FusionGen: Feature Fusion-Based Few-Shot EEG Data Generation

Yuheng Chen[1†] **Dingkun Liu**[1,2†] **Xinyao Yang**[1] **Xinping Xu**[1] **Baicheng Chen**[1] **Dongrui Wu**[1,2*]

[1]Huazhong University of Science and Technology, Wuhan, 430074, China
[2]Zhongguancun Academy, Beijing, 100094, China

`chenyuheng@hust.edu.cn`   `liudingkun@hust.edu.cn`   `drwu09@gmail.com`

Abstract

Brain-computer interfaces (BCIs) provide potential for applications ranging from medical rehabilitation to cognitive state assessment by establishing direct communication pathways between the brain and external devices via electroencephalography (EEG). However, EEG-based BCIs are severely constrained by data scarcity and significant inter-subject variability, which hinder the generalization and applicability of EEG decoding models in practical settings. To address these challenges, we propose FusionGen, a novel EEG data generation framework based on disentangled representation learning and feature fusion. By integrating features across trials through a feature matching fusion module and combining them with a lightweight feature extraction and reconstruction pipeline, FusionGen ensures both data diversity and trainability under limited data constraints. Extensive experiments on multiple publicly available EEG datasets demonstrate that FusionGen significantly outperforms existing augmentation techniques, yielding notable improvements in classification accuracy.

1. Introduction

Brain-computer interfaces (BCIs) establish direct communication pathways between the human brain and external devices, holding great promise for medical rehabilitation [5, 31], intelligent control systems [22], and sleep stage detection [1]. Figure 1 shows the pipeline of closed-loop BCIs. Among various neuroimaging modalities, electroencephalography (EEG) is particularly prominent due to its high temporal resolution, non-invasive acquisition, and cost-effectiveness. Despite these advantages, EEG-based BCIs confront two primary challenges: data scarcity and
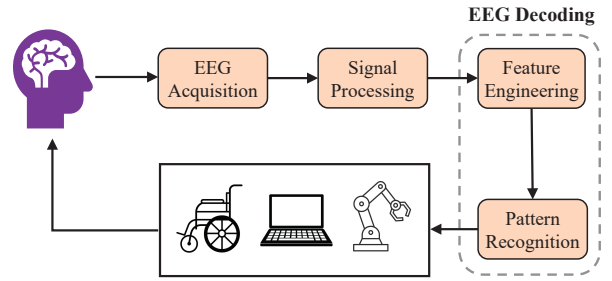

Figure 1. Closed-loop brain-computer interface system.

significant inter-subject variability. These issues considerably hinder the generalization and applicability of EEG decoding models in practical, real-world settings.

Representation learning [14, 18, 19, 30] and data generation has emerged as a vital strategy. Representation learning techniques can extract high-order latent features to disentangle underlying neural factors and guide realistic EEG data generation. Conventional data augmentation methods are typically classified into two categories: transformation-based methods and generative model-based methods. Transformation-based approaches employ signal manipulation techniques such as noise addition, amplitude scaling, and frequency shifting. These methods are computationally efficient but often yield limited diversity and physiological realism in the augmented data. Conversely, generative model-based approaches, including generative adversarial networks (GANs) [1], variational autoencoders (VAEs) [11], and diffusion models [3], can produce richer and more diverse synthetic data. However, these models typically require large training datasets and significant computational resources, limiting their effectiveness in few-shot scenarios common in BCIs.

Moreover, EEG data inherently possess substantial variability across subjects due to physiological and neuroanatomical discrepancies. Traditional data augmentation methods, predominantly tailored for single-subject con-
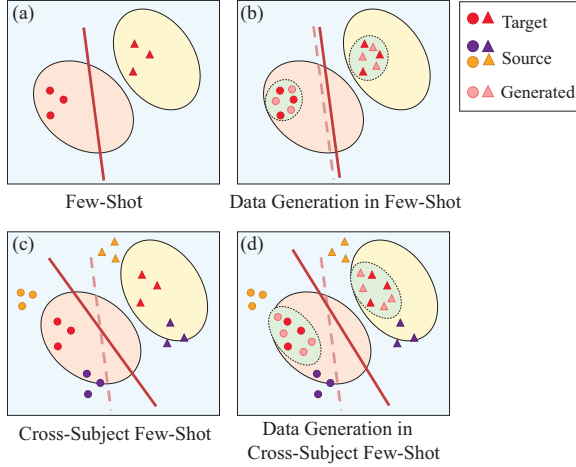

Figure 2. Scenarios of data distributions and generation strategies.

texts, exhibit sharply reduced efficacy when deployed in cross-subject scenarios. Consequently, the development of augmentation strategies specifically designed for cross-subject, few-shot EEG scenarios is critical to advancing practical BCI applications.

As shown in Figure 2, in (a), the few-shot setting with few target samples (red) produces a limited decision boundary. In (b), conventional within-domain augmentation adds synthetic target samples, expanding and smoothing that boundary. In (c), the cross-subject few-shot scenario shows source samples lying outside the target distribution, making transfer ineffective. In (d), data generation approach in cross-subject few-shot scenario, bridging the distribution gap and yielding a refined decision boundary.

In this paper, we propose FusionGen, a novel feature fusion-based EEG data generation framework that efficiently synthesizes diverse EEG signals from scarce samples. By integrating features across samples through a dedicated fusion module and combining them with a lightweight feature extraction and reconstruction pipeline, our architecture ensures both data diversity and trainability under severe data constraints. Comprehensive experiments on multiple publicly available EEG datasets demonstrate that our approach significantly outperforms existing augmentation techniques, yielding notable improvements in classification accuracy.

1. We propose FusionGen, a few-shot EEG data generation framework that enhances generalization and scalability in brain–computer interface applications.
2. We introduce a feature matching fusion module that integrates cross-sample features in the latent representation space and reconstructs high-fidelity EEG signals from these fused embeddings.
3. We validate FusionGen on multiple EEG datasets on MI and SSVEP paradigms, showing consistent accuracy im-

provements in few-shot scenarios.

