# SEAM: SUPPORT-ENCODED ADAPTATION WITH META-LEARNED PROMPTING FOR FEW-SHOT CROSS-SUBJECT EEG FOUNDATION MODEL ADAPTATION

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

Recent advances in deep learning have improved EEG decoding by learning task-relevant representations from neural recordings. Large-scale pretraining has further enabled EEG foundation models that capture reusable structure across subjects and tasks. However, adapting these models to an unseen subject from only a few labeled trials remains unreliable. Fine-tuning on small, subject-biased support sets can distort pretrained representations and induce negative transfer. We propose SEAM (Support-Encoded Adaptation with Meta-Learned Prompting), a target-time optimization-free framework for few-shot cross-subject EEG foundation model adaptation. SEAM jointly meta-trains the Transformer backbone, classifier head, and a lightweight prompt generator on source subjects, then freezes their weights for target adaptation. From a target support set, it constructs a global subject prototype and class residual prototypes; each query routes over these residuals to generate a personalized prompt. This preserves task-generic knowledge while expressing subject- and query-specific variations without target-time updates. Experiments on different EEG benchmarks show that SEAM better exploits limited support data, mitigates negative transfer from 2.4% to -4.8% and update-induced forgetting, and achieves a favorable accuracy–efficiency trade-off.

## 1 INTRODUCTION

EEG-based brain-computer interfaces have been widely investigated for neurorehabilitation, affective computing, and human-computer interaction (Wolpaw et al., 2002; Daly & Wolpaw, 2008; Koelstra et al., 2011). A central barrier is that EEG signals vary strongly across subjects. Electrode impedance, cortical anatomy, mental state, fatigue, and session-specific artifacts all change the observed signal distribution (Wu et al., 2020). Calibrating deep learning-based EEG decoding models specifically on new subjects can improve decoding performance, but collecting many labeled calibration trials is expensive and often unrealistic (He & Wu, 2019). This makes few-shot cross-subject adaptation a practical requirement rather than a secondary convenience.

Recent EEG foundation models change the landscape of this problem (Kostas et al., 2021; Jiang et al., 2024; Wang et al., 2024). By pretraining on large-scale neural recordings, they can learn reusable temporal, spatial, and task-related structure. In principle, such models should be a strong starting point for a new subject. In practice, however, the adaptation mechanism becomes the bottleneck. A large model with useful general representations is also easy to damage when its parameters are updated with a tiny, noisy, and subject-biased support set (Kumar et al., 2022; Wortsman et al., 2022).

Few-shot target fine-tuning can degrade, rather than improve, a source-pretrained EEG foundation model, suggesting that limited target supervision may be insufficient for stable parameter updates. To examine this failure mode, a preliminary study on two motor-imagery datasets compares zero-shot cross-subject inference with few-shot fine-tuning of the same model across varying source-cohort sizes. As shown in Fig. 1 (left), fine-tuning can yield negative gains relative to the zero-shot baseline, particularly with larger source cohorts. This motivates an adaptation mechanism that uses target support trials while preserving source-trained representations.



Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

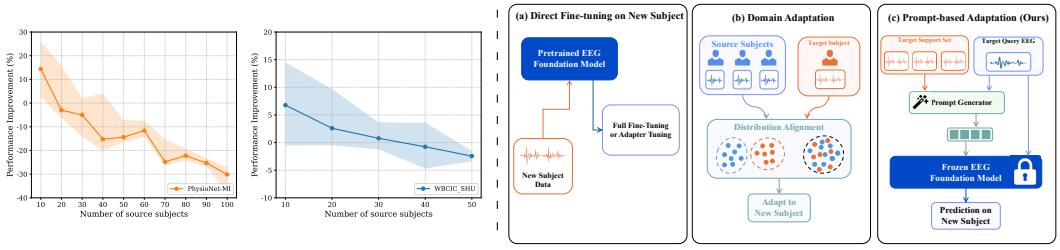


Figure 1: Left: On two motor imagery datasets, direct few-shot fine-tuning can produce negative gains over zero-shot cross-subject evaluation, especially when the number of source domain subjects increased. Right: Different cross-subject adaptation paradigms: direct fine-tuning, domain adaptation, and our prompt-based adaptation.

To address the overfitting issue, we propose SEAM, a support-encoded meta-adaptation prompting method for few-shot cross-subject EEG foundation model adaptation. During source-subject episodic meta-training, SEAM jointly trains the Transformer backbone, classifier head, and lightweight conditioning module. For a new target subject, it freezes all learned weights and generates support-conditioned prompt tokens without gradient updates.

This design follows a simple principle: the foundation model should preserve task-generic neural knowledge, while an external conditioning module should express subject-specific shifts. The prompt is not a free parameter optimized per target subject. It is generated from the support set via a class prototype aggregation module, and fused with query samples to activate the subject-specific semantics, so adaptation is fast, deterministic, and does not expose the backbone to catastrophic forgetting.

Our contributions are threefold. First, we identify a failure pattern in few-shot EEG foundation model adaptation: fine-tuning can degrade a strong zero-shot model and harm source-domain general ability. Second, we introduce SEAM, a training-free cross-subject adaptation framework that combines query-specific prompt generation. Third, we evaluate SEAM on motor imagery and affective computing datasets with cross-subject few-shot protocols. The results demonstrate that it significantly mitigates the negative transfer of EEG foundation models in few-shot scenarios.

