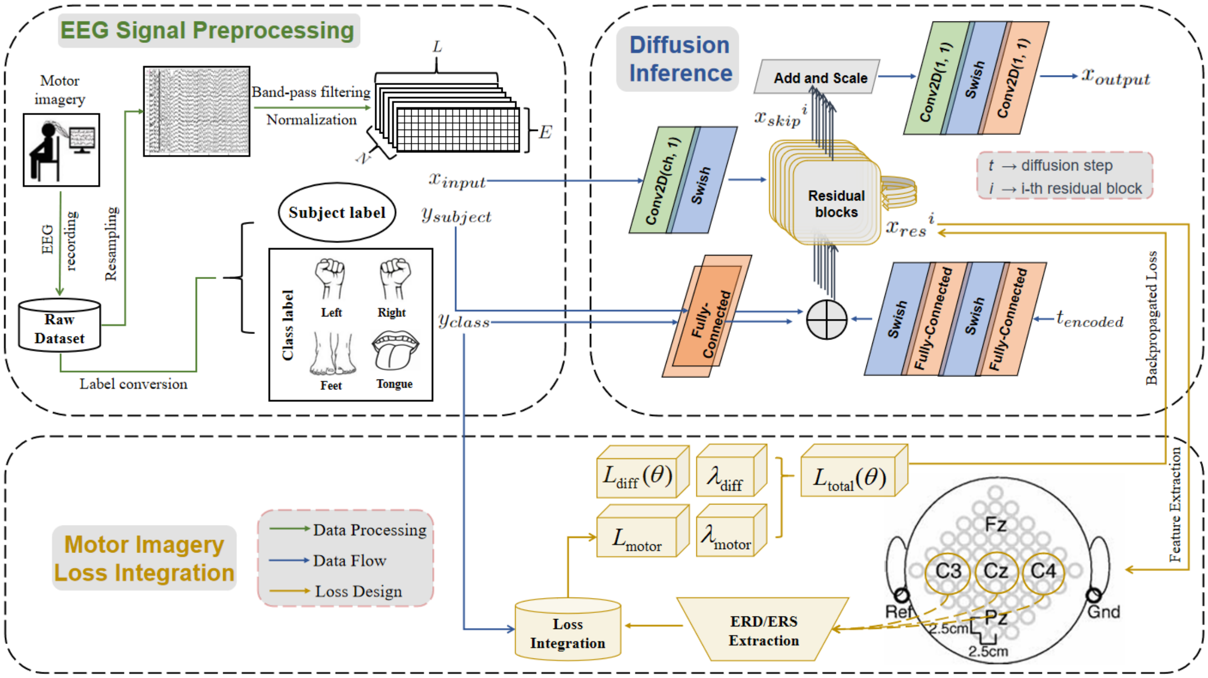

Fig. 1. The architecture of DiffEEGLossNet: A framework for motor imagery EEG signal generation integrating diffusion inference and ERD/ERS-informed loss. The diagram consists of three main modules: EEG signal preprocessing, diffusion inference with residual blocks, and motor imagery loss integration to enhance neuroscientific fidelity and improve classification performance.

patterns that appear during motor imagery tasks and designed a new loss term L_{motor}.

The extraction process begins by mapping the EEG signals to a feature space where the ERD/ERS patterns are identifiable. The ERD/ERS% is calculated using Equation (1). Based on this, we define L_{motor} as follows:

$$L_{\text{motor}} = \lambda_{\text{ERDS}} \cdot \mathbb{E}\left[\|\text{ERDS}_{\text{real}} - \text{ERDS}_{\text{generated}}\|_2^2\right] \quad (5)$$

where λ_{ERDS} is a hyperparameter that balances the weight of the ERD/ERS loss. The purpose of the loss function is to minimize the mean squared error between the real and generated ERD/ERS values, ensuring that the model can generate EEG signals that conform to the actual physiological characteristics.

