

EEG-GANet: Simulating Electroencephalogram Data to Address Sample Imbalance in P300 Speller

Xicheng Lou, Xinwei Li*, Hongying Meng, Peiyang Li, Lisha Zhong, and Zhangyong Li*

Abstract: The P300 speller is a brain–computer interface (BCI) system that enables character selection using event-related potentials (ERPs). However, the inherent imbalance between target and non-target stimuli biases deep learning models. Reducing non-target samples to mitigate this issue risks underutilizing valuable data. We propose EEG-GANet, a generative adversarial network leveraging domain transformation to augment target samples and mitigate class imbalance in P300 speller datasets. Additionally, we develop EEG-DBNet-V2, a compact dual-branch network that efficiently extracts temporal and spectral features from electroencephalography (EEG) signals, serving both as a classification model and as the discriminator within EEG-GANet. Extensive experiments with ten-fold cross-validation on three public datasets demonstrated EEG-DBNet-V2's superior classification accuracy compared to state-of-the-art models, achieving this with significantly fewer parameters. Integrating EEG-GANet's augmented data via fine-tuning further improved classification performance. EEG-GANet effectively addresses sample imbalance by generating physiologically accurate augmented data, substantially enhancing EEG-DBNet-V2's classification capability. Our study introduces a targeted generative adversarial network (GAN)-based EEG data augmentation framework, enabling balanced model training without discarding valuable non-target samples. The integration of EEG-GANet and EEG-DBNet-V2 offers a lightweight yet robust solution, advancing the practical deployment of EEG-based BCIs under data-constrained conditions. The source code is publicly available at <https://github.com/xicheng105/EEG-GANet>.

Key words: P300 speller; electroencephalography (EEG); brain–computer interface (BCI); generative adversarial network (GAN)

1 Introduction

A brain-computer interface (BCI) enables direct communication between the brain and external devices,

facilitating information transmission. Its core principle is to monitor and analyze brain activity in real time, converting neural signals into interpretable commands for device control and interaction^[1].

- Xicheng Lou is with School of Communications and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China. E-mail: louxicheng@ctbu.edu.cn.
- Xinwei Li, Peiyang Li, and Zhangyong Li are with the Research Center of Biomedical Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China. E-mail: lixinwei@cqupt.edu.cn; pyli@cqupt.edu.cn; lizy@cqupt.edu.cn.
- Hongying Meng is with Department of Electronic and Electrical Engineering, Brunel University London, London UB8 3PH, UK. E-mail: hongying.meng@brunel.ac.uk.
- Lisha Zhong is with School of Medical Information and Engineering, Southwest Medical University, Luzhou 646000, China. E-mail: zhonglisha@swmu.edu.cn.

* To whom correspondence should be addressed.

Manuscript received: 2024-12-05; revised: 2025-03-26; accepted: 2025-04-19

Among various BCI applications, the P300 speller has been widely utilized as an assistive text input system, enabling individuals with severe communication impairments, such as those with amyotrophic lateral sclerosis (ALS), to convey messages through electroencephalography (EEG)-based neural responses^[2]. It utilizes event-related potentials (ERP)^[3] to detect user-intended characters via EEG signals. The P300 is a positive ERP component that emerges approximately 300 ms after a target stimulus, reflecting cognitive processing. In the P300 speller, this response is elicited only when the user attends to the target character^[4].

Despite its potential, the P300 speller faces several challenges that hinder its real-world applicability. First, the classification of P300 signals is difficult due to the low signal-to-noise ratio (SNR) and interference from background EEG activity. Second, traditional machine learning approaches rely heavily on signal preprocessing, demanding expert knowledge to extract informative features. Third, deep learning models, while effective, require large amounts of labeled training data, which is often infeasible in P300 speller scenarios due to the labor-intensive data collection process^[5]. Finally, imbalanced sample distribution between target and non-target stimuli leads to model bias, reducing classification accuracy.

To address these challenges, this study explores both model optimization and data augmentation strategies for P300 speller classification. We propose EEG-GANet, a generative adversarial network (GAN) that augments P300 target samples by learning the intrinsic relationship between target and non-target signals. Additionally, we introduce EEG-DBNet-V2, a compact dual-branch classification model that enhances feature extraction from EEG signals while maintaining low computational complexity.

The main contributions of this study are as follows:

(1) We propose EEG-GANet, a P300 signal augmentation model based on domain transformation. By leveraging the intrinsic relationship between target and non-target signals, EEG-GANet generates pseudo-target samples from non-target samples in the source data. This approach mitigates the pattern collapse issue commonly observed when generating target signals from noise, while preserving the shared characteristics of target and non-target signals.

(2) We introduce EEG-DBNet-V2, a compact dual-branch architecture designed to enhance EEG feature

extraction. By jointly decoding temporal and spectral domain characteristics at multiple scales, EEG-DBNet-V2 significantly improves classification performance while maintaining a low computational footprint. In experiments conducted on the P300-ALS^[6] and P300-normal^[7] datasets, EEG-DBNet-V2 consistently achieved superior classification accuracy compared to state-of-the-art models, despite having only one-fifth of their parameters.

(3) We conduct a simulated spelling experiment on the P300-BCI22b dataset^[8] to evaluate the impact of sample imbalance on model performance. We further propose a fine-tuning strategy using augmented data generated by EEG-GANet, which significantly enhances the classification performance of EEG-DBNet-V2.

The remainder of this paper is structured as follows: Section 2 reviews related studies on P300 classification, deep learning models, and EEG data augmentation. Section 3 presents the proposed EEG-DBNet-V2 and EEG-GANet models in detail. Section 4 describes the experimental setup, results, and key discussions. Finally, Section 5 summarizes our findings and outlines potential future directions.

2 Related Work

2.1 Traditional machine learning approaches

Traditional machine learning techniques have been widely used for P300 classification, often relying on extensive feature engineering and preprocessing. Li et al.^[9] employed correlation coefficient matrices to capture inter-channel relationships before classifying P300 signals with a support vector machine (SVM). Qu et al.^[10] introduced a 3D character representation paradigm to enhance visual stimulation. Kundu and Ari^[11] applied principal component analysis (PCA) to reduce feature redundancy, while Akram et al.^[12] used a random forest classifier in a modified 3×3 speller interface. Despite their effectiveness, these approaches require extensive domain expertise for feature extraction and are less generalizable across subjects.

2.2 Deep learning-based methods

Deep learning has demonstrated superior performance over traditional approaches by learning high-level EEG features directly from raw signals. Cecotti and Graser^[13] investigated the impact of EEG channel count and convolutional kernel design using

convolutional neural networks (CNNs). Manor and Geva^[14] incorporated spatio-temporal regularization to mitigate overfitting in large-parameter networks. However, given the limited availability of P300 samples, excessively complex models may suffer from overfitting.

To address this issue, compact network architectures have been proposed. Liu et al.^[15] introduced BN3, which compresses the temporal dimensions of signals through convolution to reduce the parameters, incorporates batch normalization (BN)^[16] to lessen the impact of overfitting, and uses rectified linear units (ReLU)^[17] to accelerate model convergence. Shan et al.^[18] proposed OCLNN, which consists of only one convolutional layer and one fully connected layer, resulting in a parameter count an order of magnitude lower than that of BN3. Lawhern et al.^[19] introduced EEGNet, which reduces the parameter count by incorporating separable convolution layers and proposed the concept of training feature maps for each signal channel. EEGNet serves as a compact network model that achieves a relatively high accuracy with minimal training parameters, becoming an important reference in the field of EEG signal classification. Alvarado-González et al.^[20] proposed SepConv1D, which consists of only one separable convolution layer and one fully connected layer, with a parameter count even an order of magnitude lower than that of EEGNet. Compact network models exhibit significant advantages in terms of resource consumption, training time, and inference speed; however, this does not imply that such models can achieve a highly satisfactory level of classification accuracy.

Some researchers argue that increasing the diversity of extracted signal features is an effective means to enhance model classification accuracy. Santamaria-Vázquez et al.^[21] proposed EEG-Inception, which employs an inception multi-branch structure to enable multi-scale feature extraction from signals, thereby improving classification accuracy. Bagchi and Bathula^[22] introduced EEG-ConvTransformer, which reinforces global feature extraction from signals by integrating a multi-head attention mechanism.