## 2 Related Work

Large-scale self-supervised EEG foundation models learn transferable representations across subjects, datasets, and tasks. Representative models include BENDR, BIOT, LaBraM, and EEGPT (Kostas et al., 2021; Yang et al., 2023; Jiang et al., 2024; Wang et al., 2024), while BrainBERT, Brant, CBraMod, NeuroLM, BrainWave, and BrainOmni extend pretraining to intracranial signals, spatial-temporal modeling, neural-language alignment, and multimodal brain signals (Wang et al., 2023; Zhang et al., 2023a; Wang et al., 2025; Jiang et al., 2025; Yuan et al., 2024; Xiao et al., 2026). However, these models are still mainly adapted by full fine-tuning, linear probing, or task-specific heads, and stronger pretraining does not necessarily ensure label-efficient adaptation (Lee et al., 2025; Liu et al., 2026). SEAM instead uses source-subject meta-training followed by optimization-free target adaptation.

Classical few-shot learning acquires transferable adaptation rules through metric learning or episodic optimization, including Matching Networks, Prototypical Networks, Relation Networks, MAML, and MetaOptNet (Vinyals et al., 2016; Snell et al., 2017; Sung et al., 2018; Finn et al., 2017; Lee et al., 2019). These principles have been applied to few-shot EEG decoding through relation learning, attention-based comparison, and subject-level meta-learning (An et al., 2020; 2023; Ng & Guan, 2024; Han et al., 2024; Chen et al., 2025; Zhang et al., 2026). Recent approaches further introduce task-conditioned prompts or lightweight adapters for cross-subject EEG adaptation (Liu et al., 2025). Unlike conventional few-shot classification, cross-subject Brain Signal Encoding/Decoding adaptation retains the label space but changes the subject-dependent signal distribution. Several recent studies directly encode the support set as contextual information to guide target downstream brain-computer interface encoding/decoding tasks (Yu et al., 2026; Nan et al., 2026).

Domain adaptation reduces source–target discrepancies through statistical or adversarial alignment, as represented by DAN, CORAL, DANN, and CDAN (Long et al., 2015; Sun et al., 2017; Ganin et al., 2016; Long et al., 2018). EEG-specific methods align covariance geometry or deep class-aware distributions across subjects and sessions (Congedo et al., 2017; He & Wu, 2019; Zhao et al., 2020; 2021; She et al., 2023; Zhong et al., 2023). However, aggressive invariance may suppress informative subject-specific characteristics. Parameter-efficient methods instead freeze most pretrained parameters and optimize adapters, prefixes, prompts, or low-rank updates (Houlsby et al., 2019; Li & Liang, 2021; Lester et al., 2021; Liu et al., 2022; Hu et al., 2022; Wu et al., 2024a;b), with recent extensions to EEG foundation models (Ma et al., 2026).

# 3 PRELIMINARY

Let $\mathcal{D}_s = \{D_1, \ldots, D_M\}$ denote source subjects and $D_t$ denote an unseen target subject. Each trial contains an EEG segment $x \in \mathbb{R}^{C\times T}$ and label $y \in \{1, \ldots, N\}$. In a $K$-shot target episode, the target subject provides a support set $S_t = \{(x_i, y_i)\}$ with $K$ examples per class, and the model is evaluated on a disjoint query set $Q_t$. The adaptation method can use $S_t$, but it must not access query labels.

Before introducing the method, we rethink two common paradigms for cross-subject adaptation of EEG foundation models, as shown in Fig. 1. Direct fine-tuning and adapter-based tuning adapt a foundation model through target-time parameter optimization. However, in the few-shot cross-subject setting, $S_t$ is both extremely small and subject-biased. Optimizing on such limited supervision can overfit artifacts and idiosyncratic patterns, distort the pretrained representation manifold, and induce catastrophic forgetting of source knowledge. Consequently, the lack of stability may move strong zero-shot model toward a worse target decision rule, resulting in negative transfer.

In contrast, domain adaptation follows the opposite strategy by suppressing inter-subject discrepancies and enforcing invariant source–target representations. Although this improves stability, it often regards subject-specific EEG characteristics as nuisance variations. Aggressive alignment may therefore erase target-specific information that remains useful for fine-grained decoding. In this case, excessive stability limits the model's ability to accommodate the unseen subject and can likewise produce negative transfer.

This observation reveals a key factor underlying the negative transfer of EEG Foundation Models in few-shot cross-subject scenarios–the stability–plasticity dilemma. We therefore reformulate target-subject adaptation as non-invasive prompt conditioning. After source-subject meta-training updates the shared model, its weights are frozen for unseen target subjects, while support-derived prompts express subject-specific information without target-time updates.

# 4 METHOD

## 4.1 EEG FOUNDATION MODEL

Figure 2 summarizes source meta-training and target-time inference. SEAM starts from a pretrained EEG foundation model composed of a tokenizer, a transformer backbone, and a task head. Given an EEG trial $x$, the tokenizer maps raw signals into input tokens $z \in \mathbb{R}^{C\times W\times d}$, where $C$ is the channel/token group dimension, $W$ is the temporal token length, and $d$ is the hidden dimension. A transformer encoder then produces contextual representations for classification.