The final total loss function is:

$$L_{\text{total}}(\theta) = \lambda_{\text{diff}} \cdot L_{\text{diff}}(\theta) + \lambda_{\text{motor}} \cdot L_{\text{motor}} \quad (6)$$

where λ_{diff} and λ_{motor} are hyperparameters that balance the weights of different loss terms. By adjusting these weights, we can control the model's focus on different aspects of the loss during training. This loss function ensures that the synthesized EEG signals not only match the original signals in the time domain, but also strengthen the ERD/ERS features crucial for motor imagery tasks. By incorporating ERD/ERS features into the loss function, DiffEEGLossNet achieves a more accurate and plausible generation of EEG signals.

C. Progressive Distillation

Since diffusion models require multiple iterations during the sampling process to generate data, we apply progressive distillation [26] inference to reduce the number of iterations. Progressive distillation uses a teacher-student setup, where the student model approximates the teacher with half the sampling steps. The teacher is trained as a continuous-time diffusion model. Then, both the teacher and the student are assigned discrete sampling steps, T_t and $T_s = T_t/2$. The distillation is iterative, and each iteration starts with initializing the student model with the teacher's parameters. The student denoises a latent variable z_t using one reverse DDPM step, while the teacher uses two DDPM steps. The process continues with halved T_s until convergence. We eventually obtain a version of DiffEEGLossNet that requires only a single step (DiffEEGLossNet $1\times$), which will be evaluated together in subsequent experiments.

IV. Experiments

A. Experimental Setting

This section provides an overview of the datasets, comparison methods, evaluation metrics, and classification models used for data augmentation. The hyperparameter settings during training are presented in Section I-B of the **supplemental material**.

1) Datasets: The BCI Competition IV dataset 2a [27] and 2b [28] are chosen for our generative experiments. The 2a dataset contains motor imagery signals from 9 participants across four classes: left hand, right hand, feet, and tongue movements. Data are recorded from 22 EEG channels following the 10-20 standard system. We apply a zero-phase filter to band-pass filter signals between 4 and 38 Hz and down-sample them to 128 Hz. Additionally, we extract data within [0.5, 4] seconds after the onset of each trial's cue and normalize them to the range [-1, 1] using channel-wise mean subtraction and deviation division. The 2b dataset records data from three EEG channels (C3, Cz, and C4) for only two motor imagery tasks: left hand and right hand. The preprocessing steps are consistent with those used for the 2a dataset.

2) Comparison methods: Four EEG signal generation models were selected for performance comparison, including Wasserstein GAN (WGAN) [29], EEGWave [16], WaveGrad [17], and TransDiff [18], all based on deep generative models GAN and DPM.

3) Evaluation metrics: For the quantitative evaluation of the quality of the generated sample, we measure the commonly used Initial Score (IS) [30], Frechet Inception Distance (FID) [31] and the spatial FID (sFID) [32] metrics. Furthermore, Precision and Recall have been shown to not only measure classification model performance but also reflect the quality of samples generated by generative models [33].

4) EEG classification models: To further validate the effectiveness of the proposed method, we mix the generated data with the original EEG dataset and perform classification experiments using two commonly used EEG classification models, EEGNet [34] and EEGProgress [35]. This approach is used to verify that our generated EEG signals can effectively improve the classification performance of the models.

B. Signal Generation and Results

The current experiment aims to generate high-quality EEG data from simple Gaussian noise distribution samples.

1) Quantitative evaluation: For the BCI Competition IV dataset 2a [27] and 2b [28], the quantitative results are given in Table I. We use the training subset as the evaluation benchmark and provide the results of the test subset as a reference to comprehensively evaluate the performance of the proposed method and the comparison methods mentioned in Section IV-A2 across various metrics.

As shown in Table I, our method achieves state-of-the-art performance across most metrics. Specifically, IS [30] is primarily used to measure the diversity and quality of generated samples. A higher IS score indicates that the generated samples are not only of high quality but also rich in diversity. The proposed method achieves an IS score of 1.37 on the 2a dataset, which is very close to the training set's 1.40 and even higher than the test set's 1.36, demonstrating that the generated samples are both high-quality and diverse. FID [31] measures the similarity between generated and real samples in the feature space. A lower FID score indicates that the generated samples are closer to the real samples in overall distribution. The proposed method achieves the second-best FID scores of 3.47 and 4.17 on the two datasets, while the distilled one-step model achieves the best FID score on the 2b dataset. sFID [32] measures the distance between generated and real samples in the feature space while emphasizing the spatial characteristics of EEG data, making it a complementary metric to FID. The proposed method also achieves second-best and best sFID scores on the two datasets, respectively.

