

**本 科 毕 业 设 计（ 论 文 ）**



**题目: 基于特征不变性联邦学习的图像分类算法研究**

**姓 名 邓萱**

**学 院 人工智能学院**

**专 业 信息工程**

**班 级 2021219106**

**学 号 2021213606**

**指导教师 赵志诚**

**2025 年5月**

**北 京 邮 电 大 学**

**本科毕业设计（论文）诚信声明**

本人声明所呈交的毕业设计（论文），题目《基于特征不变性联邦学习的图像分类算法研究》是本人在指导教师的指导下，独立进行研究工作所取得的成果。尽我所知，除了文中特别加以标注和致谢中所罗列的内容以外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得北京邮电大学或其他教育机构的学位或证书而使用过的材料。

申请学位论文与资料若有不实之处，本人承担一切相关责任。

本人签名： 日期： 2025.5.5

**关于论文使用授权的说明**

本人完全了解并同意北京邮电大学有关保留、使用学位论文的规定，即：北京邮电大学拥有以下关于学位论文的无偿使用权，具体包括：学校有权保留并向国家有关部门或机构送交学位论文，有权允许学位论文被查阅和借阅；学校可以公布学位论文的全部或部分内容，有权允许采用影印、缩印或其它复制手段保存。汇编学位论文，将学位论文的全部或部分内容编入有关数据库进行检索。（保密的学位论文在解密后遵守此规定）

本人签名： 日期： 2025.5.5

导师签名： 日期： 2025.5.5

**基于特征不变性联邦学习的图像分类算法研究**

**摘 要**

联邦学习作为一种新兴的分布式学习框架，能够在保护数据隐私的前提下实现模型的协同训练。然而，传统联邦学习方法在异构数据分布场景中往往存在泛化能力不足的问题。领域泛化作为深度学习的重要研究方向，旨在通过多个源领域的联合训练，学习具有跨域共享性的特征，从而提升模型在未见领域中的鲁棒性与稳定性。本文聚焦于嵌入领域泛化机制的联邦学习框架设计与优化。

首先，为缓解经典联邦平均算法（FedAvg）训练过程中的“更新漂移”问题，引入L2正则项约束，在本地优化目标中加入全局模型参数与本地模型参数之间的平方差惩罚项，形成联邦近似算法（FedProx），在PACS、OfficeHome和Terra Incognita三个常用数据集上的平均准确率从63%提升至63.38%。

其次，设计了一种松耦合、可插拔的域调整（Generalization Adjustment，GA）方法，通过引入泛化差距方差正则项，动态调整客户端聚合权重，进一步增强模型鲁棒性。将GA方法融合至FedAvg和FedProx框架中后，在三个数据集上的平均准确率分别提升了1.33%和2.05%，验证了GA方法在非独立同分布数据环境下对模型领域泛化能力的有效性。

最后，针对联邦学习中参数传输导致的隐私泄露风险，设计并实现了一种以本地模型梯度代替传统模型参数进行传输的方法，在提升系统的隐私保护水平与灵活性的同时，还改善了部分场景下的分类性能。

**关键词** 神经网络 深度学习 联邦学习 领域泛化

**Research on Image Classification Algorithms Based on Feature-Invariant Federated Learning**

**ABSTRACT**

Federated learning, as an emerging distributed learning framework, enables collaborative training of models while preserving data privacy. However, traditional federated learning methods often suffer from poor generalization in heterogeneous data distribution scenarios. Domain generalization, as a significant research direction in deep learning, aims to enhance model robustness and stability in unseen domains by jointly training across multiple source domains to learn domain-shared features. This thesis focuses on designing and optimizing a federated learning framework embedded with domain generalization mechanisms.

Firstly, to mitigate the "update drift" issue in classical Federated Averaging (FedAvg) during training, we introduce an L2 regularization constraint and add a squared difference penalty between global model parameters and local model parameters in the local optimization objective, forming the Federated Proximal algorithm (FedProx). This approach improves the average accuracy from 63% to 63.38% on three commonly used datasets: PACS, OfficeHome, and Terra Incognita.

Secondly, we propose a loosely coupled, pluggable Generalization Adjustment (GA) method. By introducing a regularization term based on variance of generalization gaps and dynamically adjusting client aggregation weights, GA enhances model robustness further. Integrating GA into both FedAvg and FedProx frameworks increases average accuracy by 1.33% and 2.05% respectively across the three datasets, demonstrating its effectiveness in improving domain generalization capability in non-i.i.d. data environments.

Finally, addressing privacy risks from parameter transmission in federated learning, we design and implement a method that replaces traditional model parameters with local model gradients for transmission. This method not only enhances privacy protection and system flexibility but also improves classification performance in certain scenarios.

**KEY WORDS** Neural Network Deep Learning Federated Learning Domain Generalization

**目录**

[第一章 绪论 1](#_Toc159)

[1.1 研究背景与意义 1](#_Toc22373)

[1.2 国内外研究现状 2](#_Toc16016)

[1.3 研究内容与目标 3](#_Toc27339)

[1.4 章节安排 4](#_Toc5200)

[第二章 联邦域泛化相关技术 5](#_Toc21200)

[2.1 引言 5](#_Toc21780)

[2.2 联邦域泛化数据集 5](#_Toc2094)

[2.3 分类网络 7](#_Toc10139)

[2.4 联邦领域泛化算法 8](#_Toc11219)

[2.4.1 联邦学习 8](#_Toc4514)

[2.4.2 联邦领域泛化 9](#_Toc31487)

[2.5 FedAvg算法 10](#_Toc22518)

[2.5.1 理论推导 10](#_Toc31129)

[2.5.2 算法流程 11](#_Toc8199)

[2.5.3 优缺点 11](#_Toc11242)

[2.6 FedProx算法 12](#_Toc26412)

[2.6.1 理论推导 12](#_Toc17238)

[2.6.2 算法流程 13](#_Toc17659)

[2.6.3 优缺点 13](#_Toc32212)

[2.7 本章小结 14](#_Toc22877)

[第三章 自适应聚合的域泛化模块 15](#_Toc13876)

[3.1 引言 15](#_Toc3144)

[3.2 优化目标 15](#_Toc21359)

[3.3 理论推导 16](#_Toc13382)

[3.4 算法流程 17](#_Toc30609)

[3.5 本章小结 18](#_Toc8751)

[第四章 梯度传输的参数保护模块 19](#_Toc4875)

[4.1 引言 19](#_Toc17346)

[4.2 梯度传输理论分析 19](#_Toc26471)

[4.3 算法流程 20](#_Toc23559)

[4.4 算法优势 21](#_Toc7343)

[4.5 本章小结 21](#_Toc1937)

[第五章 实验结果与分析 22](#_Toc26379)

[5.1 实验设计 22](#_Toc13351)

[5.1.1 环境介绍 22](#_Toc26216)

[5.1.2 模型选择方法 22](#_Toc11865)

[5.2 实验描述与结果对比分析 22](#_Toc30747)

[5.2.1 基线实验 22](#_Toc15622)

[5.2.2 FedProx实验 23](#_Toc13025)

[5.2.3 FedGA实验 26](#_Toc16874)

[5.2.4 梯度传输实验 32](#_Toc18353)

[5.3 本章小结 34](#_Toc20385)

[第六章 总结与展望 36](#_Toc5684)

[6.1 研究工作总结 36](#_Toc26822)

[6.2 研究展望 36](#_Toc31197)

[参考文献 37](#_Toc31005)

[致谢 39](#_Toc15616)

[附录 40](#_Toc24801)

# 第一章 绪论

## 1.1 研究背景与意义

随着信息技术的飞速发展和深度学习的广泛应用，人工智能在图像识别、语音处理、自然语言理解等领域的表现愈发卓越。从卷积神经网络（CNN）[1]在图像分类中的突破性应用，到生成对抗网络（GAN）[2]和Transformer[3]的出现，深度学习持续推动着各领域技术的革新与进步。特别是在智慧城市、远程医疗与自动驾驶等复杂应用场景中，深度学习已超越传统方法，成为解决实际问题的核心动力。然而，随之而来的数据隐私保护需求与模型泛化能力不足等挑战，也引起了学术界与工业界的广泛关注。

传统深度学习模型通常依赖大量标注数据进行训练，并假设数据样本独立同分布（i.i.d.）。但在现实应用中，数据异质性普遍存在，这种分布差异往往导致模型训练效果和实际泛化性能大幅下降。此外，随着数据隐私保护法规日益严格，集中式收集多个机构或设备数据以进行统一训练已变得愈发困难。因此，如何在保护数据隐私的同时，提升模型在异构数据分布环境下的泛化能力，成为当前深度学习领域亟待解决的重要问题。

领域泛化（Domain Generalization, DG）[4]作为深度学习中一种重要的研究方向，旨在缓解模型在未知领域中的性能退化问题。域偏移是指用于训练的源域和用于测试的目标域之间的分布差异[5]，导致传统模型在目标域上表现不佳。领域泛化方法通过在多个源域上进行联合训练，学习具备跨域共享性的特征，使模型即便在未见过的目标域中也能保持良好的鲁棒性与稳定性。例如，在PACS数据集[12]中，训练集包括草图、卡通图像和艺术画作，领域泛化算法能够训练出在自然照片这一未见领域中仍具备较好分类能力的模型。类似地，在人体活动识别任务中，由于传感器数据来源于不同个体，数据分布存在显著差异，导致模型在新用户上的性能下降，领域泛化同样可有效缓解这一问题。

联邦学习（Federated Learning, FL）[6]作为一种新兴的分布式学习框架，在保护数据隐私的前提下实现了跨设备、跨机构的模型协同训练。其核心思想是允许各参与方仅共享本地模型参数，而非原始数据，从而有效降低隐私泄露风险。在智能医疗领域，联邦学习已被广泛应用于构建跨医院协作的医疗模型，同时保护患者隐私。例如，MediSecFed[7]系统结合了区块链与同态加密技术，不仅实现了对敏感数据的保护，还增强了对数据中毒攻击的鲁棒性。研究人员在实际肺炎数据集上的实验评估，验证了该系统在恶劣环境下的可行性与有效性。在智能交通领域，联邦学习被应用于联网车辆的数据协作，各车辆在本地采集驾驶数据（如图像和坐标）并独立训练模型，通过车联网设备（如路边单元）共享模型更新，最终在云端聚合，支持大规模交通流预测与优化。然而，由于各参与方数据分布异质性显著，联邦学习模型易出现对部分客户端数据过拟合、对其他客户端数据泛化性不足的问题。因此，提升联邦学习在多样数据环境下的鲁棒性与泛化能力，成为当前研究的热点。

将领域泛化技术引入联邦学习框架，对解决上述挑战提供了新的方向。通过学习跨域共享特征，领域泛化有望缓解数据分布不一致对模型性能的影响，提升联邦模型在多设备、多环境下的普适性与稳定性。然而，由于联邦学习的特性要求各参与方数据彼此隔离，传统领域泛化方法难以直接迁移应用。因此，如何在数据隔离的联邦环境下，设计兼顾领域泛化效果与联邦学习约束的新型训练策略，成为亟需解决的关键问题。

## 1.2 国内外研究现状

领域泛化的目标是从一个或多个不同但相关的源领域（训练集）中学习一个模型，使其在未见的测试领域上也能实现良好的泛化性能。在DG的定义中，“不同但相关”是关键，即尽管各个域的数据分布存在差异，但它们共享相同的类别标签空间。现有领域泛化方法主要可以从数据操作、表征学习和学习策略三个方面进行分类[4]。数据操作主要通过对训练数据进行增强和变化，以扩展数据分布的多样性，包括数据增强和数据生成两大类。表征学习则致力于学习领域不变特征，使模型能够适配不同领域，主要涵盖核方法、显式特征对齐、领域对抗训练以及不变风险最小化等方向。学习策略方面，则通过引入机器学习中成熟的学习模式，提升模型在多领域训练中的泛化能力，代表性方法包括基于集成学习[4]和元学习的技术[4]。

联邦学习旨在保护用户隐私的前提下，通过在各设备上本地处理数据来实现模型训练。在本地训练完成后，每个设备仅共享模型参数的更新，而非原始数据。Google于2016年提出了经典的联邦平均算法（Federated Averaging, FedAvg）。FedAvg通过本地模型的梯度下降优化，每个客户端（如手机或其他边缘设备）在本地使用自身数据训练模型，并将更新后的模型参数发送至中央服务器。服务器收集所有客户端的更新并进行平均，再将更新后的全局模型下发至各客户端，反复迭代直到模型收敛。随后，控制变量辅助联邦优化算法（Stochastic Controlled Averaging for Federated Learning，SCAFFOLD）[8]被提出，主要关注梯度稀疏性和设备间的通信。SCAFFOLD通过在每一轮训练中维护一个控制变量，调整本地模型梯度，从而减少通信频次，提高通信效率，同时保证模型性能。

