

FreqTrans-Meta: Subject-Adaptive Cross-Subject SSVEP Decoding via Frequency-Domain Transformer and Meta-Learning

Miao Sun^{*†}, Jian Zhang^{*‡}, Xiaohong Lan^{*§}

^{*} College of Computer and Information Science, Chongqing Normal University, Chongqing, China

[†]281452063@qq.com, ORCID: 0000-0002-2269-9540

[‡]3315127040@qq.com, ORCID: 0009-0009-0919-8424

[§]2446082755@qq.com, ORCID: 0009-0008-5682-7046

Abstract—Cross-subject decoding in steady-state visual evoked potential (SSVEP) brain-computer interfaces remains a significant challenge due to pronounced inter-subject variability and the scarcity of calibration data. Traditional spatial filtering methods often experience severe performance degradation in this setting, while existing deep learning models are prone to overfitting or catastrophic forgetting under few-shot calibration conditions. To address these challenges, this paper proposes FreqTrans-Meta, a unified framework designed to reconcile high-dimensional representation learning with data-efficient adaptation.

The proposed framework integrates three components, each targeting a specific deficiency in current decoding pipelines. First, to address the inherent spectral nature of SSVEP signals, a Frequency-Domain Channel-wise Transformer backbone is constructed. Unlike time-domain models, this backbone explicitly encodes stimulus frequencies and harmonic structures in the range of 5–45 Hz, capturing robust subject-invariant spectral dependencies. Second, to resolve the conflict between model capacity and data scarcity, a lightweight residual Subject Adapter is introduced. This decouples individual variability from general features, preventing overfitting by freezing the backbone parameters. Third, to enable rapid convergence with minimal data, a Reptile-based meta-learning strategy is employed to optimize the initialization of the adapter. This ensures the model can effectively personalize using fewer than four calibration trials per class.

Extensive experiments on the BETA (70 subjects) and Benchmark (35 subjects) datasets demonstrate that FreqTrans-Meta achieves average classification accuracies of 85.04% and 87.51%, respectively. By logically integrating spectral modeling, parameter-efficient adaptation, and meta-initialization, the proposed method significantly outperforms state-of-the-art baselines such as SSVEPformer and TDCA.

Index Terms—SSVEP, brain-computer interface, cross-subject decoding, Transformer, meta-learning, contrastive learning, few-shot calibration

I. INTRODUCTION

Steady-state visual evoked potential (SSVEP) based brain-computer interfaces (BCIs) are valued for their high information transfer rate (ITR) and minimal training requirements [1]– [6]. However, reliable **cross-subject** SSVEP decoding

remains a fundamental bottleneck. Pronounced inter-subject variability leads to significant domain shifts [32], [33], causing models trained on source subjects to suffer severe performance degradation when applied to unseen targets. While calibration can mitigate this, collecting ample data is tedious and fatigues the user. Thus, the core challenge lies in achieving high performance with extremely limited calibration data (e.g., fewer than five trials).

Traditional methods like TRCA [12] and TDCA [14] often struggle to capture complex non-linear inter-subject variations. Conversely, deep learning models (e.g., SSVEPformer [19]) are data-hungry; directly fine-tuning high-dimensional networks on a handful of calibration trials inevitably leads to **overfitting** or catastrophic forgetting. Although transfer learning [21]– [25] has been explored, existing techniques often lack explicit spectral modeling or require computationally expensive adaptation processes.

To circumvent these limitations, we propose **FreqTrans-Meta**, a cohesive framework designed to simultaneously address explicit spectral modeling, parameter efficiency, and rapid initialization:

- 1) **Explicit Spectral Modeling:** Generic time-series models often struggle to efficiently capture the strict harmonic structures of SSVEP signals. We propose a **Frequency-Domain Channel-wise Transformer** that transforms signals into the frequency domain to explicitly encode harmonic dependencies and learn subject-invariant spectral patterns.
- 2) **Parameter-Efficient Adaptation:** To prevent overfitting during personalization, we introduce a lightweight **Subject Adapter (SA)**. By freezing the shared backbone and updating only this low-rank module, we decouple general representation learning from individual adaptation, handling subject-specific shifts with a minimal parameter set.
- 3) **Rapid Initialization via Meta-Learning:** Simply adding an adapter does not guarantee convergence with minimal data. We employ **Reptile-based Meta-Learning** [29] to optimize the adapter’s *initialization* to be highly sensitive to gradients, enabling the model

This work was supported by the National Undergraduate Innovation and Entrepreneurship Training Program (202510637052) and the University-level Undergraduate Innovation and Entrepreneurship Training Program (X202510637037).

to rapidly adapt to new subjects using fewer than four calibration trials.