2. Related Work

2.1. Transformation-Based Approaches

Wei *et al.* [25] proposed noise addition, a simple yet effective method to simulate signal perturbations for EEG augmentation. Xu *et al.* [34] proposed amplitude scaling, a novel approach that linearly scales EEG signal amplitudes to diversify training data. Wang *et al.* [27] proposed channel reflection, a spatial-domain method exploiting hemispheric symmetry to generate new EEG samples. Wang *et al.* [28] proposed DWTaug, a discrete wavelet transform–based augmentation strategy that recombines subband signals across samples. Wang *et al.* [28] also proposed HHTaug, a Hilbert–Huang transform–based method for cross-sample signal decomposition and reconstruction.

2.2. Generative Model-Based Methods

Zhang *et al.* [33] proposed DCGAN, a deep convolutional GAN for synthesizing realistic EEG signals in motor imagery tasks. Luo and Lu [32] proposed CWGAN, a conditional Wasserstein GAN to enhance classification accuracy through label-conditioned EEG synthesis. Komolovaite *et al.* [12] performed VAE framework for EEG signal augmentation, learning latent distributions to generate clinically relevant samples. Huang *et al.* [8] adopted diffusion models, modeling EEG denoising via forward-reverse diffusion processes to produce high-fidelity synthetic data.

2.3. Few-Shot Generative Models

Wang *et al.* [4] proposed AGE, an adaptive latent-space feature transformation model for few-shot generative synthesis. Singh *et al.* [35] proposed LSO, a latent space optimization approach modeling category-specific distributions under limited data. Kim *et al.* [13] proposed HAE, a hyperbolic autoencoder capturing hierarchical semantics to improve few-shot generation diversity. Li *et al.* [6] proposed LoFGAN, a local feature fusion GAN that enhances sample diversity by matching and fusing localized features. Chen *et al.* [38] proposed F2DGAN, a feature-distribution matching GAN combined with variational feature learning for diverse EEG sample synthesis.

In summary, while transformation-based methods are limited by the diversity and quantity of generated data, and generative model-based approaches require large-scale training datasets, few-shot generative models offer a potential solution for EEG augmentation in BCI applications with limited data. Leveraging limited EEG data for generation holds great promise in BCIs.
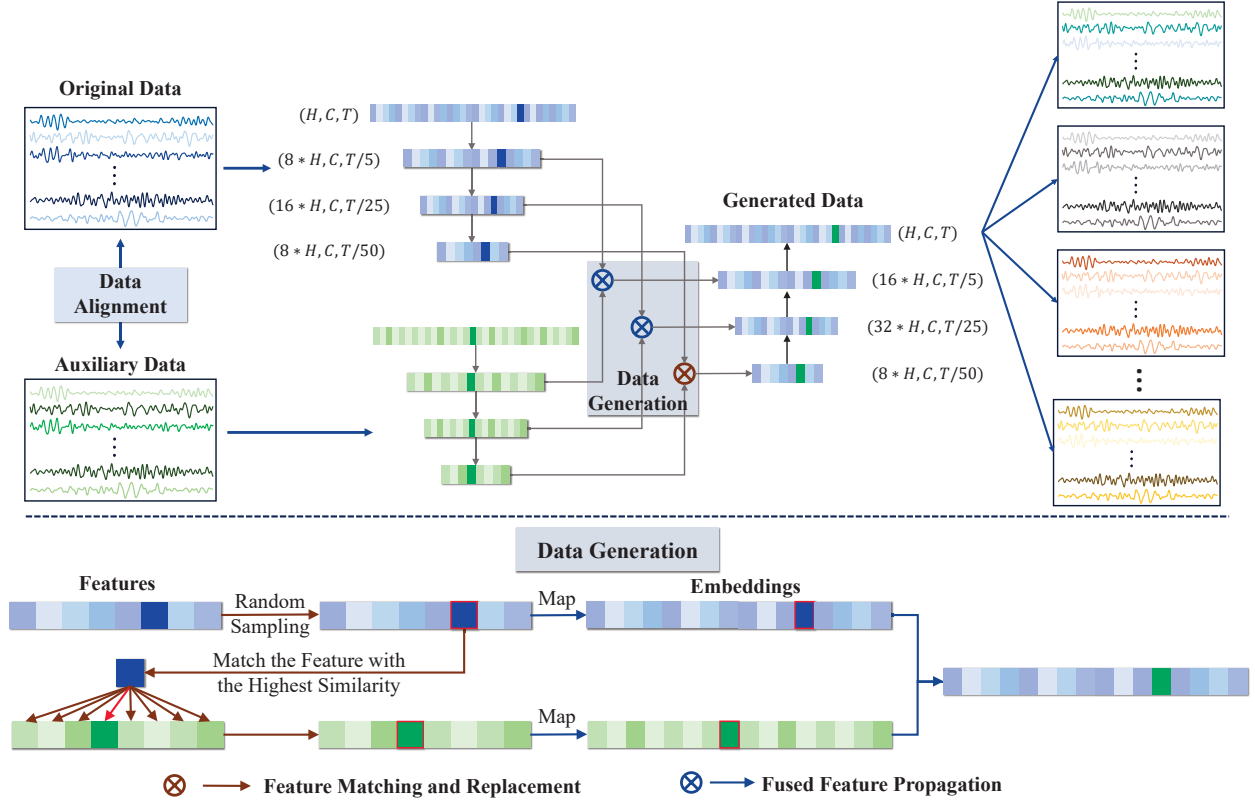
Figure 3. Architecture of proposed FusionGen. Raw and auxiliary trials are first aligned, then encoded into multi-scale features; randomly sampled target features are matched and replaced with source features, propagated through the decoder via skip connections, and finally decoded to produce realistic generated EEG trials.