在联邦领域泛化研究中，联邦域对齐[9]成为关键，它通过提升模型在多域环境下的泛化性能，提高了数据利用效率并进一步强化了隐私保护。Chen等提出了D-WFA[10]，基于最大均值差异（MMD）设计动态加权策略，针对不同客户端调整权重，并训练出多个具有更好泛化能力的局部模型。Peng[9].等人提出了联合对抗对齐方法，通过将优化过程划分为局部特征提取器训练与全局判别器训练两个阶段，有效对齐了不同域之间的特征分布。Li等人提出了模型对比联邦学习（Model-Contrastive Federated Learning，MOON[11]）框架，通过模型级别的对比学习方法，缩小全局模型与本地模型之间的分布差距，从而提升联邦学习的整体性能。尽管上述方法在联邦领域泛化方向取得了初步成效，但在特征不变性建模、通信开销与模型性能的平衡、以及泛化能力的评估机制设计等方面，仍存在诸多挑战，有待进一步研究与完善。

## 1.3 研究内容与目标

本文拟将领域泛化算法嵌入联邦学习，并完成如下主要研究内容：

（1）实现参数距离对齐的域泛化联邦学习框架FedProx

首先，构建基于FedAvg的联邦学习基线框架，在保证数据隐私不被泄露的前提下，实现跨客户端的协同模型训练。在此基础上，引入正则化机制，即在每个客户端的本地优化目标中加入全局模型参数与本地模型参数之间的平方差作为惩罚项，用以约束本地与全局模型之间的参数偏移，从而实现了一种FedProx框架，缓解本地训练过程中模型参数偏离全局模型的“更新漂移”问题。

（2）设计自适应聚合的域泛化模块

从模型聚合角度出发，设计一种基于泛化性能差异动态调整聚合权重的策略，进一步强化全局模型的泛化能力。具体地，引入泛化差距（Generalization Gap）方差作为正则项，动态调整客户端模型的聚合权重，形成FedGA（Federated learning with Generalization Adjustment）模块。该模块能够引导各客户端更合理地参与模型聚合，有效缩小全局模型在不同客户端之间的泛化误差差异，提升模型在未知测试域上的适应能力与鲁棒性。

1. 设计梯度传输的参数保护模块

在FedGA基础上，从通信机制优化的角度切入，设计并实现一种以本地模型梯度替代传统模型参数进行上传的方法，进一步提升隐私保护强度与应用灵活性，有助于降低模型信息泄露的风险，适用于异构设备环境下对隐私保护和通信效率要求更高的联邦学习场景。

## 1.4 章节安排

本文分为五章，章节安排如下：

第一章，绪论。主要讲述了本毕业设计研究课题的研究背景与意义，综述国内外在相关领域的研究现状，并明确本文的研究内容与目标。

第二章，参数距离对齐的域泛化联邦学习框架实现。首先介绍联邦领域泛化任务中常用的数据集，接着详细说明所采用的分类网络结构，然后重点阐述FedAvg框架的原理与实现。在此基础上，引入L2正则项以实现FedProx框架，通过惩罚本地模型与全局模型参数之间的偏差，缓解模型更新漂移问题。

第三章，自适应聚合的域泛化模块。从参数聚合方面进行算法优化，提出基于泛化差距方差的GA方法，用于动态调整聚合权重，提高全局模型在异构数据域上的泛化能力。

第四章，梯度传输的参数保护模块。从通信机制优化角度出发，提出以本地模型梯度代替传统模型参数进行上传的方法，旨在提升隐私保护水平，并增强全局优化过程的灵活性与鲁棒性。

第五章，实验结果与分析。首先基于FedAvg构建基准训练框架，并在此基础上实现FedProx。随后分别从聚合机制与传输内容两个角度引入优化策略，与基准框架融合，最终在多个数据集上进行实验，对比融合前后的模型性能表现，分析实验结果并得出结论。

第六章，总结与展望。对本文的研究工作进行总结，归纳主要贡献与成果，并对未来可能的研究方向和改进空间进行展望。

# 第二章 联邦域泛化相关技术

## 2.1 引言

本章首先介绍联邦领域泛化任务中常用的标准数据集，随后对所采用的分类网络结构进行详细说明。在此基础上，系统阐述了FedAvg框架的基本原理及其具体实现方式。进一步引入L2正则项以构建FedProx框架，并对其设计理论、优势与局限性进行了深入分析。

## 2.2 联邦域泛化数据集

在联邦领域泛化领域任务的研究中，为充分验证算法在多域适应性、数据异质性及隐私保护等方面的性能，所选用的数据集需具备以下关键特性：（1）包含多个不同的域，每个域内涵盖相同的类别；（2）各域间数据分布存在明显差异；（3）每个样本带有完整、准确的监督标签。

目前广泛使用的领域泛化基准数据集有PACS[12]、OfficeHome[13]、Terra Incognita[14]、Digit-five[15][16]、Office-31[17]、 DomainNet[18]、Rotated MNIST[19]和Colored MNIST[20]等，这些数据集均可划分为多个域，且各域的图像共享相同类别的标签信息。表2-1对上述常用数据集的基本情况进行了总结。

表2-1 用于联邦领域泛化的8个图像数据集

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 数据集 | 领域数量 | 分类数量 | 样本总数 | 领域描述 |
| PACS | 4 | 7 | 9,991 | Art, Cartoon, Photos, Sketches |
| Office-Home | 4 | 65 | 15,588 | Art, Clipart, Product, Real |
| Terra Incognita | 4 | 10 | 24,788 | Wild animal images taken at locations L100, L38, L43, L46 |
| Digit-five | 5 | 10 | 730,000+ | MNIST、USPS、SVHN、MNIST-M、Synthetic Digits |
| Office-31 | 3 | 31 | 4,110 | Amazon, Webcam, DSLR |
| DomainNet | 6 | 345 | 586,575 | Clipart, Infograph, Painting, Quickdraw, Real, Sketch |
| Rotated MNIST | 6 | 10 | 70,000 | Digits rotated from 0  ◦ to 90◦ with an interval of 15◦ |
| Colored MNIST | 2 | 10 | 140,000 | Train, Target |

为了对联邦领域泛化算法进行公平和有效的对比实验，所选数据集应具备以下特征：具备多个领域、涵盖多个类别、样本分布相对均衡、图像接近真实场景，并且整体数据规模适中。若领域数量过少，模型将难以学习出对领域变化不敏感的特征，从而削弱其泛化能力；类别数量不足则会限制模型的应用场景，影响其实用性；样本分布若存在严重不平衡，可能导致模型在某些类别上的预测效果明显优于其他类别，影响整体性能；图像若偏离自然图像特性，模型在面对真实环境中的输入时容易出现性能下降；而数据规模若过大，则可能对计算资源造成较大压力，导致内存溢出或训练过程难以控制在合理时间范围内。

综合考虑上述要求，这里选取PACS、OfficeHome和Terra Incognita（后文简称为TerraInc）三个典型的多领域图像分类数据集作为后续实验的数据集。

其中，PACS数据集由Li等于2020年提出，专为领域泛化任务设计，包含Photo、Art painting、Cartoon和Sketch四个风格差异显著的图像域。每个领域下均包含7个通用类别，共计9,991张227×227的RGB图像。该数据集结构简单但挑战性强，适合对算法的基本泛化能力进行快速验证。OfficeHome数据集由Venkateswara等于2017年提出，广泛应用于多领域图像识别任务。该数据集涵盖Art、Clipart、Product 和 Real-World四个域，每个域包含65个日常办公与家庭场景下的物体类别，总图像数量约为 15,500 张。相比PACS，OfficeHome类别更多、任务粒度更细，适合评估算法在复杂多类环境中的泛化表现。TerraInc数据集最初由Beery等在2018年发布，来源于野外相机陷阱对野生动物的真实拍摄图像，本文选取L38、L43、L46和L100四个地理区域作为不同领域。每个域下包含10个动物类别，总图像数有24,330张，图像背景复杂多变，分布非结构化，极大地模拟了真实应用中跨地域领域偏移的挑战，适合检验模型在自然场景下的稳健性与迁移能力。综上，这三个数据集在领域数量、类别数量、图像风格与任务复杂度上各具特色，能够从不同维度全面评估联邦领域泛化算法的性能与鲁棒性。

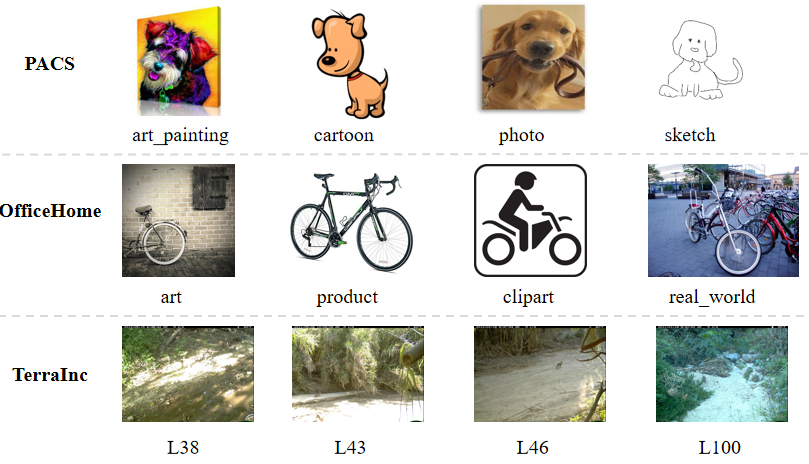


图2-1 数据集不同域样本展示图

三个数据集中不同域样本展示如上图2-1。

## 2.3 分类网络

深度学习已成为智能计算领域的重要支撑技术，在图像处理、语音分析与自然语言理解等多种任务中展现出卓越性能。其中，ResNet[23]通过引入残差连接机制，成功缓解了深层网络训练过程中梯度消失的问题，并在保持较高网络深度的同时，能够兼顾稳定性与训练效率的优良特性。本文选取ResNet-18与ResNet-50作为基础特征提取网络。

ResNet-18由五个卷积阶段与一个全连接层组成，采用如图2-2左侧所示标准残差块（BasicBlock）作为基本单元，每个残差块中包含两个连续的3×3卷积层及跨层直连路径。标准残差块通过简单高效的设计实现了浅层特征与深层特征的融合，保证了信息传递的完整性与梯度流动的顺畅性，适用于样本数量相对较少、类别数量适中、特征复杂度适中的任务场景。

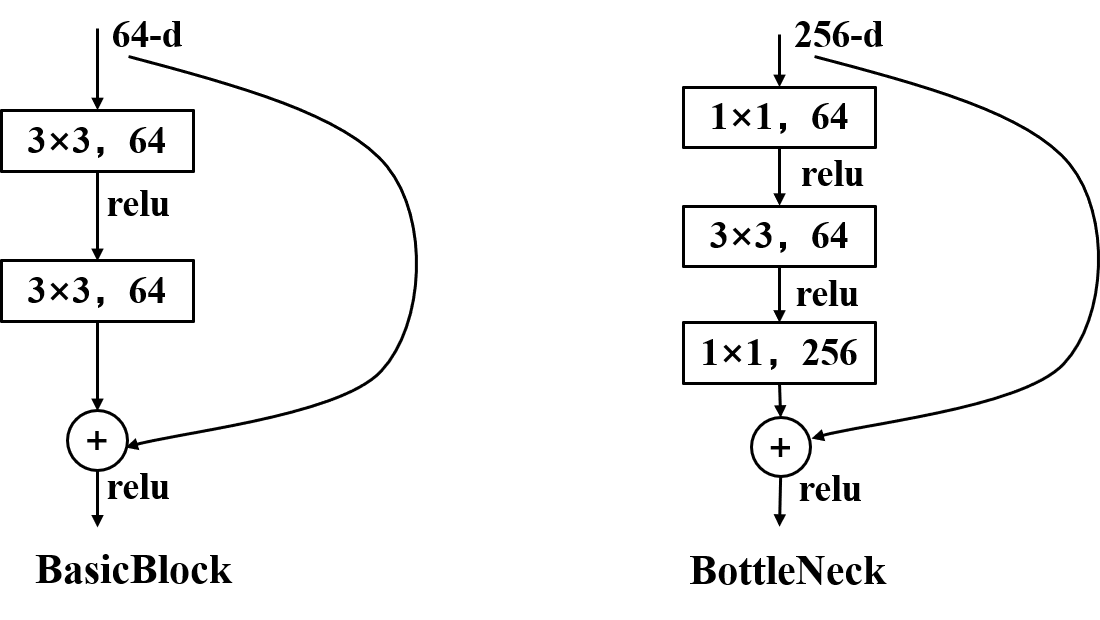


图2-2 ResNet残差结构图[23]

相比之下，ResNet-50采用如图2-2右侧所示的更为复杂的瓶颈残差结构（Bottleneck Architecture）。每个瓶颈块包含三个卷积层：首先通过1×1卷积降维以减少计算量，其后以3×3卷积进行核心特征提取，最后再以1×1卷积恢复通道维度。瓶颈结构在保证特征表达能力的同时，有效降低了计算复杂度，使得网络能够在深度进一步扩展的情况下依然保持良好的可训练性与推理效率。ResNet-50特别适用于需要处理大规模数据并具备复杂特征分布的任务，能够更充分地挖掘图像中的细粒度信息，提升分类精度。