2.3 EEG data augmentation techniques

Given the challenges in collecting large EEG datasets, data augmentation has emerged as a promising alternative. Aznan et al.^[23] explored GAN-based approaches, including deep convolutional generative

adversarial network (DCGAN)^[24], Wasserstein GAN (WGAN)^[25], and variational autoencoder (VAE)^[26], to improve classification performance. However, conventional GAN-based EEG generation often suffers from mode collapse, limiting sample diversity.

More advanced augmentation strategies have been proposed to overcome these issues. EEGANet^[27] integrates adversarial training to generate EEG samples that better preserve neural signal characteristics. MASER^[28] employs state-space modeling to enhance EEG spatial resolution, optimizing feature extraction for P300 classification. ERPENet^[29] utilizes a multi-task autoencoder to enhance feature generalization across datasets, significantly improving P300 classification accuracy. Penava and Buettner^[30] introduced an EEG augmentation model based on conditional tabular generative adversarial network (CTGAN)^[31], which simulated frequency domain features to enhance a random forest (RF)^[32] classifier's accuracy. Xu et al.^[33] applied CycleGAN^[34] to generate EEG topographies, visually confirming their similarity to real data.

Compared to downsampling the majority class, augmenting the minority class ensures optimal utilization of available data. Bhandari et al.^[35] proposed TL-P3GAN, a CNN-LSTM hybrid model, to augment minority class EEG, significantly improving classification accuracy. Pan et al.^[36] introduced TEGAN, which extends short EEG segments into longer samples through time-window segmentation, demonstrating improved generalizability. Additionally, since augmented data exhibits better cross-subject robustness, it has been employed in domain adaptation for P300 classification. Li et al.^[37] introduced OADANN, which generates subject-independent EEG features and applies transfer learning to facilitate cross-subject classification.

Despite these advancements, EEG augmentation using GANs still faces challenges, including mode collapse and the difficulty of generating task-specific EEG signals. These limitations motivate the development of EEG-GANet, which is designed to generate high-quality target samples by leveraging domain adaptation techniques.

3 Method

The EEG decoding process of the P300 speller is illustrated in Fig. 1. Firstly, EEG-DBNet-V2 is pretrained using an equal number of target and non-

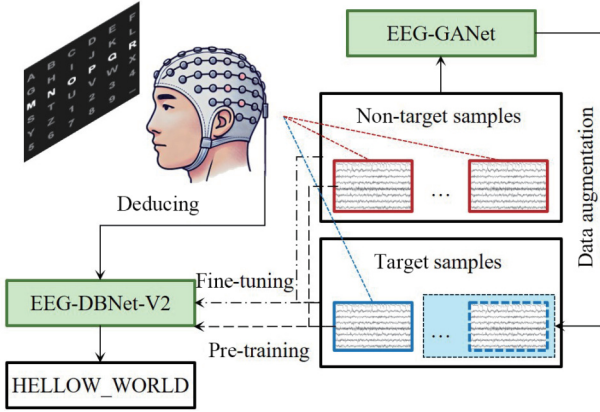


Fig. 1 EEG decoding process of the P300 speller.

target samples. Then, EEG-GANet generates target samples using non-target samples as input, ensuring that the number of target samples matches that of the non-target samples. Finally, EEG-DBNet-V2 is fine-tuned using the training set, which includes the augmented samples.

The source EEG originates from the stimulation signal emitted by the P300 speller and is segmented into epochs with a duration of 1 s. The training set is denoted as $\{X_i, y_i\}_{i=1}^m$, where $X_i \in \mathbb{R}^{C \times S}$, $i \in \{1, 2, \dots, m\}$. Here, C represents the number of EEG electrodes, S denotes the number of sampling points per epoch, and m is the total number of epochs. The specific values of sampling frequency, S , and C for each dataset are provided in Table 1. The label $y_i \in \{0, 1\}$ corresponds to non-target and target samples, respectively.

Table 1 Parameters of public P300 EEG datasets.

Dataset	SF (Hz)	S	C
P300-ALS ^[6]	256	256	8
P300-normal ^[7]	256	256	9
P300-BCI22b ^[8]	240	240	64

Note: SF: Sampling frequency (Hz), S : Number of sampling points per epoch, C : Number of EEG electrodes.

3.1 EEG-DBNet-V2

EEG-DBNet-V2 is a dual-branch compact network, as shown in Fig. 2. Based on our previously proposed EEG-DBNet^[38] designed for motor imagery EEG signal classification, EEG-DBNet-V2 was developed by modifying the convolutional approach and the trend in the number of convolutional kernels to model the global convolution and feature reconstruction functional modules of EEG-DBNet. While EEG-DBNet-V2 shows a slight decrease in accuracy, its parameter count is only one-twentieth of that of EEG-DBNet. Each branch of EEG-DBNet-V2 consists of four convolutional layers, including one depthwise convolutional layer and three separable convolutional^[39] layers. The first of these separable convolutional layers uses traditional convolution, while the latter two employ dilated convolution^[40].

The source signal X_i is first convolved with F convolutional kernels of size $(1, K)$, where K is equal to half of the signal sampling frequency, resulting in F feature maps containing characteristics above 2 Hz. In terms of the temporal branch, the first convolutional layer performs depthwise convolution with a depth of

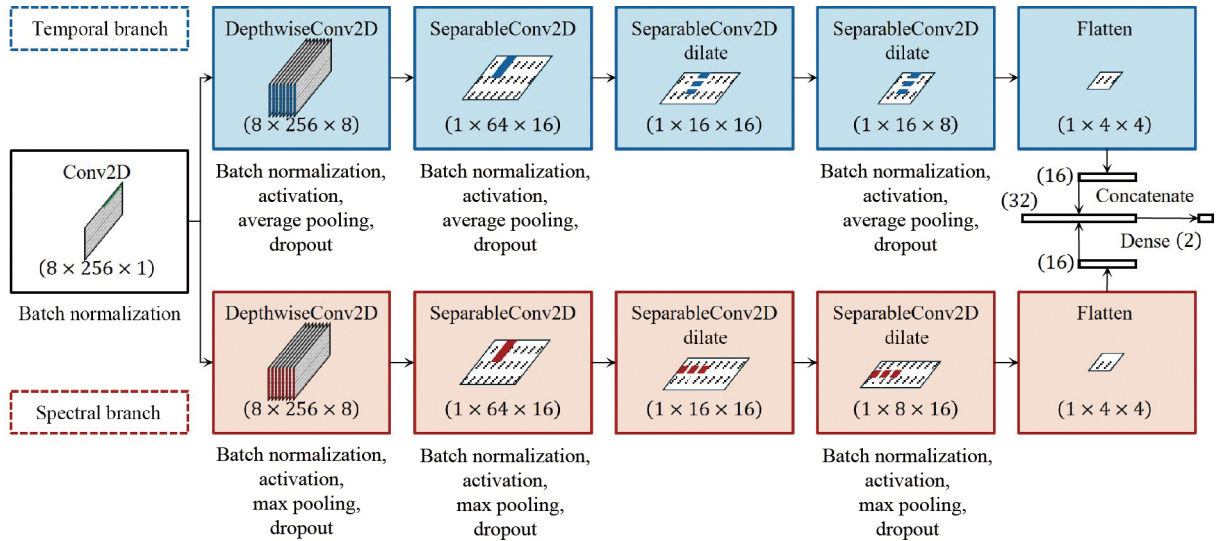


Fig. 2 Diagram of EEG-DBNet-V2 architecture. The dimensions shown in the figure correspond to the P300-ALS dataset, where $T = 256$ (number of sampling points per epoch) and $C = 8$ (number of EEG electrodes).

D and kernel size of $(C, 1)$, extracting D spatial features for each feature map. Subsequently, an average pooling layer of size $(1, 4)$ reduces the sampling frequency to $SF/4$ Hz. The second convolutional layer performs separable convolution with $2 \times F$ convolutional kernels of size $(1, K/2)$ to extract features above 4 Hz, followed by another average pooling layer of size $(1, 4)$, reducing the sampling frequency to $SF/16$ Hz. The third and fourth convolutional layers introduce dilated convolutions to extract global features. The receptive field of the dilated convolution is calculated using Eq. (1),

$$R_j = 1 + (k_j - 1) \times d_j, \\ R = \sum_{j=1}^{n-1} R_j + (R_n - 1) \times d_n \quad (1)$$

where R is the receptive field of the separable dilated convolution module, n is the number of convolutional layers in the module, R_j is the receptive field of each convolutional layer, k_j and d_j are the kernel size and dilation rate of the j -th convolutional layer, respectively. In this module, $n = 2$, $k_j = (1, K/32)$, and $d_j = D$. Therefore, the receptive field of the separable dilated convolution module, which consists of the last two convolutional layers of the temporal branch, is 19. After the fourth convolutional layer, an average pooling layer is applied, reducing the signal size to $(1, 4, 4)$.