Furthermore, we use Precision and Recall [33] to further evaluate the quality of the generated samples. Precision and Recall, traditionally used for classification tasks, have been adapted to assess generative models by Kynkäänniemi et al. [33]. Our method achieves the best results in Precision and Recall, demonstrating that the generated samples not only align closely with the real data in terms of quality but also effectively cover the diversity of the real data. This comprehensive evaluation shows that DiffEEGLossNet is capable of generating high-quality EEG signals that are both diverse and well-aligned with the original data. Furthermore, our distilled DiffEEGLossNet $1\times$ model also achieves competitive results on these metrics, with IS and Recall scores ranking second-best among all models. Additionally, as can be seen in the right half of Table I, both DiffEEGLossNet and DiffEEGLossNet $1\times$ perform exceptionally well on the 2b dataset.

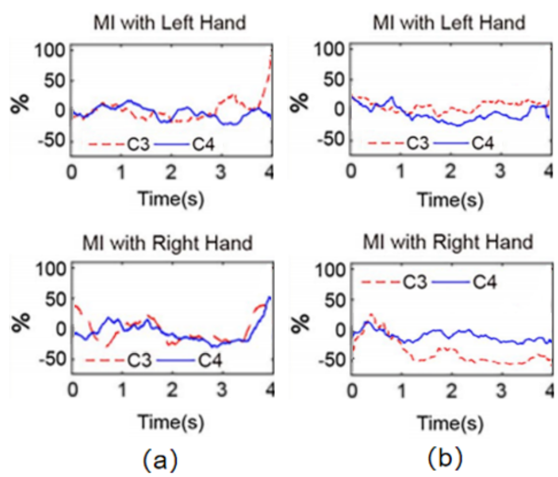

Fig. 2. ERD/ERS curves for Subject 3 during left and right hand motor imagery tasks. The C3 channel is shown as a red dashed line, and the C4 channel as a blue solid line. (a) shows the results of EEGWave [16], while (b) shows the results of DiffEEGLossNet. The ERD/ERS in this figure is calculated using Equation (1).

2) Contralateral advantage of ERD/ERS patterns: As we know, the ERD/ERS maps generated from data during left- and right-hand MI tasks always reveal a clear pattern of ERD in the contralateral motor areas, which is contralateral dominance. In Fig.2, we provide the comparisons of the ERD/ERS of the EEGWave [16] generation data and the proposed DiffEEGLossNet. According to the signal power in channels C3 and C4, the signal power in C3 was commonly higher than in C4 during left-handed MI and vice versa. In addition, we could also observe that, although there exists a pattern of ERD in both cases, the ERD of the proposed method is much more obvious, demonstrating the effectiveness of the proposed method.

In addition, we also present the sFID scores for each subject to demonstrate that the EEG data generated by the proposed method exhibit subject specificity, indicating that the model can learn and replicate subject-specific features.

TABLE I
COMPREHENSIVE ASSESSMENT OF THE QUALITY AND DIVERSITY OF GENERATED SAMPLES ON THE BCI COMPETITION IV DATASET 2A AND 2B USING IS, FID, SFID, PRECISION, AND RECALL METRICS.

	BCI Competition IV dataset 2a					**BCI Competition IV dataset 2b**				
	***IS*↑**	***FID*↓**	***sFID*↓**	***Precision*↑**	***Recall*↑**	***IS*↑**	***FID*↓**	***sFID*↓**	***Precision*↑**	***Recall*↑**
Training set	1.40	0.00	0.00	1.00	1.00	1.38	0.00	0.00	1.00	1.00
Test set	1.36	0.18	80.44	0.90	0.92	1.30	0.12	63.07	0.91	0.90
RWGAN	1.29	7.23	142.87	0.47	0.41	1.19	12.43	97.49	0.64	0.54
WaveGrad	1.17	6.59	147.79	0.76	0.47	1.20	7.64	107.42	0.73	0.56
TransDiff	1.14	5.44	174.83	0.83	0.44	1.14	6.79	96.44	0.64	0.52
EEGWave	1.19	**3.07**	**51.92**	0.97	0.58	1.19	4.31	57.42	0.89	0.64
DiffEEGLossNet(Ours)	**1.37**	3.47	76.44	**0.97**	**0.62**	**1.23**	4.17	62.77	**0.95**	**0.69**
DiffEEGLossNet $1\times$(Ours)	1.32	4.55	112.16	0.87	0.59	1.21	**3.92**	**54.85**	0.77	0.67

Detailed discussions and visualized results are provided in Section II-A of the **supplemental material**.