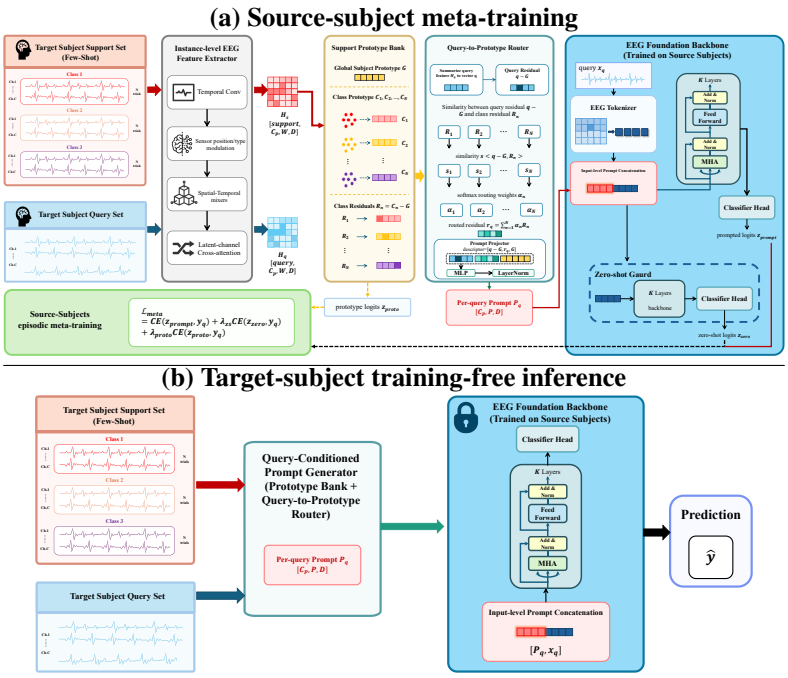


Figure 2: Overview of SEAM. Source-subject episodes update the Transformer backbone, classifier head, and prompt modules (top). For an unseen target subject, all weights are frozen and support-generated prompts condition inference without gradient updates (bottom).

The freeze applies only to target-subject adaptation. During source-subject episodic meta-training, the Transformer backbone and classifier head are updated jointly with the support feature extractor and prompt generator. For an unseen target subject, the tokenizer, backbone, head, and prompt modules are fixed; support trials condition predictions through prompts without modifying any weights. This protects source-trained task knowledge from overfitting to a small target support set.

### 4.2 Support-Conditioned Prompt Generation

For each target episode, SEAM encodes the support trials with a lightweight EEG feature extractor. We intentionally use a raw-EEG feature extractor rather than reusing embedding layers of the foundation model backbone, to prevent the attenuation of subject-specific information. Given an EEG trial, the extractor first splits the raw signal into temporal windows aligned with the backbone tokenizer, and applies multi-kernel temporal convolutions to capture EEG patterns at different time scales. The resulting electrode-window tokens are then modulated by sensor position and type embeddings, so that spatial electrode information is explicitly injected into the representation. To model temporal and inter-channel interactions, several lightweight axial convolution blocks are applied along the time and channel dimensions. Finally, a set of learnable latent channel queries attends to the electrode tokens and projects them into a fixed prompt-channel space, followed by a linear projection to the prompt hidden dimension. This design preserves enough temporal, spatial, and sensor-specific information for subject-conditioned prompt generation.

With the set of support trial features, SEAM then builds two forms of support statistics. First, a global subject prototype summarizes the average support pattern of the target subject. Second, class prototypes summarize label-conditioned support patterns. Let $e_\phi(x)$ denote the support feature of trial $x$. For class $n$, SEAM computes

$$c_n = \frac{1}{|S_n|} \sum_{(x,y)\in S_n} e_\phi(x), \quad c_g = \frac{1}{|S|} \sum_{(x,y)\in S} e_\phi(x),$$

and uses class residuals $\Delta_n = c_n - c_g$. The global prototype captures subject-level calibration factors shared by all classes, while the residuals describe how each class deviates from this common

subject state. This residual representation is useful because cross-subject EEG shifts often affect all classes, while the classifier still needs class-discriminative evidence.

The prompt is also query-conditioned. For a query trial, SEAM extracts a query summary $q = e_\phi(x)$ and forms a query residual $\Delta_q = q - c_g$. The residual is matched against the class residual prototypes:

$$\alpha_n = \frac{\exp(\tau \cdot \mathrm{sim}(\Delta_q, \Delta_n))}{\sum_j \exp(\tau \cdot \mathrm{sim}(\Delta_q, \Delta_j))}, \quad \bar{\Delta} = \sum_n \alpha_n \Delta_n,$$

where $\tau$ is a temperature and $\mathrm{sim}(\cdot, \cdot)$ is a similarity function. The routed residual $\bar{\Delta}$, query residual $\Delta_q$, and global context $c_g$ are concatenated and projected into prompt tokens:

$$p_t = G_\phi(c_g, \Delta_q, \bar{\Delta}).$$

Thus the generated prompt is not a single static subject vector. It is conditioned on both the target support set and the current query, allowing different query trials from the same subject to receive different prompt emphasis while sharing the same support-derived calibration.

The generated prompt $p_t$ is concatenated to the tokenized input before the transformer blocks:

$$h_0 = [p_t; z].$$

This input-level injection lets every transformer layer attend to both the original EEG tokens and the subject-specific prompt. Through conditional injection of prompt vectors, attention mechanism can modulate backbone features towards adapting to target subject semantics and facilitating decoding tasks.

### 4.3 Meta-Learning Objective

SEAM is trained on source-subject episodes that mirror target deployment. For a source subject $i$, a support set $S_i$ generates the support payload, and the query set $Q_i$ supplies the meta-training loss. The prompt generator is therefore optimized to solve a family of held-out subject adaptation problems rather than a single supervised classification problem. The primary loss is the prompted query classification loss:

$$\mathcal{L}_{\mathrm{prompt}} = \frac{1}{|Q_i|} \sum_{(x,y) \in Q_i} \mathrm{CE}(f_\theta(x, p_i), y).$$

To preserve the general behavior of the foundation model, SEAM also keeps an unprompted branch active during meta-training:

$$\mathcal{L}_{\mathrm{zs}} = \frac{1}{|Q_i|} \sum_{(x,y) \in Q_i} \mathrm{CE}(f_\theta(x, \varnothing), y).$$