3. Method

3.1. Problem Definition

Given m source subjects S_m each providing n_m limited labeled trials $\{X_m^i, y_m^i)\}_{i=1}^{n_m}$, where $X_m^i \in \mathbb{R}^{C \times T}$ and $y_m^i \in 1, \ldots, K$, and a target subject S_t providing n_t labeled trials $\{(X_t^i, y_t^i)\}_{i=1}^{n_t}$ and n_u unlabeled trials $\{X_u^i\}_{i=1}^{n_u}$, where $X_t^i, X_u^i \in \mathbb{R}^{C \times T}$, our goal is to learn a generator $\mathcal{G}$ that synthesizes n_a additional labeled trials $\{(\tilde{X}^j, y'^j)\}_{j=1}^{n_a}$ with $\tilde{X}^j \in \mathbb{R}^{C \times T}$ and $\tilde{y}^j \in 1, \ldots, K$, so that the augmented training set $(X_m^i, y_m^i) \cup (X_t^i, y_t^i) \cup (\tilde{X}^j, \tilde{y}^j)$ better approximates the true EEG distribution and improves downstream classification performance on the unlabeled trials X_u^i. Table 1 summarizes the main notations used throughout this paper.

3.2. Input Distribution Alignment

EEG signals are inherently non-stationary and exhibit substantial variability across sessions and subjects. To mitigate these effects and improve consistency, we adopt Euclidean alignment (EA) [7, 29], a simple yet effective whitening-

Table 1. Notations used in this paper.

Notation	Description
m	Number of source subjects.
C	Number of EEG channels.
T / T'	Original / compressed time samples.
k	Number of replaced features.
S_m / S_t	Source / target subject.
$\{(X_m^i, y_m^i)\} / \{(X_t^i, y_t^i)\}$	Labeled trials of source / target domain.
$\bar{X}_i$	Aligned EEG trial.
$\tilde{X}_i$	Generated EEG trial.
F_m^ℓ / F_t^ℓ	Feature map of source / target at layer ℓ.
f_m^p / f_t^q	Source / target latent embeddings.
$\hat{F}_t$	Fused feature map after matching and replacement.
$\bar{R}$	Mean covariance matrix of all trials.
$(c_m, x_m) / (c_t, x_t)$	Replaced feature indices of source / target.
$\mathcal{G}$	EEG trials generator.
$\mathcal{L}_{\mathrm{rec}}$	Reconstruction loss.

based preprocessing step. Given a recording session with n trials $\{X_i\}_{i=1}^n$, where each $X_i \in \mathbb{R}^{C \times T}$, EA first computes

the mean covariance matrix of all trials:

$$\bar{R} = \frac{1}{n}\sum_{i=1}^{n} X_i X_i^\top. \tag{1}$$

Next, each trial is whitened by

$$\bar{X}_i = \bar{R}^{-\frac{1}{2}} X_i. \tag{2}$$

Since

$$\frac{1}{n}\sum_{i=1}^{n} \bar{X}_i \bar{X}_i^\top = \bar{R}^{-\frac{1}{2}}\left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i^\top\right)\bar{R}^{-\frac{1}{2}} = I, \tag{3}$$

the second-order statistics of the aligned trials become identity, effectively reducing covariance shifts. The aligned set $\{\bar{X}_i\}_{i=1}^{n}$ then replaces the raw trials in all subsequent processing steps. By performing input distribution alignment, we substantially suppress marginal distribution differences among subjects, facilitating downstream feature extraction and classification.

3.3. Feature Matching Fusion

As shown in Figure 3, we employ a U-Net-shaped encoder–decoder network and perform pairwise feature fusion between source and target feature representations in the latent space at each latent space, injecting the fused features into the model via skip connections. Specifically, after EA and initial feature extraction, we obtain source and target feature maps:

$$F_m, F_t \in \mathbb{R}^{C \times T'}. \tag{4}$$

To inject cross-sample diversity while preserving class semantics, we perform feature matching fusion in four steps:

1. **Vectorization**: Reshape each map into $N = C \times T'$ column vectors:

$$\{f_m^p \in \mathbb{R}^d\}_{p=1}^{N}, \quad \{f_t^q \in \mathbb{R}^d\}_{q=1}^{N}, \tag{5}$$

where d is the channel depth.

2. **Index Selection**: Randomly sample k target positions:

$$\mathcal{Q} = \{q_1, \ldots, q_k\} \subset \{1, \ldots, N\}. \tag{6}$$

3. **Cosine Matching**: For each $q \in \mathcal{Q}$ compute similarities to all source vectors:

$$s_{pq} = \frac{\langle f_t^q, f_m^p \rangle}{\|f_t^q\| \, \|f_m^p\|}, \tag{7}$$

$$p_q^* = \arg\max_p s_{pq}. \tag{8}$$

We then substitute each target representation embeddings with its most similar source feature:

$$f_t^q \leftarrow f_m^{p_q^*}, \quad \forall q \in \mathcal{Q}. \tag{9}$$
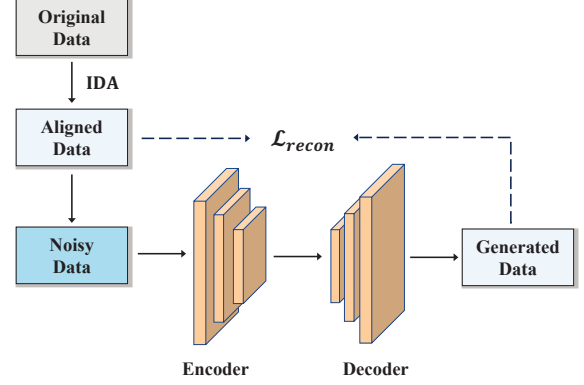

Figure 4. Overview of the proposed EEG data generation pipeline.

4. **Map Reconstruction**: Reassemble the modified set of vectors into the fused feature map:

$$\hat{F}_t = \text{reshape}\left(\{f_t^q\}_{q=1}^{N}\right) \in \mathbb{R}^{C \times T'}. \tag{10}$$

This procedure injects rich, cross-subject patterns into the target features without discarding their original label-specific information.