The first convolution extracts F temporal-spatial feature maps from the source signal. In the first convolutional layer of the branch, the spatial dimension is eliminated, allowing the input signal to the second convolutional layer of the branch to be viewed as virtual temporal-spectral feature maps. Therefore, unlike the temporal branch, the spectral branch performs convolution on the spectral dimension of the signal in the third and fourth convolutional layers. Additionally, for the spectral dimension, the importance of specific sampling points outweighs the relationships between sampling points^[38], so the pooling layer in the temporal branch employs max pooling. The feature outputs from the two branches are flattened into vectors, concatenated together, and then passed through a fully connected (FC) layer to complete the signal classification, as Eq. (2),

$$\hat{y}_i = \sigma(W(\hat{x}_i \hat{\sim} \tilde{x}_i)) \quad (2)$$

where $\hat{y}_i$ is the predicted label, $\hat{\sim}$ denotes concatenate, W is parameters of the FC layer, and $\hat{x}_i$ and $\tilde{x}_i$

represent the feature vectors output by the temporal branch and spectral branch, respectively. The parameter description of EEG-DBNet-V2 is presented in Table 2.

3.2 EEG-GANet

EEG-GANet is designed to enhance P300 target sample generation by learning the transformation between non-target and target EEG signals. Since target samples contain P300 waveforms that are absent in non-target samples, our approach aims to decode the underlying relationship between these two domains and generate pseudo-target samples that preserve essential P300 characteristics while maintaining background EEG features.

To achieve this, EEG-GANet adopts a domain-transfer framework, similar to DiscoGAN^[41]. Specifically, we use EEG-DBNet-V2 as the discriminator and employ a U-Net^[42] architecture for the generator, forming a generative model capable of transforming non-target samples into P300 target samples. The model consists of two generators (G_{BA} and G_{AB}) and two discriminators (D_A and D_B). The generators learn the mapping between the non-target and target EEG domains, while the discriminators evaluate the authenticity of the generated samples.

3.2.1 EEG-GANet architecture

The structure of EEG-GANet is shown in Fig. 3a. The generators, G_{AB} and G_{BA} , establish bidirectional mappings: from the non-target domain to the target domain and vice versa. The discriminators, D_A and D_B , are responsible for determining whether the samples belong to the real or generated distributions. Both generators share an identical network structure, which is detailed in Fig. 3b. To enhance training stability and generalization, the generator network incorporates

Table 2 EEG-DBNet-V2 parameters and settings.

Branch	Parameter	Setting
Temporal branch	Filters (F)	8
	Kernel size (K)	SF/2
	Depth (D)	2
	Average pooling size	(1, 4)
	Dropout rate	0.5
Spectral branch	Filters (F)	8
	Kernel size (K)	SF/2
	Depth (D)	2
	Max pooling size	(1, 4)
	Dropout rate	0.5

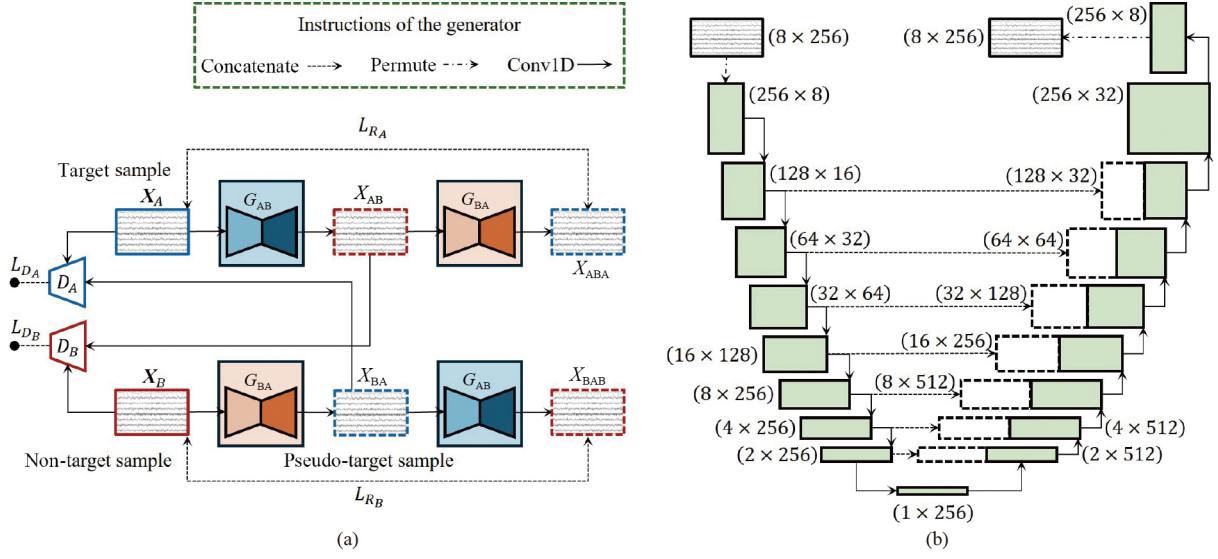


Fig. 3 EEG-GANet architecture. (a) Layout of each module in EEG-GANet, with the discriminator being EEG-DBNet-V2. (b) Generator architecture in EEG-GANet. The dimensions shown in the figure correspond to the P300-ALS dataset, where $T = 256$ (number of sampling points per epoch) and $C = 8$ (number of EEG electrodes).

group normalization^[43] and leaky ReLU activation^[44]. The discriminators adopt the EEG-DBNet-V2 architecture, as introduced in Section 3.1.

In the target sample generation process, a non-target EEG sample X_B serves as input to G_{BA} , which generates the corresponding pseudo-target sample X_{BA} . To enforce structural consistency, the pseudo-target sample is then passed through G_{AB} to reconstruct the original non-target sample X_{BAB} , ensuring that the transformation preserves key EEG features,

$$X_{BA} = G_{BA}(X_B) \quad (3)$$

$$X_{BAB} = G_{AB}(X_{BA}) \quad (4)$$

3.2.2 Domain-specific loss functions

The generator is optimized using a combination of validity loss and reconstruction loss, which together ensure that the generated EEG samples are both realistic and physiologically meaningful.

The validity loss (adversarial loss) encourages the generator to produce samples that are indistinguishable from real target EEG signals. It is formulated using the binary cross-entropy^[45] loss,

$$L_{CB} = -E_{X_B \sim P_B} [\ln D_A(G_{BA}(X_B))] \quad (5)$$

where $E_{X_B \sim P_B}[\cdot]$ denotes the mathematical expectation over the real non-target EEG data distribution P_B , and P_B represents the real non-target data distribution. A similar formulation applies for L_{CA} when generating non-target samples.

Since EEG signals contain complex spectral and temporal dependencies, it is crucial to preserve individual subject characteristics when transforming data between domains. To achieve this, we introduce a cycle-consistency constraint, formulated as Eq. (6),

$$L_{RB} = E_{X_B \sim P_B} [\|G_{AB}(G_{BA}(X_B)) - X_B\|_1] \quad (6)$$

where $\|\cdot\|_1$ denotes the ℓ_1 -norm.

This constraint ensures that when a generated pseudo-target sample is mapped back to the non-target domain, it remains consistent with the original input. Such a constraint prevents the generator from producing unrealistic EEG structures and preserves intra-subject variations in the generated samples. The total loss for the generator is given by Eq. (7),

$$L_G = L_{CA} + L_{CB} + \lambda(L_{RA} + L_{RB}) \quad (7)$$

where λ balances the adversarial loss and reconstruction loss. Since these losses operate on different scales, we empirically set $\lambda = 0.1$ to maintain training stability.

