

基于MAML的少样本跨被试事件相关电位分类快速适应方法

陈朝琛¹, 季洪飞¹, 白钟飞^{2*}, 李洁^{1*}

¹同济大学计算机科学与技术学院, 上海

²上海市养志康复医院, 上海

收稿日期: 2026年4月23日; 录用日期: 2026年5月22日; 发布日期: 2026年5月29日

摘要

针对跨被试事件相关电位(ERP)解码中个体间差异大、新用户校准样本极少导致传统方法泛化性能严重下降的问题, 文章提出一种基于模型无关元学习(MAML)的少样本快速适应方法DF-MAML。该方法以预训练的域不变特征提取器为初始参数, 通过元学习框架在源被试构建的大量模拟任务上进行双层优化: 内循环利用支持集模拟快速校准, 外循环优化元参数使其具备“易于快速适应”的特性。在此基础上, 进一步提出基于任务难度的动态分类器冻结策略, 根据支持集与查询集的分布偏移自适应地决定是否冻结分类器梯度, 从而稳定特征学习并提升元参数质量。在公开ERN数据集与自采中文语义句法违反数据集上的实验表明, 在5-shot设定下, 本方法的平均AUC分别达到0.7563和0.6705, 显著优于EEGNet、原型网络及标准MAML等基线方法; 消融实验验证了动态冻结策略的有效性; 少样本性能曲线显示, 提供5个校准样本可使性能较零样本提升约9%。本方法为低校准负担、高鲁棒性的脑机接口系统提供了关键技术支撑。

关键词

事件相关电位, 跨被试分类, 少样本学习, 模型无关元学习

MAML-Based Rapid Adaptation for Few-Shot Cross-Subject ERP Classification

Chaochen Chen¹, Hongfei Ji¹, Zhongfei Bai^{2*}, Jie Li^{1*}

¹School of Computer Science and Technology, Tongji University, Shanghai

²Yangzhi Rehabilitation Hospital, Shanghai

Received: April 23, 2026; accepted: May 22, 2026; published: May 29, 2026

*共同通讯作者。

文章引用: 陈朝琛, 季洪飞, 白钟飞, 李洁. 基于 MAML 的少样本跨被试事件相关电位分类快速适应方法[J]. 计算机科学与应用, 2026, 16(5): 350-364. DOI: 10.12677/csa.2026.165189

Abstract

To address the challenge in cross-subject event-related potential (ERP) decoding, where large inter-subject variability and extremely limited calibration samples for new users severely degrade the generalization performance of conventional methods, this paper proposes a MAML-based few-shot rapid adaptation approach. The proposed method initializes with a pre-trained domain-invariant feature extractor and performs bi-level optimization over a large number of simulated tasks constructed from source subjects within a meta-learning framework. Specifically, the inner loop leverages the support set to emulate rapid calibration, while the outer loop optimizes meta-parameters to facilitate fast adaptation. Furthermore, a task-difficulty-aware dynamic classifier freezing strategy is introduced, which adaptively decides whether to freeze classifier updates based on the distribution shift between the support and query sets, thereby stabilizing feature learning and improving the quality of the meta-learned parameters. Experiments conducted on a public ERN dataset and a self-collected Chinese semantic-syntax violation dataset demonstrate that, under the 5-shot setting, the proposed method achieves average AUC scores of 0.7563 and 0.6705, respectively, significantly outperforming baseline methods including EEGNet, Prototypical Networks, and standard MAML. Ablation studies further validate the effectiveness of the proposed dynamic freezing strategy. The few-shot performance analysis shows that using only five calibration samples yields an improvement of approximately 9% over the zero-shot setting. The proposed method offers a promising solution for developing brain-computer interface systems with low calibration cost and high robustness.

Keywords

Event-Related Potential, Cross-Subject Classification, Few-Shot Learning, Model-Agnostic Meta-Learning

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

脑机接口(Brain-Computer Interface, BCI)技术,特别是基于事件相关电位(Event-Related Potential, ERP)的范式,已在神经康复、认知增强与人机交互等领域展现出巨大的应用潜力。然而,从实验室研究走向实际广泛部署的过程中,系统的高校准负担与低用户友好性成为核心障碍之一。由于个体间在神经解剖结构、认知策略及生理状态等方面存在显著差异,在源域数据上训练出的高性能解码模型,在目标域上往往出现显著的性能下降,即被试间变异性。传统的解决方案是为每位新用户收集大量脑电数据,进行完整的重新校准。但该过程耗时冗长,给用户带来沉重的认知与时间负担,严重影响了BCI系统的实用性和用户接受度。

为降低校准成本,域适应方法将预训练的通用模型在新用户的少量数据上继续进行梯度下降训练。然而,在极少量样本的约束下,直接微调面临严峻挑战。首先,存在严重的过拟合风险:深度模型参数众多,少量样本无法约束其优化方向,模型极易记住特定样本的噪声而非学习有意义的个体化模式,导致在后续新样本上泛化性能崩溃。其次,优化轨迹不稳定:基于随机梯度的优化在极小数据集上波动极大,可能导致模型性能不升反降,甚至破坏预训练模型中学到的宝贵通用知识。因此,迫切需要一种新的范式——不仅能利用少量样本,更能高效、稳健地利用它们,实现快速且可靠的个性化适应。

元学习(Meta-Learning)为解决这一难题提供了新思路。其核心思想是通过在大量相关任务(例如, 从不同源被试构建的众多少样本分类任务)上进行元训练, 使模型掌握快速适应新任务的能力。具体到基于优化的元学习, 如模型无关元学习(Model-Agnostic Meta-Learning, MAML), 其目标是找到一个优良的模型初始化参数, 使得模型从初始点出发, 仅需极少的梯度更新步数和少量样本就能在新任务中达到优异性能。这使得 BCI 获得快速校准的能力, 将一个预先训练好的、具备良好泛化能力的模型, 通过元学习进一步改造, 使其成为一个易适应模型。当新用户提供少量数据时, 该模型能以此为原点, 迅速调整从而捕捉该用户的独特神经特征, 避免发生过拟合。

基于以上分析, 本文提出一种面向 ERP 特征的元学习快速适应框架。该框架以预训练的域不变特征提取器为元初始化起点, 构建双层优化过程: 内循环利用新被试的极少量支持数据进行快速适配, 外循环通过查询集上的损失回传优化元参数。同时, 针对少样本适配中分类器决策边界振荡的问题, 进一步设计了基于任务难度的动态分类器冻结策略, 以稳定训练并提升元参数质量。训练完成后, 模型支持零样本与少样本两种应用模式, 可灵活应对不同校准数据量的现实需求。在公开 ERN 数据集与自采中文语义违反数据集上的实验验证了本方法的有效性。

2. 相关工作

2.1. ERP 信号分析

事件相关电位(ERP)作为与特定认知事件锁时的脑电信号, 为研究大脑高级功能提供了毫秒级时间分辨率窗口[1]。传统的 ERP 分析以叠加平均法为基石, 利用 ERP 的锁时性与锁相性, 通过多次试次对齐叠加来增强信号、抑制背景噪声[2] [3]。然而, 该方法无法分析单试次变异信息。为更全面刻画 ERP 的动态特性, 短时傅里叶变换(STFT) [4]、小波变换(WT) [5]、希尔伯特-黄变换(HHT) [6]等时频分析方法被广泛引入。

在特征提取后的分类环节, 传统机器学习方法曾长期处于主导地位。支持向量机(SVM) [7]和线性判别分析(LDA) [8] [9]通过寻找最优分类边界实现高效解码。为提升对复杂时空模式的刻画能力, 研究者发展了基于空间滤波的 xDawn 算法[10] [11]、基于黎曼几何的流形分类方法[10] [12]以及判别典型模式匹配(DCPM) [13]等改进策略。然而这些方法普遍存在特征工程与分类器训练相割裂的局限, 性能高度依赖手工特征的有效性。