由于PACS与OfficeHome数据集规模适中且类别清晰，本文采用ResNet-18可以保证训练效率与性能的良好平衡；而针对图像数量庞大、分布复杂的TerraInc数据集，则选用ResNet-50以更好地适应复杂任务需求，确保联邦领域泛化实验的可靠性与科学性。

## 2.4 联邦领域泛化算法

### 2.4.1 联邦学习

联邦学习是一种分布式机器学习框架，其核心目标是在保护数据隐私的前提下，实现跨多个设备或客户端的协同模型训练。与传统的集中式训练方式不同，联邦学习不需要将本地数据上传到服务器，而是仅共享模型参数或梯度，从而有效缓解了隐私泄露和数据孤岛问题。联邦学习的框架示意图如图2-3所示。

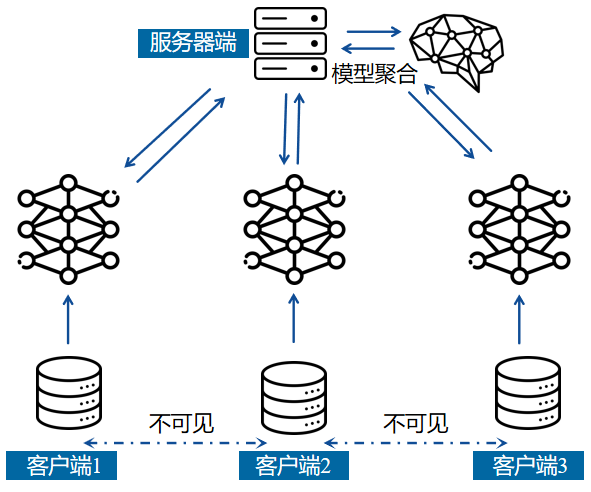


图2-3 联邦学习框架示意图

在典型的联邦学习流程中，服务器首先初始化全局模型参数，并将其下发给选中的部分客户端。每个客户端随后在本地数据上对模型进行若干轮训练，完成后将本地更新的模型参数发送回服务器。服务器对这些客户端上传的模型参数进行加权聚合，更新全局模型，并进入下一轮训练。通过不断迭代这一过程，联邦学习模型逐步收敛。

然而，在实际应用中，联邦学习面临诸多挑战。首先，不同客户端的数据通常来自于不同的用户或场景，其分布往往存在显著差异，呈现非独立同分布（Non-IID）的特征，这会导致模型性能下降或收敛缓慢。其次，客户端的计算能力和网络带宽存在差异，系统在资源调度和同步时需要考虑设备的异构性。此外，由于模型训练过程在分布式环境下进行，通信成本较高，特别是在模型复杂或客户端数量较多时，异步通信时延成为了阻碍模型聚合的关键因素。最后，由于各客户端无法访问全局数据，模型在未参与训练的客户端或新环境中往往缺乏良好的泛化能力。因此，如何在保障隐私的同时提升模型泛化性能，是联邦学习研究中的关键问题之一。

### 2.4.2 联邦领域泛化

领域泛化旨在解决训练数据和测试数据分布不一致的问题，其核心目标是在多个源域数据上训练出一个能够适应未知目标域的模型。即便在目标域完全不可见的情况下，模型仍能保持良好的泛化性能。领域泛化的主流思路是学习领域不变的特征表示，即使源域之间的分布不同，模型也能够提取出在多个领域中共享的本质信息。为实现这一目标，已有的研究方法从多个角度展开：从数据层面进行风格迁移或数据增强[24]，以提升模型的鲁棒性；从特征层面引入显式或对抗性特征对齐方法（如最大均值差异、领域判别器），以最小化不同域之间的分布差距；从优化层面采用元学习、结构正则化、不变风险最小化等策略，以增强模型的泛化能力。

然而，现有领域泛化方法大多基于集中式学习范式，需要将来自不同源域的训练数据集中至同一服务器。这种方式在实验环境中可行，但在实际应用中却面临严重的隐私风险和法律限制。例如，医疗、金融等领域的数据具有极高的敏感性，不同机构之间无法直接共享原始数据；即便是跨公司合作进行模型训练，也需要遵循严格的数据保护条例，如GDPR等隐私法规。这种数据孤岛问题在实际应用中普遍存在，成为阻碍领域泛化技术落地的主要瓶颈。

为了解决上述问题，研究者尝试将领域泛化与联邦学习相结合，提出联邦领域泛化[25]（Federated Domain Generalization, FDG）的新型范式。该方法保留了联邦学习的隐私保护优势，即模型训练过程中仅在各个客户端本地进行，服务器仅聚合模型参数或梯度，无需传输原始数据；同时结合领域泛化策略，以提升模型在未见域上的适应能力。

如图2-4所示，领域泛化与联邦领域泛化之间存在一些相似之处：源域之间具有不同的分布，目标域完全不可见。然而，联邦领域泛化因起隐私保护的特性，各源域之间相互隔离。通过在不同源域设备上分别训练本地模型，并设计有效的参数聚合与领域对齐机制，联邦领域泛化成功解决了领域泛化中的隐私泄露问题，在不暴露数据的前提下实现了模型的泛化学习。通过这种方式，FDG不仅能有效应对跨域的泛化问题，还能在保障隐私的同时，提升模型的适应性和鲁棒性，尤其适用于数据隐私要求较高的实际应用场景。

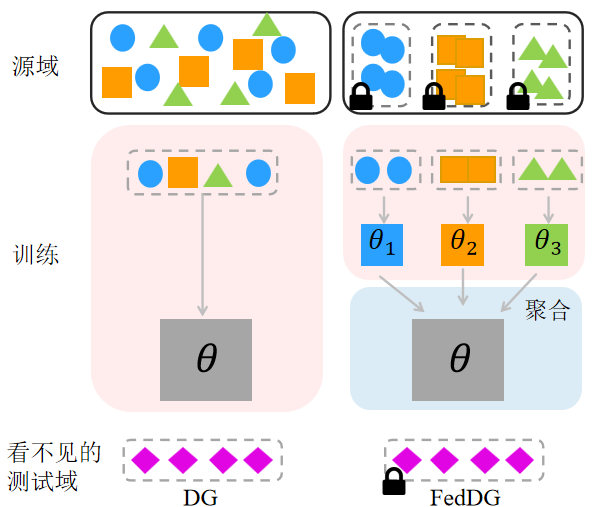


图2-4 领域泛化和联邦领域泛化示意图对比

## 2.5 FedAvg算法

### 2.5.1 理论推导

在联邦学习框架下，系统通常由M个客户端组成，每个客户端m拥有自己的本地数据集。设客户端m的数据集大小为，而全局系统的总样本数为N=，模型的参数则用来表示。每个客户端在本地数据上进行训练，并通过梯度下降优化方法对其模型参数进行更新，而这些更新最终将通过全局聚合形成一个统一的全局模型。

FedAvg的目标是通过最小化全局损失函数来达到全局模型的优化，具体的全局目标函数可以表示为：

式（2-1）

其中：是该客户端上的经验损失。

式（2-2）

该目标函数的关键在于通过局部优化和全局聚合的方式，使得每个客户端能够在保证数据隐私的前提下，依据其本地数据对全局模型进行贡献。每个客户端独立地在本地进行训练，更新其模型参数后将更新后的模型参数发送到服务器，服务器再将所有客户端的更新进行聚合，形成更新后的全局模型。

假设每个客户端的模型参数在第r轮更新时为，则全局模型参数在第r+1轮更新的公式为：

式（2-3）

FedAvg的目标函数和更新规则提供了一种通过局部训练和全局聚合相结合的方法，使得联邦学习能够在不泄露客户端数据的情况下进行高效的模型训练。该过程既保证了数据的隐私性，又能够通过多轮的迭代更新来逐步优化全局模型的性能。

### 2.5.2 算法流程

FedAvg算法流程如下：

算法2-1 FedAvg算法流程

|  |
| --- |
| 输入：M，N，，E，，，， |
| 输出：全局模型 |
| 服务器随机选择一个包含K个设备的子集（每个设备k以概率被选中）  服务器将发送给所有选中的设备  每个设备将返回给服务器  服务器将所有聚合为 |

### 2.5.3 优缺点

FedAvg算法因其结构简单、易于实现且适配性强，成为联邦学习领域的重要方法。该算法能够在保护隐私的前提下实现跨设备的协同建模，且具有较低的通信开销，使其在多种神经网络架构中得到了广泛应用。然而，FedAvg在实际应用过程中面临两大主要挑战：数据异构性和系统异构性。

数据异构性指的是不同客户端上的数据分布存在显著差异，客户端的数据往往呈现非独立同分布（Non-IID）的特性。由于FedAvg使用统一的加权平均策略对各客户端模型进行聚合，这种静态聚合方法可能使得一些低质量甚至误导性的本地模型对全局模型产生负面影响，从而影响整体性能。尤其是在数据非IID的情况下，各客户端的本地训练过程容易偏向于局部最优，而缺乏对本地更新过程的有效约束，导致全局模型无法有效兼顾所有客户端的数据需求。

此外，FedAvg未能充分考虑系统异构性带来的影响。在实际部署中，客户端可能由于计算资源、存储能力、网络带宽等方面的差异而表现出不同的训练能力。FedAvg规定每个客户端在固定时间内完成本地训练轮数，但在某些设备上，资源限制可能导致训练未完成，这时未完成训练的客户端要么被丢弃，要么与其他设备的更新进行聚合。这样一来，丢弃的设备可能导致模型产生偏差，而减少的设备数量也可能降低模型的多样性，从而影响模型的收敛性和准确性。

## 2.6 FedProx算法

### 2.6.1 理论推导

FedProx[26]（Federated Proximal）在FedAvg的基础上进行了改进。FedProx在每个客户端的本地优化目标中引入了一个额外的proximal正则项。该正则项的核心思想是：在进行本地更新时，通过一个额外的惩罚项，使得本地模型的更新与全局模型保持一致，从而减小本地更新对全局模型的偏离程度。

具体而言，FedAvg 的本地优化目标为最小化全局损失函数：

式（2-4）

而FedProx在此基础上引入正则项，优化目标为：

式（2-5）

其中，为当前轮次的全局模型参数，为正则项权重。

将原来的(·)变为(·)使得每个客户端在进行本地更新时，除了最小化本地损失函数外，还需要考虑本地模型与全局模型之间的差异。这样，客户端的本地更新就不会完全依赖于本地数据的特点，而是会参考全局模型的结构，避免在本地数据的局部最优解上过度优化。

各客户端基于本地数据计算本地更新模型，需要满足以下条件：

式（2-6）

同时定义了-inexact solution,通过对的非精确求解，动态调整本地迭代次数，保证对异构系统的容忍度。如果满足下式则称为的-inexact solution

式（2-7）

式（2-8）

则此时，:→,然后聚合得到新一轮的

式（2-9）

衡量第k个客户端在全局第t 次迭代的本地计算量，因此可以根据该变量代理本地轮次。而不需要手动设置本地epochs的数量，一定程度上也解决了统计异质性的问题。

### 2.6.2 算法流程

FedProx框架流程如下所示：

算法2-2 FedProx算法流程

|  |
| --- |
| 输入：K，T，，，N，， |
| 输出：全局模型 |
| 服务器随机选择一个包含K个设备的子集（每个设备k以概率被选中）  服务器将发送给所有选中的设备  每个设备找到一个-不精确最小化者，使得：  arg  每个设备将返回给服务器  服务器将所有聚合为 |

### 2.6.3 优缺点

FedProx优于FedAvg的地方在于它引入了正则化项，能够更好地容忍设备间的异构性和数据异质性，保证了全局模型的稳定性和收敛性。

FedProx相较于FedAvg的缺点在于计算开销略高，超参数调节更加复杂，尤其是正则化项的权重需要根据数据异质性和设备特性进行调整。过强的正则化可能会抑制本地数据的有效利用，导致模型性能下降。特别在数据分布差异较小的情况下，过度约束本地更新可能影响模型的灵活性和准确性。

## 2.7 本章小结

本章围绕联邦领域泛化任务展开，系统介绍了所采用的实验数据集、分类网络结构以及基础联邦学习框架的实现方法。通过对FedAvg框架的原理解析及其局限性的分析，引出了FedProx框架，并详细阐述了其改进思路与优缺点。上述内容为后续算法的设计与性能优化奠定了理论基础和实现依据。

# 第三章 自适应聚合的域泛化模块

## 3.1 引言

本章围绕参数聚合策略的优化展开，提出一种基于泛化差距方差的自适应聚合方法——Generalization Adjustment（GA），通过动态调整各客户端的聚合权重，引导全局模型更加关注具有较强泛化能力的本地模型，从而提升联邦学习在异构数据分布下的泛化性能。随后依次介绍该方法的优化目标、理论推导过程以及完整的算法流程。

## 3.2 优化目标

传统的联邦学习使用固定权重聚合，该方法面对领域泛化场景存在明显局限，主要体现在三个方面。其一，面对数据异构场景，其无法根据客户端数据规模、质量及分布差异，自适应调节更新权重，易引发模型对优势数据的过拟合，削弱泛化性能；其二，设备性能的多样性使训练效率受制于低性能节点，固定权重机制难以动态平衡计算与通信负载，导致全局模型收敛效率与质量下降；其三，在非独立同分布数据环境下，该方法无法有效捕捉数据特征变化，延缓收敛进程甚至陷入次优解。