The main contributions of this work are summarized as follows:

- We propose a frequency-domain channel-wise Transformer that explicitly captures the harmonic structure and spatial dependencies of SSVEP signals.
- We design a parameter-efficient Subject Adapter with a Reptile-optimized initialization, specifically targeting the instability of few-shot calibration.
- Extensive evaluations on large-scale public datasets (BETA [13] and Benchmark [27]) under strict leave-one-subject-out protocols demonstrate consistent improvements over state-of-the-art methods.

II. RELATED WORK

A. Traditional SSVEP Decoding Methods

Early SSVEP-based BCIs primarily relied on correlation analysis and spatial filtering techniques, such as canonical correlation analysis (CCA), filter-bank CCA (FBCCA), task-related component analysis (TRCA), and task-discriminant component analysis (TDCA) [10]–[14]. These methods effectively exploit spectral peaks and harmonic information of SSVEP signals and achieve competitive performance in within-subject settings. However, their reliance on linear assumptions and subject-specific calibration limits robustness and scalability in cross-subject scenarios.

B. Deep Learning and Transformer-Based Approaches

With the advancement of deep learning, convolutional neural networks and Transformer-based models have been introduced for SSVEP decoding to automatically learn discriminative representations from EEG signals [15]–[20]. Frequency-domain modeling and self-attention mechanisms have demonstrated improved robustness in short time-window and high-target settings. Nevertheless, most deep learning approaches involve large parameter spaces and remain sensitive to limited training data, making them prone to overfitting under few-shot calibration or online adaptation conditions. Moreover, many existing studies focus on within-subject or weak cross-subject settings, offering limited reduction in calibration cost.

C. Cross-Subject Adaptation and Meta-Learning

To alleviate inter-subject variability, transfer learning, domain adaptation, and meta-learning strategies have been actively explored [21]–[27]. Domain adaptation methods aim to align feature distributions across subjects, while meta-learning approaches, such as MAML and Reptile, learn model initializations that can be rapidly adapted using a few calibration trials. Despite these advances, existing methods often lack explicit modeling of SSVEP spectral structures and channel-wise dependencies. Furthermore, subject adaptation typically requires fine-tuning high-dimensional model parameters, leading to instability and unreliable performance under extremely limited calibration data.

In contrast, our method integrates frequency-domain priors, channel-wise Transformer learning, and parameter-efficient meta-learned subject adaptation within a unified framework, enabling robust and data-efficient cross-subject SSVEP decoding.

III. METHOD

A. Overall Framework

We propose a cross-subject SSVEP decoding framework that integrates frequency-domain modeling, a channel-wise Transformer backbone, and meta-learning-driven subject adaptation. As illustrated in Fig. 1, the framework consists of four components: (i) a frequency-domain channel-wise Transformer (FreqTrans), (ii) a lightweight Subject Adapter (SA), (iii) Reptile-based meta-learning for few-shot calibration, and (iv) robust training and inference strategies, including supervised contrastive learning, Mixup, and test-time adaptation.

Given a multi-channel EEG trial

$$\mathbf{X} \in \mathbb{R}^{C \times T}, \quad (1)$$

the decoding objective is to learn a mapping

$$f_{\theta} : \mathbb{R}^{C \times T} \rightarrow \mathbb{R}^K, \quad (2)$$

where C is the number of channels, T the temporal length, and K the number of stimulus targets.

Each trial is standardized and transformed into the frequency domain using FFT. Spectral components within 5–45 Hz are retained, and real and imaginary parts are concatenated to form a frequency-domain representation $\mathbf{F}$. Treating channels as sequence elements, $\mathbf{F}$ is fed into the channel-wise Transformer to learn a cross-subject shared representation $\mathbf{H}_{\text{shared}}$. A residual Subject Adapter then produces subject-specific features, which are used for final classification. Meta-learning is applied to optimize the adapter initialization for stable few-shot adaptation. During inference, performance-safe model selection, time-shift test-time augmentation, and entropy minimization are jointly applied to enhance robustness.

B. Frequency-Domain Transformer Backbone

1) *Frequency-Domain Feature Construction*: SSVEP discriminative information is mainly concentrated at stimulus frequencies and their harmonics. After preprocessing, FFT is applied to each channel:

$$\mathbf{Z}_c = \text{FFT}(\mathbf{x}_c), \quad (3)$$

and frequency bins corresponding to 5–45 Hz are retained. The complex spectrum is decomposed into real and imaginary parts:

$$\mathbf{f}_c = [\text{Re}(\mathbf{Z}_c), \text{Im}(\mathbf{Z}_c)] \in \mathbb{R}^{2S}, \quad (4)$$

yielding the frequency-domain feature matrix

$$\mathbf{F} = [\mathbf{f}_1; \dots; \mathbf{f}_C] \in \mathbb{R}^{C \times 2S}. \quad (5)$$