深度学习的兴起为 ERP 分析带来了范式转变。以卷积神经网络(CNN)为代表的模型能够端到端地学习原始 EEG 中的时空特征[14] [15]。经典架构如 Deep ConvNet [16]、EEGNet [17]通过深度可分离卷积实现了高效的特征提取与分类。EEG-Inception 通过多尺度卷积核并行提取特征[18]。近年来, 以自注意力为核心的 Transformer 架构进一步突破了 CNN 在长程依赖建模上的局限, 如 EEGConformer 通过卷积与 Transformer 的融合实现了对局部精细模式和全局功能连接的综合建模[19]。总体而言, ERP 分析方法正从依赖专家知识的手工特征工程, 向数据驱动的端到端统一学习范式演进。

2.2. 跨被试识别

跨被试识别旨在利用多个源被试的脑电数据构建模型, 使其能够对未见过的目标被试进行准确解码, 是实现“即插即用”脑机接口的关键[20]。然而, 脑电信号存在显著的被试间变异, 导致基于特定个体训练的模型在新被试上性能严重下降。

域泛化(Domain Generalization, DG)旨在通过提取域不变特征使得模型无需校准即拥有强泛化能力。域泛化早期工作侧重于显式的分布对齐: 利用最大均值差异(MMD)约束不同被试特征分布[21], 或通过相关性对齐(CORAL)对齐二阶统计量[22]。数据增强与正则化策略如 Mixup [23]以及不变风险最小化

(IRM) [24]从不同角度提升泛化能力。域对抗神经网络引入对抗博弈机制,通过特征提取器与域判别器的对抗训练生成域混淆特征[25]。Shen 等人将神经科学的“一致性”假设转化为对比学习的优化目标,使模型直接学习跨被试的情绪不变特征[26]。

域适应(Domain Adaptation, DA)通过最小化源域与目标域之间的分布差异,使训练好的模型能够推广至目标域。域适应通常假设在训练过程中可以获得带有稀疏标签或无标签的目标域数据[27]。典型代表是引入域对抗神经网络(DANN)来学习跨域不变特征表示[25]。沿着这一方向, Li 等人开发了双半球域对抗神经网络(BiDANN)用于情绪识别[28], Zhong 等人则将域对抗训练用于模型正则化[29]。Chen 等人采用域判别器与多注意力适配模块,在跨会话运动想象/运动执行场景中提取深度表征[30]。除域对抗策略外, Wang 等人提出了无监督域适应框架用于跨被试运动想象解码[31]。Li 等人还开发了一种半监督元学习方法用于被试迁移校准[32],这与本研究采用的基于元学习的快速适应思路密切相关。

尽管已有诸多进展,但跨被试识别仍面临根本性挑战:任务相关信号与个体特异性噪声在时、空、频域高度混杂,深度模型容易过拟合源被试中的偶然模式而缺乏真正的泛化能力;同时,实际场景中往往仅能获得极少量甚至零个目标被试标注样本,这对模型的少样本/零样本适应能力提出了更高要求。

2.3. 元学习方法

元学习(Meta-Learning),即“学会学习”,旨在通过在大量相关任务上训练,使模型获得快速适应新任务的能力[33]。根据适应机制不同,元学习主要分为三类:基于优化的方法、基于度量的方法和基于模型的方法。

基于优化的元学习的代表是模型无关元学习(MAML) [34]及其变体 Reptile [35]。其核心思想是学习一个良好的参数初始化点,使得从该点出发针对任何新任务仅需少数几步梯度下降即可达到最优性能。MAML 通过内循环(任务特定更新)和外循环(元参数更新)的双层优化实现这一目标。基于度量的元学习则学习一个通用度量空间,例如原型网络(Prototypical Networks),以每个类别支持集样本的特征均值作为类别原型,依靠最近距离原则完成分类[36]。基于模型的元学习通常采用循环神经网络等结构直接生成任务专用参数,适应速度快但模型复杂度高。

近年来,元学习被广泛应用于脑电信号跨被试快速校准。在情绪识别领域, MetaEmotionNet 融合三维注意力与元学习,显著提升对新被试的适应速度[37];部分工作探索了零校准场景,通过学习被试无关的元知识实现直接泛化[38] [39]。在运动想象解码中,结合预训练与元学习(如 Reptile)被证明能有效减少校准时间[35] [40]。此外,元学习还被拓展至神经退行性疾病 EEG 生物标志物识别及隐私保护场景[41]。上述研究一致表明,元学习范式为克服脑电个体差异大、标注数据稀缺的难题提供了极具前景的解决路径。

3. 方法

3.1. 问题形式化与双模元任务构建

为将跨被试 ERP 少样本适应问题纳入元学习框架,首先对学习阶段与任务进行形式化定义:设有包含 M 个被试的完整数据集 $D = \{D_i\}_{i=1}^M$,其中 D_i 表示第 i 个被试的数据,包含 C 个类别的标记样本。采用留一被试交叉验证(Leave-One-Subject-Out, LOSO)策略来划分元训练与元测试阶段,每轮选定一个被试 D_j 作为元测试目标域,用以模拟新用户;其余 $M-1$ 个被试的数据 $D_{\text{source}} = D \setminus D_j$ 构成元训练源域。目标是利用源域 D_{source} 进行元训练,使模型能够在目标域 D_j 上,仅使用其 K 个样本就能快速学习。

在源域中,采用嵌套留一法构建元任务以模拟跨被试分布偏移:依次将其中每一个被试 D_s 视为一个“伪目标域”,并将其数据作为查询集 Q_s 。其余 $M-2$ 个被试的数据共同作为支持集 S_s 。由此可构建 $M-1$ 个元任务 $\{T_i\}$,每个任务强制模型学习处理被试间分布差异,以增强跨被试泛化与快速适应能力。

每个元任务 T_i 都包含一个独特的支持集 - 查询集对 (S_i, Q_i) 。元训练阶段的目标并非直接学习在 D_{source} 上的分类, 而是学会如何利用任意任务的支持集 S_i 快速调整自身参数, 以最小化在该任务查询集 Q_i 上的损失。在元训练中, 参数为 θ 的模型在大量元任务 T_i 上进行训练, 先用支持集 S_i 进行内循环的快速适应, 得到适应后的参数 θ'_i ; 随后用 θ'_i 在查询集 Q_i 上计算损失, 该损失的梯度被用于外循环, 以更新原始参数 θ 。此阶段的目标是优化 θ , 使其成为一个能快速适应任何新任务的优秀“元初始化点”。

元测试阶段, 定义两种应用模式: 1) 少样本场景: 固定元模型参数 θ 。对新目标被试 D_j , 从其数据中随机抽取 K 个样本构成最终的支持集 S_{target} , 剩余为测试集。模型基于 S_{target} 对 θ 进行几步梯度更新, 得到适应后的模型, 并在测试集上评估其性能。2) 零样本场景: 在无法获取目标被试数据时, 先使用全部源域被试的数据对元模型的 θ 进行整体微调。随后, 将适应后的模型直接应用于目标被试进行测试。

通过以上形式化, 将跨被试适应问题转化为一个可优化的元学习问题, 严格模拟了现实中使用极少量用户数据完成快速校准并投入使用的全过程, 为后续具体算法框架的设计奠定基础。

3.2. 模型框架

为实现从预训练特征到快速适应的演进, 本章构建了以模型无关元学习(MAML)为核心, 使用动态梯度冻结策略的统一框架 DF-MAML (Dynamic Freezing MAML)。该框架将基于对比学习的预训练双视图编码器(Multi-View Contrastive Learning Domain Generalization, MVCLDG)作为元学习特征提取网络的强初始化起点, 在 MAML 元训练过程中使用基于任务难度的动态梯度冻结策略。整个框架涵盖元训练、元适应及少样本测试三个阶段, 形成从学习通用适应策略到快速校准的完整闭环, 模型整体架构见图 1:

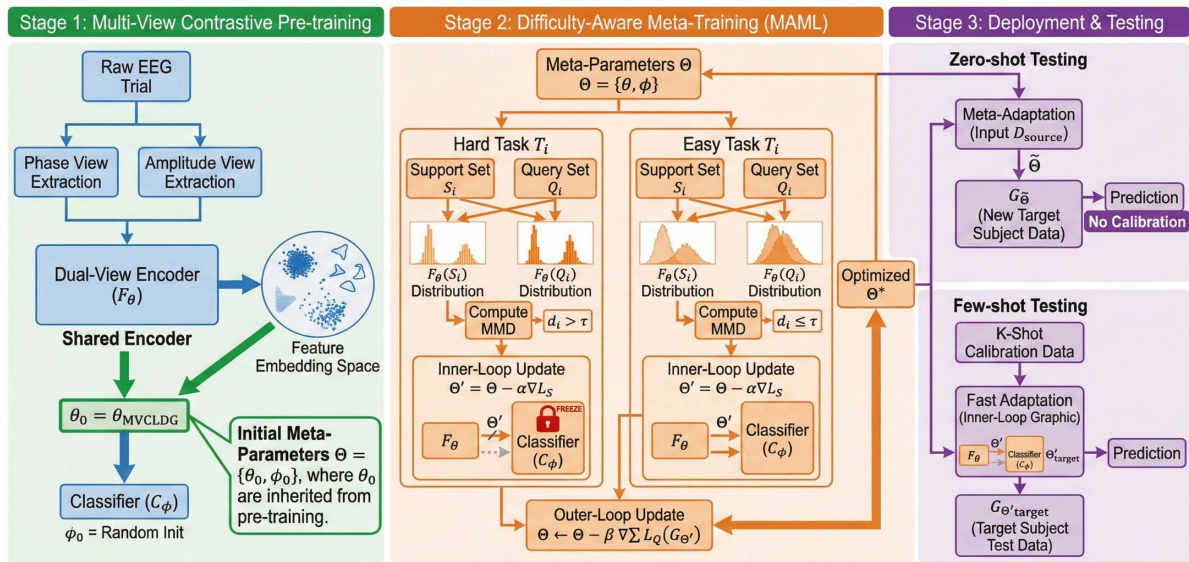


Figure 1. Model framework based on MAML
图 1. 基于 MAML 的模型框架

框架的初始化是后续所有优化的基石。待优化的元学习模型定义如下:

$$G_{\Theta} = C_{\phi} \circ F_{\theta} \tag{1}$$

其中, F_{θ} 为特征提取器, C_{ϕ} 为分类器, $\Theta = \{\theta, \phi\}$ 代表所有待优化的元参数。特征提取器的初始参数 θ_0 直接继承自预训练的编码器参数 θ_{MVCLDG} , 而分类器参数 ϕ 保持随机初始化。 θ_{MVCLDG} 在多被试数据上通过对比学习与域对齐损失进行预训练, 模型结构如图 2 所示, 以原始 EEG 及其希尔伯特变换信号作为输

入，分别建模时域幅值与瞬时相位特征。同时通过引入对比学习损失，在特征空间中约束跨被试同类特征的对齐与聚集，从而引导模型学习判别性强、对域变化鲁棒的域不变特征。其产生的特征空间能有效解耦幅值与相位信息，并抑制被试间变异。这一策略使元训练专注于学习“如何以最少步骤将已有特征快速适配到新被试”，而非从零学习判别性特征。

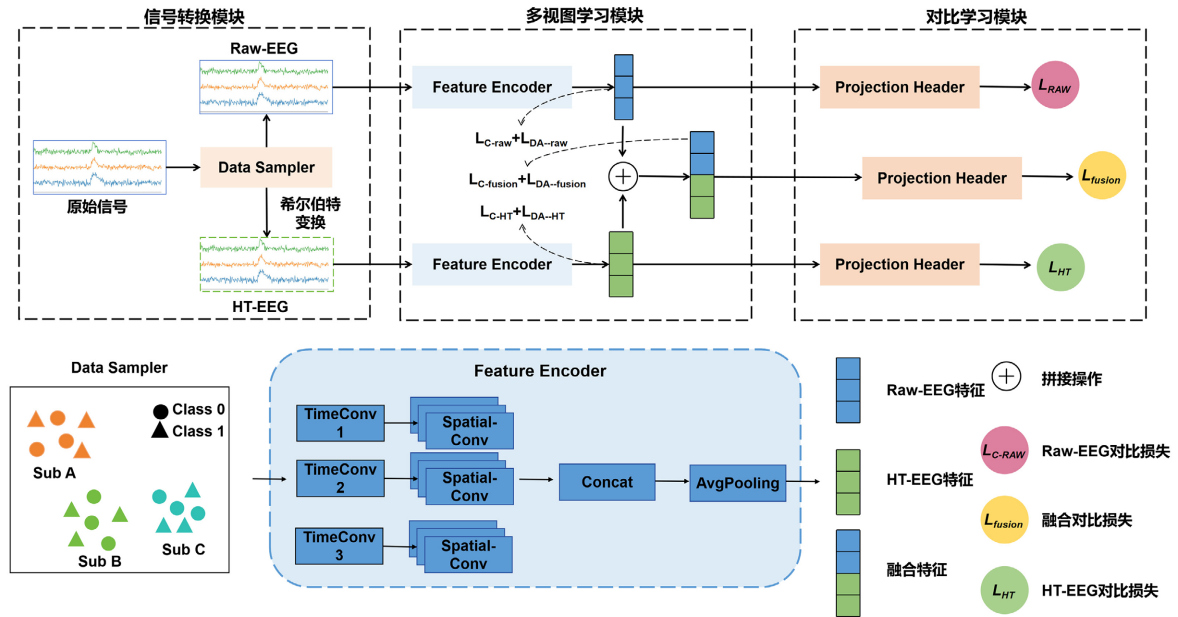


Figure 2. Model framework of MVCLDG
图 2. MVCLDG 模型框架

元训练严格遵循之前定义的嵌套留一法，利用元任务集合 $\{T_i\}$ 驱动双层优化过程。元训练通过 MAML 的双层优化机制锻造出具备易适应特性的元参数 Θ^* 。对于每个元任务 T_i ，内循环模拟快速个性化校准：模型从当前元参数 Θ 出发，仅基于该任务的支持集 S_i 计算损失并执行若干步梯度下降，得到一个针对该任务调整后的参数。

$$\Theta'_i = \Theta - \alpha \nabla \Theta \mathcal{L}_{S_i}(G_\Theta) \quad (2)$$

其中， α 为内循环学习率， \mathcal{L}_{S_i} 为支持集上的损失。外循环评估这次快速校准的泛化效能，它计算每个临时模型 $G_{\Theta'_i}$ 在其对应查询集 Q_i 上的损失，并汇总所有任务的查询损失以更新初始参数：

$$\Theta \leftarrow \Theta - \beta \nabla \Theta \sum_{i=1}^{M-1} \mathcal{L}_{Q_i}(G_{\Theta'_i}) \quad (3)$$

其中， β 为外循环学习率，梯度计算需通过链式法则回溯至初始参数 Θ 。通过大量元任务的迭代，初始元参数 Θ 收敛至一个独特的最优点 Θ^* ：从这个点出发，向任意新被试数据分布方向的微小梯度更新，都能高效地引导模型达到对该被试优异的表现。

获得 Θ^* 后，框架分化为两条路径。第一条路径是“元适应 - 零样本测试”模式，利用全部源被试数据 D_{source} 对 Θ^* 进行几轮整体微调，得到增强参数 $\tilde{\Theta}$ 。此步骤可视为利用集体经验对模型的通用性进行最终强化。随后，将模型直接部署于全新目标被试，实现零校准。第二条路径是“少样本快速适应 - 测试”模式，当获得目标被试提供的极少量(K-shot)校准样本后，直接利用 Θ^* 的易适应性，启动与元训练内循环相同的快速适应过程，生成个性化参数 Θ'_{target} ，随后进行测试。该框架在算法层面统一了零样本与少样本