The discriminator loss function follows a standard binary cross-entropy formulation, applied to both real and generated samples,

$$L_{DB} = -E_{X_B \sim P_B} [\ln D_B(X_B)] - E_{X_{AB} \sim P_{G_{AB}}} [\ln(1 - D_B(X_{AB}))] \quad (8)$$

where $E_{X_{AB} \sim P_{G_{AB}}}[\cdot]$ denotes the mathematical expectation with respect to the model distribution induced by the generator G_{AB} , and $P_{G_{AB}}$ represents the

distribution of generated target EEG samples. The same formulation applies to L_{D_A} in the reverse transformation. The total discriminator loss is formulated as Eq. (9),

$$L_D = L_{D_A} + L_{D_B} \quad (9)$$

4 Experiment and Discussion

The experimental data are obtained from three publicly available P300 speller datasets: P300-ALS^[6], P300-normal^[7], and P300-BCI22b^[8]. The P300-ALS dataset contains EEG recordings from eight ALS patients using a standard 6×6 P300 speller. The P300-normal dataset includes ten subjects performing a visual P300 task, but only the last eight subjects used a standard 6×6 speller, so we only utilize their data. The P300-BCI22b dataset (2nd Wadsworth BCI dataset) is part of the BCI Classification Contest 2002, in which participants are required to train models on the training set to predict the spelled characters in the test set.

For all three datasets, the training and test sets are maintained as in the original datasets. The training set is further divided into training and validation sets using stratified tenfold cross validation^[46], with 90% used for training and 10% for validation. No preprocessing steps, such as filtering or artifact removal, are applied to the raw EEG signals.

The experiments are conducted using the TensorFlow framework, running on an Intel(R) Xeon(R) Gold 6248R CPU @ 3.00 GHz and six NVIDIA A10 GPUs. The experiment workflow consisted of two stages: Pre-training and fine-tuning.

4.1 Pre-training

We train and validate the classification model with an equal number of target and non-target samples by downsampling the non-target signals to match the number of target samples, while maintaining the sample size disparity between different classes in the test set. The training is conducted using the Adam optimizer with a learning rate of 0.0009 and a batch size of 64, with cross-entropy as the loss function. The model is trained for 1000 epochs, with early stopping set to a patience of 300 epochs.

To evaluate model performance, we employ stratified tenfold cross-validation, training ten separate models and testing each on the corresponding test set. The final result is obtained by averaging the test performance

across all ten models. Evaluation metrics included classification accuracy P and the F1-score.

Classification accuracy P is defined as Eq. (10),

$$P = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

where TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives, respectively.

Precision and Recall are defined as Eqs. (11) and (12),

$$\text{Precision} = \frac{TP}{TP + FP} \quad (11)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (12)$$

The F1-score is calculated as Eq. (13),

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (13)$$

The F1-score is the harmonic mean of Precision and Recall, providing a balanced measure of a model's performance when dealing with imbalanced datasets.

4.1.1 Comparison with other models

The proposed EEG-DBNet-V2 is compared with five other models: BN3^[15], OCLNN^[18], EEGNet^[19], EEGInception^[21], and SepConv1D^[20]. Among them, SepConv1D is an ultra-compact network, EEGNet and OCLNN are compact networks, while BN3 and EEGInception are complex networks.

Table 3 presents the classification performance of these models on the P300-ALS and P300-normal datasets. The results demonstrate that EEG-DBNet-V2 achieves the highest classification accuracy and F1-score across both datasets while maintaining a relatively low parameter count.

On the P300-ALS dataset, EEG-DBNet-V2 attains the highest classification accuracy ($P = 76.45\%$) and F1-score ($F_1 = 0.6859$), surpassing all other models. Compared to EEGInception, the second-best model in terms of accuracy ($P = 75.09\%$), EEG-DBNet-V2 still outperforms it while requiring only one-fifth of the trainable parameters.

On the P300-normal dataset, EEG-DBNet-V2 achieves the highest classification accuracy ($P = 69.54\%$) and F1-score ($F_1 = 0.6125$), outperforming all baseline models. Notably, while EEGInception achieves a competitive accuracy ($P = 68.58\%$), EEG-DBNet-V2 still surpasses it, maintaining its advantage in both classification performance and computational efficiency.

Table 3 Pre-training classification performance (in P and F_1) on P300-ALS and P300-normal: A comparison of the proposed EEG-DBNet-V2 with other models.

Subject	BN3 ^[15]		OCLNN ^[18]		EEGNet ^[19]		EEGInception ^[21]		SepConv1D ^[20]		EEG-DBNet-V2		
	<i>P</i> (%)	<i>F</i> ₁	<i>P</i> (%)	<i>F</i> ₁	<i>P</i> (%)	<i>F</i> ₁	<i>P</i> (%)	<i>F</i> ₁	<i>P</i> (%)	<i>F</i> ₁	<i>P</i> (%)	<i>F</i> ₁	
P300-ALS	1	63.56*	0.5578*	64.99*	0.5692*	69.01	0.6112	71.32	0.6294	57.83*	0.5128*	72.12	0.6410
	2	65.65*	0.5765*	64.88*	0.5742*	68.02	0.6008*	69.64	0.6225	57.91*	0.5137*	70.78	0.6307
	3	74.43*	0.6622*	76.22	0.6785*	74.70*	0.6672*	76.54	0.6912	69.62*	0.6201*	77.93	0.7029
	4	62.60*	0.5599*	61.32*	0.5405*	66.28*	0.5947*	69.16	0.6130	58.07*	0.5124*	70.65	0.6240
	5	70.76*	0.6241*	67.75*	0.6016*	75.65*	0.6626*	75.92*	0.6736*	59.42*	0.5297*	78.16	0.6944
	6	72.07*	0.6417*	73.06*	0.6466*	73.98*	0.6565*	77.02	0.6913	63.84*	0.5650*	77.96	0.7001
	7	69.70*	0.6092*	70.67*	0.6214*	75.13*	0.6656*	76.00	0.6812*	61.10*	0.5385*	79.20	0.7111
	8	82.12	0.7511	83.46*	0.7625*	83.12	0.7642	85.09	0.7857	76.01*	0.6931*	84.81	0.7828
StD [†]	6.40	0.0642	7.18	0.0713	5.39	0.0544	5.16	0.0554	6.62	0.0648	4.92	0.0528	
Mean	70.11	0.6228	70.29	0.6243	73.24	0.6529	75.09	0.6735	62.98	0.5607	76.45	0.6859	
NumP [‡]	48 721		2097		1777		26 597		289		5633		
P300-normal	3	64.96*	0.5595*	54.11*	0.4882*	65.47	0.5702	71.38	0.6273	51.54*	0.4500*	72.60	0.6261
	4	72.17*	0.6239*	61.19	0.5220	66.22	0.5786	66.67	0.5836	58.82*	0.5122*	63.06	0.5643
	5	61.02*	0.5493*	47.94*	0.4320*	67.62	0.5542*	73.71	0.6315	49.29*	0.4269*	73.76	0.6376
	6	53.89	0.4664	54.58	0.4661*	51.17	0.4426	55.18	0.4842	49.82*	0.4393*	56.41	0.4891
	7	56.93*	0.5008*	54.80*	0.4816*	59.72*	0.5276*	64.18	0.5646	55.48*	0.4656*	65.36	0.5674
	8	68.74*	0.6074*	66.03*	0.5739*	72.19	0.6398	68.27*	0.6109*	61.42*	0.5173*	72.80	0.6482
	9	75.04	0.6745	68.36*	0.6034*	69.92	0.6195	72.52	0.6459	59.69*	0.5091*	73.28	0.6619
	10	73.22*	0.6504*	68.24*	60.12*	74.26*	0.6514*	76.69	0.6825	61.77*	0.5440*	79.07	0.7056
StD [†]	7.87	0.0728	7.65	0.0650	7.41	0.0679	6.74	0.0605	5.18	0.0429	7.31	0.0684	
Mean	65.75	0.5790	59.41	0.5211	65.82	0.5730	68.58	0.6038	55.98	0.4831	69.54	0.6125	
NumP [‡]	48 739		2321		1793		26 645		309		5665		

Note: * $p < 0.05$. [†] Standard deviation. [‡] Number of trainable parameters. Bold values indicate the best performance.

Additionally, the p -values indicate that EEG-DBNet-V2 achieves statistically significant improvements ($p < 0.05$) over most baseline models across subjects, further validating its effectiveness in P300 classification.

Overall, these results confirm that the improved EEG-DBNet-V2 effectively enhances classification performance while maintaining computational efficiency, making it a robust solution for practical BCI applications.