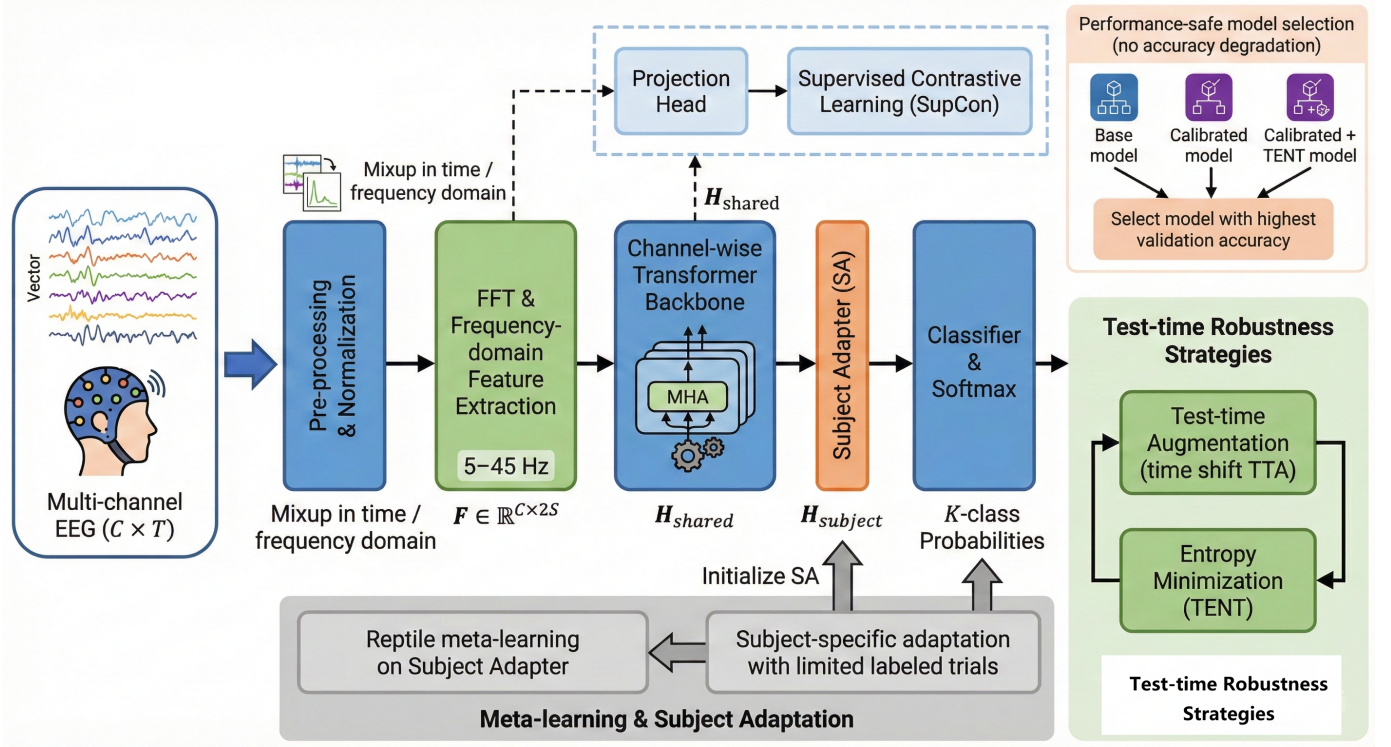


Fig. 1. Overall architecture of the proposed FreqTrans-Meta framework for cross-subject SSVEP decoding.

2) *Channel-wise Transformer and Classification*: A channel-combination module first linearly mixes spectral features:

$$\mathbf{H}^{(0)} = \text{Dropout}(\phi(\text{LN}(\text{Conv}_{1 \times 1}(\mathbf{F}))))). \quad (6)$$

The features are then processed by a two-layer channel-wise Transformer:

$$\mathbf{H}'_l = \mathbf{H}_{l-1} + \text{Dropout}(\text{MHA}(\text{LN}(\mathbf{H}_{l-1}))), \quad (7)$$

$$\mathbf{H}_l = \mathbf{H}'_l + \text{FFN}(\text{LN}(\mathbf{H}'_l)), \quad (8)$$

producing the shared representation

$$\mathbf{H}_{\text{shared}} = \mathbf{H}^{(2)}. \quad (9)$$

The classification head outputs

$$\mathbf{p} = \text{Softmax}(\text{FC}(\text{Flatten}(\text{LN}(\mathbf{H}_{\text{shared}})))). \quad (10)$$

3) *Supervised Contrastive Learning*: A projection head maps $\mathbf{H}_{\text{shared}}$ to a normalized embedding $\mathbf{g}$. The training objective combines cross-entropy and supervised contrastive loss:

$$\mathcal{L} = \mathcal{L}_{\text{CE}} + \lambda \mathcal{L}_{\text{SupCon}}. \quad (11)$$

Frequency-consistent Mixup is applied to improve robustness.