学习范式，为脑机接口系统提供了兼具性能与灵活性的解决方案。

3.3. 基于被试差异的动态分类器梯度冻结策略

在 MAML 元学习框架中，内循环快速适应过程的稳定性与效率直接影响元参数 Θ^* 的质量。特征提取器与分类器的优化目标存在内在张力：分类器致力于构建清晰的决策边界，而特征提取器需产生对分布变化鲁棒的特征。当不同元任务的数据分布差异显著时，分类器的快速调整可能导致决策边界剧烈振荡，进而干扰特征提取器学习跨任务一致的特征表示。引入间歇性固定周期冻结分类器，可为特征提取器提供相对静止的决策参考系，促进稳定的特征学习，如图 3 所示。

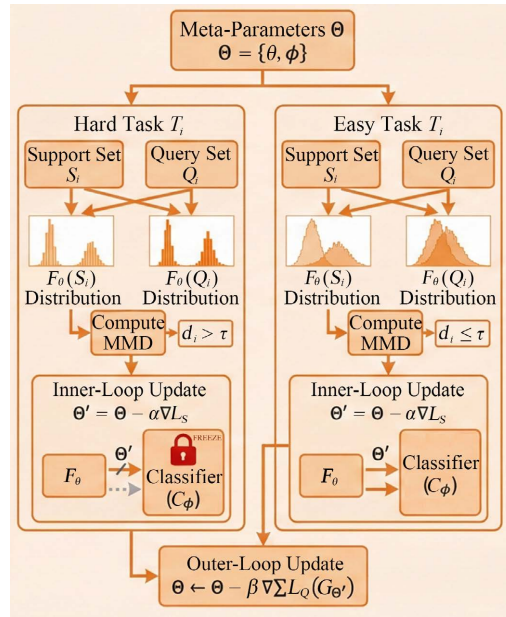


Figure 3. Dynamic classifier gradient freezing strategy based on subject differences

图 3. 基于被试差异的动态分类器梯度冻结策略

然而，固定周期冻结是与数据无关的启发式方法，未考量任务的实际难度。跨被试脑电解码中，不同源被试组合构成的元任务难度差异显著：对于支持集与查询集分布相似的“简单”任务，同步更新两者高效无害；对于分布差异巨大的“困难”任务，盲目更新分类器易引发决策边界失稳。因此，本研究提出基于任务难度的动态分类器冻结策略，在线评估每个元任务的适配难度，自适应调整优化重心。

定义任务难度系数 d_i 为度量支持集 S_i 和查询集 Q_i 在经过当前特征提取器 F_θ 映射后的特征分布之间的差异。采用最大均值差异(Maximum Mean Discrepancy, MMD)作为分布距离的度量标准。MMD 计算两个分布样本在再生核希尔伯特空间(RKHS)中均值嵌入的距离，其定义如下：

$$d^i = \text{MMD}^2[\mathcal{F}, F_\theta(S_i), F_\theta(Q_i)] = \left\| \frac{1}{|S_i|} \sum_{x \in S_i} \Phi(F_\theta(x)) - \frac{1}{|Q_i|} \sum_{x \in Q_i} \Phi(F_\theta(x)) \right\|_{\mathcal{H}}^2 \quad (4)$$

其中， $F_\theta(\cdot)$ 表示由当前特征提取器参数 θ 生成的特征映射函数， Φ 是将特征映射 RKHS \mathcal{H} 的核映射。MMD 距离 d_i 的值越大，支持集与查询集的分布差异越显著，二者间的迁移任务越“困难”。

获得难度评估 d_i 后，动态冻结决策模块根据预设的阈值 τ 来裁定分类器的更新状态。阈值 τ 设定为所有元任务难度系数的分位数(如中位数)。若 $d_i > \tau$ ，则判定任务 T_i 为困难任务。在此情况下，系统将冻结分类器参数 ϕ ，仅更新特征提取器参数 θ ，迫使模型在面对高分布偏移时，将全部的优化能力集中于特

征提取器，学习弥合分布差距的特征。若 $d_i \leq \tau$ ，则判定为容易任务，允许 θ 与 ϕ 同步更新，以追求更优的任务特定性能。

将动态冻结策略集成到元学习的双层优化框架中，需要对内循环和外循环的参数更新规则进行条件化重构。以内循环为例，对于任务 T_i ，其参数适应过程修改如下：

$$\theta'_i = \theta - \alpha_\theta \nabla_{\theta} \mathcal{L}_{S_i}(G_{\Theta}) \quad (5)$$

$$\phi'_i = \begin{cases} \phi - \alpha_\phi \nabla_{\phi} \mathcal{L}_{S_i}(G_{\Theta}), & \text{if } d_i \leq \tau \\ \phi, & \text{if } d_i > \tau \end{cases} \quad (6)$$

其中， α_θ 和 α_ϕ 分别为特征提取器和分类器在内循环的学习率。外循环中，查询损失梯度向初始元参数 Θ 的传播也遵循相同的条件逻辑： $\partial \mathcal{L}_{\Theta} / \partial \phi$ 的计算图依赖于内循环中 ϕ 是否被更新。

3.4. 零样本/少样本双模适应流程

在获得最优元参数 Θ^* 之后，模型支持两种部署模式：“元适应 - 零样本测试”与“少样本快速测试”，二者共享元训练优化的特征提取基础。

“元适应 - 零样本测试”模式面向的是对用户友好性要求极高、不允许任何形式校准的即插即用场景。以最优元参数 $\Theta^* = \{\theta^*, \phi^*\}$ 为起点。由于 Θ^* 的目标是追求快速适应的潜力，而非在完全未知被试上的直接分类精度，因此引入一个元适应步骤，利用全部源域数据 D_{source} 进行数轮额外的整体微调，得到一组增强的模型参数 $\Theta^- = \{\theta^-, \phi^-\}$ 。该模式消除了用户准备时间，但其性能上限在很大程度上依赖于源域数据的多样性与泛化强度。

“少样本快速测试”模式适用于可以容忍短暂校准环节，并追求更高个性化性能的场景。当系统获得目标用户提供的极少量标记样本(即 K-shot 支持集 $S_{\text{target}} = \{(x_k, y_k)\}_{k=1}^{K \times C}$ ，其中 C 为类别数)后，立即启动一个与元训练内循环完全同构的快速适应过程：

$$\Theta'_{\text{target}} = \Theta^* - \alpha \nabla_{\Theta^*} \mathcal{L}_{S_{\text{target}}}(G_{\Theta^*}) \quad (7)$$

其中， α 为快速适应学习率。通过几步梯度更新得到个性化参数 $G_{\Theta'_{\text{target}}}$ ，随后冻结模型并在测试集上评估。该模式校准效率高，能将通用模型快速塑造为个性化版本，将校准时间与数据需求降至极低。

4. 实验结果与分析

4.1. 数据集

本节通过事件相关电位(ERP)分类任务，对所提出方法在两个数据集上的有效性进行评估，其中一个为自采数据集。数据集情况如下：

从 ERN (Feedback Error-Related Negativity)数据集[42]中选取 16 名健康受试者的脑电信号，该数据集采用 56 通道扩展 10~20 系统记录，采样率为 600 Hz。实验通过 6×6 字符矩阵诱发 P300 反应。每名被试完成 340 个试次，其中错误反馈试次中位数为 115.5 (范围 20~199)，正确反馈试次中位数为 237.5 (范围 141~320)。预处理时，使用 EEGLAB 软件的反 FIR 滤波器对 1~40 Hz 范围内的信号进行带通滤波，降采样至 128 Hz，并提取反馈后[0, 1.25]秒时间窗内的信号作为二分类任务的特征数据。

CSSV-ERP (Chinese Semantic-Syntactic Violation ERP)数据集为自采数据，包含 12 名健康受试者(10 男 2 女，年龄 22~26 岁)，采用 Brain Products actiChamp Plus 放大器和 64 通道 actiCAP slim/snap 电极(国际 10~20 系统)记录，采样率为 500 Hz。实验以逐词形式呈现中文句子，包含三种条件：正确句子(CORR)、局部句法违反(SYN-P)和语义违反(SEM)，刺激呈现由 S1/S2/S4 标记(分别对应正确/语义违反