4.1.2 Feature extraction analysis of EEG-DBNet-V2

The design of EEG-DBNet-V2 incorporates a dual-branch architecture to enhance EEG feature extraction by jointly capturing complementary neural representations. To assess the contribution of each branch, we analyze the scalp topographies of the depthwise convolutional layer weights, which indicate the spatial importance assigned to different EEG electrodes. Figures 4b and 4c illustrate the weight distributions learned by the two branches, while Fig. 4d

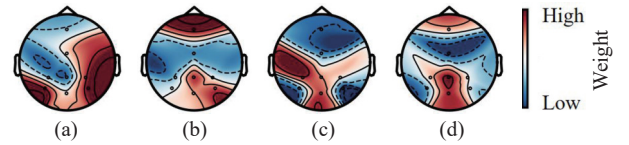


Fig. 4 Scalp topographies of depthwise convolutional layer weights across different models. (a) EEGNet; (b) temporal branch of EEG-DBNet-V2; (c) spectral branch of EEG-DBNet-V2; (d) averaged weights of (b) and (c).

represents their averaged contribution. Notably, the two branches exhibit a complementary activation pattern, suggesting that they emphasize distinct but synergistic characteristics of EEG signals. Figure 4d, which integrates both branches, highlights Pz, P3, P4, and Fz electrodes, corresponding to central parietal, lateral parietal, and central frontal regions—key areas for detecting P300 signals^[47]. This observation confirms that EEG-DBNet-V2 effectively extracts spatially relevant features, reinforcing its capacity to capture critical neural activity associated with P300 responses.

4.1.3 Simulated spelling performance analysis

Sample imbalance significantly affects the classification performance of P300 speller models. To investigate this, we compare two training strategies: (1) Downsampling the non-target samples to balance the dataset and (2) maintaining the original 1 : 5 target-to-non-target sample ratio. The results are illustrated in Fig. 5.

For the P300-ALS dataset, the impact of sample imbalance is evident. In Fig. 5a where non-target samples are downsampled, the model achieves a balanced classification performance with TN and TP rates of 76.13% and 78.04%, respectively. However, in Fig. 5b, where the original 1 : 5 ratio is maintained, the model becomes biased toward non-target samples, achieving a TN rate of 94.98% but a drastically reduced TP rate of 51.27%. This confirms that when non-target samples dominate the training data, the model prioritizes non-target classifications at the cost of target detection accuracy.

For the P300-normal dataset, a similar trend is observed. Figure 5c shows that maintaining the original imbalance (1 : 5 ratio) results in a TN rate of 95.54%, but the TP rate drops to 37.81%, further demonstrating that a heavily imbalanced dataset skews

the model's classification decision toward the majority class.

To assess the impact of sample imbalance on the practical application of P300 spellers, we conduct a simulated spelling experiment using the P300-BCI22b dataset. The experimental setup followed the two training strategies mentioned earlier.

In this experiment, the test set contains eight target words. Since the models are trained using tenfold cross-validation, we obtain ten trained models. Each model is used to predict the target characters in the test set. For each character, 15 trials are conducted, and the model's predictions are recorded. The predictions are ranked based on their frequency in descending order. If the correct target character ranks first, it is considered successfully identified. The final result for each target word is obtained by averaging the rankings from the ten models.

Table 4 presents the results of this simulated spelling experiment. A mean sequence index of 1.0 indicates that all ten models correctly identified the target character, whereas a higher index suggests that either more models misclassified the target or that the degree of misclassification was more severe. The results show that maintaining the original sample imbalance significantly reduces the accuracy of target character detection. Under the original imbalance condition, none of the target words are correctly identified by all ten models, highlighting the detrimental effect of sample imbalance on P300 speller performance.

These findings highlight a key challenge in P300

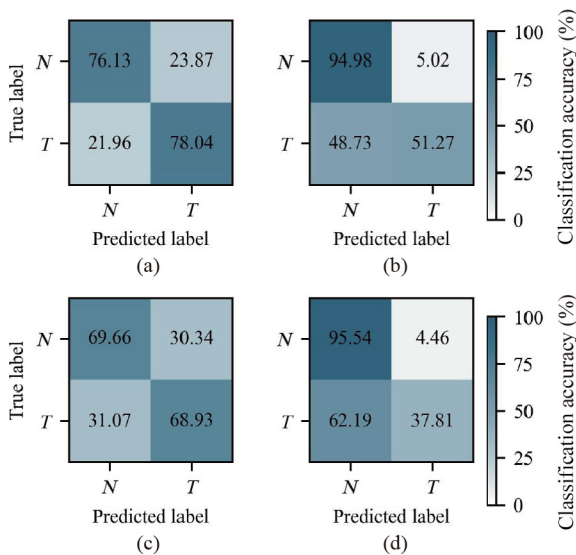


Fig. 5 Confusion matrices illustrating the effect of sample imbalance on EEG-DBNet-V2 classification performance. (a) Model trained with an equal number of target (T) and non-target (N) samples on the P300-ALS dataset. (b) Model trained with a 1:5 target-to-non-target ratio on the P300-ALS dataset. (c) and (d) show the corresponding results for the P300-normal dataset. The color bar indicates classification accuracy (%).

Table 4 Mean sequence index for each target word in the simulated spelling experiment using P300-BCI22b. Lower values indicate better model performance, with 1.0 meaning perfect identification by all ten models.

Target word	Undersample [☆]					Original [©]			
FOOD	—	1.4	1.0	1.0	1.0	—	1.9	1.8	1.6
MOOT	—	1.0	1.0	1.0	1.0	—	1.0	1.0	1.3
HAM	—	—	1.0	1.0	1.0	—	—	1.1	1.7
PIE	—	—	1.0	1.0	1.0	—	—	1.0	1.0
CAKE	—	1.1	1.0	1.0	1.0	—	1.4	1.0	1.0
TUNA	—	1.0	1.0	1.0	1.0	—	1.0	1.7	1.0
ZYGOT	1.0	1.0	1.0	1.0	1.0	1.2	1.1	1.2	1.0
4567	—	1.0	1.0	1.1	1.0	—	1.0	1.0	2.5

Note: [☆] The model was trained with an equal number of target and non-target samples. [©] The model was trained while maintaining the original 1:5 target-to-non-target sample ratio. — indicates that the corresponding spelling trial does not exist because the target word contains fewer characters.

classification: Imbalanced training data leads to biased classification performance, favoring the majority class (non-target samples) and severely degrading the detection of target samples, which are crucial for effective BCI applications.

4.2 Fine-tuning

To further refine and optimize the parameters of the pre-trained EEG-DBNet-V2, enhancing its ability to learn data characteristics specific to the target samples, we conduct fine-tuning using surplus non-target samples and generated pseudo-target samples.

EEG-GANet is first trained with an equal number of target and non-target samples using the Adam optimizer (learning rate: 0.0001, batch size: 64) for ten training rounds, each consisting of 5000 epochs. During fine-tuning, we gradually augment the target samples by dividing the difference between the number of non-target and target samples into ten equal parts, as illustrated in Fig. 6.

In each iteration, a specific number of non-target samples are input into the generator to produce an equal number of pseudo-target samples. These generated pseudo-target samples, along with an equal number of previously unused non-target samples, are concatenated to the original training set to form an augmented training set. This updated training set is subsequently used to fine-tune EEG-DBNet-V2, following the same Adam optimizer settings as in the pre-training phase.

Since EEG-GANet underwent ten separate training iterations, ten different generators are obtained. Each generator's output is used for a corresponding fine-tuning iteration, resulting in ten fine-tuning cycles of 100 epochs each. For each pre-trained model obtained from the tenfold cross-validation, we select the best-performing model from its corresponding ten fine-tuning iterations based on validation set performance. These selected models are then evaluated on the test set.

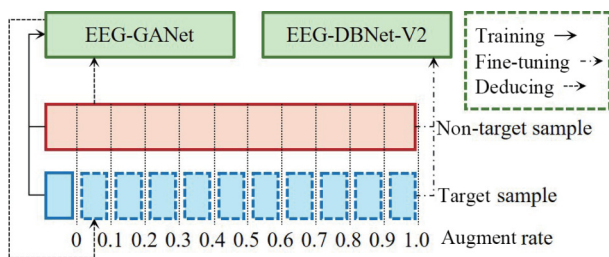


Fig. 6 Fine-tuning process of EEG-DBNet-V2.