针对以上问题，本章设计了一种自适应聚合方法（Generalization Adjustment，GA）旨在联邦学习训练过程中引入泛化误差方差正则项，通过动态调整聚合权重，优化全局模型在不同域上的泛化能力。

传统FedAvg框架的优化目标为最小化全局经验损失函数，即：

式（3-1）

其中，表示全局模型参数，表示第i个客户端的本地数据集，是按样本数加权的比例。

为了提高模型的跨域泛化能力，GA方法引入了泛化误差方差（generalization gap variance）作为正则项，新的优化目标函数被定义为：

式（3-2）

其中，为第i个客户端的自适应聚合权重，为控制主损失与正则项平衡的超参数。泛化差距定义为：

式（3-3）

其中，表示第i个客户端在本地数据集上的最优模型参数。

为了动态更新每一轮的聚合权重，GA在每轮训练结束后根据各客户端的泛化差距计算新的聚合系数，更新规则为：

式（3-4）

式（3-5）

其中，μ是第r轮各客户端泛化差距的均值，是随轮数递减的调整步长。最终，服务器使用更新后的权重对各客户端模型进行聚合。

式（3-6）

## 3.3 理论推导

在联邦领域泛化问题中，模型的目标是具备跨域鲁棒性，即在多个源域联合训练后，能够在未知目标域上保持良好性能。由于目标域不可访问，无法直接对其进行监督优化，因此如何基于源域信息提升对目标域的泛化能力，成为理论分析的关键。

本节以泛化误差为切入点，推导目标域上模型性能的理论上界，并阐述引入泛化误差方差正则项的合理性。

设T表示目标域，表示第i个客户端（即源域）的经验分布，为全局模型参数，为目标域上的最优模型参数，则目标域上的泛化误差上界可被刻画为：

式（3-7）

其中，=表示全局模型在客户端i上相较于本地最优模型的性能损失，即该域上的泛化差距；是基于假设空间的分布散度指标，用于度量源域与目标域之间的分布差异；表示模型容量引起的误差项；是由于有限样本带来的误差项，通常与样本数量相关，满足Hoeffding不等式：

式（3-8）

由于目标域分布不可见，无法直接计算。本文引入如下核心假设，即域间分布差异与源域之间的泛化差距方差正相关：

式（3-9）

该假设基于如下直觉：若源域之间的泛化差距较大，表明各域对全局模型的适应能力不一致，从而推测这些域间与目标域的距离也存在波动。因此，减少源域之间的泛化误差差异，有助于间接降低源域与目标域的平均散度。从而可以间接减小目标域上的泛化误差上界，提升模型在未知域的泛化能力。

在此理论基础上，GA方法通过引入源域泛化差距方差作为正则项，使得最终优化目标不仅最小化经验风险，同时约束不同域上的泛化差距趋于一致：

式（3-10）

该目标函数构成了经验风险最小化（ERM）与结构正则化之间的联合优化机制。一方面，全局模型在源域上的性能维持较低误差；另一方面，方差正则项确保模型在所有源域上具有一致性和稳健性，从而提升对未知目标域的泛化能力。

综上，GA方法通过动态聚合机制引入泛化误差方差，理论上可降低目标域上的风险上界，实质上构建了一个结构感知的聚合策略，为联邦领域泛化问题提供了可解释且有效的优化路径。

## 3.4 算法流程

将GA算法与FedAvg融合后的FedGA框架流程如下：

算法3-1 FedGA算法流程

|  |
| --- |
| 输入：  全局模型；个客户端：；  初始权重：；  超参数：本地迭代次数 E，总通信轮次 R，GA的步长d |
| 输出：全局模型： |
| 1.服务器初始化：使用全局模型初始化本地模型 ：  2**.for all in do:**  客户端训练：  计算每个客户端上的  在领域 上训练本地模型 ：  获取本地模型的经验损失  服务器更新：  根据 和 更新 :  使用聚合以获得新的全局模型:  广播:将更新后的全局模型 广播给所有客户端:.  3.**end for** |

## 3.5 本章小结

本章在FedAvg和FedProx基础上，在参数聚合阶段，提出自适应聚合的域泛化模块GA，与前文框架结合形成FedGA，用于动态调整聚合权重。详细阐述了其目标函数。理论推导和算法流程。

# 第四章 梯度传输的参数保护模块

## 4.1 引言

本章从通信机制优化的视角出发，提出一种基于梯度传输的参数保护方法，采用本地模型梯度替代传统的模型参数进行上传。随后，将围绕该方法展开理论分析，详述算法流程，并进一步阐明其在安全性、通信开销及模型性能等方面的综合优势。

## 4.2 梯度传输理论分析

在第三章提出的FedGA框架基础上，结合梯度下降优化算法的基本原理，梯度作为损失函数对模型参数的导数，指示了模型参数更新的方向与幅度。因此，本章进一步引入梯度传输机制，即客户端在本地训练完成后，仅上传本地训练产生的梯度信息，服务器基于各客户端上传的梯度动态聚合后更新全局模型参数，以在保持训练性能的同时，提升隐私保护水平。

在客户端本地训练阶段，设第轮通信时，全局模型参数为 ,客户端接收到后，在本地数据集上以初 始模型：

式（4-1）

开始本地训练。客户端根据本地经验损失最小化目标函数，执行轮梯度下降更新，得到本地模型

式（4-2）

其中，表示本地迭代轮数；是客户端本地学习率；

经过E轮更新后，得到本地最终模型 。

客户端随后计算模型更新量，即初始模型与本地训练后模型的差异，近似为累积的本地梯度更新：

= 式（4-3）

各客户端将和根据第三章GA方法计算的一起上传至服务器。

每个客户端将本地模型更新量与根据第三章GA模块计算得到的源域泛化差距一同上传至服务器。

在服务器端，首先根据GA方法动态调整聚合权重，更新策略为：

式（4-4）

其中，为权重调整步长，GA( ) 表示自适应权重调整策略。

服务器随后根据更新后的权重，聚合各客户端上传的本地梯度，得到全局更新方向： 式（4-5）

并使用聚合的全局梯度近似进行一次全局模型更新，相当于在服务器端执行一轮全局的随机梯度下降（SGD）步骤：

式（4-6）

最终，梯度传输方法在保持高隐私性的同时，能够实现与传统参数传输方式相当甚至更优的最终模型性能。

## 4.3 算法流程

算法4-1 梯度传输结合FedGA算法流程

|  |
| --- |
| 输入：  全局模型；  个客户端：；  初始权重：；  超参数：本地迭代次数 E，总通信轮次 R，GA的步长d |
| 输出：全局模型： |
| 1.服务器初始化：使用全局模型初始化本地模型 ：  2**.for all in do:**  客户端训练：  计算每个客户端上的  在领域 上训练本地模型 ：  计算本地模型与初始模型的差值即梯度近似：  获取本地模型的经验损失  服务器更新：  根据 和 更新 :  根据各客户端上传的，聚合得到全局更新的梯度近似  使用聚合的差值更新全局模型:  广播:将更新后的全局模型 广播给所有客户端:.  3.**end for** |

## 4.4 算法优势

在保证模型性能的基础上，梯度传输方法相较于传统参数传输还具备以下优势：

第一，系统安全性得到增强。与直接上传模型参数相比，上传梯度在传输过程中暴露的信息量更少，难以直接反推出本地数据分布或完整模型结构，在一定程度上抵抗了模型反演攻击与属性推理攻击。特别是在本章引入的 FedGA 动态聚合机制下，不同客户端梯度被赋予自适应权重，进一步破坏了攻击者从聚合结果中分离单个客户端信息的可能性，从而提升了系统的隐私保护能力。

第二，系统灵活性显著提升。传统方法需在服务器端存储所有客户端的完整模型参数，带来存储压力并增加管理复杂度，尤其在客户端数量动态变化场景下，维护成本高昂。而梯度传输模式仅需上传局部梯度，无需长期保存客户端模型，极大地简化了服务器端管理流程，并提升了系统在大规模、动态环境下的扩展性和适应性。

第三，通信成本有效降低。结合梯度量化、稀疏化、哈夫曼编码等压缩技术，梯度传输可以显著减少每轮通信的数据量。量化通过降低精度压缩梯度表示，稀疏化筛选关键更新元素，哈夫曼编码进一步实现无损压缩，这些技术综合应用可有效缓解带宽受限环境下的通信负担，提升联邦学习系统的训练效率与响应速度。

综上，梯度传输方法不仅能够在理论上保持甚至提高模型性能，还在安全性、灵活性与通信效率等方面展现出明显优势，为大规模、动态环境下的联邦学习系统提供了更优的解决方案。

## 4.5 本章小结

本章在第三章提出的FedGA基础上，在通信上传阶段，提出梯度传输的参数保护模块，提升隐私保护水平和全局优化灵活性。详细阐述了其理论分析、算法流程和算法优势。

# 第五章 实验结果与分析

## 5.1 实验设计

### 5.1.1 环境介绍

#### （1）实现方法

本文选用ResNet18和ResNet50两种经典的深度学习架构作为特征提取网络，选用PACS、OfficeHome和TerraInc数据集，采用FedProx、FedAvg +GA、FedProx + GA，三种领域泛化算法与单纯使用FedAvg的基线方法进行对比分析。

#### **（2）**硬件配置与软件环境

本章实验全部在服务器上完成，其CPU为Intel(R) Xeon(R) Gold 5218R CPU，内存251G，本实验在实验室服务器上完成，服务器配备 4 块 NVIDIA Tesla T4 显卡，每块显存为 15GB，支持 CUDA 11.4。运行系统为 Ubuntu 16.04.6 LTS。

在软件方面，运行的实验环境统一为Python3.8.20，Nvidia Driver=470.161.03，CUDA Version=11.4。主要的库文件版本为Numpy = 1.24.3，Torch = 1.11.0， Torchvision = 0.12.0。编译器采用Pycharm2023.1.4通过SSH连接对服务器文件进行远程操作。

### 5.1.2 模型选择方法

为了评估不同模型在各个领域上的表现，本章采用留一域交叉验证（Leave-One-Domain Cross-Validation，LODV）方法。留一域交叉验证是一种在领域泛化任务中常用的模型选择策略，其基本思想是：在多个领域的数据集上进行训练和评估，每次保留一个领域作为验证集，剩余领域作为训练集。在不同的数据集上进行实验时，使用该方法能选择最具鲁棒性的模型，并获得更为可靠的泛化性能评估。

## 5.2 实验描述与结果对比分析

### 5.2.1 基线实验

本节使用ResNet18搭配FedAvg在PACS、OfficeHome数据集上训练，使用ResNet18搭配 FedAvg在TerraInc数据集上训练。在训练阶段，图像首先会被随机裁剪，裁剪区域的面积占原图的70%到100%之间，并缩放至224×224的输入尺寸；接着以50%的概率对图像进行水平翻转，以增强模型的鲁棒性。此外，还对图像的亮度、对比度、饱和度和色调进行颜色抖动处理，从而增加图像的多样性；并且进行随机灰度化。每次将一个域作为测试集，其他域作为训练集，批量大小为16，客户端本地每次训练5轮，通信聚合40轮，学习率为0.001。到第32次聚合即本地第160轮训练时学习率下降到原来的10%，以更加精确地接近最优解。优化器使用SGD。实验结果分类准确率如表5-1、表5-2、表5-3，准确率和分类损失值变化如图5-1、图5-2所示。

表5-1 FedAvg在PACS数据集上的测试结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 方法 | PACS | | | | |
| Photo | Art painting | Cartoon | Sketch | 平均准确率 |
| FedAvg | 92.28 | 77.25 | 78.63 | 81.17 | 82.33 |

表5-2 FedAvg在OfficeHome数据集上的测试结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 方法 | OfficeHome | | | | |
| Product | Art | Clipart | Real | 平均准确率 |
| FedAvg | 73.55 | 59.21 | 51.64 | 74.91 | 64.83 |

表5-3 FedAvg在TerraInc数据集上的测试结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 方法 | Terra Incognita | | | | |
| L38 | L100 | L43 | L46 | 平均准确率 |
| FedAvg | 35.76 | 55.83 | 44.26 | 31.53 | 41.85 |

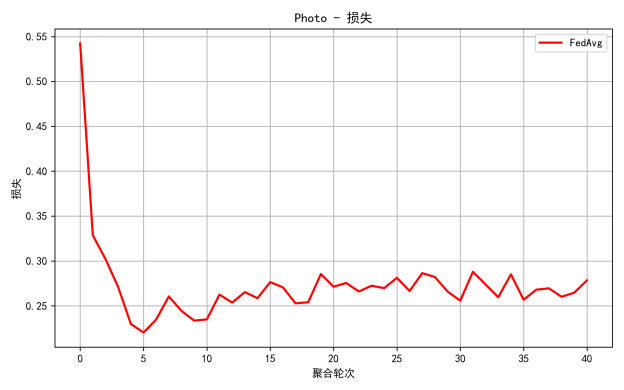
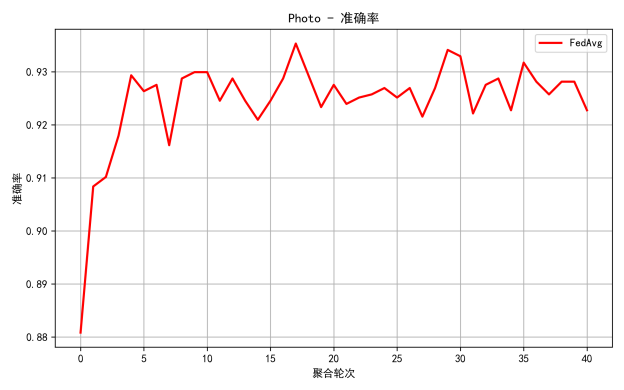