To ensure the fused features permeate every decoder skip-connection, we propagate them hierarchically across all L encoder–decoder layers. Denote the cumulative downsampling factor at layer ℓ by r_ℓ. A base-layer coordinate (c, t) then corresponds to a region in layer ℓ:

$$R_\ell(c, t) = \left[\lfloor c/r_\ell \rfloor : \lceil c/r_\ell \rceil\right] \times \left[\lfloor t/r_\ell \rfloor : \lceil t/r_\ell \rceil\right]. \tag{11}$$

Within each region $R_\ell(c, t)$ of the ℓ-th target feature map F_t^ℓ, we perform:

$$F_t^\ell(u, v) \leftarrow F_m^\ell\left(\lfloor u \cdot r_\ell \rfloor, \lfloor v \cdot r_\ell \rfloor\right), \quad (u, v) \in R_\ell(c, t). \tag{12}$$

By applying Eqs. (11) and (12) for all selected (c, t) across $\ell = 1, \ldots, L$, every decoder skip-connection integrates the fused features, preserving semantic alignment and enabling high-fidelity reconstruction under scarce data conditions.

3.4. Network Architecture and Training

Figure 4 illustrates the overall generator architecture, which follows a U-Net–shaped encoder–decoder design with integrated fusion modules. Given an aligned input trial $\bar{X} \in \mathbb{R}^{C \times T}$, the encoder comprises three convolutional blocks that reduce the temporal dimension by factors $r_1 = 5$, $r_2 = 5$, and $r_3 = 2$. Each bottleneck feature map $F_t \in \mathbb{R}^{C \times T'}$ is then passed through our Feature Matching Fusion modules to produce a fused map $\hat{F}_t$.

The decoder mirrors the encoder with three transposed convolutions that upsample by (r_3, r_2, r_1) and channel

depths reversed. A final transposed convolution restores the output $\tilde{X} \in \mathbb{R}^{C \times T}$.

To train the network, we adopt a denoising autoencoder strategy: each aligned trial X is perturbed with additive Gaussian noise to yield $\tilde{X}$, and the model is optimized to reconstruct the clean signal. We minimize the mean squared error over the aligned dataset:

$$\mathcal{L}_{\text{recon}} = \frac{1}{n} \sum_{i=1}^{n} \left\| \bar{X}_i - \tilde{X}_i \right\|_2^2. \tag{13}$$

After convergence, the trained generator $\mathcal{G}$ synthesizes augmented trials by applying the fusion modules to new inputs, enabling diverse and physiologically plausible EEG data generation in few-shot scenarios.

4. Experiment

4.1. Datasets

We consider the motor imagery (MI) and steady-state visual evoked potentials (SSVEP) paradigm in our experiment, the most widely adopted protocol in BCI research. MI [21] is the cognitive process of imagining the movement of different body parts without actually moving them. Three MI datasets and one SSVEP dataset, all from the mother of all BCI benchmark (MOABB) [9] and summarized in Table 2, were utilized in the experiments.

The three MI datasets used in this study are described below:

1. BNCI2014001 [24]: BNCI2014001 dataset contains EEG data from 9 subjects performing four MI tasks: left hand, right hand, both feet, and tongue. Each subject participated in two sessions, with each session consisting of 6 runs, yielding a total of 288 trials per session.
2. BNCI2014002 [23]: BNCI2014002 dataset includes EEG data from 13 participants performing sustained MI of the right hand and feet. The session consists of eight runs, with 50 trials per class for training and 30 trials for validation. EEG was recorded at 512 Hz from 15 electrodes, including C3, Cz, and C4, with a biosignal amplifier and active Ag/AgCl electrodes.
3. Zhou2016 [37]: Zhou2016 dataset includes EEG data from 4 subjects performing three MI tasks: left hand, right hand, and feet. Each subject participated in three sessions, with each session consisting of two runs of 75 trials (25 trials per class).

The SSVEP-based Nakanishi2015 dataset used in this study are described below:

1. Nakanishi2015 [20]: Nakanishi2015 dataset is an SSVEP-based EEG benchmark comprising recordings from 9 subjects with 8 channels across 12 visual stimulation classes (frequencies ranging from 9.25 Hz to 14.75 Hz in 0.5 Hz steps), each with 15 trials of 4.15 s duration sampled at 256 Hz in a single session.

4.2. Settings

All data were preprocessed using the standard MOABB pipeline. Specifically, EEG recordings were downsampled to 250 Hz, bandpass-filtered to 8–32 Hz, and each trial was truncated to the first 4 s (1000 samples).

To prevent data leakage, we split training and test sets in temporal order rather than at random. In the cross-subject evaluation, we adopt a leave-one-subject-out (LOSO) protocol: for each target subject, the first n_{train} continuous trials per class are used for augmentation and model training, and the remaining trials for testing. In the within-subject scenario, we similarly split each subject's trials by selecting the first n_{train} trials per class for training and the remainder for testing.

The generator network is trained as a denoising autoencoder using the Adam optimizer with learning rate 0.01, batch size 64. Feature matching fusion selects $k = 0.2 \times N$ vectors (where $N = C \times T'$) and uses a noise intensity coefficient of 5. To mitigate randomness from the sampling and fusion steps, each experiment is repeated 10 times and results are averaged.

For downstream classification, we apply the classic CSP-LDA pipeline [2], extracting 10 spatial filters via common spatial patterns (CSP) and classifying with a linear discriminant analysis (LDA) model.

4.3. Main Results

Figure 5 visualizes the comparison between original and generated EEG signals by various augmentation methods. Table 3 reports cross-subject few-shot accuracy for three MI datasets under varying numbers of calibration trials. FusionGen consistently outperforms both no-augmentation and most baselines. On BNCI2014001, FusionGen achieves 57.78% (vs. 53.62% without augmentation), a 4.16% enhancement; on BNCI2014002 it reaches 73.97% (vs. 72.46%), a 1.51% boost; and on Zhou2016 it attains 65.84% (vs. 62.04%), a 3.80% improvement.

Table 4 shows within-subject few-shot accuracy. On BNCI2014001, FusionGen achieves 55.86% (vs. 52.53% without augmentation), a 3.33% enhancement; on BNCI2014002 it reaches 70.47% (vs. 65.83%), a 4.64% boost; and on Zhou2016 it attains 62.53% (vs. 54.63%), a 7.90% improvement.