C. Classification Experiments with Data Augmentation

For generative tasks, an important metric of generative model quality is whether generated samples can effectively improve the performance of classification models. Therefore, we conduct DA experiments on the BCI Competition IV 2a and 2b datasets using the generated samples. Specifically, we use EEGNet [34] and EEGProgress [35] as classification models. The original BCI Competition IV 2a and 2b datasets are split into training, validation, and test subsets with a ratio of 70%, 15%, 15%. We generate the same number of samples as in the training subsets using DiffEEGLossNet. DA experiments on EEGNet and EEGProgress are conducted using the following two training strategies:

- Mixed training: We double the size of the training subset by mixing the same amount of signals generated as the number in the original subset.
- Pre-training: Pre-train EEGNet and EEGProgress on the generated data and then fine-tune the initialized models with the real EEG data.

In both cases, we stop the training only if overfitting is detected via monitoring the validation loss.

The baseline model is established by training it without any DA (i.e., 0% synthetic data). Similarly, the data generated by EEGWave [16], WaveGrad [17], and TransDiff [18] are used for mixed augmentation as comparisons to the two proposed models. The classification results of the EEGNet [34] and EEGProgress [35] models on the two datasets are shown in Table II. It can be observed that both the mixed training and pre-training strategies improve the classification accuracy of the models, with DiffEEGLossNet achieving the best accuracy. These results demonstrate that the EEG data generated by our proposed method can effectively augment the original dataset and improve the performance of the classification models.

V. CONCLUSION

In this paper, we address the limitation of the availability of EEG data by proposing a novel generative framework based on diffusion models. Specifically, DiffEEGLossNet integrates features specific to motor imagery and incorporates them into the diffusion process through the design of a unique loss function that accounts for ERD/ERS patterns. This loss function effectively enhances the fidelity of the generated EEG signals, ensuring that they better reflect the neural dynamics underlying motor imagery tasks. This integration represents a significant step towards producing EEG signals that are more aligned with neuroscience, which is particularly valuable for data augmentation in EEG-based applications.

Experimental results show that the proposed method significantly improves the quality of generated EEG data. Both DiffEEGLossNet and its distilled and accelerated single-step model, DiffEEGLossNet $1\times$, demonstrate strong competitiveness on traditional metrics for evaluating the quality of generated samples, consistently outperforming existing models. Furthermore, the method achieves effective data augmentation results through mixed training and pre-training with the generated samples and the original dataset, thereby improving the performance of EEG classification models. Future work could explore applying this method to EEG data from other paradigms, such as emotion recognition and seizure detection, to further investigate the potential of combining generative models with specific neuroscientific knowledge.

REFERENCES

[1] Amardeep Singh, Ali Abdul Hussain, Sunil Lal, and Hans W Guesgen, "A comprehensive review on critical issues and possible solutions of motor imagery based electroencephalography brain-computer interface," *Sensors*, vol. 21, no. 6, pp. 2173, 2021.

[2] Ravikiran Mane, Tushar Chouhan, and Cuntai Guan, "Bci for stroke rehabilitation: motor and beyond," *Journal of neural engineering*, vol. 17, no. 4, pp. 041001, 2020.

[3] Neethu Robinson, Ravikiran Mane, Tushar Chouhan, and Cuntai Guan, "Emerging trends in bci-robotics for motor control and rehabilitation," *Current Opinion in Biomedical Engineering*, vol. 20, pp. 100354, 2021.

[4] Javier Pascual, Francisco Velasco-Álvarez, Klaus-Robert Müller, and Carmen Vidaurre, "First study towards linear control of an upper-limb neuroprosthesis with an eeg-based brain-computer interface," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2012, pp. 3269–3273.