/句法违反), 被试反应由 S7/S8 标记(正确/错误)。每名被试完成 192 个试次(每类 64 个)。预处理阶段, 对 0.1~60 Hz 范围内的信号进行带通滤波, 使用独立成分分析(ICA)去除眼电和肌电伪迹, 以 TP9 和 TP10 为参考进行重参考。提取关键词出现后[-0.15, 1]秒时间窗内的信号作为多分类特征, 并进行基线校正, 试次保留率中位数为 94.2% (范围 82%~100%)。

4.2. 基于公开数据集的算法性能对比

为验证方法的普适性, 首先在 ERN 公开数据集上进行了评测。在基线模型方面, 选取了六种代表性的基线方法进行对比, 包括经典的 EEGNet、基线模型、基于优化的元学习方法(MAML、Reptile)、基于度量的方法(原型网络), 以及近年提出的专门针对 EEG 的域适应方法(MFA-LR)。本文采用 AUC (Area under the ROC Curve)作为主要评估指标, 因为其对类别不平衡不敏感, 能够更全面地衡量模型在二分类及多分类任务中的排序性能与跨被试泛化能力。为模拟基于极少量目标被试校准数据环境, 本实验对于所有域适应与元学习方法, 均仅使用从目标被试数据中随机采样的 5 个带标签样本(即 5-shot)构成支持集, 以进行模型适配或原型计算。

定量结果如表 1 所示, 本文方法在 ERN 数据集上取得了最优的综合性能。

Table 1. Comparison of AUC between this model and the Baseline model on the ERN dataset
表 1. 本模型与 Baseline 模型在 ERN 数据集上的 AUC 对比

方法	sub1	sub2	sub3	sub4	sub5	sub6	sub7	sub8	sub9
EEGNet	0.5623	0.5397	0.7082	0.6872	0.5686	0.7585	0.5996	0.6917	0.8068
MVCLDG	0.6304	0.6973	0.8261	0.754	0.5386	0.7986	0.6524	0.7062	0.8394
MAML	0.7205	0.7129	0.7996	0.7622	0.6318	0.8085	0.6292	0.7015	0.8240
Reptile	0.7401	0.6908	0.7789	0.7379	0.6281	0.8050	0.6510	0.6970	0.8285
原型网络	0.6947	0.6523	0.7502	0.7426	0.5833	0.7667	0.5766	0.6892	0.7999
MLA-LR	0.6959	0.519	0.7131	0.7307	0.6099	0.7831	0.5752	0.687	0.8261
本模型	0.7377	0.7166	0.7768	0.7744	0.6599	0.8056	0.6605	0.7331	0.8362
方法	sub10	sub11	sub12	sub13	sub14	sub15	sub16	均值	方差
EEGNet	0.6745	0.6848	0.6717	0.8302	0.7596	0.681	0.5268	0.6720	0.0920
MVCLDG	0.6658	0.7311	0.8611	0.8414	0.7871	0.6862	0.5779	0.7246	0.0974
MAML	0.6715	0.7070	0.9429	0.8966	0.7509	0.7172	0.5847	0.7413	0.0961
Reptile	0.6664	0.6921	0.9242	0.8683	0.7295	0.7204	0.5963	0.7347	0.0883
原型网络	0.6954	0.6212	0.8543	0.7755	0.7187	0.7096	0.5334	0.6977	0.0871
MLA-LR	0.6369	0.6802	0.8016	0.8504	0.7422	0.7048	0.543	0.6937	0.0979
本模型	0.6966	0.7458	0.9505	0.8934	0.7452	0.7502	0.6185	0.7563	0.0856

总体而言, 具备参数化快速适应能力的基于梯度的元学习方法(包括 MAML、Reptile 及本方法)在 5-shot 设定下均展现出对基线模型 MVCLDG 的帮助。这一结果验证了: 通过元学习获得对任务分布敏感、且易于快速调整的模型初始化, 是应对极少量校准样本的有效范式, 但该范式内部存在明显的性能分化。本方法取得了最佳结果, 在 AUC 这一核心指标上, 所提方法达到了 0.7563。而经典的 MAML 与 Reptile 基线表现相对有限, 仅仅达到 0.7413 和 0.7347。

另一方面, 基于度量的原型网络与基于域特征对齐的 MFA-LR 方法在本实验中表现不佳。原型网络的性能在处理 ERN 这类存在较大被试间波形变异的数据时, 其依赖的“原型”中心难以稳定建立, AUC 达到 0.6977, 仅略高于传统的 EEGNet 算法。ERN 信号固有的高被试间变异性与 5-shot 支持集所提供的

有限样本量, 共同导致估计出的“原型”点统计显著性不足, 无法可靠地代表整个类的分布, 从而使得基于距离的分类决策失效。MFA-LR 由于其依赖于对全局特征分布的强假设或特定的互信息估计, 在 ERN 数据集复杂的试次间波动面前泛化能力不足, AUC 达到 0.6937。该类方法通常需要对源域与目标域的联合分布或互信息进行精确估计, 而在仅有 5 个目标域样本的情况下, 基于统计的估计容易产生偏差。这种估计偏差会误导对齐过程, 使其无法捕捉真实的域不变特征, 最终导致跨被试泛化的失败。

同时, 为了评估方法在真实场景中的可行性, 在统一硬件平台(NVIDIA RTX 5000, Ubuntu 22.04)上对比了各模型的计算成本, DF-MAML 元训练总耗时 3.9 小时(含预训练 2.05 小时与元训练 1.85 小时), MAML 为 2.1 小时, EEGNet 单被试训练 0.5 小时。5-shot 部署阶段, DF-MAML 与 MAML 单次适应耗时分别为 0.15 秒和 0.13 秒, 均小于 1 秒, 满足实时 BCI 响应要求。EEGNet 无需适应。结果表明, 本方法在保持性能优势的同时, 计算成本与 MAML 相当, 适应极快, 具备实用潜力。

上述结果表明, 本研究提出的模型避免了直接在少样本上估计复杂的全局分布差异, 而是通过对比学习隐式地塑造不变特征, 并通过轻量级适配器进行针对性微调, 从而在机制上更为稳健, 其输出的类别概率估计具有较好的校准性与判别力。

4.3. 基于自采数据集的算法性能对比

为进一步验证方法在更复杂认知任务上的有效性, 在自采的中文语义句法违反数据集上进行了实验。该数据集诱发 N400/P600 等经典 ERP 成分, 其信号相比公开的 ERN 数据集具有更高的试次间变异性和背景噪声, 且涉及的认知神经过程更为复杂。实验遵循与公开数据集评测一致的方法, 所有需要适应目标被试的元学习及域适应方法, 均仅使用随机抽取的 5 个带标签样本(5-shot)作为支持集进行快速适配或原型计算。

定量结果如表 2 所示, 在这一更具挑战性的数据集上, 所提方法的优势更为明显:

Table 2. Comparison of AUC between this model and the Baseline model on the CSSV-ERP dataset

表 2. 本模型与 Baseline 模型在中文语义句法违反数据集上的 AUC 对比

方法	sub1	sub2	sub3	sub4	sub5	sub6	sub7
EEGNet	0.5398	0.5243	0.6919	0.5752	0.5235	0.5713	0.5452
MVCLDG	0.5277	0.6051	0.7219	0.6908	0.671	0.5822	0.622
Reptile	0.5997	0.7444	0.7105	0.6367	0.6382	0.6197	0.6768
MAML	0.5856	0.7354	0.7195	0.7002	0.6610	0.6380	0.6643
原型网络	0.5581	0.6596	0.6448	0.5591	0.5568	0.5519	0.5746
MFA-LR	0.5950	0.6562	0.7028	0.6823	0.6284	0.6159	0.6455
本模型	0.5983	0.7354	0.7206	0.6708	0.6568	0.6179	0.6903
方法	sub8	sub9	sub10	sub11	sub12	平均值	标准差
EEGNet	0.6464	0.6149	0.6547	0.5862	0.5029	0.5814	0.0595
MVCLDG	0.6725	0.7006	0.7335	0.6951	0.6045	0.6522	0.0630
Reptile	0.6701	0.6759	0.6877	0.5921	0.5782	0.6525	0.0504
MAML	0.6698	0.6595	0.7253	0.5989	0.5716	0.6608	0.0546
原型网络	0.5985	0.6198	0.5623	0.4960	0.5616	0.5786	0.0451
MFA-LR	0.6321	0.6524	0.6130	0.6019	0.5139	0.6283	0.0481
本模型	0.6956	0.6622	0.7107	0.5882	0.6991	0.6705	0.0479