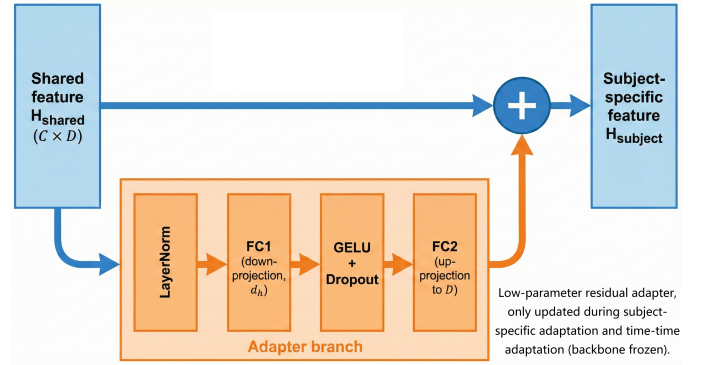


Fig. 2. Internal structure of the proposed Subject Adapter (SA).

C. Subject Adapter and Meta-Learning

1) *Residual Subject Adapter*: To model inter-subject variability with limited calibration data, a parameter-efficient residual Subject Adapter is introduced:

$$\mathbf{U} = \text{LN}(\mathbf{H}_{\text{shared}}), \quad (12)$$

$$\Delta \mathbf{H} = \phi(\mathbf{U} \mathbf{W}_1^T) \mathbf{W}_2^T, \quad (13)$$

$$\mathbf{H}_{\text{subject}} = \mathbf{H}_{\text{shared}} + \Delta \mathbf{H}. \quad (14)$$

2) *Reptile-Based Few-Shot Calibration*: Under the LOSO protocol, only a few labeled trials are available for a new subject. Reptile meta-learning is used to optimize the adapter

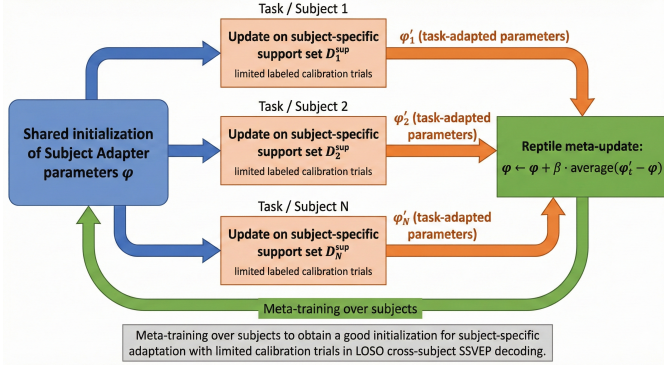


Fig. 3. Reptile-based meta-training process for the Subject Adapter.

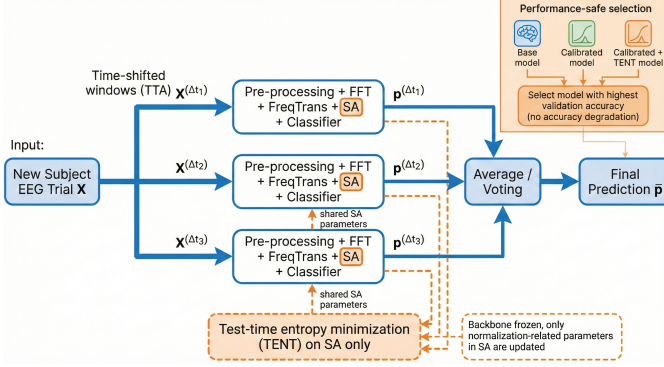


Fig. 4. Robust inference strategies used in the proposed framework.

initialization:

$$\phi \leftarrow \phi + \beta \frac{1}{|T|} \sum_{t \in T} (\phi'_t - \phi), \quad (15)$$

enabling rapid and stable few-shot adaptation.

D. Robust Inference

As shown in Fig. 4, robustness is further enhanced during inference using three strategies. A performance-safe constraint selects the calibrated model only when it outperforms the baseline *on the labeled calibration set* of the target subject (denoted as D_{cal}), i.e., $\text{Acc}_{\text{sel}}(\cdot)$ is computed using D_{cal} only:

$$\text{Acc}_{\text{sel}}(M_{\text{cal}}) \geq \text{Acc}_{\text{sel}}(M_{\text{base}}). \quad (16)$$

Time-shift test-time augmentation averages predictions across temporal offsets:

$$\mathbf{p} = \frac{1}{L} \sum_{l=1}^L \mathbf{p}^{(\delta_l)}, \quad (17)$$

while test-time entropy minimization minimizes prediction uncertainty using *only unlabeled test data* without accessing any ground-truth labels:

$$\mathcal{L}_{\text{TENT}} = \mathbb{E}_{\mathbf{X}} \left[- \sum_{k=1}^K p_k(\mathbf{X}) \log p_k(\mathbf{X}) \right]. \quad (18)$$

IV. EXPERIMENTS AND RESULTS

To verify the effectiveness of the proposed cross-subject SSVEP decoding method based on a frequency-domain Transformer and meta-learning-driven subject-adaptive calibration, we conducted systematic evaluations on two large-scale public cross-subject SSVEP datasets.