图5-1 基线测试集准确率变化示例 图5-2 基线分类损失值变化示例

### 5.2.2 FedProx实验

在Baseline的基础上，本文引入FedProx框架对联邦领域泛化过程进行改进。FedProx在本地优化目标中增加了额外的proximal正则项，以限制本地模型在训练过程中偏离全局模型过远，从而缓解由数据异构性带来的局部模型发散问题。为了验证在本地损失函数上引入正则项是否能优化模型性能，出于控制变量的角度，本节的FedProx实验未定义来动态调整本地训练轮数，仍与其他实验一样指定本地训练轮数为5。

具体地，分别在PACS和OfficeHome数据集上使用ResNet18搭配FedProx进行训练，在TerraInc数据集上使用ResNet50搭配FedProx进行对比实验。实验中，数据预处理方法和训练参数等均与基线方法保持一致，主要差异在于优化器的设置：FedAvg使用标准优化器SGD，而FedProx引入专门设计的FedProx优化器，并在初始化阶段通过update\_old\_init方法保存当前全局模型参数，以用于计算proximal项。在FedProx中，正则项强度由超参数mu控制，本文统一设置mu=0.1，以确保实验对比的公平性。实验结果如表5-4、表5-4、表5-6所示

表5-4 FedProx在PACS数据集上的测试结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 方法 | PACS | | | | |
| Photo | Art painting | Cartoon | Sketch | 平均准确率 |
| FedProx | 93.17 | 77.15 | 78.41 | 81.73 | 82.45 |

表5-5 FedProx在OfficeHome数据集上的测试结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 方法 | OfficeHome | | | | |
| Product | Art | Clipart | Real | 平均准确率 |
| FedProx | 74.34 | 58.43 | 52.26 | 74.89 | 64.98 |

表5-6 FedProx在TerraInc数据集上的测试结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 方法 | TerraInc | | | | |
| L38 | L100 | L43 | L46 | 平均准确率 |
| FedProx | 33.64 | 60.03 | 45.49 | 31.70 | 42.72 |

从表5-1至表5-6的实验结果可观察到，无论在FedAvg还是FedProx框架下，模型在PACS数据集上取得的分类准确率最高，OfficeHome次之，TerraInc最低。该现象反映出不同数据集在任务复杂度和领域间差异方面的本质差异。

PACS数据集包含四个视觉风格差异显著的子域，但每个子域仅涵盖七类通用物体。尽管图像风格各异，域间存在视觉特征层面的差异，但其类别语义一致性较高，且任务目标相对清晰，使得模型更容易在训练中学习到可迁移的跨域共享特征。因此，该数据集上的分类任务相对简单，模型在此数据集上通常能够达到最优性能。相比之下，OfficeHome数据集同样包括四个域，但每个子域包含多达65个与日常生活和办公场景相关的类别，任务粒度显著增加。类别数目的扩大显著提高了模型识别的难度，且由于类别之间存在更多细粒度的语义差异，模型在该数据集上泛化能力受到一定限制，因此整体表现介于PACS与TerraInc之间。TerraInc数据集源自真实野外相机陷阱图像，本文选取L38、L43、L46和L100四个不同地理区域作为领域划分。尽管该数据集每个域的类别数量相较OfficeHome较少，仅为10类野生动物，但由于其图像背景高度复杂且呈非结构化分布，不同地理区域之间的图像差异显著，导致域间偏移显著大于PACS和OfficeHome。模型难以捕捉稳定的领域不变特征，进而在此数据集上的性能最为不理想。实验结果显示了数据集复杂性与领域间异质性对模型性能的显著影响

为更加直观地对比FedAvg、FedProx的性能，本文汇总出FedAvg、FedProx在三个数据集上的平均测试结果如表5-7所示。

表5-7 FedAvg、FedProx在三个数据集上的平均测试结果

|  |  |
| --- | --- |
| 方法 | 平均准确率 |
| FedAvg | 63.00 |
| FedProx | 63.38 |

从表5-7的整体平均准确率来看，FedProx相较于FedAvg在三个数据集上的平均性能提升了0.38个百分点，表明引入正则化项对模型泛化能力具有一定的增强作用。进一步从单个数据集的结果来看，在PACS数据集上，FedProx的平均准确率为82.45%，较FedAvg的82.33%提升0.12%；在OfficeHome上，FedProx为64.98%，较FedAvg的64.83%提升0.15%；而在TerraInc数据集上，FedProx的提升最为显著，从FedAvg的41.85%提高至42.72%，提升了0.87%。

可见，尽管在PACS和OfficeHome两个数据集上性能提升幅度有限，但FedProx在TerraInc数据集上的显著提升说明其在应对高度异质数据时更具优势。这是由于TerraInc数据集不同领域之间的分布差异远大于PACS与OfficeHome。在域间差异相对较小的情况下，FedAvg已能够较好地完成建模任务，而在存在显著分布偏移的TerraInc中，FedProx通过引入proximal正则项，有效约束了本地模型的偏移程度，增强了模型对异构性的适应能力，从而带来了更明显的性能提升。

从数据集内各域的准确率分析来看，以PACS数据集为例，从表5-1和表5-4可看出，两种框架在Photo域上均取得了最高的分类准确率，超过90%；而在Art painting和Cartoon两个域上，准确率均相对较低，未达80%。这一差异主要源于照片图像风格更接近自然分布，图像结构清晰、背景干扰少，且目标位置较为固定，利于模型提取稳定特征；而艺术画和卡通图像则因风格多样、形态变化大、主体位置不稳定，造成域间偏移增强，从而增加了模型学习领域不变特征的难度。此外，PACS各域样本分布不均亦会对模型性能造成影响，例如Art painting域包含2048张图像，Cartoon域2344张，Photo域仅1670张，后者更容易产生过拟合，导致模型泛化能力受限。

尽管FedProx在多数域上相较FedAvg表现更优，但在个别情况下其性能可能下降。例如，当以TerraInc中的L38域作为测试集时，FedProx的准确率反而由FedAvg的35.76%降至33.64%。这可能是由于L38域与其他源域间的分布差异过大，使得正则化项未能有效缩小域间差距，反而在一定程度上抑制了关键特征的学习。此外，也不排除正则项权重设置或优化超参数（如学习率）未调节得当，影响了模型的泛化表现。

综上，实验结果验证了FedProx在引入正则约束后，能够在多数情形下增强模型在非独立同分布数据下的泛化能力，尤其在高异质数据环境中更具鲁棒性，具有一定的实用价值。

### 5.2.3 FedGA实验

本小节分别使用FedAvg和FedProx与GA模块融合后，搭配ResNet18在PACS和OfficeHome数据集上训练，搭配ResNet50在TerraInc数据集上训练。数据预处理、训练基本参数等都与baseline相同，FedAvg with GA使用SGD优化器，FedProx with GA引入专门设计的FedProx优化器。统一设置GA模块的参数step size=0.05。

#### （1）分类准确率

实验结果如表5-8、表5-9、表5-10所示。

表5-8 FedAvg、FedProx融合GA前后在PACS数据集上的测试结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 方法 | PACS | | | | |
| Photo | Art painting | Cartoon | Sketch | 平均准确率 |
| FedAvg | 92.28 | 77.25 | 78.63 | 81.17 | 82.33 |
| +GA | 94.97 | 79.79 | 78.37 | 83.15 | 84.07 |
| FedProx | 93.17 | 77.15 | 78.41 | 81.73 | 82.45 |
| +GA | 95.27 | 82.08 | 77.52 | 82.59 | 84.37 |

表5-9 FedAvg、FedProx融合GA前后在OfficeHome数据集上的测试结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 方法 | OfficeHome | | | | |
| Product | Art | Clipart | Real | 平均准确率 |
| FedAvg | 73.55 | 59.21 | 51.64 | 74.91 | 64.83 |
| +GA | 73.64 | 58.96 | 51.82 | 75.37 | 64.95 |
| FedProx | 74.34 | 58.43 | 52.26 | 74.89 | 64.98 |
| +GA | 74.68 | 57.81 | 52.62 | 74.57 | 64.92 |

表5-10 FedAvg、FedProx融合GA前后在TerraInc数据集上的测试结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 方法 | TerraInc | | | | |
| L38 | L100 | L43 | L46 | 平均准确率 |
| FedAvg | 35.76 | 55.83 | 44.26 | 31.53 | 41.85 |

续表5-10 FedAvg、FedProx融合GA前后在TerraInc数据集上的测试结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| +GA | 41.98 | 54.67 | 46.32 | 32.89 | 43.97 |
| FedProx | 33.64 | 60.03 | 45.49 | 31.70 | 42.72 |
| +GA | 43.24 | 58.62 | 47.51 | 37.43 | 46.70 |

表5-11 FedAvg、FedProx融合GA前后在三个数据集上的平均测试结果

|  |  |
| --- | --- |
| 方法 | 平均准确率 |
| FedAvg | 63.00 |
| +GA | 64.33 |
| FedProx | 63.38 |
| +GA | 65.43 |

从表5-11中可以看出，将GA模块引入后，无论是基于FedAvg还是FedProx框架，三个数据集上的平均准确率均有明显提升。其中，FedAvg融合GA后平均准确率由63.00%提升至64.33%，提升幅度为1.33%；FedProx融合GA后则从63.38%提高至65.43%，提升幅度达到2.05%。

进一步分析各数据集结果，GA在TerraInc数据集上的提升最为显著，分别提升了2.12%（FedAvg+GA）和4.02%（FedProx+GA）。这表明，在领域间分布差异较大的情况下，GA模块通过动态调整聚合权重，能够有效缓解各域间性能差异，提升全局模型的鲁棒性和泛化能力。相较之下，在OfficeHome数据集上，GA带来的提升较为有限。FedAvg融合GA后仅提升了0.12%，而FedProx融合GA后甚至出现了0.06%的轻微下降。这可能是由于OfficeHome数据集本身类别数量多、分类任务复杂，且不同域之间的分布差异相对较小，使得GA方法在优化聚合权重时优势难以充分发挥，甚至在部分情形下引入了额外的扰动，导致模型性能微弱下降。

总体而言，实验结果表明，GA方法在领域异质性显著的数据集上（如TerraInc）能够有效提升模型的领域泛化能力，而在域间差异较小且任务本身具有较高复杂度的数据集上，其性能提升受限，甚至存在一定的不确定性。但综合三大数据集的平均表现，可以验证GA方法整体上能够促进联邦学习框架下模型的跨域泛化能力，具有良好的应用前景。

#### **（2）**准确率和损失变化曲线图

以PACS数据集为例，本文绘制了FedAvg、FedAvg + GA、FedProx、FedProx + GA 这四种算法分别在以Photo、Art、Cartoon、Sketch四个域作测试集时的准确率和损失变化曲线图。