从方法范式的角度进行分析, 实验结果呈现出与公开数据集一致的趋势。首先, 基于梯度的元学习方法(Reptile、MAML 及本模型)整体上依然展现出对基础模型(EEGNet)及 MVCLDG 模型的性能提升, 这进一步证实了元学习初始化策略在应对极少量校准样本问题上的普适价值。然而, 在该范式内部, 本方法(0.6705)相较于 MAML (0.6608)与 Reptile (0.6525)保持了稳定的优势。这一优势在噪声更高、机制更复杂的语言 ERP 数据中得以维持, 模型在复杂场景下的微调效率与稳定性优于直接在原始信号空间进行元优化的经典方法。

其次, 基于度量的原型网络在本数据集上的表现不佳(0.5786), 本研究推测这一结果是其在任务复杂度升级(如从二分类扩展至三分类)时所面临的新挑战。在类别数量增加而每个类别的支持样本数(5-shot)保持不变的情况下, 原型网络需要为更多类别建立有区分度的原型中心, 这使得本就有限的样本所承载的统计信息被进一步稀释, 导致其基于距离的度量机制近乎失效。MFA-LR 方法在本数据集上的表现亦不理想, 该结果强化了本研究在公开数据集实验中的分析: 在高噪声、高个体差异且支持样本极少的条件下, 进行复杂信息论估计的域对齐方法, 其估计过程极易产生偏差, 从而限制了其跨被试泛化效能。

综上所述, 在噪声更高、神经机制更复杂的自采中文语义句法违反数据集上, 本方法所采用的轻量元学习适配器快速微调策略再次展现出显著的优越性。该方法避免了对少量样本进行不稳定的全局分布估计或原型计算, 而是通过预训练凝聚跨被试的判别性特征先验, 再通过高效的参数化机制进行精准调整, 从而在极具挑战的少样本跨被试脑电解码任务中实现了更可靠的鲁棒性与泛化能力。

4.4. 消融实验

为深入剖析所提框架中各个关键组件的贡献, 并验证其设计选择的必要性, 在 ERN 公开数据集上设计并实施了一系列系统的消融实验。并构建了三个逐步接近完整模型的变体进行对比分析:

1) w/o Meta: 移除整个元学习适配器, 仅使用预训练的多视图特征提取器 + 一个简单的分类器进行零样本测试。

2) w/ Meta (No-Freeze): 保留元学习适配器, 但在针对新被试的少样本(5-shot)内循环适配过程中, 不实施任何参数冻结策略, 即采用类似标准 MAML 的方式同时更新特征提取器与分类器的所有权重。

3) w/ Meta (Static-Freeze): 在内循环适配时, 采取静态冻结(Static-Freeze)策略, 即以固定周期冻结分类器参数, 同时更新训练特征提取器的全部参数。

Table 3. Ablation experiment

表 3. 消融实验

方法	sub1	sub2	sub3	sub4	sub5	sub6	sub7	sub8	sub9
w/o Meta	0.6304	0.6973	0.8261	0.754	0.5386	0.7986	0.6524	0.7062	0.8394
w/ Meta (No-Freeze)	0.7365	0.6592	0.7948	0.7404	0.6224	0.7901	0.6604	0.6681	0.8089
w/ Meta (Static-Freeze)	0.7413	0.7129	0.7804	0.7586	0.6235	0.8180	0.6175	0.7194	0.8378
本模型	0.7377	0.7166	0.7768	0.7744	0.6599	0.8056	0.6605	0.7331	0.8362
方法	sub10	sub11	sub12	sub13	sub14	sub15	sub16	均值	方差
w/o Meta	0.6658	0.7311	0.8611	0.8414	0.7871	0.6862	0.5779	0.7246	0.0974
w/ Meta (No-Freeze)	0.6750	0.7119	0.9212	0.8374	0.7150	0.7170	0.6235	0.7301	0.0827
w/ Meta (Static-Freeze)	0.6671	0.7182	0.9348	0.8847	0.7603	0.7258	0.5925	0.7433	0.0946
本模型	0.6966	0.7458	0.9505	0.8934	0.7452	0.7502	0.6185	0.7563	0.0856

定量结果如表 3 所示, 完整模型取得了平均 AUC 为 0.7563 的最佳性能, 且其标准差(0.0856)相对较低, 表明性能跨被试较为稳定。“w/o meta-learning”变体性能最差(0.7246), 这确凿地证实了元学习适配环节不可或缺, 否则模型无法利用极少数目标样本进行任何个性化调整, 其泛化能力受到根本性限制。其次, “w/ Meta (No-Freeze)”变体性能(0.7301)虽优于零样本版本, 但显著低于完整模型。这一结果表明, 盲目更新所有网络参数极易导致灾难性遗忘或对少量支持样本的过拟合, 从而损害模型的跨被试泛化基础。最后, “w/ Meta (Static-Freeze)”(固定周期冻结)变体性能(0.7433)介于前述两者之间。它虽然通过固定周期冻结防止决策边界失稳问题的发生, 但也忽视了新被试特异性进行微调的可能性, 限制了模型适应能力的上限。本文提出的方法中阈值 τ 分别取 1/4 分位数、中位数和 3/4 分位数进行对比, 当 $\tau=3/4$ 分位数时模型性能最优。MMD 核函数采用高斯核($\sigma=1.0$)时效果最佳, σ 在 0.5~2 范围内 AUC 波动 $<1\%$ 。元训练中约 25% 的任务被判定为困难并触发冻结, 冻结后支持集与查询集的特征分布中心距离平均缩短约 10%, 表明该策略有效稳定了特征学习。

综上所述, 消融实验验证了本方法所采用的动态间歇性冻结策略的有效性。该策略使模型在元训练阶段学会任务感知的参数更新模式: 适配新被试时, 动态决定各层参数的更新与否, 保留通用特征稳定泛化, 同时允许部分参数微调以吸纳个性化信息。这种兼具稳定性和灵活性的适配机制, 是实现从通用先验到特定被试高效迁移的关键。

4.5. 少样本适应性实验

为评估所提框架在贴近脑机接口实际部署场景下的实用性, 本节系统探究了模型在不同校准数据规模下的适应性能。本实验在 ERN 数据集上模拟少样本校准场景: 对于每一个作为目标域的新被试, 限定仅使用其 N 个随机抽样的标记试次作为支持集 $\mathcal{S}_{\text{target}} = \{(x_i, y_i)\}_{i=1}^N$ 进行快速适配, 随后在其余查询集上评估性能。设置了 $N \in \{0, 1, 2, 5\}$ 四种条件, 分别对应零样本、单样本、双样本及五样本适应场景。其中, $N=0$ 代表不进行任何基于目标被试的参数更新。