The unprompted branch depends on $\theta$ but not $\phi$. Thus, during source meta-training, $\mathcal{L}_{\mathrm{zs}}$ back-propagates through the trainable backbone and classifier head, while $\nabla_\phi \mathcal{L}_{\mathrm{zs}} = 0$. It preserves the shared unprompted predictor rather than directly regularizing the prompt modules. The prompted loss trains $\phi$, and prototype supervision further shapes its support representation. SEAM uses the following auxiliary prototype loss based on the query-to-prototype routing logits:

$$\mathcal{L}_{\mathrm{proto}} = \frac{1}{|Q_i|} \sum_{(x,y) \in Q_i} \mathrm{CE}(r(x, S_i), y),$$

where $r(x, S_i)$ denotes the class-routing logits induced by support prototypes. This term does not replace the foundation-model classifier. It regularizes the prompt generator so that the support representation remains class-aware. The final objective is

$$\mathcal{L} = \mathcal{L}_{\mathrm{prompt}} + \lambda_{\mathrm{zs}} \mathcal{L}_{\mathrm{zs}} + \lambda_{\mathrm{proto}} \mathcal{L}_{\mathrm{proto}}.$$

At test time, no parameters are updated. SEAM computes support statistics once for the target episode, generates prompt tokens for query trials, and performs prompted forward passes through the frozen foundation model. This makes the adaptation deterministic for a fixed support set and removes the need to tune learning rates, iteration counts, or early-stopping rules for each new subject.

# 5 EXPERIMENTS

## 5.1 EXPERIMENT SETUP

We evaluate SEAM on two EEG classification benchmarks that stress different aspects of cross-subject transfer. PhysioNet-MI (Schalk et al., 2004) contains motor imagery trials and tests adaptation under a multi-class MI protocol. FACED (Chen et al., 2023) contains affective EEG trials and introduces stronger class imbalance and subject variation. For each dataset, we adopt a 5-fold cross-validation protocol, subjects are split into source training, validation, and held-out target subject groups. Target evaluation uses 2-shot, 4-shot, and 8-shot support sets, with query labels hidden during adaptation.

We report accuracy, macro F1, and Cohen's kappa. Accuracy measures overall correctness, macro F1 emphasizes class-balanced recognition, and kappa discounts chance agreement. This combination is important for EEG because target subjects may have different class distributions and a method can improve raw accuracy while weakening class-balanced behavior.

Baselines cover following categories: foundation EEG encoding model (Xiao et al., 2026; Lawhern et al., 2018), parameter-efficient tuning (Hu et al., 2022; Li & Liang, 2021; Liu et al., 2022), domain adaptation (She et al., 2023; Zhang et al., 2023b; Li et al., 2022), and few-shot/meta-learning methods (Finn et al., 2017; Snell et al., 2017; Chen et al., 2025; An et al., 2023; Zhang et al., 2026).

## 5.2 IMPLEMENTATION DETAILS

We use the pretrained BrainOmni-tiny (Xiao et al., 2026) model as the EEG foundation backbone. The backbone contains a BrainTokenizer followed by a 12-layer spatial-temporal Transformer with hidden dimension 256 and 8 attention heads. The generated prompt contains 10 prompt steps and is injected at the input token level before the first Transformer block.

For the prompt module, the Instance-level EEG feature extractor uses hidden dimension 128, dropout 0.1 for latent channel projection. The query-conditioned prompt generator uses hidden dimension 256 and dropout 0.1. We jointly update the Transformer backbone, classifier head, and prompt modules on source-subject episodes for 10 meta-epochs, with 100 training steps per epoch, a global meta-batch size of 16, and query batch size 256. The objective combines prompted classification, zero-shot preservation, and prototype losses, with $\lambda_{\mathrm{zs}} = 0.5$ and $\lambda_{\mathrm{proto}} = 0.2$. All main experiments use 8 x A100 GPU distributed training and 2-shot, 4-shot, and 8-shot cross-subject evaluation protocols; no parameters are updated for target subjects.

## 5.3 MAIN RESULTS

The result on PhysioNet-MI is shown in Tab. 1, SEAM improves as the support set grows and reaches the best 2shot, 4-shot and 8-shot results among all adaptation-designed methods except EEGNet. Moreover, across all few-shot experimental rounds, it mitigates negative transfer and achieves positive gain. The pattern is important. BrainOmni zero-shot is already strong, but several target-time adaptation methods reduce performance below it. Head-only tuning, prompt tuning, prefix tuning, LoRA, and prototype classification all fail to reliably exploit the support set. Full fine-tuning becomes competitive at 8-shot, yet it remains weaker than SEAM and requires updating the foundation model. SEAM is the only method in the table that combines positive shot usage with no target-time gradient update. As a lightweight EEG model, EEGNet achieved the best performance across all three shot settings. This corroborates our assessment of negative transfer: due to its small number of parameters, EEGNet is less susceptible to overfitting and catastrophic forgetting during target domain adaptation.

The result on FACED is shown in Tab. 2, SEAM provides the best accuracy at 4-shot and 8-shot. The result is more nuanced for macro F1 and kappa, where compact EEGNet remains strong. SEAM is designed for the harder case where a large EEG foundation model already has useful general capability and direct adaptation can damage it. The FACED results therefore show both the benefit and the boundary of the proposed approach: SEAM improves foundation-model adaptation, while compact specialized models can still be competitive on some class-sensitive metrics. The remaining domain adaptation methods perform poorly in few-shot cross-subject scenarios. Because the limited

Table 1: Main comparison on PhysioNet-MI. Each entry reports Acc/F1/Kappa in percent.