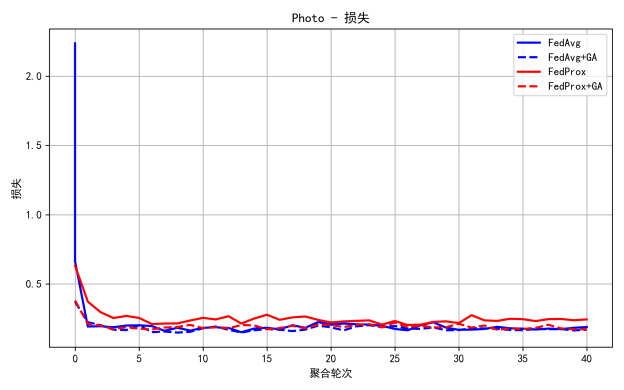
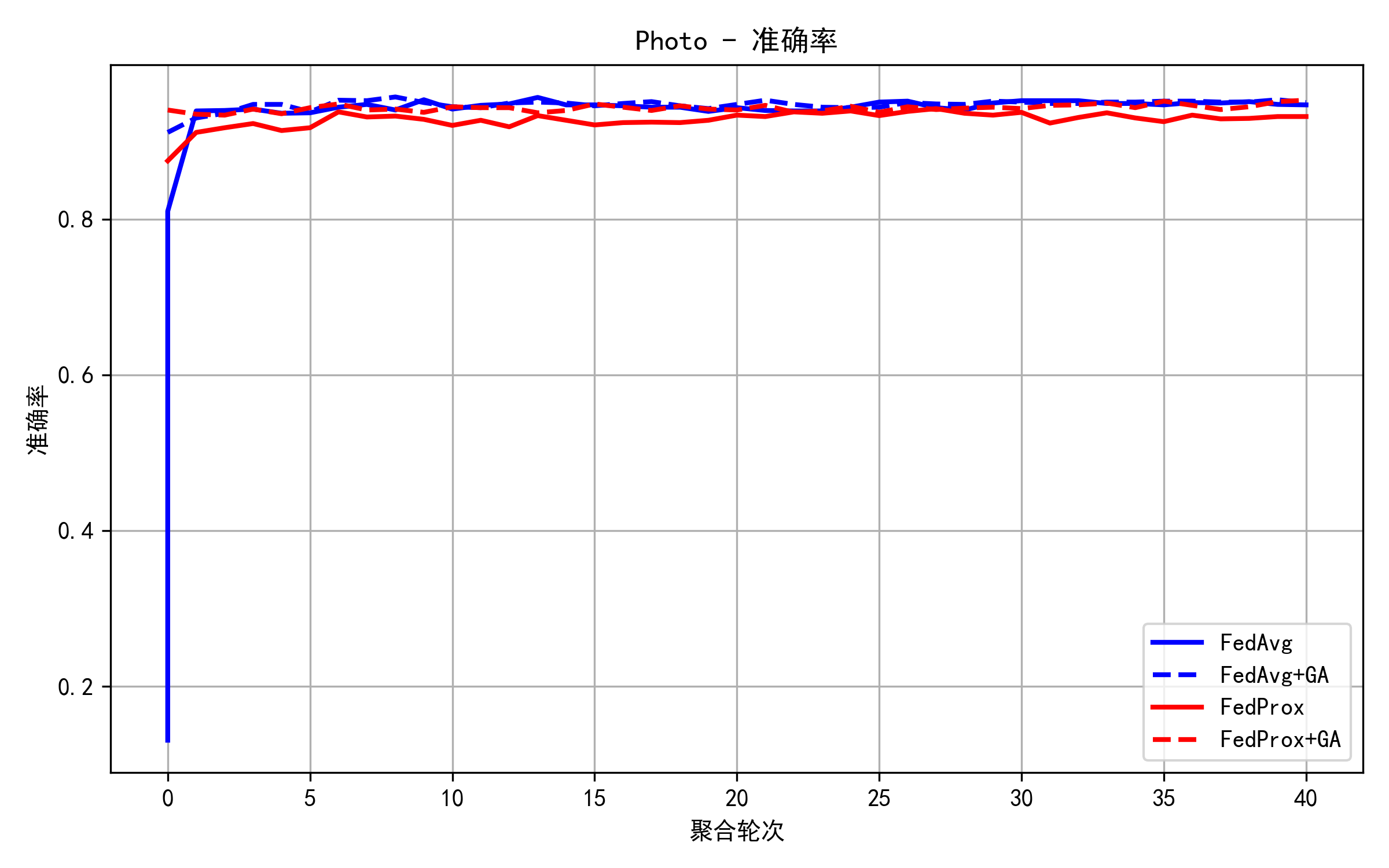


图5-3 Photo作测试集准确率变化示例 图5-4 Photo作测试集损失值变化示例

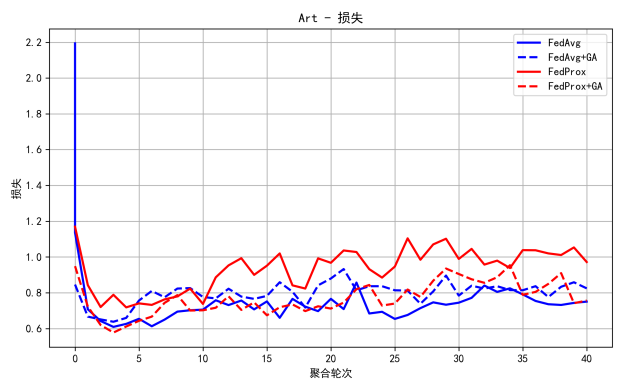
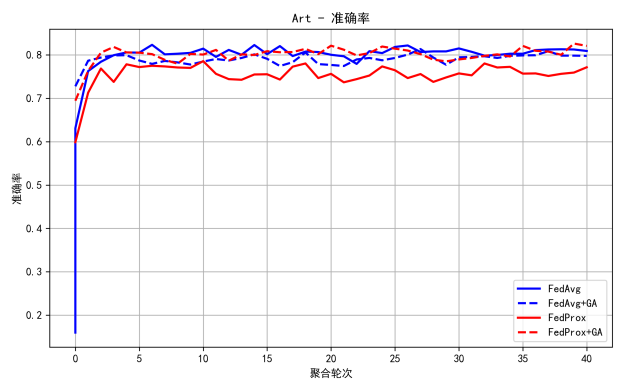


图5-5 Art作测试集准确率变化示例 图5-6 Art作测试集损失值变化示例

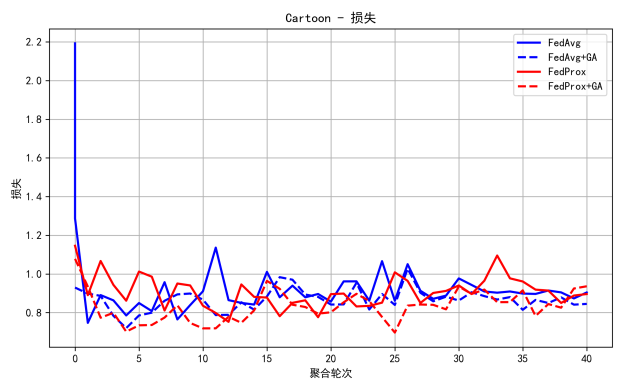
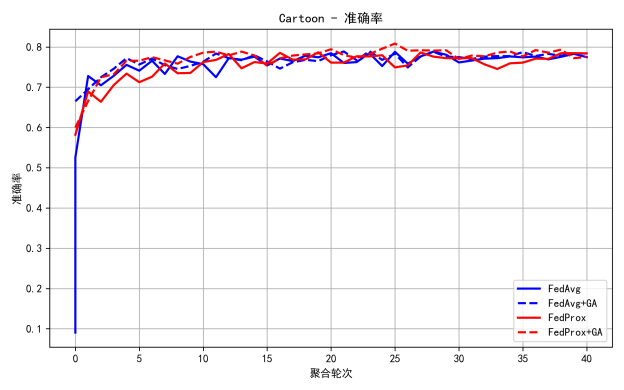


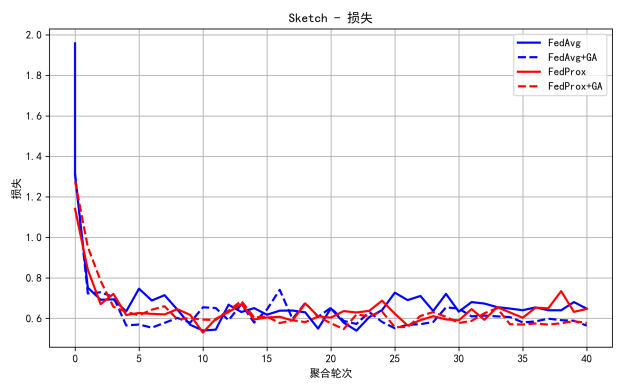
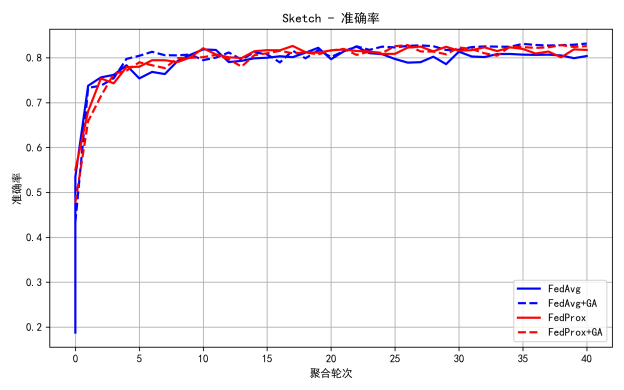
图5-7 Cartoon作测试集准确率变化示例 图5-8 Cartoon作测试集损失值变化示例

图5-9 Sketch作测试集准确率变化示例 图5-10 Sketch作测试集损失值变化示例

从图中可以观察到，在四个测试域中，各模型的准确率曲线均随着训练轮数的增加而上升，最终趋于收敛。总体而言，FedProx相较于FedAvg表现出小幅性能提升，而引入GA模块后，在Photo与Art两个测试域上的准确率提升尤为显著；但在Cartoon与Sketch两个域上，融合GA带来的性能改进相对有限。

结合损失曲线进一步分析，Photo与Sketch两个域的损失值随训练轮数的增加逐步下降，并最终收敛，表明模型训练过程较为稳定。而在Art与Cartoon两个域中，损失曲线波动明显，未能完全收敛，可能与任务本身复杂度较高、域间分布偏移显著以及训练参数设置不当等因素有关。在此背景下，GA模块通过动态调整聚合权重，有效缓解了训练过程的不稳定性，显著降低了各域的损失值并减小了波动幅度，尤其在Art域中，准确率与收敛性均获得显著改善。

然而，在Cartoon域中，FedProx结合GA后相较于未融合时并未展现出性能优势。分析认为，FedProx与GA均依赖正则化机制以抑制局部模型与全局模型的偏移，若正则化权重设置不合理，可能限制了模型对该域特征的有效学习，导致泛化能力受限，损失波动幅度较大，准确率提升亦不明显。

综上所述，GA模块在多数情形下能够提升模型的领域泛化能力，尤其在域间差异显著或训练波动较大的任务中表现优异。然而，其性能在特定域上仍受任务复杂性与正则项权重设定的影响，因而在实际应用中需结合具体数据分布特征对模型进行针对性调整与优化。