In most few-shot settings (7, 10, 15), FusionGen achieves good performance, validating its ability to bridge source and target distributions under extreme data scarcity.

Furthermore, we also validate the performance of FusionGen on the SSVEP-based Nakanishi2015 dataset, as reported in Table 5, demonstrating consistent improvements across different BCI paradigms.

Table 2. Summary of the three MI datasets.

BCI Paradigm	Dataset	Number of Subjects	Number of Channels	Sampling Rate (Hz)	Trial Length (seconds)	Number of Total Trials	Class Labels
MI	BNCI2014001	9	22	250	4	2592	left & right hand, tongue, feet
	BNCI2014002	14	15	512	5	1400	left hand, right hand
	Zhou2016	4	14	250	5	1842	left hand, right hand, feet
SSVEP	Nakanishi2015	9	8	256	4.15	1620	12 different stimuli

Table 3. Classification accuracy (%) of various data augmentation approaches on three MI datasets under a cross-subject few-shot evaluation.

Dataset	Trials	None	Noise	Scale	Flip	Cut&Resize	Fshift	CR	DWTaug	HHTaug	FusionGen
BNCI2014001	7	52.37	52.26	51.19	52.13	52.96	52.15	53.16	**54.97**	54.17	<u>54.85</u>
	10	53.56	53.29	52.73	53.38	52.84	53.14	54.64	57.00	<u>57.43</u>	**58.02**
	15	54.93	53.41	53.65	54.26	54.95	54.88	55.18	57.10	<u>57.16</u>	**60.48**
	Avg.	53.62	52.99	52.52	53.26	53.58	53.39	54.33	<u>56.36</u>	56.25	**57.78**
BNCI2014002	7	71.61	72.56	71.25	70.89	72.26	71.73	67.08	<u>72.62</u>	71.79	**73.56**
	10	72.80	73.27	73.27	68.75	72.20	72.80	70.36	**73.75**	<u>73.57</u>	73.45
	15	72.98	72.44	73.63	69.82	72.80	73.69	71.25	74.17	<u>74.64</u>	**74.89**
	Avg.	72.46	72.76	72.72	69.82	72.42	72.74	69.56	<u>73.51</u>	73.33	**73.97**
Zhou2016	7	62.31	62.46	63.50	57.38	58.74	61.76	63.34	<u>63.97</u>	**64.00**	63.71
	10	59.56	57.92	60.84	56.69	58.30	64.52	64.43	64.62	**68.88**	<u>64.95</u>
	15	64.26	64.14	66.45	68.78	64.77	64.84	68.40	68.50	**69.72**	<u>68.85</u>
	Avg.	62.04	61.51	63.60	60.95	60.60	63.71	65.39	65.70	**67.53**	<u>65.84</u>

4.4. Effectiveness of FusionGen Integration

To assess the compatibility of FusionGen with standard augmentation techniques, we appended our feature-fusion generator to each baseline approach and evaluated the combined performance on BNCI2014001 with 10 calibration trials. As shown in Figure 6, all FusionGen–augmented pipelines outperform their standalone counterparts. Notably, combining FusionGen with DWTaug yields the highest accuracy (58.96%), representing a further enhancement over DWTaug alone. Even methods that individually provided modest improvements, such as Noise or Scale—benefit from integration with FusionGen, demonstrating consistent boosts of 0.5-1.0%. These results confirm that FusionGen can be seamlessly integrated with diverse augmentation strategies to deliver complementary enhancements, underscoring its general applicability in EEG-based BCI.

4.5. Analysis of Generated Sample Distribution

Different subjects exhibit substantial EEG variability, creating a large gap between source and target domains. In practical BCI applications, we aim to minimize, or even

eliminate the calibration, resulting in very few target trials. As Figure 7 shows, FusionGen generates abundant synthetic trials (green) that closely follow the true target distribution (red), effectively filling the sparse region around the scarce target samples. This demonstrates that FusionGen can faithfully mimic subject-specific EEG characteristics and achieve high-quality data generation under extreme few-shot conditions.

5. Discussion

5.1. Applications

In transfer-learning scenarios for BCIs, reducing or eliminating per-subject calibration is highly desirable. However, most existing methods require target domain trials to align or adapt source models, limiting their practicality when only limited calibration trials are available [17, 36]. FusionGen addresses this bottleneck by synthesizing large volumes of target-like EEG data from minimal samples, enabling robust source–target alignment and downstream model adaptation with few calibration.

Furthermore, EEG data collection remains costly and

Table 4. Classification accuracy (%) of various data augmentation approaches on three MI datasets under a within-subject few-shot evaluation.

Dataset	Trials	None	Noise	Scale	Flip	Cut&Resize	Fshift	CR	DWTaug	HHTaug	FusionGen
BNCI2014001	10	47.56	49.82	48.56	49.80	48.39	46.88	**52.55**	47.71	44.41	_51.93_
	14	53.55	54.01	53.60	_55.32_	54.76	54.33	53.43	53.88	45.39	**56.83**
	20	56.48	_57.08_	56.73	56.53	54.79	56.15	55.09	56.20	45.94	**58.81**
	Avg.	52.53	53.64	52.96	_53.88_	52.65	52.45	53.69	52.60	45.25	**55.86**
BNCI2014002	10	59.74	65.13	62.37	66.09	66.35	62.44	61.47	60.45	_67.95_	**68.15**
	14	66.67	69.49	67.50	69.29	67.88	66.99	68.59	66.35	**71.54**	_71.21_
	20	71.09	69.87	70.13	70.90	71.22	70.71	68.53	71.22	_71.86_	**72.04**
	Avg.	65.83	68.16	66.67	68.76	68.48	66.71	66.20	66.01	_70.45_	**70.47**
Zhou2016	10	47.98	53.49	54.98	52.46	50.44	49.01	**63.59**	49.12	_61.61_	58.31
	14	52.98	55.75	53.85	57.86	55.86	55.30	**65.55**	52.44	_65.31_	63.00
	20	62.92	65.62	68.15	67.25	67.37	63.39	**71.11**	62.00	_68.27_	66.27
	Avg.	54.63	58.29	58.99	59.19	57.89	55.90	**66.75**	54.52	_65.06_	62.53

Table 5. Classification accuracy (%) of various data augmentation approaches on Nakanishi2015 datasets under a cross-subject 1-shot evaluation.