| Type | Method | 2-shot | 4-shot | 8-shot |
|---|---|---|---|---|
| Foundation | BrainOmni zero-shot | 61.4/60.6/48.2 | 59.8/59.5/46.5 | 58.8/58.3/45.1 |
| | BrainOmni Head FT | 49.1/49.1/32.2 | 52.1/51.9/36.1 | 56.4/56.4/41.9 |
| | BrainOmni Full FT | 59.3/59.2/45.7 | 58.0/57.8/44.0 | 61.9/61.8/49.2 |
| | EEGNet | **63.5/63.2/51.3** | **63.4/63.2/51.2** | **64.6/64.3/52.8** |
| PEFT | Prompt tuning | 39.7/39.1/19.7 | 42.0/41.2/22.7 | 44.3/43.7/25.8 |
| | Prefix tuning | 29.5/29.0/6.0 | 32.2/32.1/9.6 | 34.3/34.4/12.4 |
| | LoRA | 45.3/45.3/27.1 | 50.4/50.5/33.8 | 52.8/52.7/37.1 |
| Few-shot | ProtoNet | 37.3/37.2/16.3 | 41.0/40.9/21.4 | 43.7/43.7/24.9 |
| | DA-RelationNet | 49.6/49.5/32.7 | 56.2/55.7/41.6 | 56.7/55.7/42.2 |
| DA | IDAN | 60.7/60.6/47.6 | 58.2/58.2/44.2 | 60.0/59.8/46.6 |
| | MI-CAT | 41.3/39.9/21.8 | 41.8/41.1/22.5 | 41.4/40.6/21.9 |
| Ours | SEAM | 61.5/61.0/48.4 | 62.5/61.3/50.0 | 63.6/63.0/51.5 |

support set cannot adequately model the full target subject EEG distribution, direct distribution alignment methods suffer from significant drift.

Table 2: Main comparison on FACED. Each entry reports Acc/F1/Kappa in percent.

| Type | Method | 2-shot | 4-shot | 8-shot |
|---|---|---|---|---|
| Foundation | BrainOmni zero-shot | **51.8**/36.7/14.7 | 52.8/37.3/15.1 | 55.9/38.3/17.4 |
| | Head-only fine-tuning | 43.0/40.4/11.2 | 45.2/41.8/13.7 | 47.4/41.8/16.4 |
| | Full fine-tuning | 47.2/43.7/16.4 | 48.9/45.1/18.6 | 50.4/43.3/20.3 |
| | EEGNet | 51.3/**48.8/25.6** | 50.6/47.6/**23.9** | 51.2/45.2/**23.4** |
| PEFT | Prompt tuning | 47.8/42.9/15.7 | 48.9/44.6/18.9 | 46.7/40.4/16.0 |
| | Prefix tuning | 42.7/39.9/11.1 | 43.9/40.7/13.3 | 48.7/42.5/17.9 |
| | LoRA | 42.7/40.1/11.3 | 43.5/40.9/14.3 | 44.7/39.6/15.2 |
| Few-shot | ProtoNet | 41.3/37.8/7.4 | 42.1/39.0/9.5 | 46.7/40.9/15.0 |
| | MAML | 44.1/41.4/14.3 | 44.9/42.1/15.8 | 44.8/39.8/14.7 |
| | STS-ML | 43.3/39.9/10.4 | 49.3/45.0/18.0 | 50.6/44.2/20.3 |
| DA | DDA-SDDA | 50.4/45.1/16.6 | 50.6/45.3/17.4 | 50.7/43.6/16.0 |
| Ours | SEAM | 51.4/46.0/20.9 | **54.3/47.7**/23.4 | **56.9/45.8**/23.2 |

### 5.4 Ablation Study

Tab. 3 evaluates the contribution of each component on PhysioNet-MI under the 4-shot setting. Removing the global prototype context causes the largest degradation, reducing BAcc from 62.6% to 56.8%, which demonstrates that subject-level calibration information is essential for modeling the common distribution shift of a new subject. Removing the zero-shot loss also decreases BAcc by 2.3 percentage points, consistent with preserving the shared backbone and classifier pathway during source meta-training. Query-conditioned generation and the support prototype branch provide consistent gains, indicating that both instance-specific prompt modulation and class-aware support statistics contribute to adaptation. Finally, removing the prototype loss reduces BAcc and Acc to 62.0% and 61.2%, respectively, showing that explicit class-aware supervision improves the discriminability of the generated prompts. Overall, the full SEAM achieves the best performance, validating the complementary roles of subject-level, class-level, and query-specific conditioning.

Table 3: Ablation on PhysioNet-MI 4-shot. BAcc denotes balanced accuracy. Values are in percent.

| Variant | BAcc | Acc |
|---|---|---|
| Full SEAM | 62.6 | 62.5 |
| w/o query-conditioned generation | 62.1 | 62.0 |
| w/o support prototype branch | 62.3 | 62.3 |
| w/o global prototype context | 56.8 | 56.8 |
| w/o $\lambda_{\text{zs}}$ | 60.3 | 60.2 |
| w/o $\lambda_{\text{proto}}$ | 62.0 | 61.2 |

### 5.5 Forgetting and General Ability Preservation

The forgetting diagnostic further explains the failure of direct adaptation. As shown in Fig. 3, we split the source domain subjects into train and test splits, and evaluate the extent to which different methods retain their capability on source subjects after adaptation to new subjects. Head-only and full fine-tuning reduce source-domain accuracy after target adaptation, while SEAM does not modify target-time weights and therefore has no such update-induced forgetting. This matters because a foundation model is valuable precisely because it contains general knowledge learned from many subjects. A target adaptation method that improves one subject by erasing this knowledge is risky, especially if the system is reused across subjects or sessions.

We also evaluate source held-out behavior under prompted inference. Prompted SEAM reaches 70.2% accuracy compared with 47.4% for the vanilla source-trained baseline. This suggests that the conditioning module does more than prevent damage; it can activate useful source-domain behavior when support information is available. The result supports the central separation in SEAM: the backbone stores reusable task information, while the prompt generator selects a subject-conditioned operating point without changing the backbone.