[5] Jonathan R. Wolpaw, "Chapter 6 - brain–computer interfaces," in *Neurological Rehabilitation*, Michael P. Barnes and David C. Good, Eds., vol. 110 of *Handbook of Clinical Neurology*, pp. 67–74. Elsevier, 2013.

TABLE II
PERFORMANCE COMPARISON OF EEGNET AND EEGPROGRESS ON THE BCI COMPETITION IV 2A AND 2B DATASETS UNDER DIFFERENT TRAINING CONDITIONS WITH DATA AUGMENTATION (ACCURACY, %).

		EEGNet		**EEGProgress**	
Training Condition	**Method**	**BCI-IV-2a**	**BCI-IV-2b**	**BCI-IV-2a**	**BCI-IV-2b**
Original datasets	**Baseline**	46.50	72.17	73.67	76.64
Mixed training	EEGWave	<u>50.49</u>	73.14	<u>75.34</u>	75.46
	WaveGrad	49.67	73.89	74.32	74.22
	TransDiff	46.72	66.42	72.98	69.73
	DiffEEGLossNet(Ours)	**50.72**	**74.58**	**76.28**	**77.41**
	DiffEEGLossNet 1×(Ours)	49.92	<u>74.26</u>	74.69	<u>76.89</u>
Pre-training	EEGWave	49.13	70.42	71.29	<u>77.76</u>
	WaveGrad	48.27	64.32	70.21	73.64
	TransDiff	47.92	67.37	70.46	71.33
	DiffEEGLossNet(Ours)	**51.32**	**73.43**	**75.92**	77.25
	DiffEEGLossNet 1×(Ours)	<u>50.41</u>	<u>71.64</u>	<u>73.48</u>	**78.22**

[6] Tianhua Xu, Sheng-hua Zhong, and Zhijiao Xiao, "Protecting intellectual property of eeg-based model with watermarking," in *2023 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2023, pp. 37–42.

[7] Shuaibin Wang, Jinglong Wu, Deyu Zhang, Dingjie Suo, and Tianyi Yan, "An algorithm for motor imagery classification based on transfer learning and feature fusion," in *2022 16th ICME International Conference on Complex Medical Engineering (CME)*. IEEE, 2022, pp. 246–251.

[8] Elnaz Lashgari, Dehua Liang, and Uri Maoz, "Data augmentation for deep-learning-based electroencephalography," *Journal of Neuroscience Methods*, vol. 346, pp. 108885, 2020.

[9] Jonathan Ho, Ajay Jain, and Pieter Abbeel, "Denoising diffusion probabilistic models," *Advances in neural information processing systems*, vol. 33, pp. 6840–6851, 2020.

[10] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.

[11] Prafulla Dhariwal and Alexander Nichol, "Diffusion models beat gans on image synthesis," *Advances in neural information processing systems*, vol. 34, pp. 8780–8794, 2021.

[12] Yun Luo, Li-Zhen Zhu, and Bao-Liang Lu, "A gan-based data augmentation method for multimodal emotion recognition," in *Advances in Neural Networks–ISNN 2019: 16th International Symposium on Neural Networks, ISNN 2019, Moscow, Russia, July 10–12, 2019, Proceedings, Part I 16*. Springer, 2019, pp. 141–150.

[13] Zhi Zhang, Sheng-hua Zhong, and Yan Liu, "Ganser: A self-supervised data augmentation framework for eeg-based emotion recognition," *IEEE Transactions on Affective Computing*, 2022.

[14] Suman Ravuri and Oriol Vinyals, "Classification accuracy score for conditional generative models," *Advances in neural information processing systems*, vol. 32, 2019.

[15] Zhifeng Kong, Wei Ping, Jiaji Huang, Kexin Zhao, and Bryan Catanzaro, "Diffwave: A versatile diffusion model for audio synthesis," *arXiv preprint arXiv:2009.09761*, 2020.

[16] Szabolcs Torma and Luca Szegletes, "Brain signal generation and data augmentation with a single-step diffusion probabilistic model," 2023.

[17] Nutapol Soingern, Akraradet Sinsamersuk, Itthi Chatnuntawech, and Chaklam Silpasuwanchai, "Data augmentation for eeg motor imagery classification using diffusion model," in *International Conference on Data Science and Artificial Intelligence*. Springer, 2023, pp. 111–126.