Paradigm	Approaches	S1	S2	S3	S4	S5	S6	S7	S8	S9	Avg.
SSVEP	None	53.65	34.11	77.08	80.73	83.85	84.90	79.43	71.09	86.20	72.34
	Noise	64.84	_52.08_	88.28	84.64	**96.09**	94.27	85.68	84.64	_95.31_	82.87
	Scale	_69.01_	47.92	90.36	89.58	89.06	**99.22**	**91.41**	85.68	93.23	83.94
	Flip	40.62	18.23	45.57	36.72	71.61	39.06	59.11	39.06	42.19	43.58
	Cut&Resize	49.22	26.04	67.45	61.20	76.30	75.52	73.18	56.77	72.66	62.04
	Fshift	57.55	45.83	91.41	89.84	91.93	96.09	87.24	_86.46_	87.76	81.57
	DWTaug	66.67	48.70	**95.31**	_95.31_	_94.01_	93.23	_90.36_	82.03	93.49	_84.35_
	HHTaug	66.93	47.40	_92.97_	81.77	91.93	_96.88_	85.94	81.77	91.15	81.86
	FusionGen	**76.56**	**69.01**	82.29	**97.66**	72.92	_96.88_	80.21	**86.72**	**100.0**	**84.69**

time-consuming, and publicly available datasets often lack the scale needed for training large models. While researchers are progressing toward foundation models for BCIs [10, 15, 26], the scarcity of large-scale, high-quality datasets remains a fundamental constraint. CLEAN-MI [16] attempts to filter out high-quality motor imagery data, but is still limited by insufficient data. This paper proposed FusionGen, providing a scalable solution by generating physiologically plausible EEG signals that preserve the statistical properties of real data distributions. This synthetic data generation capability holds significant potential for advancing BCI research, as it can supply unlimited, diverse, and distribution-matched training samples, thereby facilitating the development of more robust and generalizable foundation models.

5.2. Hyperparameter Analysis

We analyze the impact of the feature selection rate α on classification accuracy (Figure 8). As α varies from 0.1 to 0.5, accuracy remains stable around 57.7-58.1%, indicating that FusionGen is robust to the ratio of replaced features. However, when α exceeds 0.6, performance begins to decline (57.0% at 0.6, 56.3% at 0.7), suggesting that overly aggressive feature replacement can degrade the representation learning capacity of the latent space.

We revisit the compression ratio experiment to assess its effect on the learned latent space. Figure 9 shows the reconstructed signal spectra at 5× and 10× compression. While 5× compression preserves both low- and high-frequency EEG components, 10× compression severely attenuates higher-frequency bands, degrading signal fidelity.

Results indicate that a moderate compression ratio (5×) and a feature selection rate in the range $[0.1, 0.5]$ strike the
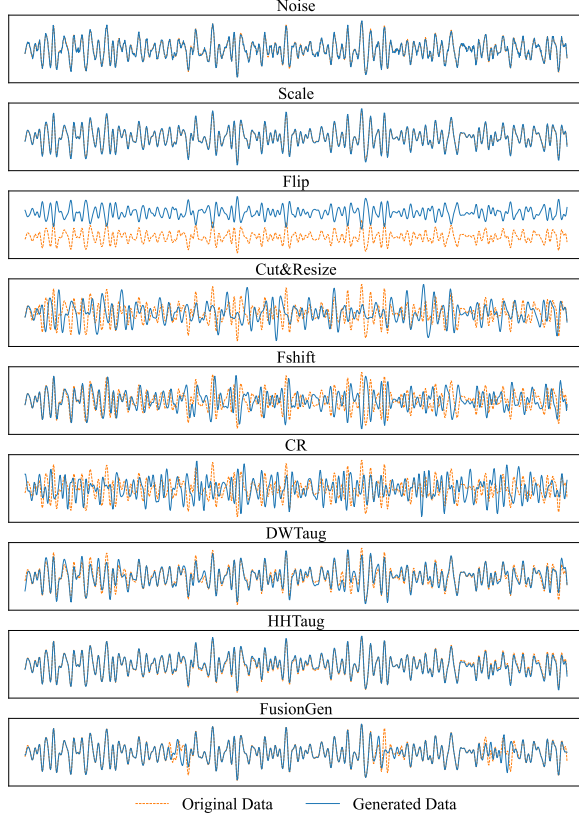
Figure 5. Visualization of generated EEG signals by various augmentation methods.
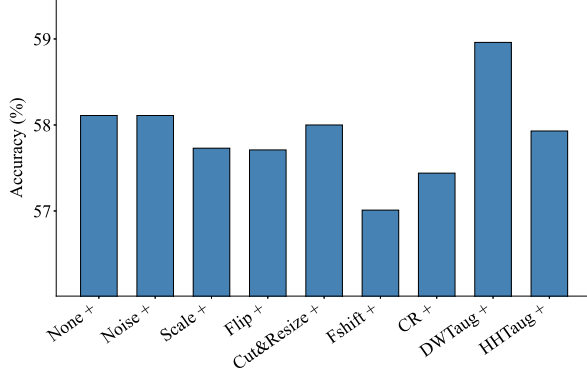

Figure 6. Accuracy of baseline augmentations combined with FusionGen on BNCI2014001 with 10 calibration trials.

best balance between diversity injection and preservation of target-specific EEG features.