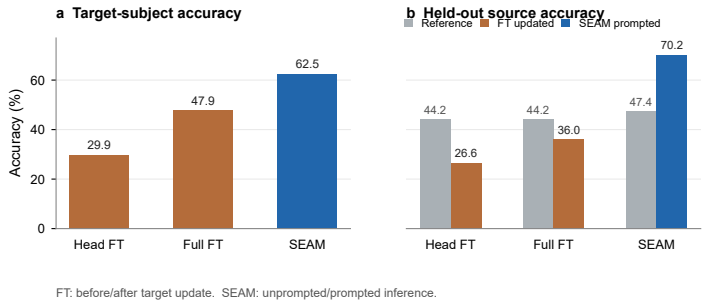


Figure 3: Target-time forgetting diagnostic on PhysioNet-MI 4-shot. Direct adaptation can reduce source-domain behavior, while SEAM avoids update-induced forgetting.

### 5.6 Prompt Utility

We replaced the prompts used during meta-testing on PhysioNet-MI to verify whether the support-conditioned prompts are truly meaningful for the target subjects. As shown in Fig. 4 (left), prompt controls help separate useful conditioning from arbitrary prompt perturbation. On PhysioNet-MI 4-shot, the real support prompt reaches 62.5% accuracy, compared with 57.3% for train-mean prompts (average prompts on source subjects), 55.7% for subject-shuffled prompts, 33.7% for Gaussian prompts, and 41.2% for zero-shot inference. The gap to Gaussian prompts shows that arbitrary prompt tokens are not sufficient. The gap to train-mean prompts shows that a dataset-level average prompt cannot replace target support evidence. The gap to subject-shuffled prompts shows that using support trials from the wrong subject weakens adaptation, which is a direct test of subject specificity.

The query-anchored prompt similarity matrix in Fig. 4 (right) gives complementary evidence: The prompt generator can generate diverse prompt vectors for different subjects' queries and supports to

represent specific subject semantics, rather than collapsing into a constant prompt that yields generic gains on the source domain.

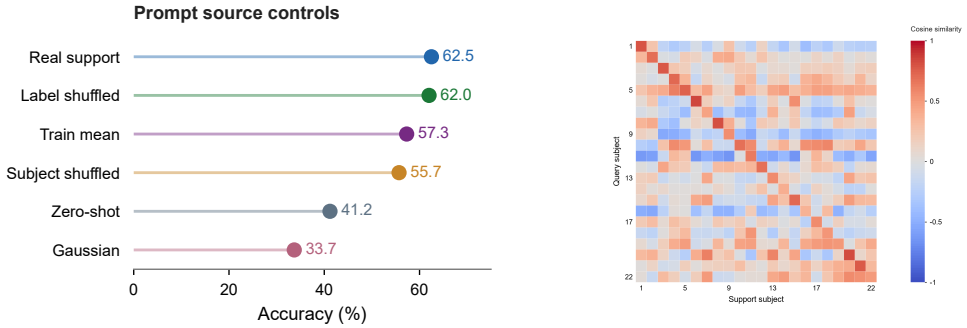


Figure 4: Left: Prompt source controls on PhysioNet-MI 4-shot. Real support prompts outperform train-mean, subject-shuffled, Gaussian, and zero-shot controls. Right: Cross-subject prompt similarity. Generated prompts vary across subjects, supporting their role as subject-specific conditioning signals.

### 5.7 Parameter Sensitivity

Figure 5 shows that the zero-shot guard is important for stable adaptation. A moderate guard weight produces the best accuracy in this setting. Without it, source-meta-training updates to the shared backbone and head can erode the zero-shot decision rule; if it is too strong, the model becomes conservative and leaves less room for support-conditioned adaptation. Prototype supervision has a similar trade-off. Moderate prototype weight helps the prompt generator preserve class-aware support structure, while excessive weight can over-constrain the generator and reduce final prompted accuracy.

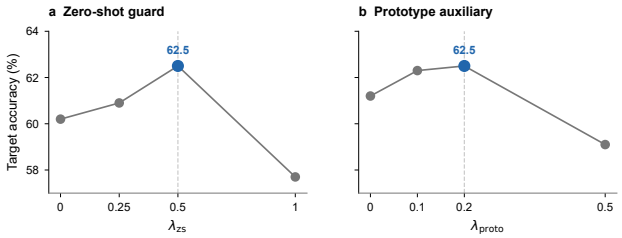


Figure 5: Loss coefficient sensitivity on PhysioNet-MI 4-shot.

## 6 Conclusion

We presented SEAM, a support-encoded prompting framework for few-shot cross-subject EEG foundation model adaptation. Limited target trials can damage the general ability of a strong source-trained model. SEAM jointly meta-trains the backbone, classifier, and prompt generator on source subjects, then freezes their weights for target adaptation. Experiments on PhysioNet-MI and FACED show that SEAM improves few-shot adaptation, reduces target-time cost, and avoids update-induced forgetting. The results also point to a clear future direction: stronger label-grounded prompt generation under imbalanced support sets.

## AI USE STATEMENT

We used large language models to assist with writing and language refinement of the manuscript and with the preparation of some figures and tables. All other work, including the research design, experiments, analysis, and interpretation, was carried out by the authors. We reviewed the AI-assisted material and take responsibility for the final content.

## REFERENCES

Sion An, Soopil Kim, Philip Chikontwe, and Sang Hyun Park. Few-shot relation learning with attention for eeg-based motor imagery classification. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020.

Sion An, Soopil Kim, Philip Chikontwe, and Sang Hyun Park. Dual attention relation network with fine-tuning for few-shot eeg motor imagery classification. *IEEE transactions on neural networks and learning systems*, 2023.

Cheng Chen, Hao Fang, Yuxiao Yang, and Yi Zhou. Model-agnostic meta-learning for eeg-based inter-subject emotion recognition. *Journal of Neural Engineering*, 2025.