A. Datasets and Evaluation Protocols

Experiments were conducted on two public large-scale SSVEP benchmark datasets.

BETA Dataset [13]: This dataset contains EEG recordings from 70 subjects performing SSVEP tasks with 40 stimulus frequencies. Signals were acquired using a 64-channel EEG cap at a sampling rate of 250 Hz.

Benchmark Dataset [27]: Released by Tsinghua University, this dataset includes EEG data from 35 healthy subjects under 40 stimulus frequencies, with a collection configuration similar to the BETA dataset and 64 EEG channels.

To balance steady-state response convergence and online real-time performance, a unified windowing strategy was adopted for both datasets. Specifically, a 2 s EEG segment starting 0.5 s after stimulus onset was extracted as the model input for all methods.

Evaluation Protocol: A strict leave-one-subject-out (LOSO) protocol was employed to assess cross-subject generalization. For the BETA dataset, 70-fold experiments were conducted, each using one subject for testing and the remaining 69 subjects for training. Similarly, 35-fold experiments were performed for the Benchmark dataset.

Calibration and Test Split: In calibration experiments, up to four trials per class are randomly selected as the labeled calibration set for each test subject, while the remaining trials are reserved for evaluation. The calibration set is used *only* for subject-level adaptation and model selection; importantly, no labels from the held-out evaluation trials are used at any stage. To ensure determinism and reproducibility, we fix the random seed and will release the split indices along with the code upon acceptance.

B. Comparative Methods

To comprehensively evaluate performance, we selected representative baselines from three categories. All methods were evaluated under the same 2 s window and LOSO protocol.

1) *Traditional Spectral Correlation and Spatial Filtering Methods:* **CCA**: Canonical Correlation Analysis classifies stimuli by computing the correlation between EEG signals and sine-cosine reference templates.

FBCCA: Filter Bank CCA introduces multi-band decomposition and weighted fusion to enhance harmonic utilization.

TRCA: Task-Related Component Analysis learns spatial filters by maximizing the reproducibility of same-class trials.

TDCA: Task-Discriminant Component Analysis incorporates discriminative information to jointly optimize spatial filters and is regarded as a strong baseline for low-frequency SSVEP decoding.

2) *Deep Learning Baselines*: **EEGNet**: A lightweight CNN model using depth-wise separable convolutions to extract spatiotemporal EEG features.

SSVEPformer: A Transformer-based SOTA model that applies self-attention to complex spectral representations.

CNN Baseline: A 1D CNN implemented in this work using time-domain convolutions, channel mixing, and a fully connected classification head.

3) *Proposed Method and Variants*: To evaluate the contribution of meta-learning, two variants were defined.

FreqTrans (w/o Meta): Includes the frequency-domain Transformer backbone and Subject Adapter, trained with cross-entropy, supervised contrastive learning, and Mixup. Inference uses time-shift TTA without meta-learning, TENT, or performance-safe selection.

FreqTrans-Meta (w/ Meta): Extends FreqTrans by introducing Reptile meta-learning for SA initialization, few-shot calibration, TENT adaptation, and performance-safe model selection. Unless otherwise specified, "Our Method" refers to FreqTrans-Meta.

Fair Comparison: All methods are evaluated under the same preprocessing and the same 2 s windowing strategy, using the identical LOSO splits. For each baseline, we follow the original paper and/or official implementations when available; any necessary hyperparameters are determined using only the training subjects (source domain) and are never tuned on the held-out test subject.

C. Overall Cross-Subject Decoding Performance

Fig. 5 compares cross-subject decoding performance on the BETA and Benchmark datasets.

On the BETA dataset, traditional spatial filtering methods achieved accuracies between 43% and 64%, with large standard deviations, indicating strong sensitivity to inter-subject variability. Deep learning baselines such as EEGNet and SSVEPformer achieved accuracies of 71.50% and 81.10%, respectively.

The proposed FreqTrans baseline achieved $83.41\% \pm 16.71\%$, demonstrating the superiority of the frequency-domain Transformer backbone. After incorporating meta-learning, TENT, and performance-safe calibration, FreqTrans-Meta further improved accuracy to 85.04% on BETA and 87.51% on Benchmark, achieving state-of-the-art performance on both datasets.