6. Conclusion

This paper proposed FusionGen, a novel feature fusion-based EEG data generation framework for addressing data
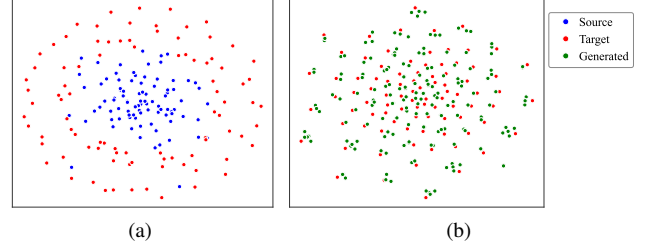

Figure 7. t-SNE visualization of the latent distributions on BNCI2014001. (a) Source data and target data. (b) Generated data and target data.
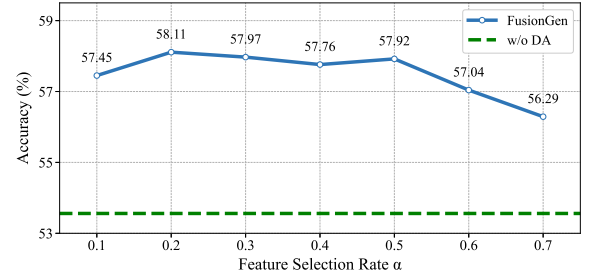

Figure 8. Sensitivity of classification accuracy to feature selection rate α
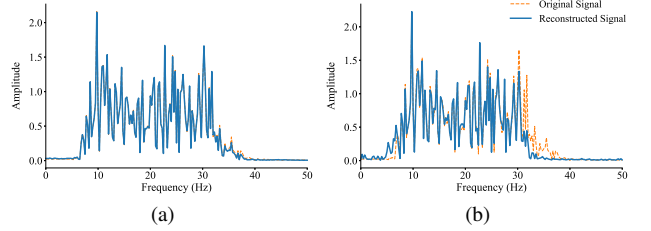

Figure 9. Frequency-domain reconstruction under different compression ratios. (a) 5× compression; (b) 10× compression.

scarcity and inter-subject variability in BCIs. Existing augmentation methods either lack diversity or require large datasets, leading to suboptimal performance in few-shot scenarios. FusionGen employs disentangled representation learning to capture essential features and integrates them through a fusion module, enhancing data diversity while preserving physiological realism. Extensive experiments on various EEG datasets and paradigms demonstrate that FusionGen achieves superior performance in cross-subject and within-subject scenarios. FusionGen leverage disentangled representation learning for EEG data generation in BCIs, may serve as a powerful data engine for BCI large models.

References

[1] Marco Altini and Hannu Kinnunen. The promise of sleep: A multi-sensor approach for accurate sleep stage detection using the oura ring. *Sensors*, 21(13):4302, 2021. 1

[2] Benjamin Blankertz, Ryota Tomioka, Steven Lemm, Motoaki Kawanabe, and Klaus-Robert Muller. Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal Processing Magazine*, 25(1):41–56, 2007. 5

[3] Florinel-Alin Croitoru, Vlad Hondru, Radu Tudor Ionescu, and Mubarak Shah. Diffusion models in vision: A survey. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 45(9):10850–10869, 2023. 1

[4] Guanqi Ding, Xinzhe Han, Shuhui Wang, Shuzhe Wu, Xin Jin, Dandan Tu, and Qingming Huang. Attribute group editing for reliable few-shot image generation. In *Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, pages 11194–11203, New Orleans, USA, 2022. 2

[5] Ruyi Foong, Kai Keng Ang, Chai Quek, Cuntai Guan, Kok Soon Phua, Christopher Wee Keong Kuah, Vishwanath Arun Deshmukh, Lester Hon Lum Yam, Deshan Kumar Rajeswaran, Ning Tang, et al. Assessment of the efficacy of EEG-based MI-BCI with visual feedback and EEG correlates of mental fatigue for upper-limb stroke rehabilitation. *IEEE Trans. on Biomedical Engineering*, 67(3):786–795, 2019. 1

[6] Zheng Gu, Wenbin Li, Jing Huo, Lei Wang, and Yang Gao. Lofgan: Fusing local representations for few-shot image generation. In *Proc. of the IEEE/CVF Int'l Conf. on Computer Vision*, pages 8463–8471, Montreal, Canada, 2021. 2

[7] He He and Dongrui Wu. Transfer learning for brain–computer interfaces: A euclidean space data alignment approach. *IEEE Trans. on Biomedical Engineering*, 67(2):399–410, 2019. 3

[8] Xiaoyang Huang, Chang Li, Aiping Liu, Ruobing Qian, and Xun Chen. EEGDfus: A conditional diffusion model for fine-grained EEG denoising. *IEEE Journal of Biomedical and Health Informatics*, 2024. 2

[9] Vinay Jayaram and Alexandre Barachant. MOABB: trustworthy algorithm benchmarking for BCIs. *Journal of Neural Engineering*, 15(6):066011, 2018. 5

[10] Wei-Bang Jiang, Li-Ming Zhao, and Bao-Liang Lu. Large brain model for learning generic representations with tremendous EEG data in BCI. In *Proc. Int'l Conf. on Learning Representations*, Vienna, Austria, 2024. 7

[11] Diederik P Kingma, Max Welling, et al. Auto-encoding variational bayes, 2013. 1

[12] Dovilė Komolovaitė, Rytis Maskeliūnas, and Robertas Damaševičius. Deep convolutional neural network-based visual stimuli classification using electroencephalography signals of healthy and alzheimer's disease subjects. *Life*, 12(3):374, 2022. 2

[13] Lingxiao Li, Yi Zhang, and Shuhui Wang. The euclidean space is evil: Hyperbolic attribute editing for few-shot image generation. In *Proc. of the IEEE/CVF Int'l Conf. on Computer Vision*, pages 22714–22724, Paris, France, 2023. 2

[14] Shijie Lian, Hua Li, Runmin Cong, Suqi Li, Wei Zhang, and Sam Kwong. Watermask: Instance segmentation for underwater imagery. In *Proc. of the IEEE/CVF Int'l Conf. on Computer Vision*, pages 1305–1315, Paris, France, 2023. 1

[15] Dingkun Liu, Zhu Chen, Jingwei Luo, Shijie Lian, and Dongrui Wu. Mirepnet: A pipeline and foundation model for eeg-based motor imagery classification. *arXiv preprint arXiv:2507.20254*, 2025. 7