Jingjing Chen, Xiaobin Wang, Chen Huang, Xin Hu, Xinke Shen, and Dan Zhang. A large finer-grained affective computing eeg dataset. *Scientific Data*, 2023.

Marco Congedo, Alexandre Barachant, and Rajendra Bhatia. Riemannian geometry for eeg-based brain-computer interfaces; a primer and a review. *Brain-Computer Interfaces*, 2017.

Janis J Daly and Jonathan R Wolpaw. Brain–computer interfaces in neurological rehabilitation. *The Lancet Neurology*, 2008.

Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *International conference on machine learning*, 2017.

Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario March, and Victor Lempitsky. Domain-adversarial training of neural networks. *Journal of machine learning research*, 2016.

Ji-Wung Han, Soyeon Bak, Jun-Mo Kim, Woohyeok Choi, Dong-Hee Shin, Young-Han Son, and Tae-Eui Kam. Meta-eeg: Meta-learning-based class-relevant eeg representation learning for zero-calibration brain–computer interfaces. *Expert Systems with Applications*, 2024.

He He and Dongrui Wu. Transfer learning for brain–computer interfaces: A euclidean space data alignment approach. *IEEE Transactions on Biomedical Engineering*, 2019.

Neil Houlsby, Andrei Giurgiu, Stanislaw Jastrzebski, Bruna Morrone, Quentin De Laroussilhe, Andrea Gesmundo, Mona Attariyan, and Sylvain Gelly. Parameter-efficient transfer learning for nlp. In *International conference on machine learning*, 2019.

Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Liang Wang, Weizhu Chen, et al. Lora: Low-rank adaptation of large language models. *Iclr*, 2022.

Wei-Bang Jiang, Liming Zhao, and Bao-Liang Lu. Large brain model for learning generic representations with tremendous eeg data in bci. In *International Conference on Learning Representations*, 2024.

Wei-Bang Jiang, Yansen Wang, Bao-Liang Lu, and Dongsheng Li. Neurolm: A universal multi-task foundation model for bridging the gap between language and eeg signals. In *International conference on learning representations*, volume 2025, pp. 55436–55457, 2025.

Sander Koelstra, Christian Muhl, Mohammad Soleymani, Jong-Seok Lee, Ashkan Yazdani, Touradj Ebrahimi, Thierry Pun, Anton Nijholt, and Ioannis Patras. Deap: A database for emotion analysis; using physiological signals. *IEEE transactions on affective computing*, 2011.

Demetres Kostas, Stephane Aroca-Ouellette, and Frank Rudzicz. Bendr: Using transformers and a contrastive self-supervised learning task to learn from massive amounts of eeg data. *Frontiers in Human Neuroscience*, 2021.

Ananya Kumar, Aditi Raghunathan, Robbie Matthew Jones, Tengyu Ma, and Percy Liang. Fine-tuning can distort pretrained features and underperform out-of-distribution. In *International Conference on Learning Representations*, 2022.

Vernon J Lawhern, Amelia J Solon, Nicholas R Waytowich, Stephen M Gordon, Chou P Hung, and Brent J Lance. Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces. *Journal of neural engineering*, 2018.

Kwonjoon Lee, Subhransu Maji, Avinash Ravichandran, and Stefano Soatto. Meta-learning with differentiable convex optimization. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019.

Na Lee, Konstantinos Barmpas, Yannis Panagakis, Dimitrios Adamos, Nikolaos Laskaris, and Stefanos Zafeiriou. Are large brainwave foundation models capable yet? insights from fine-tuning. *arXiv preprint arXiv:2507.01196*, 2025.

Brian Lester, Rami Al-Rfou, and Noah Constant. The power of scale for parameter-efficient prompt tuning. In *Proceedings of the 2021 conference on empirical methods in natural language processing*, 2021.

Xiang Lisa Li and Percy Liang. Prefix-tuning: Optimizing continuous prompts for generation. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 2021.

Zhunan Li, Enwei Zhu, Ming Jin, Cunhang Fan, Huiguang He, Ting Cai, and Jinpeng Li. Dynamic domain adaptation for class-aware cross-subject and cross-session eeg emotion recognition. *IEEE Journal of Biomedical and Health Informatics*, 2022.

Dingkun Liu, Yuheng Chen, Zhu Chen, Zhenyao Cui, Yaozhi Wen, Jiayu An, Jingwei Luo, and Dongrui Wu. Eeg foundation models: Progresses, benchmarking, and open problems. *arXiv preprint arXiv:2601.17883*, 2026.

Haiqi Liu, CL Chen, and Tong Zhang. Face: few-shot adapter with cross-view fusion for cross-subject eeg emotion recognition. *arXiv preprint arXiv:2503.18998*, 2025.

Xiao Liu, Kaixuan Ji, Yicheng Fu, Weng Tam, Zhengxiao Du, Zhilin Yang, and Jie Tang. P-tuning: Prompt tuning can be comparable to fine-tuning across scales and tasks. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 2022.

Mingsheng Long, Yue Cao, Jianmin Wang, and Michael Jordan. Learning transferable features with deep adaptation networks. In *International conference on machine learning*, pp. 97–105. PMLR, 2015.

Mingsheng Long, Zhangjie Cao, Jianmin Wang, and Michael Jordan. Conditional adversarial domain adaptation. *Advances in neural information processing systems*, 2018.

Jingying Ma, Feng Wu, Yucheng Xing, Qika Lin, Tianyu Liu, Chenyu Liu, Ziyu Jia, and Mengling Feng. Structured prototype-guided adaptation for eeg foundation models. *arXiv e-prints*, 2026.

Mu Nan, Muquan Yu, Weijian Mai, Jacob S Prince, Hossein Adeli, Rui Zhang, Jiahang Cao, Benjamin Becker, John A Pyles, Margaret M Henderson, et al. Meta-learning in-context enables training-free cross subject brain decoding. 2026.