4.2.1 Comparison with alternative data augmentation methods

To assess the effectiveness of EEG-GANet for data augmentation, we compare its performance with five alternative augmentation methods: DCGAN^[24], VAE^[26], oversampling, synthetic minority oversampling technique (SMOTE)^[48], and direct training. DCGAN employs a traditional generative adversarial network structure identical to EEG-GANet's generator and discriminator; VAE utilizes random noise for generating augmented samples; oversampling replicates existing target samples to balance class distributions; and SMOTE generates synthetic target samples by interpolating between existing samples. The direct training method involves training EEG-DBNet-V2 directly on augmented data generated by EEG-GANet without pre-training and fine-tuning.

Table 5 summarizes the classification performance of each augmentation approach on the P300-ALS and P300-normal datasets. All augmentation ratios are selected by prioritizing the highest TP rates achieved on the test set.

The results show that EEG-GANet achieves the highest classification accuracy ($P = 81.88\%$) and F1-score ($F_1 = 0.7225$) on the P300-ALS dataset. On the P300-normal dataset, EEG-GANet again demonstrates superior classification accuracy ($P = 76.26\%$), slightly outperforming other methods, although oversampling achieves a marginally higher F1-score ($F_1 = 0.6595$). Notably, EEG-GANet consistently outperforms DCGAN and VAE, indicating its ability to generate more realistic and diverse EEG samples. Furthermore, EEG-GANet surpasses the direct training approach,

Table 5 Performance comparison of different data augmentation methods on the P300-ALS and P300-normal datasets.

Data augmentation method	P300-ALS		P300-normal	
	P (%)	F_1	P (%)	F_1
DCGAN ^[24]	81.06	0.7208	75.93	0.6505
VAE ^[26]	79.61	0.6752	71.71	0.5401
Oversample [§]	78.36	0.7037	74.63	0.6595
SMOTE	78.51	0.7044	72.31	0.6365
Direct training [◇]	79.68	0.7090	75.60	0.6582
EEG-GANet	81.88	0.7225	76.26	0.6474

Note: [§] Replicating existing target samples to balance class distributions. [◇] Training EEG-DBNet-V2 directly on augmented data generated by EEG-GANet without pre-training and fine-tuning. Bold values indicate the best performance.

confirming the advantage of the proposed fine-tuning strategy in enhancing EEG-DBNet-V2's generalization capability.

4.2.2 Quality evaluation of generated EEG samples

To evaluate the quality of augmented EEG samples generated by EEG-GANet, we visually compare the pseudo-target samples with real samples (target and non-target) using averaged EEG waveforms, scalp topographies, and time-frequency representations. Figure 7 illustrates these comparisons for Subject 1 on the P300-ALS dataset.

The averaged EEG waveforms in Figs. 7a, 7b, and 7c depict real target samples, real non-target samples, and pseudo-target samples, respectively. It is evident that

pseudo-target samples exhibit waveform characteristics closely aligned with real target samples, especially the prominent P300 oscillation around 300 ms. Furthermore, the scalp topographies from 293 ms to 387 ms (Figs. 7d, 7e, and 7f) show that pseudo-target samples (Fig. 7f) share spatial activation patterns highly similar to real target samples (Fig. 7d), whereas real non-target samples (Fig. 7e) differ markedly. Additionally, the time-frequency maps (Figs. 7g, 7h, and 7i) further confirm that pseudo-target samples reproduce the frequency-domain features of real target samples, with notable activity concentrated around 15 Hz during the critical P300 interval (200–400 ms).

Beyond visual comparison, we further assess the quality of generated EEG samples using the root mean

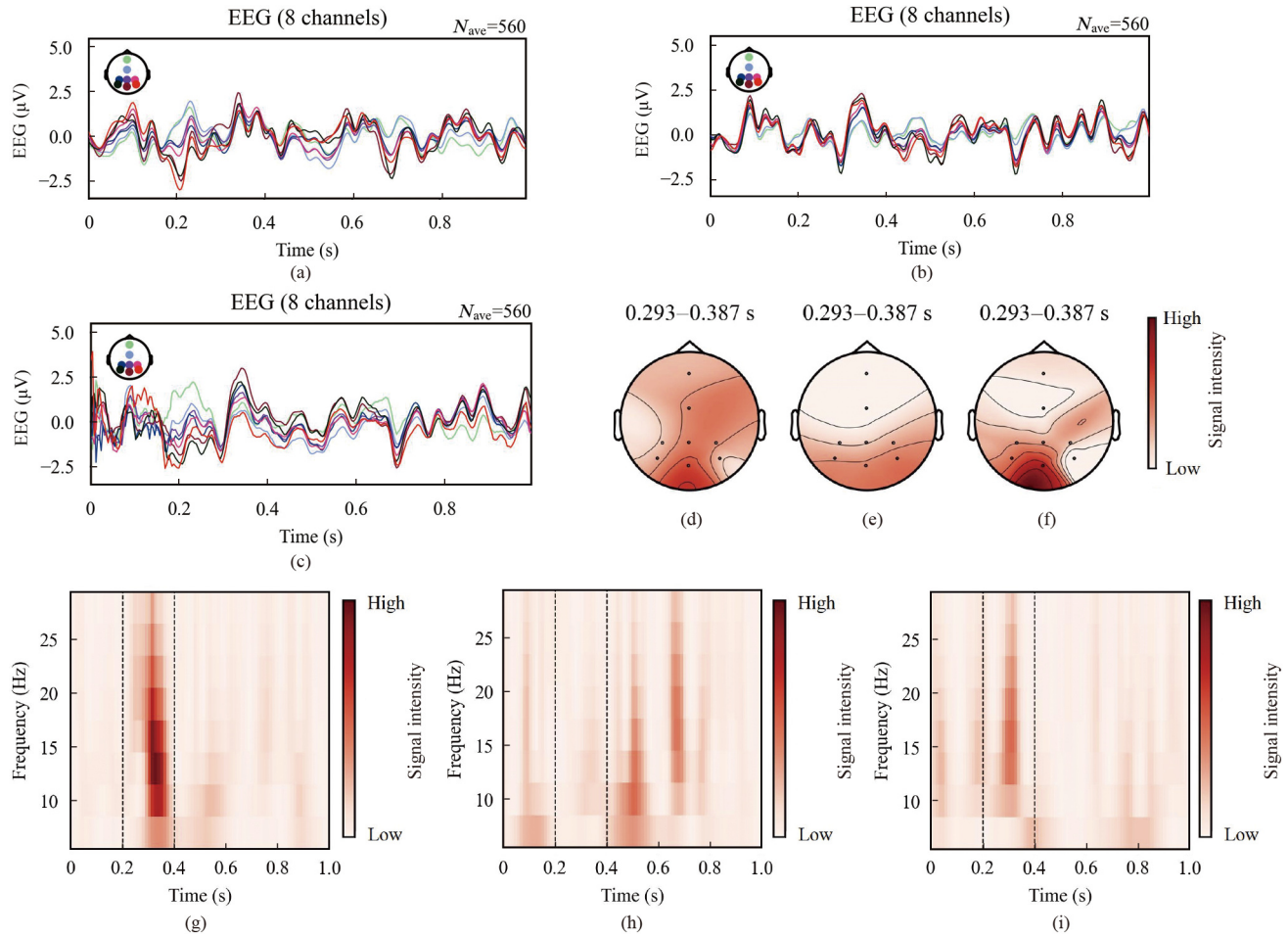


Fig. 7 Comparison of pseudo-target samples and real samples for the first subject. (a) EEG of real target samples. (b) EEG of real non-target samples. (c) EEG of pseudo-target samples. (d) Topographic map of real target samples. (e) Topographic map of real non-target samples. (f) Topographic map of pseudo-target samples. (g) Time–frequency map of real target samples. (h) Time–frequency map of real non-target samples. (i) Time–frequency map of pseudo-target samples. The dimensions shown in the figure correspond to the P300-ALS dataset, where $T = 256$ (number of sampling points per epoch) and $C = 8$ (number of EEG electrodes). The color bars in (d)–(i) indicate signal intensity, where warmer colors represent higher potential amplitude or power and lighter colors indicate lower values. The colored dots in (a)–(c) indicate the positions of the eight EEG electrodes.

square error (RMSE), which quantifies the similarity between generated pseudo-target samples and real target samples. Table 6 presents the RMSE results across eight subjects in the P300-ALS dataset for three augmentation methods: DCGAN, VAE, and EEG-GANet.