[18] Mingzhi Chen, Yiyu Gui, Yuqi Su, Yuesheng Zhu, Guibo Luo, and Yuchao Yang, "Improving eeg classification through randomly reassembling original and generated data with transformer-based diffusion models," *arXiv preprint arXiv:2407.20253*, 2024.

[19] Zhiwen Zhang, Feng Duan, Jordi Sole-Casals, Josep Dinares-Ferran, Andrzej Cichocki, Zhenglu Yang, and Zhe Sun, "A novel deep learning approach with data augmentation to classify motor imagery signals," *IEEE Access*, vol. 7, pp. 15945–15954, 2019.

[20] Yun Luo, Li-Zhen Zhu, Zi-Yu Wan, and Bao-Liang Lu, "Data augmentation for enhancing eeg-based emotion recognition with deep generative models," *Journal of Neural Engineering*, vol. 17, no. 5, pp. 056021, 2020.

[21] Nanxin Chen, Yu Zhang, Heiga Zen, Ron J Weiss, Mohammad Norouzi, and William Chan, "Wavegrad: Estimating gradients for waveform generation," *arXiv preprint arXiv:2009.00713*, 2020.

[22] Gert Pfurtscheller and FH Lopes Da Silva, "Event-related eeg/meg synchronization and desynchronization: basic principles," *Clinical neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999.

[23] Gert Pfurtscheller and Christa Neuper, "Motor imagery and direct brain-computer communication," *Proceedings of the IEEE*, vol. 89, no. 7, pp. 1123–1134, 2001.

[24] Christa Neuper, Gernot R Müller-Putz, Reinhold Scherer, and Gert Pfurtscheller, "Motor imagery and eeg-based control of spelling devices and neuroprostheses," *Progress in brain research*, vol. 159, pp. 393–409, 2006.

[25] Diederik Kingma, Tim Salimans, Ben Poole, and Jonathan Ho, "Variational diffusion models," *Advances in neural information processing systems*, vol. 34, pp. 21696–21707, 2021.

[26] Tim Salimans and Jonathan Ho, "Progressive distillation for fast sampling of diffusion models," *arXiv preprint arXiv:2202.00512*, 2022.

[27] Clemens Brunner, Robert Leeb, Gernot Müller-Putz, Alois Schlögl, and Gert Pfurtscheller, "Bci competition 2008–graz data set a," *Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces), Graz University of Technology*, vol. 16, pp. 1–6, 2008.

[28] R Leeb, C Brunner, G Müller-Putz, A Schlögl, and GJGUOT Pfurtscheller, "Bci competition 2008–graz data set b," *Graz University of Technology, Austria*, vol. 16, pp. 1–6, 2008.

[29] Sharaj Panwar, Paul Rad, Tzyy-Ping Jung, and Yufei Huang, "Modeling eeg data distribution with a wasserstein generative adversarial network to predict rsvp events," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 8, pp. 1720–1730, 2020.

[30] Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen, "Improved techniques for training gans," *Advances in neural information processing systems*, vol. 29, 2016.

[31] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter, "Gans trained by a two time-scale update rule converge to a local nash equilibrium," *Advances in neural information processing systems*, vol. 30, 2017.

[32] Charlie Nash, Jacob Menick, Sander Dieleman, and Peter W Battaglia, "Generating images with sparse representations," *arXiv preprint arXiv:2103.03841*, 2021.

[33] Tuomas Kynkäänniemi, Tero Karras, Samuli Laine, Jaakko Lehtinen, and Timo Aila, "Improved precision and recall metric for assessing generative models," *Advances in neural information processing systems*, vol. 32, 2019.

[34] Vernon J Lawhern, Amelia J Solon, Nicholas R Waytowich, Stephen M Gordon, Chou P Hung, and Brent J Lance, "Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces," *Journal of neural engineering*, vol. 15, no. 5, pp. 056013, 2018.

[35] Zhige Chen, Rui Yang, Mengjie Huang, Fumin Li, Guoping Lu, and Zidong Wang, "Eegprogress: A fast and lightweight progressive con-

volution architecture for eeg classification," *Computers in Biology and Medicine*, vol. 169, pp. 107901, 2024.