Notably, FreqTrans-Meta maintained lower standard deviations compared with both traditional and deep learning baselines, indicating improved stability and robustness in cross-subject scenarios.

D. Ablation Study

To analyze the contribution of individual components, seven ablation configurations (A1–A7) were evaluated on the BETA dataset. Only one factor was modified in each configuration, while all other settings were kept consistent. The results are summarized in Table I.

Note that in A5, the Subject Adapter is inserted but *not* updated because no subject-specific calibration is performed in the uncalibrated setting; thus its performance is expected to remain close to A4. Improvements appear once the adapter is calibrated with few-shot data (A6) and further enhanced by meta-initialization (A7).

The ablation results demonstrate that the frequency-domain Transformer backbone provides the primary performance gain, while normalization and contrastive learning improve robustness. The Subject Adapter enables effective individualized correction, and meta-learning combined with test-time adaptation further enhances reliability under few-shot calibration.

E. Summary of Experiments

Overall, the experimental results demonstrate that the proposed FreqTrans-Meta framework consistently outperforms traditional methods, CNN-based models, and existing Transformer-based approaches in cross-subject SSVEP decoding. The ablation study further confirms the complementary contributions of frequency-domain modeling, subject-adaptive calibration, and meta-learning-driven initialization.

V. CONCLUSION

This paper proposed FreqTrans-Meta, a cross-subject SSVEP decoding framework that integrates frequency-domain priors, a channel-wise Transformer backbone, and meta-learning-driven subject adaptation. Extensive experiments on the BETA and Benchmark datasets demonstrated that the proposed method achieves state-of-the-art cross-subject decoding accuracy while requiring only a small number of calibration trials for new subjects.

The frequency-domain Transformer effectively captures stimulus-related spectral and spatial dependencies, while the Subject Adapter and Reptile-based meta-learning provide a stable pathway for few-shot personalization. Robust training and inference strategies further improve system reliability under realistic conditions.

Future work will explore the extension of this framework to asynchronous SSVEP, multimodal BCI paradigms, and real-world closed-loop control tasks, as well as the integration of more advanced meta-learning and self-supervised pre-training strategies.

CODE AVAILABILITY

The source code is available at: <https://github.com/z1z23w/FreqTrans-Meta.git>.

ACKNOWLEDGMENT

This work was supported by the National Undergraduate Innovation and Entrepreneurship Training Program under Grant 202510637052 and the University-level Undergraduate Innovation and Entrepreneurship Training Program under Grant X202510637037. The authors also acknowledge support from the Joint Training Base for Master of Education Postgraduates of Chongqing Normal University and Leju (Chongqing) Robot Technology Co., Ltd.