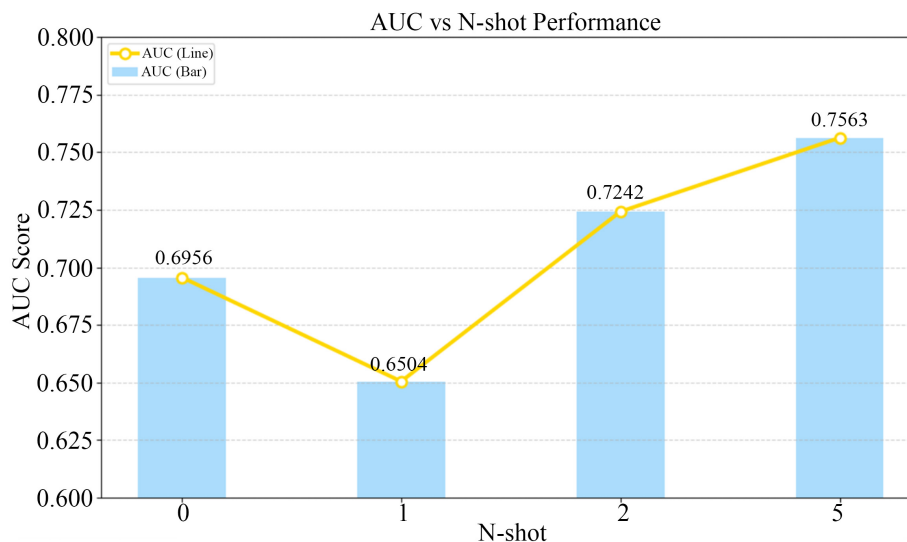


Figure 4. Results of few-shot performance experiments

图 4. 少样本适应性实验结果

如图 4 定量实验结果所示, 模型性能总体上随着校准样本数量 N 的增加而呈现出单调提升的趋势。在 $N=5$ 的条件下, 模型达到了最优的跨被试分类, AUC 为 0.7563。这一结果表明, 即便仅提供极少数

量的校准试次，元学习适配机制也能有效捕捉目标被试的个体特异性模式，对模型参数实现有针对性的微调。值得注意的是，在 $N = 0$ 的零样本条件下，模型仍能保持 0.6956 的基础性能，归功于框架中设计的元适应步骤利用全部源域数据进行整体微调，使模型参数收敛到在源域数据分布上更稳健的状态。

然而，实验结果也出现了一个关键的非单调性现象：在 $N = 1$ 的单样本适应设置下，模型性能(AUC = 0.6504)显著低于 $N = 2$ 和 $N = 5$ 的条件，甚至低于 $N = 0$ 的零样本性能。其根源在于，单样本适配过程从稳健的初始点出发，基于单个样本的损失梯度进行更新，但由于脑电的高试次间变异性，单一样本极有可能无法代表目标被试的真实数据分布，极易将模型参数引导至一个偏离稳健泛化区域的方向，甚至破坏经过元适应步骤获得的、在源域上已充分优化的特征表示，导致性能发生退化。

该结果为实际系统设计提供明确指导：若追求稳定基线，应优先采用零样本策略；若需更高精度，建议提供 5 个以上校准样本以保证支持集多样性，使梯度估计更稳定。本实验验证了所提框架在少样本及零样本条件下的有效性，揭示了校准数据数量与质量对性能的复杂影响，并凸显了元适应步骤对保障零样本稳健性的关键价值。

5. 结束语

本文针对跨被试 ERP 解码中个体差异大、校准样本极少导致传统方法泛化性能严重下降的问题，提出了一种基于模型无关元学习(MAML)的少样本快速适应方法——DF-MAML。该方法以预训练特征提取器为元初始化起点，通过嵌套留一法构建模拟训练任务，通过双层优化框架实现从“通用特征”到“个性化适配”的快速迁移。此外，设计了基于任务难度的动态分类器梯度冻结策略，量化分布偏移，自适应冻结分类器参数来稳定特征学习。同时模型支持零样本与少样本两种部署模式，可灵活应对不同校准数据量的现实需求。

在公开 ERN 数据集与自采中文语义句法违反数据集上的实验表明，在 5-shot 设定下，本方法的平均 AUC 显著优于 EEGNet、原型网络及标准 MAML 等基线方法；消融实验验证了动态冻结策略相较于无冻结和静态冻结的优越性；少样本性能曲线显示，提供 5 个校准样本可使性能较零样本提升约 9%。上述结果充分证明了所提方法在低校准负担、高鲁棒性跨被试 ERP 解码中的有效性与实用性。

本文方法仍存在一定局限性：首先，动态冻结策略中阈值需基于经验分布设定，未来可探索自适应阈值学习方法；其次，当前方法在更多类别场景下的泛化能力有待进一步测试。未来工作将围绕上述问题展开，包括实现更轻量的网络设计、结合半监督学习利用更多无标定目标域数据，进一步推动脑机接口系统的实用化进程。

声明

本研究自采数据集的实验流程经上海养志康复医院伦理审查委员会批准(批准号：养志伦审字(2023) 036 号)，所有受试者均签署知情同意书，研究遵循《赫尔辛基宣言》。

参考文献

- [1] Sur, S. and Sinha, V. (2009) Event-Related Potential: An Overview. *Industrial Psychiatry Journal*, **18**, 70-73. <https://doi.org/10.4103/0972-6748.57865>
- [2] Woldorff, M.G. (1993) Distortion of ERP Averages Due to Overlap from Temporally Adjacent ERPs: Analysis and Correction. *Psychophysiology*, **30**, 98-119. <https://doi.org/10.1111/j.1469-8986.1993.tb03209.x>
- [3] Mouraux, A. and Iannetti, G.D. (2008) Across-Trial Averaging of Event-Related EEG Responses and Beyond. *Magnetic Resonance Imaging*, **26**, 1041-1054. <https://doi.org/10.1016/j.mri.2008.01.011>
- [4] Allen, J. (1977) Short Term Spectral Analysis, Synthesis, and Modification by Discrete Fourier Transform. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **25**, 235-238. <https://doi.org/10.1109/tassp.1977.1162950>