Han Wei Ng and Cuntai Guan. Subject-independent meta-learning framework towards optimal training of eeg-based classifiers. *Neural Networks*, 2024.

Gerwin Schalk, Dennis J McFarland, Thilo Hinterberger, Niels Birbaumer, and Jonathan R Wolpaw. Bci2000: a general-purpose brain-computer interface (bci) system. *IEEE Transactions on biomedical engineering*, 2004.

Qingshan She, Tie Chen, Feng Fang, Jianhai Zhang, Yunyuan Gao, and Yingchun Zhang. Improved domain adaptation network based on wasserstein distance for motor imagery eeg classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2023.

Jake Snell, Kevin Swersky, and Richard Zemel. Prototypical networks for few-shot learning. *Advances in neural information processing systems*, 2017.

Baochen Sun, Jiashi Feng, and Kate Saenko. Correlation alignment for unsupervised domain adaptation. In *Domain adaptation in computer vision applications*. Springer, 2017.

Flood Sung, Yongxin Yang, Li Zhang, Tao Xiang, Philip HS Torr, and Timothy M Hospedales. Learning to compare: Relation network for few-shot learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018.

Oriol Vinyals, Charles Blundell, Timothy Lillicrap, Daan Wierstra, et al. Matching networks for one shot learning. *Advances in neural information processing systems*, 2016.

C Wang, V Subramaniam, A Yaari, G Kreiman, B Katz, I Cases, and A Barbu. Brainbert: self-supervised representation learning for intracranial electrodes. In *International Conference on Learning Representations*, 2023.

Guagnyu Wang, Wenchao Liu, Yuhong He, Cong Xu, Lin Ma, and Haifeng Li. Eegpt: Pretrained transformer for universal and reliable representation of eeg signals. *Advances in Neural Information Processing Systems*, 2024.

Jiquan Wang, Sha Zhao, Zhiling Luo, Yangxuan Zhou, Haiteng Jiang, Shijian Li, Tao Li, and Gang Pan. Cbramod: A criss-cross brain foundation model for eeg decoding. In *International conference on learning representations*, 2025.

Jonathan R Wolpaw, Niels Birbaumer, Dennis J McFarland, Gert Pfurtscheller, and Theresa M Vaughan. Brain–computer interfaces for communication and control. *Clinical neurophysiology*, 2002.

Mitchell Wortsman, Gabriel Ilharco, Jong Wook Kim, Mike Li, Simon Kornblith, Rebecca Roelofs, Raphael Gontijo Lopes, Hannaneh Hajishirzi, Ali Farhadi, Hongseok Namkoong, et al. Robust fine-tuning of zero-shot models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022.

Dongrui Wu, Yifan Xu, and Bao-Liang Lu. Transfer learning for eeg-based brain–computer interfaces: A review of progress made since 2016. *IEEE Transactions on Cognitive and Developmental Systems*, 2020.

Jiamin Wu, Xin Liu, Xiaotian Yin, Tianzhu Zhang, and Yongdong Zhang. Task-adaptive prompted transformer for cross-domain few-shot learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024a.

Jiamin Wu, Tianzhu Zhang, and Yongdong Zhang. Hybridprompt: Domain-aware prompting for cross-domain few-shot learning. *International Journal of Computer Vision*, 2024b.

Qinfan Xiao, Ziyun Cui, Chi Zhang, Siqi Chen, Wen Wu, Andrew Thwaites, Alexandra Woolgar, Bowen Zhou, and Chao Zhang. Brainomni: A brain foundation model for unified eeg and meg signals. *Advances in Neural Information Processing Systems*, 38:41179–41212, 2026.

Chaoqi Yang, M Brandon Westover, and Jimeng Sun. BIOT: Biosignal transformer for cross-data learning in the wild. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023.

Muquan Yu, Mu Nan, Hossein Adeli, Jacob Prince, John A Pyles, Leila Wehbe, Maggie Henderson, Michael Tarr, and Andrew Luo. Meta-learning an in-context transformer model of human higher visual cortex. *Advances in Neural Information Processing Systems*, 2026.

Zhizhang Yuan, Fanqi Shen, Meng Li, Yuguo Yu, Chenhao Tan, and Yang Yang. Brainwave: A brain signal foundation model for clinical applications. *arXiv preprint arXiv:2402.10251*, 2024.

Daoze Zhang, Zhizhang Yuan, Yang Yang, Junru Chen, Jingjing Wang, and Yafeng Li. Brant: Foundation model for intracranial neural signal. *Advances in Neural Information Processing Systems*, 2023a.

Dongxue Zhang, Huiying Li, and Jingmeng Xie. Mi-cat: A transformer-based domain adaptation network for motor imagery classification. *Neural networks*, 2023b.

Zhiqi Zhang, Siqi Cai, Xueyi Zhang, and Xian Tang. Explainable spatial-temporal-spectral meta-learning for subject-independent eeg-based emotion recognition. *Expert Systems with Applications*, 2026.

He Zhao, Qingqing Zheng, Kai Ma, Huiqi Li, and Yefeng Zheng. Deep representation-based domain adaptation for nonstationary eeg classification. *IEEE Transactions on Neural Networks and Learning Systems*, 2020.

Li-Ming Zhao, Xu Yan, and Bao-Liang Lu. Plug-and-play domain adaptation for cross-subject eeg-based emotion recognition. In *Proceedings of the AAAI conference on artificial intelligence*, 2021.

Xiao-Cong Zhong, Qisong Wang, Dan Liu, Jing-Xiao Liao, Runze Yang, Sanhe Duan, Guohua Ding, and Jinwei Sun. A deep domain adaptation framework with correlation alignment for eeg-based motor imagery classification. *Computers in biology and medicine*, 2023.