#### **（3）**泛化差距均值与方差对比

本文绘制了训练结束时各源域的泛化差距均值和方差，融合GA前后的结果分别以灰色和蓝色显示在每个图里。

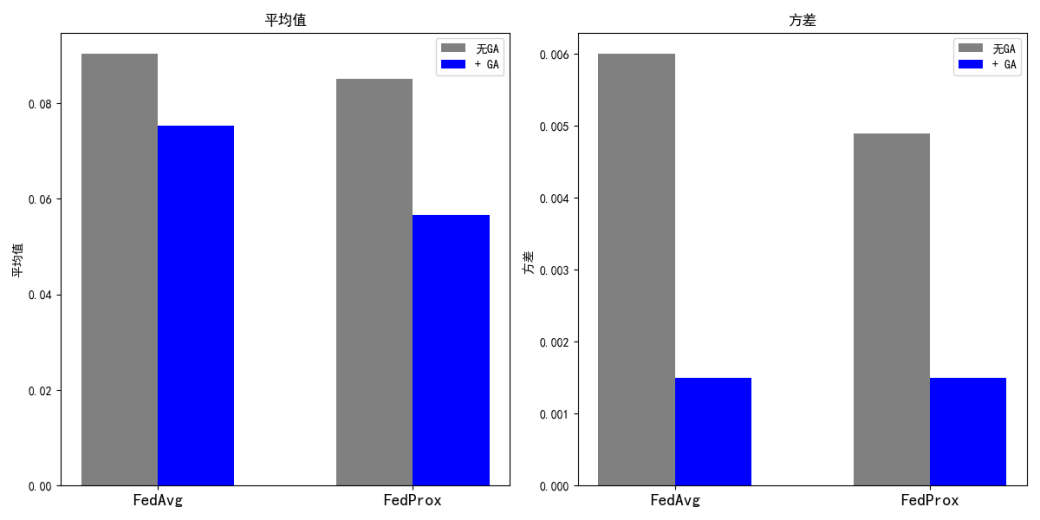


图5-11 PACS数据集上泛化差距对比图

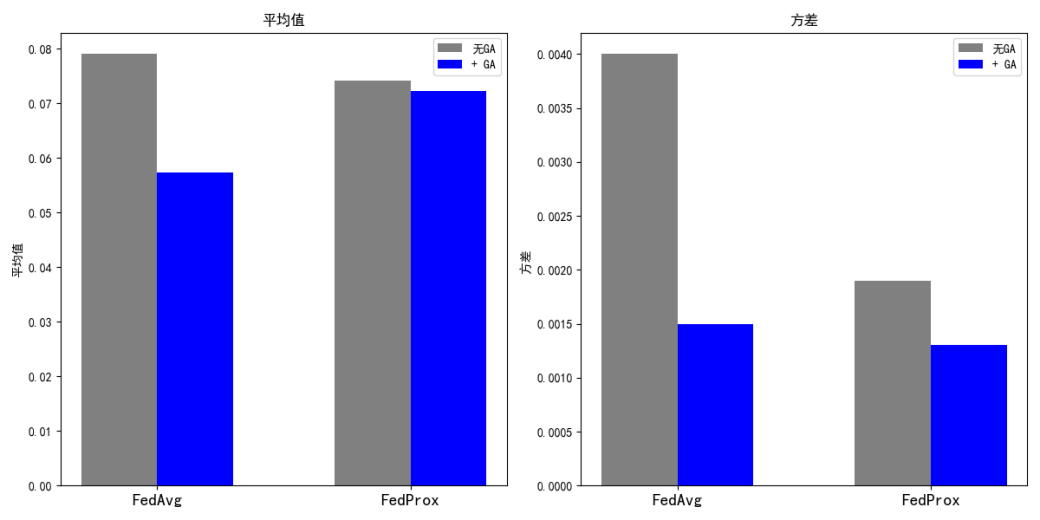


图5-12 OfficeHome数据集上泛化差距对比图

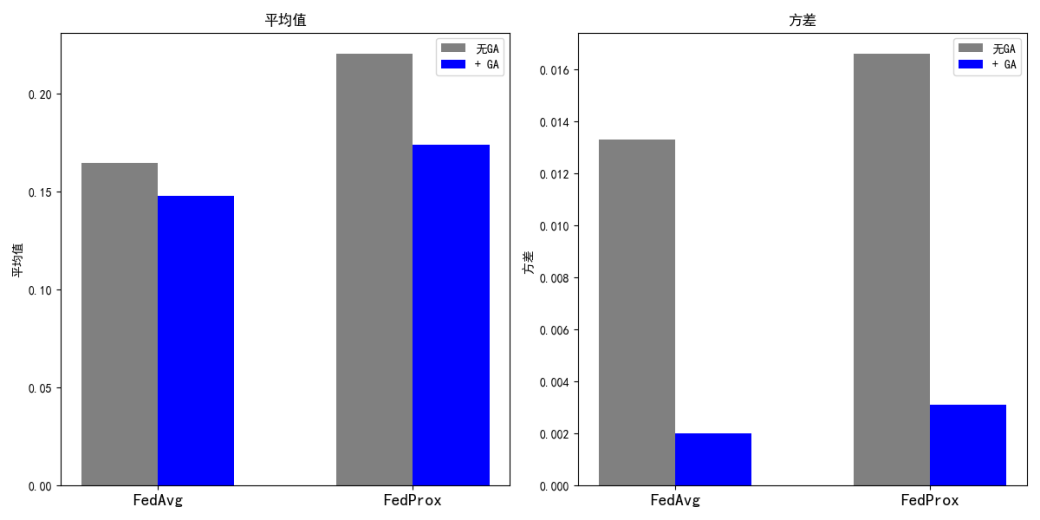


图5-13 TerraInc数据集上泛化差距对比图

通过图中结果可以明显看出，融合GA模块后，无论是泛化差距的均值还是方差均显著下降。这一现象验证了上一章的理论推导，即在联邦领域泛化的聚合过程中，引入对泛化差距的动态调整是一种有效策略，能够在全局优化过程中有效缓解领域偏差。根据第三章提出的理论公式，减小源域间泛化差距的方差有助于间接缩小目标域上的泛化误差上界，从而提升模型在未知测试域上的表现。

这一结果也与直觉一致：当各源域之间的泛化表现更加一致时，模型在面对未见领域时能够更加稳健地迁移已有知识，体现出更强的领域不变性与泛化能力。特别是在图5-13中，可以观察到TerraInc数据集的泛化差距在融合GA后大幅降低，进一步佐证了该方法在高异质性环境下的有效性。对应地，该数据集在实验中的性能提升也最为显著，体现出泛化差距方差控制对提升跨域性能的实际贡献。

#### **（4）**损失曲面图

为进一步探究GA模块在联邦领域泛化中的优化效果，本文对训练结束时的全局模型与各客户端本地模型的参数进行了二维投影，并将其嵌入到损失函数的等高面中进行可视化。以全局模型参数作为参考原点，左侧表示基础的FedAvg方法，右侧为融合 GA 模块后的 FedAvg with GA。以PACS数据集为例，图5-14 至图5-17 分别展示了四个域（Photo、Art、Cartoon、Sketch）作为测试集时的损失曲面与模型点投影分布。

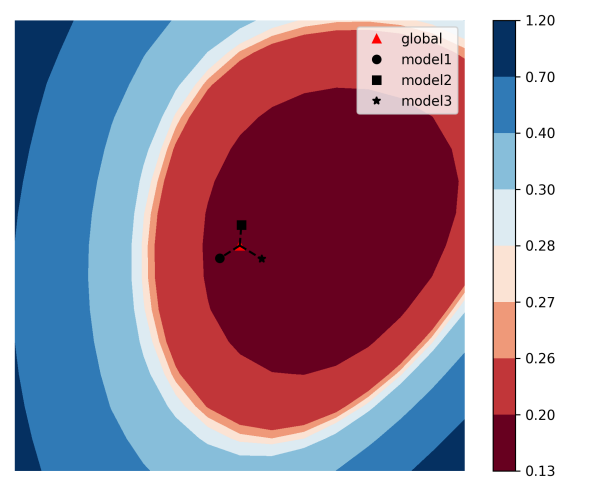
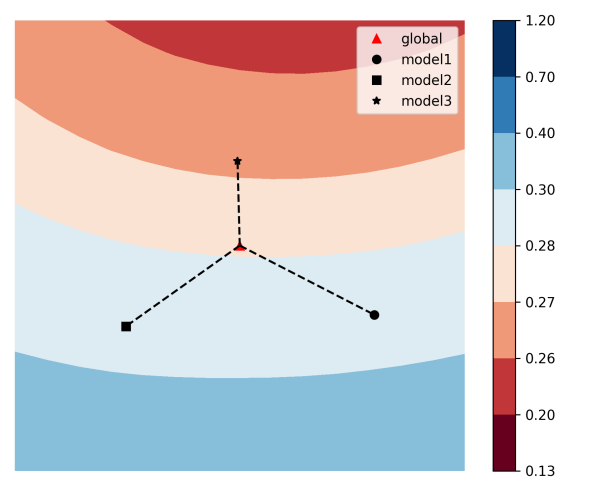


图5-14 Photo域上损失曲面图

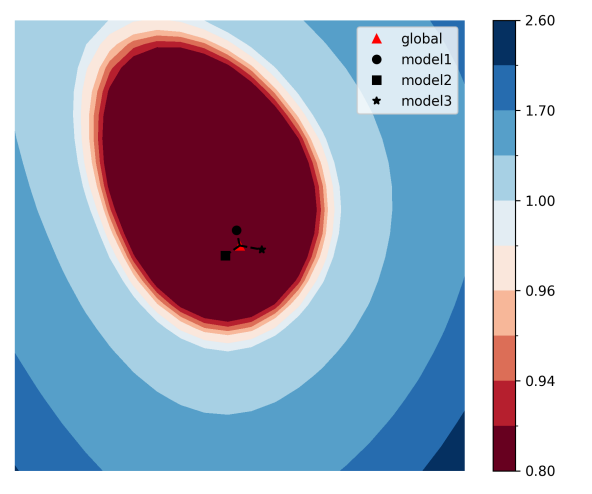
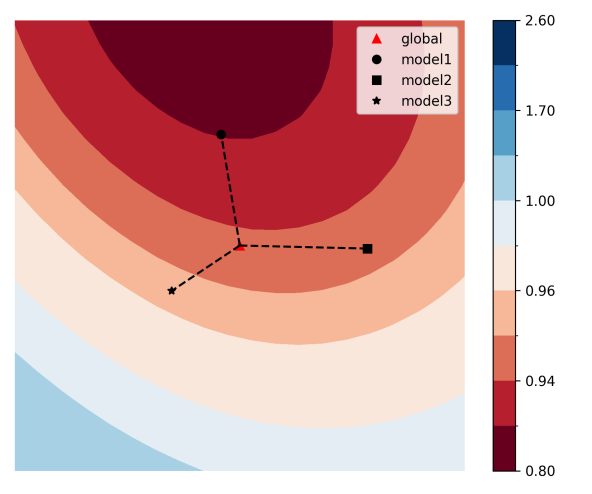


图5-15 Art域上损失曲面图

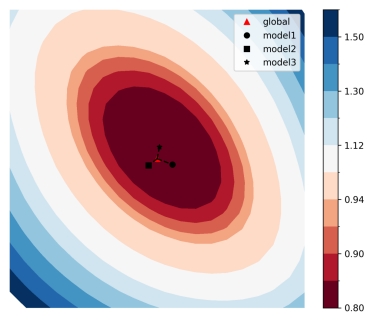
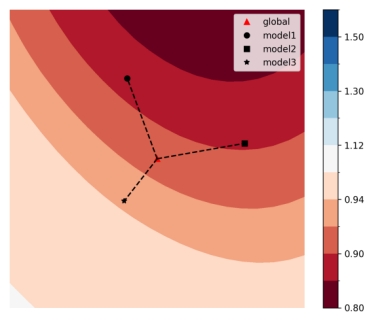


图5-16 Cartoon域上损失曲面图

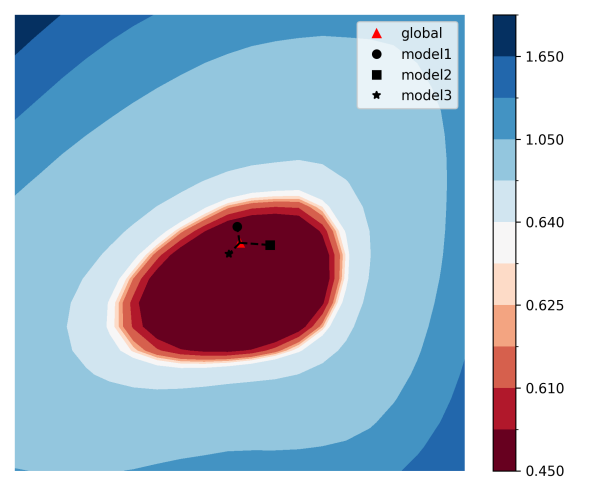
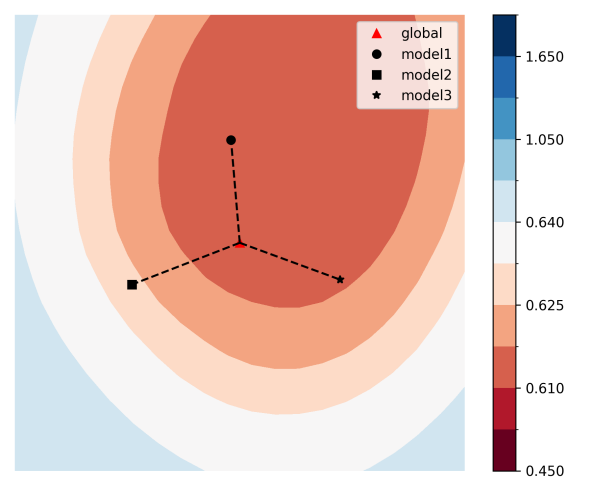


图5-17 Sketch域上损失曲面图

从图中可以直观地观察到，在引入GA模块后，全局模型与各本地模型均趋于收敛于损失表面中数值更低的区域。这一现象表明，各客户端在局部优化后的模型更新方向与全局优化目标之间的一致性明显增强，从而实现了更稳定且具有鲁棒性的模型聚合结果。该观察结果也与第三章提出的理论动机相一致：通过最小化源域间的泛化差距方差，可有效压缩目标域上的泛化误差上界，提升模型在未知域中的适应能力。

值得注意的是，以上损失曲面均基于测试阶段未曾出现的目标域绘制，进一步验证了GA方法在实际领域泛化任务中的有效性。GA模块不仅增强了模型的收敛性与一致性，也在泛化能力方面展现出显著优势。

此外，图中虚线所表示的是全局模型与各客户端本地模型之间在参数空间中的相对距离。可见，在融合GA模块后，虚线显著缩短，说明局部模型与全局模型之间的参数差异得到有效缩小。这意味着GA方法通过动态调整聚合权重，在客户端间实现了更加协调一致的优化方向，有效缓解了由数据异构性所带来的模型偏移问题，提升了局部与全局模型间的协同程度。

总体来看，损失曲面分析进一步从几何视角印证了GA方法的有效性。通过在联邦优化过程中引入自适应聚合机制，GA不仅提高了模型在训练域间的一致性，还显著增强了模型在未知目标域上的泛化性能，为处理复杂异构环境下的领域泛化任务提供了更加稳健的解决方案。

### 5.2.4 梯度传输实验

由于梯度传输模块的引入带来的主要优势是增加系统安全性和灵活性，故只在PACS数据集上进行实验，验证融合梯度传输的参数保护模块后是否能够保持或提高原有的模型性能表现。本文使用梯度传输的参数保护模块与FedGA融合，搭配ResNet18在PACS数据集上训练。数据预处理、训练基本参数、超参数等都与FedGA实验相同，实验结果如表5-12所示。

表5-12 FedGA结合梯度传输前后在PACS数据集上的测试结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 方法 | PACS | | | | |
| Photo | Art painting | Cartoon | Sketch | 平均准确率 |
| FedGA | 94.97 | 79.79 | 78.41 | 80.71 | 83.47 |
| +梯度传输 | 94.67 | 80.91 | 78.28 | 80.38 | 83.56 |

从表5-12可以看出，FedGA融合梯度传输模块后，在PACS数据集上的平均准确率与未融合时基本持平，仅提高了0.09%。这一结果验证了在提升模型安全性和灵活性的同时，梯度传输模块未显著降低模型的性能。