- [5] Samar, V.J., Swartz, K.P. and Raghuvver, M.R. (1995) Multiresolution Analysis of Event-Related Potentials by Wavelet Decomposition. *Brain and Cognition*, **27**, 398-438. <https://doi.org/10.1006/brcg.1995.1028>
- [6] Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., *et al.* (1998) The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, **454**, 903-995. <https://doi.org/10.1098/rspa.1998.0193>
- [7] Müller, K., Mika, S., Tsuda, K. and Schölkopf, K. (2018) An Introduction to Kernel-Based Learning Algorithms. In: Yu Hu, H. and Hwang, J.N., Eds., *Handbook of Neural Network Signal Processing*, CRC Press, 4-1-4-40. <https://doi.org/10.1201/9781315220413-4>
- [8] Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F. and Arnaldi, B. (2007) A Review of Classification Algorithms for EEG-Based Brain-Computer Interfaces. *Journal of Neural Engineering*, **4**, R1-R13. <https://doi.org/10.1088/1741-2560/4/2/r01>
- [9] Krusienski, D.J., Sellers, E.W., McFarland, D.J., Vaughan, T.M. and Wolpaw, J.R. (2008) Toward Enhanced P300 Speller Performance. *Journal of Neuroscience Methods*, **167**, 15-21. <https://doi.org/10.1016/j.jneumeth.2007.07.017>
- [10] Barachant, A., Bonnet, S., Congedo, M. and Jutten, C. (2012) Multiclass Brain-Computer Interface Classification by Riemannian Geometry. *IEEE Transactions on Biomedical Engineering*, **59**, 920-928. <https://doi.org/10.1109/tbme.2011.2172210>
- [11] Rivet, B., Souloumiac, A., Attina, V. and Gibert, G. (2009) xDAWN Algorithm to Enhance Evoked Potentials: Application to Brain-Computer Interface. *IEEE Transactions on Biomedical Engineering*, **56**, 2035-2043. <https://doi.org/10.1109/tbme.2009.2012869>
- [12] Congedo, M., Barachant, A. and Bhatia, R. (2017) Riemannian Geometry for EEG-Based Brain-Computer Interfaces; a Primer and a Review. *Brain-Computer Interfaces*, **4**, 155-174. <https://doi.org/10.1080/2326263x.2017.1297192>
- [13] Xiao, X., Xu, M., Jin, J., Wang, Y., Jung, T. and Ming, D. (2020) Discriminative Canonical Pattern Matching for Single-Trial Classification of ERP Components. *IEEE Transactions on Biomedical Engineering*, **67**, 2266-2275. <https://doi.org/10.1109/tbme.2019.2958641>
- [14] LeCun, Y., Bengio, Y. and Hinton, G. (2015) Deep learning. *Nature*, **521**, 436-444. <https://doi.org/10.1038/nature14539>
- [15] Craik, A., He, Y. and Contreras-Vidal, J.L. (2019) Deep Learning for Electroencephalogram (EEG) Classification Tasks: A Review. *Journal of Neural Engineering*, **16**, Article ID: 031001. <https://doi.org/10.1088/1741-2552/ab0ab5>
- [16] Schirrmester, R.T., Springenberg, J.T., Fiederer, L.D.J., Glasstetter, M., Eggenesperger, K., Tangermann, M., *et al.* (2017) Deep Learning with Convolutional Neural Networks for EEG Decoding and Visualization. *Human Brain Mapping*, **38**, 5391-5420. <https://doi.org/10.1002/hbm.23730>
- [17] Lawhern, V.J., Solon, A.J., Waytowich, N.R., Gordon, S.M., Hung, C.P. and Lance, B.J. (2018) EEGNet: A Compact Convolutional Neural Network for EEG-Based Brain-Computer Interfaces. *Journal of Neural Engineering*, **15**, Article ID: 056013. <https://doi.org/10.1088/1741-2552/aace8c>
- [18] Santamaría-Vázquez, E., Martínez-Cagigal, V., Vaquerizo-Villar, F. and Hornero, R. (2020) EEG-Inception: A Novel Deep Convolutional Neural Network for Assistive ERP-Based Brain-Computer Interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **28**, 2773-2782. <https://doi.org/10.1109/tnsre.2020.3048106>
- [19] Song, Y., Zheng, Q., Liu, B. and Gao, X. (2023) EEG Conformer: Convolutional Transformer for EEG Decoding and Visualization. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **31**, 710-719. <https://doi.org/10.1109/tnsre.2022.3230250>
- [20] Li, X., Song, D., Zhang, P., Zhang, Y., Hou, Y. and Hu, B. (2018) Exploring EEG Features in Cross-Subject Emotion Recognition. *Frontiers in Neuroscience*, **12**, Article 162. <https://doi.org/10.3389/fnins.2018.00162>
- [21] Li, H., Pan, S.J., Wang, S. and Kot, A.C. (2018) Domain Generalization with Adversarial Feature Learning. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 5400-5409. <https://doi.org/10.1109/cvpr.2018.00566>
- [22] Sun, B. and Saenko, K. (2016) Deep CORAL: Correlation Alignment for Deep Domain Adaptation. In: Hua, G. and Jégou, H., Eds., *Computer Vision—ECCV 2016 Workshops*, Springer, 443-450. https://doi.org/10.1007/978-3-319-49409-8_35
- [23] Zhang, H., Cisse, M., Dauphin, Y.N., *et al.* (2017) mixup: Beyond Empirical Risk Minimization. arXiv: 1710.09412.
- [24] Arjovsky, M., Bottou, L., Gulrajani, I., *et al.* (2019) Invariant Risk Minimization. arXiv: 1907.02893.
- [25] Ganin, Y., Ustinova, E., Ajakan, H., *et al.* (2016) Domain-Adversarial Training of Neural Networks. *Journal of Machine Learning Research*, **17**, 1-35.
- [26] Shen, X., Liu, X., Hu, X., Zhang, D. and Song, S. (2023) Contrastive Learning of Subject-Invariant EEG Representations for Cross-Subject Emotion Recognition. *IEEE Transactions on Affective Computing*, **14**, 2496-2511.

- <https://doi.org/10.1109/taffc.2022.3164516>
- [27] Zhou, K., Liu, Z., Qiao, Y., Xiang, T. and Loy, C.C. (2022) Domain Generalization: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **45**, 4396-4415. <https://doi.org/10.1109/tpami.2022.3195549>
- [28] Li, Y., Zheng, W., Zong, Y., Cui, Z., Zhang, T. and Zhou, X. (2021) A Bi-Hemisphere Domain Adversarial Neural Network Model for EEG Emotion Recognition. *IEEE Transactions on Affective Computing*, **12**, 494-504. <https://doi.org/10.1109/taffc.2018.2885474>
- [29] Zhong, P., Wang, D. and Miao, C. (2022) EEG-Based Emotion Recognition Using Regularized Graph Neural Networks. *IEEE Transactions on Affective Computing*, **13**, 1290-1301. <https://doi.org/10.1109/taffc.2020.2994159>
- [30] Chen, P., Gao, Z., Yin, M., Wu, J., Ma, K. and Grebogi, C. (2022) Multiattention Adaptation Network for Motor Imagery Recognition. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, **52**, 5127-5139. <https://doi.org/10.1109/tsmc.2021.3114145>
- [31] Wang, H., Chen, P., Zhang, M., Zhang, J., Sun, X., Li, M., *et al.* (2024) EEG-Based Motor Imagery Recognition Framework via Multisubject Dynamic Transfer and Iterative Self-Training. *IEEE Transactions on Neural Networks and Learning Systems*, **35**, 10698-10712. <https://doi.org/10.1109/tnnls.2023.3243339>
- [32] Li, J., Wang, F., Huang, H., Qi, F. and Pan, J. (2023) A Novel Semi-Supervised Meta Learning Method for Subject-Transfer Brain-Computer Interface. *Neural Networks*, **163**, 195-204. <https://doi.org/10.1016/j.neunet.2023.03.039>
- [33] Hospedales, T.M., Antoniou, A., Micaelli, P. and Storkey, A.J. (2021) Meta-Learning in Neural Networks: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **44**, 5149-5169. <https://doi.org/10.1109/tpami.2021.3079209>
- [34] Finn, C., Abbeel, P. and Levine, S. (2017) Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks. *International Conference on Machine Learning*. PMLR 2017, Sydney, 6-11 August 2017, 1126-1135.
- [35] Nichol, A., Achiam, J. and Schulman, J. (2018) On First-Order Meta-Learning Algorithms. arXiv: 1803.02999.
- [36] Snell, J., Swersky, K. and Zemel, R. (2017) Prototypical Networks for Few-Shot Learning. arXiv: 1703.05175.
- [37] Ning, X., Wang, J., Lin, Y., Cai, X., Chen, H., Gou, H., *et al.* (2024) Metaemotionnet: Spatial-Spectral-Temporal-Based Attention 3-D Dense Network with Meta-Learning for EEG Emotion Recognition. *IEEE Transactions on Instrumentation and Measurement*, **73**, 1-13. <https://doi.org/10.1109/tim.2023.3338676>
- [38] Ng, H.W. and Guan, C. (2024) Subject-Independent Meta-Learning Framework Towards Optimal Training of EEG-Based Classifiers. *Neural Networks*, **172**, Article ID: 106108. <https://doi.org/10.1016/j.neunet.2024.106108>
- [39] Han, J., Bak, S., Kim, J., Choi, W., Shin, D., Son, Y., *et al.* (2024) META-EEG: Meta-Learning-Based Class-Relevant EEG Representation Learning for Zero-Calibration Brain-Computer Interfaces. *Expert Systems with Applications*, **238**, Article ID: 121986. <https://doi.org/10.1016/j.eswa.2023.121986>
- [40] Pati, A., Mewada, D. and Samanta, D. (2023) Meta-Learning for Subject Adaptation in Low-Data Environments for EEG-Based Motor Imagery Brain-Computer Interfaces. *ICLR 2023*, Kigali, 1-5 May 2023, 1-6.
- [41] Li, S., Wu, H., Ding, L. and Wu, D. (2022) Meta-Learning for Fast and Privacy-Preserving Source Knowledge Transfer of EEG-Based BCIs. *IEEE Computational Intelligence Magazine*, **17**, 16-26. <https://doi.org/10.1109/mci.2022.3199622>
- [42] Margaux, P., Emmanuel, M., Sébastien, D., Olivier, B. and Jérémie, M. (2012) Objective and Subjective Evaluation of Online Error Correction during P300-Based Spelling. *Advances in Human-Computer Interaction*, **2012**, Article ID: 578295. <https://doi.org/10.1155/2012/578295>