[16] Dingkun Liu, Zhu Chen, and Dongrui Wu. CLEAN-MI: A scalable and efficient pipeline for constructing high-quality neurodata in motor imagery paradigm. *arXiv preprint arXiv:2506.11830*, 2025. 7

[17] Dingkun Liu, Siyang Li, Ziwei Wang, Wei Li, and Dongrui Wu. Spatial distillation based distribution alignment (SDDA) for cross-headset EEG classification. *arXiv preprint arXiv:2503.05349*, 2025. 6

[18] Dingkun Liu, Hongjie Zhou, Yilu Qu, Huimei Zhang, and Yongdong Xu. UMMAN: Unsupervised multi-graph merge adversarial network for disease prediction based on intestinal flora. *IEEE Trans. on Computational Biology and Bioinformatics*, 22:1023–1033, 2025. 1

[19] Jinming Liu, Ruoyu Feng, Yunpeng Qi, Qiuyu Chen, Zhibo Chen, Wenjun Zeng, and Xin Jin. Rate-distortion-cognition controllable versatile neural image compression. In *European Conference on Computer Vision*, pages 329–348. Springer, 2024. 1

[20] Masaki Nakanishi, Yijun Wang, Yu-Te Wang, and Tzyy-Ping Jung. A comparison study of canonical correlation analysis based methods for detecting steady-state visual evoked potentials. *PloS one*, 10(10):e0140703, 2015. 5

[21] Gert Pfurtscheller and Christa Neuper. Motor imagery and direct brain-computer communication. *Proc. of the IEEE*, 89(7):1123–1134, 2001. 5

[22] Audrey S Royer, Alexander J Doud, Minn L Rose, and Bin He. EEG control of a virtual helicopter in 3-dimensional space using intelligent control strategies. *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, 18(6):581–589, 2010. 1

[23] David Steyrl, Reinhold Scherer, Josef Faller, and Gernot R Müller-Putz. Random forests in non–invasive sensorimotor rhythm brain-computer interfaces: a practical and convenient non-linear classifier. *Biomedical Engineering/Biomedizinische Technik*, 61(1):77–86, 2016. 5

[24] Michael Tangermann, Klaus-Robert Müller, Ad Aertsen, Niels Birbaumer, Christoph Braun, Clemens Brunner, Robert Leeb, Carsten Mehring, Kai J Miller, Gernot R Müller-Putz, et al. Review of the BCI competition IV. *Frontiers in neuroscience*, 6:55, 2012. 5

[25] Fang Wang, Sheng-hua Zhong, Jianfeng Peng, Jianmin Jiang, and Yan Liu. Data augmentation for EEG-based emotion recognition with deep convolutional neural networks. In *MultiMedia Modeling: 24th Int'l Conf. MMM*, pages 82–93, Bangkok, Thailand, 2018. 2

[26] Guangyu Wang, Wenchao Liu, Yuhong He, Cong Xu, Lin Ma, and Haifeng Li. EEGPT: Pretrained transformer for universal and reliable representation of EEG signals. pages 39249–39280, Vancouver, Canada, 2025. 7

[27] Ziwei Wang, Siyang Li, Jingwei Luo, Jiajing Liu, and Dongrui Wu. Channel reflection: Knowledge-driven data augmentation for EEG-based brain–computer interfaces. *Neural Networks*, 176:106351, 2024. 2

[28] Ziwei Wang, Siyang Li, Xiaoqing Chen, and Dongrui Wu. Time-frequency transform based EEG data augmentation for brain-computer interfaces. *Knowledge-Based Systems*, page 113074, 2025. 2

[29] Dongrui Wu. Revisiting euclidean alignment for transfer learning in EEG-based brain-computer interfaces. *Journal of Neural Engineering*, 2025. 3

[30] Baao Xie, Qiuyu Chen, Yunnan Wang, Zequn Zhang, Xin Jin, and Wenjun Zeng. Graph-based unsupervised disentangled representation learning via multimodal large language models. *Advances in Neural Information Processing Systems*, 37:103101–103130, 2024. 1

[31] Haiyang Yang, Jiacheng Wan, Ying Jin, Xixia Yu, and Yinfeng Fang. EEG-and EMG-driven poststroke rehabilitation: A review. *IEEE Sensors Journal*, 22(24):23649–23660, 2022. 1

[32] Aiming Zhang, Lei Su, Yin Zhang, Yunfa Fu, Liping Wu, and Shengjin Liang. EEG data augmentation for emotion recognition with a multiple generator conditional wasserstein GAN. *Complex & Intelligent Systems*, pages 1–13, 2021. 2

[33] Kai Zhang, Guanghua Xu, Zezhen Han, Kaiquan Ma, Xiaowei Zheng, Longting Chen, Nan Duan, and Sicong Zhang. Data augmentation for motor imagery signal classification based on a hybrid neural network. *Sensors*, 20(16):4485, 2020. 2

[34] Wen Zhang, Ziwei Wang, and Dongrui Wu. Multi-source decentralized transfer for privacy-preserving BCIs. *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, 30:2710–2720, 2022. 2

[35] Chenxi Zheng, Bangzhen Liu, Huaidong Zhang, Xuemiao Xu, and Shengfeng He. Where is my spot? few-shot image generation via latent subspace optimization. In *Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, pages 3272–3281, Vancouver, Canada, 2023. 2

[36] Shitao Zheng and Dongrui Wu. Semi-supervised domain adaptation for EEG-based sleep stage classification. In *IEEE Int'l Conf. on Acoustics, Speech and Signal Processing*, pages 1776–1780, Seoul, Korea, 2024. IEEE. 6

[37] Bangyan Zhou, Xiaopei Wu, Zhao Lv, Lei Zhang, and Xiaojin Guo. A fully automated trial selection method for optimization of motor imagery based brain–computer interface. *PloS One*, 11(9):e0162657, 2016. 5

[38] Yingbo Zhou, Yutong Ye, Pengyu Zhang, Xian Wei, and Mingsong Chen. Exact fusion via feature distribution matching for few-shot image generation. In *Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, pages 8383–8392, Seattle, USA, 2024. 2