Cross-subject SSVEP decoding accuracy on BETA and Benchmark datasets

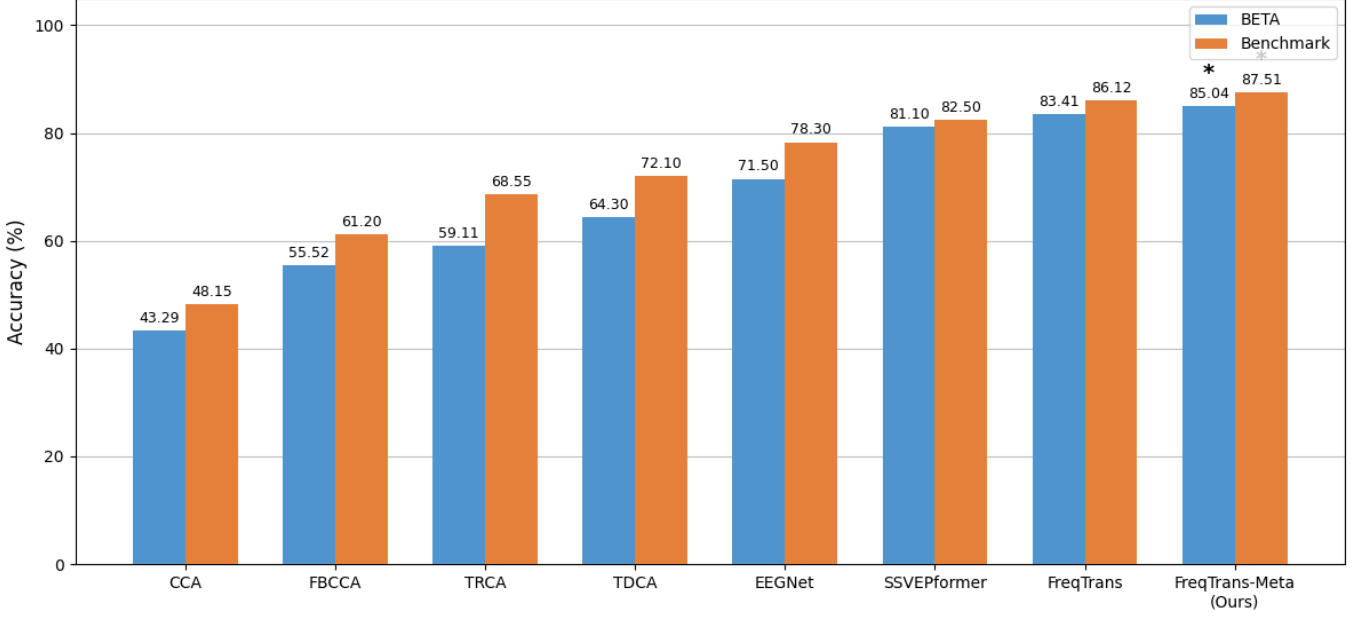


Fig. 5. Comparison of cross-subject decoding performance on the BETA and Benchmark datasets.

TABLE I
ABLATION STUDY RESULTS ON THE BETA DATASET (MEAN $\pm$ STD, %).

Configuration	Description	Uncalibrated	Calibrated
A1	Frequency Features + MLP	33.31 $\pm$ 19.23	33.31 $\pm$ 19.23
A2	Frequency Features + CNN	74.80 $\pm$ 18.50	74.80 $\pm$ 18.50
A3	Frequency Features + Transformer	79.29 $\pm$ 19.85	79.29 $\pm$ 19.85
A4	A3 + Normalization	83.55 $\pm$ 16.96	83.55 $\pm$ 16.96
A5	A4 + Subject Adapter	83.57 $\pm$ 16.69	83.57 $\pm$ 16.69
A6	A5 + Few-shot Calibration	83.80 $\pm$ 16.81	84.22 $\pm$ 16.45
A7	Full Model (Ours)	84.10 $\pm$ 16.50	85.04 $\pm$ 16.37

REFERENCES

- [1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, Jun. 2002.
- [2] S. Gao, Y. Wang, X. Gao, and B. Hong, "Visual and auditory brain-computer interfaces," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 5, pp. 1436–1447, May 2014.
- [3] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, "A review of classification algorithms for EEG-based brain-computer interfaces: A 10-year update," *Journal of Neural Engineering*, vol. 15, no. 3, p. 031005, Jun. 2018.
- [4] Y. Wang, R. Wang, X. Gao, B. Hong, and S. Gao, "A practical VEP-based brain-computer interface," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 234–240, Jun. 2006.
- [5] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, "An online multi-channel SSVEP-based brain-computer interface using canonical correlation analysis," *Journal of Neural Engineering*, vol. 6, no. 4, p. 046002, Aug. 2009.
- [6] D. Zhu, J. Bieger, G. García Molina, and R. M. Aarts, "A survey of stimulation methods used in SSVEP-based brain-computer interfaces," *Computational Intelligence and Neuroscience*, vol. 2010, Art. ID 702357, 2010.
- [7] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T. P. Jung, and S. Gao, "High-speed spelling with a noninvasive brain-computer interface," *Proceedings of the National Academy of Sciences*, vol. 112, no. 44, pp. E6058–E6067, Nov. 2015.
- [8] M. Nakanishi, Y. Wang, Y.-T. Wang, Y. Mitsukura, and T. P. Jung, "A high-speed brain speller using steady-state visual evoked potentials," *International Journal of Neural Systems*, vol. 24, no. 6, p. 1450019, Sep. 2014.
- [9] B. Xu, W. Li, D. Liu, *et al.*, "Continuous hybrid BCI control for robotic arm using noninvasive EEG, computer vision, and eye tracking," *Mathematics*, vol. 10, no. 4, p. 618, 2022.
- [10] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 6, pp. 1172–1176, Jun. 2007.
- [11] X. Chen, Y. Wang, S. Gao, *et al.*, "Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain-computer interface," *Journal of Neural Engineering*, vol. 12, no. 4, p. 046008, Aug. 2015.
- [12] M. Nakanishi, Y. Wang, X. Chen, Y.-T. Wang, X. Gao, and T.-P. Jung, "Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 1, pp. 104–112, Jan. 2018.
- [13] B. Liu, X. Chen, N. Shi, Y. Wang, S. Gao, and X. Gao, "BETA: A large benchmark database toward SSVEP-BCI application," *Frontiers in Neuroscience*, vol. 14, p. 627, Jun. 2020.
- [14] B. Liu, X. Chen, N. Shi, Y. Wang, S. Gao, and X. Gao, "Improving the performance of individually calibrated SSVEP-BCI by task-discriminant component analysis," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 1998–2007, 2021.