The results indicate that EEG-GANet consistently outperforms DCGAN and VAE across multiple subjects, achieving the lowest RMSE for five out of eight subjects. The mean RMSE across all subjects is also the lowest for EEG-GANet (7.6794 μV), compared to DCGAN (10.3130 μV) and VAE (9.3924 μV). This suggests that EEG-GANet generates pseudo-target samples that are more similar to real target samples in terms of overall signal characteristics. While VAE achieves the lowest RMSE for two subjects (Subjects 3 and 8), its overall mean RMSE is slightly higher than EEG-GANet, indicating that its performance is more dataset-dependent.

In summary, both qualitative (visual) and quantitative (RMSE) analyses confirm that EEG-GANet produces the most physiologically relevant and accurate augmented EEG signals among the evaluated methods.

5 Conclusion

This study address critical challenges in P300 speller systems, including low SNR, sample imbalance, and computational inefficiency. We introduce EEG-GANet, a generative adversarial network designed to augment P300 target samples by exploiting the intrinsic relationships between target and non-target EEG signals. Compared to traditional GANs, EEG-GANet

Table 6 RMSE comparison of generated EEG samples using different augmentation methods on the P300-ALS dataset.

(Unit: μV)			
Subject	DCGAN	VAE	EEG-GANet
1	7.5827	10.9310	10.1819
2	10.2990	10.4452	7.5827
3	8.1328	7.2168	8.2962
4	12.7283	9.0853	5.2909
5	9.2225	10.2232	7.1101
6	11.5275	7.8822	7.3300
7	11.0787	13.1548	9.3047
8	11.7513	6.2006	6.3385
Mean	10.3130	9.3924	7.6794

Note: Bold values indicate the best performance.

effectively mitigates mode collapse, generating pseudo-target samples with physiological fidelity closely aligns to genuine EEG characteristics.

Moreover, we propose EEG-DBNet-V2, a compact dual-branch classification network combining temporal and spectral feature extraction. Despite its minimal computational footprint, EEG-DBNet-V2 achieves superior or comparable accuracy against existing models, confirming its effectiveness in practical BCI applications.

Extensive experiments demonstrates that combining EEG-GANet's data augmentation with EEG-DBNet-V2 significantly improves classification performance, particularly under conditions of severe class imbalance. Fine-tuning EEG-DBNet-V2 with EEG-GANet-generated samples consistently yields higher accuracy and robustness compared to conventional methods such as DCGAN, VAE, oversampling, and SMOTE.

Overall, our approach highlights the importance of targeted data augmentation and computationally efficient model design, offering a robust and practical solution for enhancing real-world P300 speller performance. Future research directions include exploring transfer learning strategies to further reduce user-specific calibration efforts and extending our methods to other EEG-based BCI paradigms.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (Nos. 62171073, 62311530103, and 62106032); the Natural Science Foundation of Chongqing, China (No. CSTB2023NSCQ-LZX0064); Chongqing Chuying Project (Nos. CY240610 and CY240605); Sichuan Science and Technology Program (No. 2024NSFSC1662); Chongqing Scientific Research Innovation Project for Postgraduate Students (No. CYB23240); and the Doctoral Training Program of Chongqing University of Posts and Telecommunications (No. BYJS202317).

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

References

- [1] R. A. Ramadan and A. V. Vasilakos, Brain computer interface: Control signals review, *Neurocomputing*, vol. 223, pp. 26–44, 2017.
- [2] Y. Yu, Y. Liu, E. Yin, J. Jiang, Z. Zhou, and D. Hu, An

- asynchronous hybrid spelling approach based on EEG-EOG signals for Chinese character input, *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 6, pp. 1292–1302, 2019.
- [3] K. Won, M. Kwon, M. Ahn, and S. C. Jun, EEG dataset for RSVP and P300 speller brain-computer interfaces, *Sci. Data*, vol. 9, p. 1, 2022.
- [4] E. W. Sellers and E. Donchin, A P300-based brain-computer interface: Initial tests by ALS patients, *Clinical neurophysiol.*, vol. 117, no. 3, pp. 538–548, 2006.
- [5] C. F. Blanco-Díaz, C. D. Guerrero-Méndez, and A. F. Ruiz-Olaya, Enhancing P300 detection using a band-selective filter bank for a visual P300 speller, *IRBM*, vol. 44, no. 3, p. 100751, 2023.
- [6] A. Riccio, L. Simione, F. Schettini, A. Pizzimenti, M. Inghilleri, M. O. Belardinelli, D. Mattia, and F. Cincotti, Attention and P300-based BCI performance in people with amyotrophic lateral sclerosis, *Front. Hum. Neurosci.*, vol. 7, p. 732, 2013.
- [7] C. Guger, S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabalona, F. Gramatica, and G. Edlinger, How many people are able to control a P300-based brain-computer interface (BCI)? *Neurosci. Lett.*, vol. 462, no. 1, pp. 94–98, 2009.
- [8] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, BCI2000: A general-purpose brain-computer interface (BCI) system, *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1034–1043, 2004.
- [9] Y. Li, H. Liu, and S. Wang, Exploiting EEG channel correlations in P300 speller paradigm for brain-computer interface, *IEICE Trans. Inf. Syst.*, vol. E99-D, no. 6, pp. 1653–1662, 2016.
- [10] J. Qu, F. Wang, Z. Xia, T. Yu, J. Xiao, Z. Yu, Z. Gu, and Y. Li, A novel three-dimensional P300 speller based on stereo visual stimuli, *IEEE Trans. Hum.-Mach. Syst.*, vol. 48, no. 4, pp. 392–399, 2018.
- [11] S. Kundu and S. Ari, P300 detection with brain-computer interface application using PCA and ensemble of weighted SVMs, *IETE J. Res.*, vol. 64, no. 3, pp. 406–414, 2018.
- [12] F. Akram, S. M. Han, and T. S. Kim, An efficient word typing P300-BCI system using a modified T9 interface and random forest classifier, *Comput. Biol. Med.*, vol. 56, pp. 30–36, 2015.
- [13] H. Cecotti and A. Graser, Convolutional neural networks for P300 detection with application to brain-computer interfaces, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 3, pp. 433–445, 2011.
- [14] R. Manor and A. B. Geva, Convolutional neural network for multi-category rapid serial visual presentation BCI, *Front. Comput. Neurosci.*, vol. 9, p. 146, 2015.
- [15] M. Liu, W. Wu, Z. Gu, Z. Yu, F. F. Qi, and Y. Li, Deep learning based on batch normalization for P300 signal detection, *Neurocomputing*, vol. 275, pp. 288–297, 2018.
- [16] S. Ioffe and C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, arXiv preprint arXiv: 1502.03167, 2015.
- [17] A. F. Agarap, Deep learning using rectified linear units (ReLU), arXiv preprint arXiv: 1803.08375, 2019.
- [18] H. Shan, Y. Liu, and T. P. Stefanov, A simple convolutional neural network for accurate P300 detection and character spelling in brain computer interface, in *Proc. 27th Int. Joint Conf. Artificial Intelligence*, Stockholm, Sweden, 2018, pp. 1604–1610.
- [19] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, EEGNet: A compact convolutional neural network for EEG-based brain-computer interfaces, *J. Neural Eng.*, vol. 15, no. 5, p. 056013, 2018.
- [20] M. Alvarado-González, G. Fuentes-Pineda, and J. Cervantes-Ojeda, A few filters are enough: Convolutional neural network for P300 detection, *Neurocomputing*, vol. 425, pp. 37–52, 2021.
- [21] E. Santamaría-Vázquez, V. Martínez-Cagigal, F. Vaquerizo-Villar, and R. Hornero, EEG-inception: A novel deep convolutional neural network for assistive ERP-based brain-computer interfaces, *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, p. 12, 2020.
- [22] S. Bagchi and D. R. Bathula, EEG-ConvTransformer for single-trial EEG-based visual stimulus classification, *Pattern Recogn.*, vol. 26, p. 108757, 2022.
- [23] N. K. N. Aznan, A. Atapour-Abarghouei, S. Bonner, J. D. Connolly, N. Al Moubayed, and T. P. Breckon, Simulating brain signals: Creating synthetic eeg data via neural-based generative models for improved ssvep classification, in *Proc. 2019 Int. Joint Conf. Neural Networks*, Budapest, Hungary, 2019, pp. 1–8.
- [24] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, Generative adversarial networks, *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020.
- [25] M. Arjovsky, S. Chintala, and L. Bottou, Wasserstein generative adversarial networks, in *Proc. 34th Int. Conf. Machine Learning*, Sydney, Australia, 2017, pp. 214–223.
- [26] L. P. Cinelli, M. A. Marins, E. A. B. da Silva, and S. L. Netto, Variational autoencoder, in *Variational Methods for Machine Learning with Applications to Deep Networks*, Cham, Switzerland: Springer, 2021, pp. 111–149.
- [27] P. Sawangjai, M. Trakulruangroj, C. Boonnag, M. Piriyaikitakonkij, R. K. Tripathy, T. Sudhawiyangkul, and T. Wilaiprasitporn, EEGANet: Removal of ocular artifacts from the EEG signal using generative adversarial networks, *IEEE J. Biomed. Health Inform.*, vol. 26, no. 10, pp. 4913–4924, 2022.
- [28] Y. Zhang, Y. Yu, H. Li, A. Wu, L.-L. Zeng, and D. Hu, MASER: Enhancing EEG spatial resolution with state space modeling, *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 32, pp. 3858–3868, 2024.
- [29] A. Dithaporn, N. Banluesombatkul, S. Kettrat, E. Chuangsuwanich, and T. Wilaiprasitporn, Universal joint feature extraction for P300 EEG classification using multi-task autoencoder, *IEEE Access*, vol. 7, pp. 68415–68428, 2019.

- [30] P. Penava and R. Buettner, A novel small-data based approach for decoding yes/no-decisions of locked-in patients using generative adversarial networks, *IEEE Access*, vol. 11, pp. 118849–118864, 2023.
- [31] L. Xu, M. Skoularidou, A. Cuesta-Infante, and K. Veeramachaneni, Modeling tabular data using conditional gan, in *Proc. 33rd Int. Conf. Neural Information Processing Systems*, Vancouver, Canada, 2019, pp. 7335–7345.
- [32] A. Cutler, D. R. Cutler, and J. R. Stevens, Random forests, in *Ensemble Machine Learning: Methods and Applications*, C. Zhang and Y. Ma, eds. New York, NY, USA: Springer, 2012, pp. 157–175.
- [33] F. Xu, F. Rong, J. Leng, T. Sun, Y. Zhang, S. Siddharth, and T. P. Jung, Classification of left-versus right-hand motor imagery in stroke patients using supplementary data generated by CycleGAN, *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 2417–2424, 2021.
- [34] C. Chu, A. Zhmoginov, and M. Sandler, Cyclegan, a master of steganography, arXiv preprint arXiv: 1712.02950, 2017.
- [35] V. Bhandari, N. D. Londhe, and G. B. Kshirsagar, TL-P3GAN: An efficient temporal-learning-based generative adversarial network for precise P300 signal generation for P300 spellers, *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 16, no. 2, pp. 692–705, 2024.
- [36] Y. Pan, N. Li, Y. Zhang, P. Xu, and D. Yao, Short-length SSVEP data extension by a novel generative adversarial networks based framework, *Cogn. Neurodyn.*, vol. 18, no. 5, pp. 2925–2945, 2024.
- [37] S. Li, I. Daly, C. Guan, A. Cichocki, and J. Jin, Inter-participant transfer learning with attention based domain adversarial training for P300 detection, *Neural Netw.*, vol. 180, p. 106655, 2024.
- [38] X. Lou, X. Li, H. Meng, J. Hu, M. Xu, Y. Zhao, J. Yang, and Z. Li, EEG-DBNet: A dual-branch network for temporal-spectral decoding in motor-imagery brain-computer interfaces, arXiv preprint arXiv: 2405.16090, 2024.
- [39] F. Chollet, Xception: Deep learning with depthwise separable convolutions, in *Proc. IEEE 2017 Conf. Computer Vision and Pattern Recognition*, Honolulu, HI, USA, 2017, pp. 1800–1807.
- [40] A. Van Den Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, and K. Kavukcuoglu, Wavenet: A generative model for raw audio, arXiv preprint arXiv: 1609.03499, 2016.
- [41] T. Kim, M. Cha, H. Kim, J. K. Lee, and J. Kim, Learning to discover cross-domain relations with generative adversarial networks, in *Proc. 34th Int. Conf. Machine Learning*, Sydney, Australia, 2017, pp. 1857–1865.
- [42] O. Ronneberger, P. Fischer, and T. Brox, U-net: Convolutional networks for biomedical image segmentation, in *Proc. 18th Int. Conf. Medical Image Computing and Computer-Assisted Intervention MICCAI 2015*, Munich, Germany, 2015, pp. 234–241.
- [43] Y. Wu and K. He, Group normalization, in *Proc. 15th European Conf. Computer Vision*, Munich, Germany, 2018, pp. 3–19.
- [44] J. Xu, Z. Li, B. Du, M. Zhang, and J. Liu, Reluplex made more practical: Leaky ReLU, in *Proc. 2020 IEEE Symp. Computers and Communications*, Rennes, France, 2020, pp. 1–7.
- [45] A. U. Ruby, P. Theerthagiri, I. J. Jacob, and Y. Vamsidhar, Binary cross entropy with deep learning technique for image classification, *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 4, pp. 5393–5397, 2020.
- [46] S. Purushotham and B. K. Tripathy, Evaluation of classifier models using stratified tenfold cross validation techniques, in *Proc. 4th Int. Conf. Global Trends in Information Systems and Software Application*, Vellore, India, 2011, pp. 680–690.
- [47] R. van Dinteren, M. Arns, M. L. A. Jongsma, and R. P. C. Kessels, Combined frontal and parietal P300 amplitudes indicate compensated cognitive processing across the lifespan, *Front. Aging Neurosci.*, vol. 6, p. 294, 2014.
- [48] T. Zhu, Y. Lin, and Y. Liu, Synthetic minority oversampling technique for multiclass imbalance problems, *Pattern Recogn.*, vol. 72, pp. 327–340, 2017.



Xicheng Lou received the MS degree from Chongqing Technology and Business University, Chongqing, China, in 2022. He is currently a PhD candidate at School of Communications and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing, China. His research interests include deep learning and electroencephalography analysis.



Xinwei Li received the BS and PhD degrees from Beihang University, Beijing, China, in 2011 and 2018, respectively. From 2018 to 2021, she served as an assistant professor at the Research Center of Biomedical Engineering, Chongqing University of Posts and Telecommunications, Chongqing, China.

From 2022 to 2023, she was promoted to an associate professor. In 2024, she became a Wenfeng Professor at School of Life Health Information Science and Engineering, Chongqing University of Posts and Telecommunications, Chongqing, China. She is currently a professor at the Research Center of Biomedical Engineering, Chongqing University of Posts and Telecommunications, Chongqing, China. Her research focuses on medical image processing, brain science, and artificial intelligence.



Hongying Meng received the PhD degree in communication and electronic systems from Xi'an Jiaotong University, Xi'an, China, in 1998. He is currently a reader at Department of Electronic and Electrical Engineering, Brunel University London, London, UK. He has authored more than 130 publications. His research interests include digital signal processing, affective computing, machine learning, human–computer interaction, and computer vision.



Peiyang Li received the PhD degree in biomedical engineering from University of Electronic Science and Technology of China, Chengdu, China, in 2018. He is currently an associate professor at the Research Center of Biomedical Engineering, Chongqing University of Posts and Telecommunications, Chongqing, China. His research focuses on brain-computer interaction, convex optimization, machine learning, and pattern recognition.



Lisha Zhong received the PhD degree from Chongqing University of Posts and Telecommunications, Chongqing, China, in 2022. She is currently an associate professor at Southwest Medical University, Luzhou, China. Her research interests include image processing and signal analysis.



Zhangyong Li received the PhD degree from Chongqing Medical University, Chongqing, China, in 2004. He is currently a professor at Chongqing University of Posts and Telecommunications, Chongqing, China. His main research interests include biomedical information processing, digital medical instrument design, electromagnetic radiation protection technology, environmental assessment, and reliability of medical instruments.