- [15] V. J. Lawhern, A. J. Solon, N. R. Waytowich, *et al.*, “EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces,” *Journal of Neural Engineering*, vol. 15, no. 5, p. 056013, 2018.
- [16] N. Waytowich, V. J. Lawhern, J. O. Garcia, *et al.*, “Compact convolutional neural networks for classification of asynchronous SSVEPs,” *Journal of Neural Engineering*, vol. 15, no. 6, p. 066031, 2018.
- [17] O.-Y. Kwon, M.-H. Lee, C. Guan, and S.-W. Lee, “Subject-independent brain–computer interfaces based on deep convolutional neural networks,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 10, pp. 3839–3852, Oct. 2020.
- [18] Z. Gao, T. Yuan, X. Zhou, *et al.*, “A deep learning method for improving the classification accuracy of SSVEP-based BCIs,” *IEEE Transactions on Circuits and Systems II*, vol. 67, no. 12, pp. 3447–3451, Dec. 2020.
- [19] J. Chen, Y. Zhang, Y. Pan, P. Xu, and C. Guan, “A Transformer-based deep neural network for SSVEP classification,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 2425–2435, 2023.
- [20] J. Hong and X. Qin, “Signal processing algorithms for SSVEP-based brain–computer interfaces: State of the art and recent developments,” *Journal of Intelligent & Fuzzy Systems*, vol. 40, no. 6, pp. 10559–10573, 2021.
- [21] H. He and D. Wu, “Transfer learning for brain–computer interfaces: A Euclidean space data alignment approach,” *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 2, pp. 399–410, Feb. 2020.
- [22] D. Wu, X. Jiang, and R. Peng, “Transfer learning for motor imagery based brain–computer interfaces: A tutorial,” *Neural Networks*, vol. 153, pp. 235–253, 2022.
- [23] S.-Y. Chen, C.-M. Chang, K.-J. Chiang, *et al.*, “SSVEP-DAN: Cross-domain data alignment for SSVEP-based brain–computer interfaces,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 32, pp. 2027–2037, 2024.
- [24] Y. Deng, Z. Ji, Y. Wang, and S. K. Zhou, “CSDuDoFN: A cross-subject dual-domain fusion network for SSVEP classification,” *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–12, 2023.
- [25] K. Zhang, N. Robinson, S.-W. Lee, and C. Guan, “Adaptive transfer learning for EEG motor imagery classification with deep convolutional neural networks,” *Neural Networks*, vol. 136, pp. 1–10, Apr. 2021.
- [26] W. Yan, B. He, J. Zhao, *et al.*, “Frequency-domain filtering method for SSVEP-EEG preprocessing,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 2079–2089, 2023.
- [27] Y. Wang, X. Chen, X. Gao, and S. Gao, “A benchmark dataset for SSVEP-based brain–computer interfaces,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 10, pp. 1746–1752, Oct. 2017.
- [28] A. Vaswani, N. Shazeer, N. Parmar, *et al.*, “Attention is all you need,” in *Advances in Neural Information Processing Systems*, 2017, pp. 5998–6008.
- [29] A. Nichol, J. Achiam, and J. Schulman, “On first-order meta-learning algorithms,” arXiv:1803.02999, 2018.
- [30] C. Finn, P. Abbeel, and S. Levine, “Model-agnostic meta-learning for fast adaptation of deep networks,” in *Proceedings of the 34th International Conference on Machine Learning*, 2017, pp. 1126–1135.
- [31] J. L. Ba, J. R. Kiros, and G. E. Hinton, “Layer normalization,” arXiv:1607.06450, 2016.
- [32] A. M. Norcia, L. G. Appelbaum, J. M. Ales, B. R. Cottareau, and B. Rossion, “The steady-state visual evoked potential in vision research: A review,” *Journal of Vision*, vol. 15, no. 6, p. 4, 2015.
- [33] F. Di Russo, S. Pitzalis, T. Aprile, *et al.*, “Spatiotemporal analysis of the cortical sources of the steady-state visual evoked potential,” *Human Brain Mapping*, vol. 28, no. 4, pp. 323–334, 2007.
- [34] J. Li, H. Hua, Z. Xu, *et al.*, “Cross-subject EEG emotion recognition combined with connectivity features and meta-transfer learning,” *Computers in Biology and Medicine*, vol. 145, p. 105519, 2022.
- [35] E. Rahimian, S. Zabihi, A. Asif, S. F. Atashzar, and A. Mohammadi, “Few-shot learning for decoding surface electromyography for hand gesture recognition,” in *Proc. IEEE ICASSP*, 2021, pp. 1300–1304.
- [36] P. Khosla, P. Teterwak, C. Wang, *et al.*, “Supervised contrastive learning,” in *Advances in Neural Information Processing Systems*, 2020.
- [37] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, “Mixup: Beyond empirical risk minimization,” arXiv:1710.09412, 2017.
- [38] Wang D, Shelhamer E, Liu S, et al. Tent: Fully test-time adaptation by entropy minimization[J]. arXiv preprint arXiv:2006.10726, 2020.