为进一步验证梯度传输模块的效果，本文绘制了FedGA融合梯度传输模块后，在PACS数据集以各域作为测试集时的准确率和损失曲线，结果如图5-18至5-25所示。通过引入梯度传输机制，模型不仅能有效保护数据隐私，还能够在此基础上完成全局模型的有效优化。

从图中可以分析到，初期由于仅依赖客户端上传的局部梯度进行模型更新，相比于传统的直接传输模型参数的方法，模型的初期性能较低。然而，由于不同客户端之间梯度方向具有较强的一致性，模型能够在早期阶段快速收敛，准确率明显提升。随着训练轮数的增加，客户端本地模型的差异逐渐加大，导致聚合梯度的方向一致性逐渐降低，进而使得模型收敛速度减缓，但整体收敛过程趋于平稳。

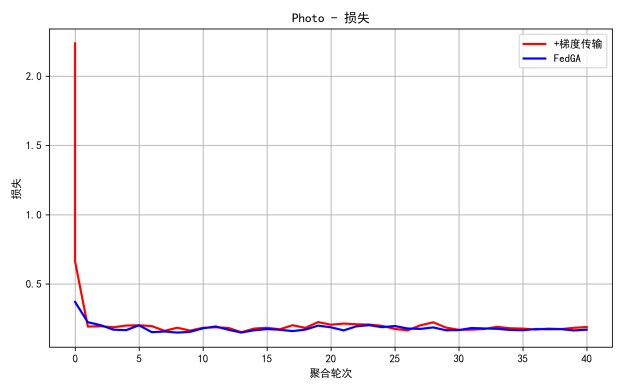
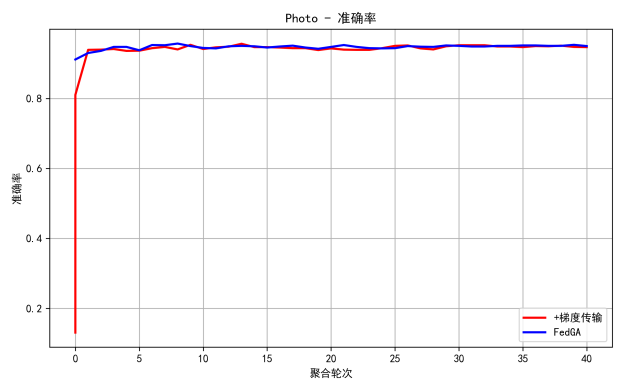


图5-18 Photo作测试集准确率变化示例 图5-19 Photo作测试集损失值变化示例

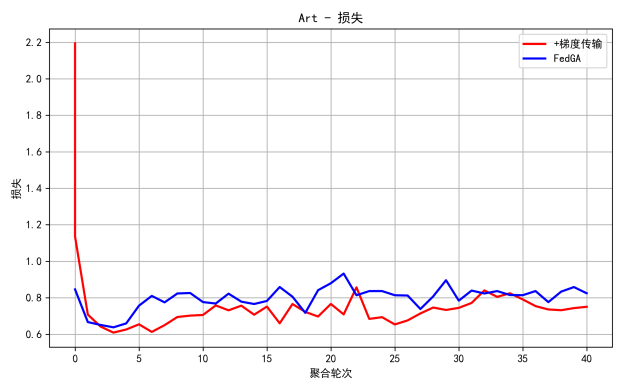
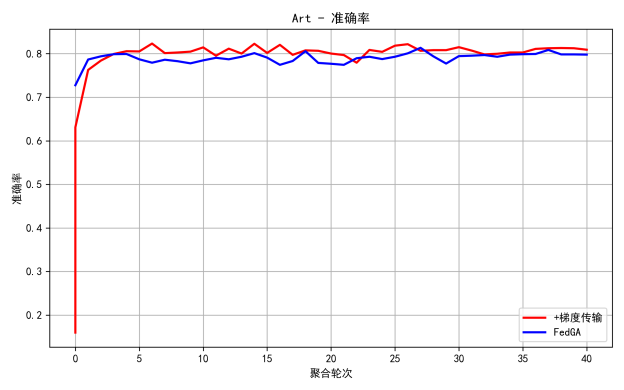


图5-20 Art作测试集准确率变化示例 图5-21 Art作测试集损失值变化示例

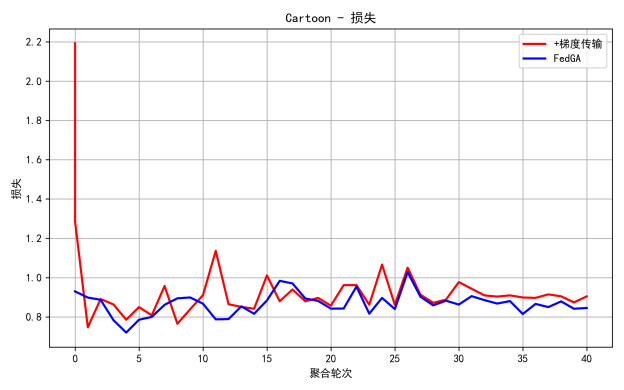
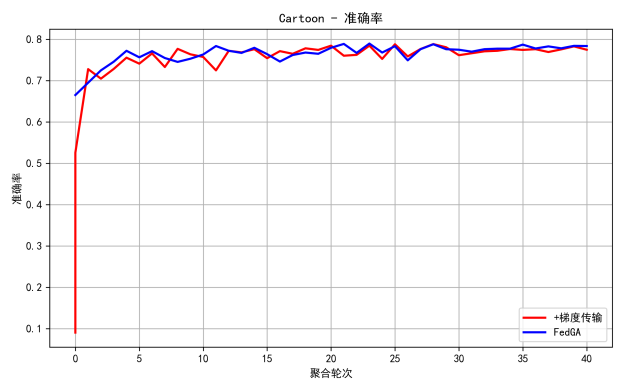


图5-22 Cartoon作测试集准确率变化示例 图5-23 Cartoon作测试集损失值变化示例

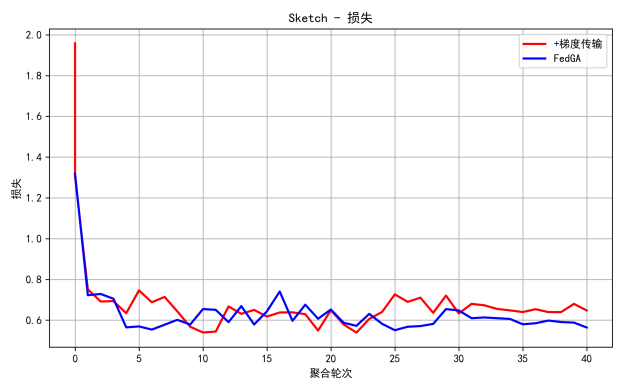
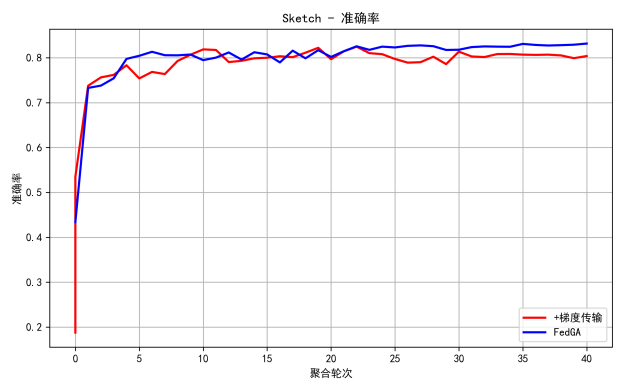


图5-24 Sketch作测试集准确率变化示例 图5-25 Sketch作测试集损失值变化示例

最终，梯度传输方法在保证高隐私性的同时，能够实现与传统参数传输方法相媲美的性能，甚至在Art域上表现更为优越。此实验结果表明，梯度传输不仅保持了全局模型的优化效果，还通过降低传输风险，为联邦学习系统提供了更加安全且高效的解决方案。

## 5.3 本章小结

本章旨在验证第二章、第三章和第四章所提出的FedProx框架、GA模块与梯度传输模块在领域泛化能力方面相较于传统FedAvg的性能提升。为此，设计了多组对比实验，系统评估各算法在不同场景下的优势与局限。

首先，在三个数据集上分别进行留一域交叉验证，比较FedAvg、FedProx、FedAvg + GA以及FedProx + GA四种算法的性能差异。其次，进一步在PACS数据集上探究引入梯度传输模块后的模型表现。本章主要关注以下三个问题：(1)Fedprox是否能够提升模型的领域泛化能力；(2)融合GA模块后是否能够进一步增强模型性能；(3)在引入梯度传输模块后是否仍能保持甚至提升模型的泛化能力。

实验结果表明，FedProx在大多数情况下能够有效提升模型的领域泛化能力，但提升幅度相对有限。融合GA动态聚合模块后，模型在各测试域上表现出更优的领域适应性，并在鲁棒性和稳定性方面具有明显优势。在此基础上，进一步引入梯度传输模块进行参数保护，模型的泛化能力得以保持，且在部分场景下取得了微小的性能提升。

# 第六章 总结与展望

## 6.1 研究工作总结

本文提出了基于特征不变性的联邦学习框架FedAvg，并在此基础上提出FedProx框架，改进本地优化目标，缓解数据异构性问题。结合自适应聚合（GA）方法，动态调整聚合权重，提升了模型的领域泛化能力。同时，探索了梯度传输优化策略以增强数据安全性。

FedProx通过引入L2正则项，有效减缓了数据异构性导致的局部模型发散问题。实验表明，FedProx能显著提升模型在目标域的识别能力，尤其在目标域未知的情况下。

结合GA方法后，动态调整聚合权重增强了客户端协同优化效果。实验结果表明，GA方法显著提升了多个数据集上的泛化能力，特别在TerraInc数据集上，性能有显著改善，展现了GA方法在异构数据环境中的优势。

最后，结合梯度传输策略进一步增强了隐私保护和全局优化灵活性。

综上，本文提出的特征不变性联邦学习算法在提升跨域图像分类性能的同时，兼顾了通信效率与隐私保护，为实际应用中的联邦学习系统部署提供了支持。

## 6.2 研究展望

尽管本文验证了FedProx与GA算法在领域泛化中的有效性，但仍有提升空间。当前实验基于PACS、OfficeHome和TerraInc数据集，未来可在更大规模、复杂度更高的数据集上验证算法性能。

FedProx与GA通过正则项提升了模型稳定性，但过强的正则化可能限制性能。未来可探索领域自适应策略，以更好平衡模型复杂度与泛化能力。梯度传输模块可结合剪枝、量化等技术减少通信量，特别适用于大规模联邦环境。提升算法在异构设备上的适配性，制定动态训练与聚合策略，将有助于增强系统稳定性与普适性。

未来研究可深入探讨聚合机制、正则化策略和数据异构性对性能的影响，并推动算法在医疗、金融等敏感领域的应用与落地。

# 参考文献

1. Kim Y. Convolutional neural networks for sentence classification[EB/OL]. arXiv:1408.5882, 2014[2024-04-29]. https://arxiv.org/abs/1408.5882.
2. I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., “Generative adversarial nets,” \*Advances in Neural Information Processing Systems\*, vol. 27, 2014.
3. A. Vaswani, N. Shazeer, N. Parmar, et al., “Attention is all you need,” \*Advances in Neural Information Processing Systems\*, vol. 30, 2017.
4. K. Zhou, Z. Liu, Y. Qiao, et al., “Domain generalization: A survey,” \*IEEE Transactions on Pattern Analysis and Machine Intelligence\*, 2022.
5. G. Blanchard, A. A. Deshmukh, Ü. Dogan, et al., “Domain generalization by marginal transfer learning,” \*Journal of Machine Learning Research\*, vol. 22, no. 1, pp. 46-100, 2021.
6. H. B. McMahan, E. Moore, D. Ramage, et al., “Federated learning of deep networks using model averaging,” \*arXiv preprint arXiv:1602.05629\*, 2016.
7. A. Kumar, V. Purohit, V. Bharti, et al., “MediSecFed: Private and secure medical image classification in the presence of malicious clients,” \*IEEE Transactions on Industrial Informatics\*, vol. 18, no. 8, pp. 5648-5657, 2021.
8. S. P. Karimireddy, S. Kale, M. Mohri, et al., “SCAFFOLD: Stochastic controlled averaging for federated learning,” in \*Proc. International Conference on Machine Learning (ICML)\*, 2020, pp. 5132-5143.
9. X. Peng, Z. Huang, Y. Zhu, et al., “Federated adversarial domain adaptation,” \*arXiv preprint arXiv:1911.02054\*, 2019.
10. J. Chen, J. Li, R. Huang, et al., “Federated transfer learning for bearing fault diagnosis with discrepancy-based weighted federated averaging,” \*IEEE Transactions on Instrumentation and Measurement\*, vol. 71, pp. 1-11, 2022.
11. Q. Li, B. He, D. Song, “Model-contrastive federated learning,” in \*Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)\*, 2021, pp. 10713-10722.
12. D. Li, Y. Yang, Y. Z. Song, et al., “Deeper, broader and artier domain generalization,” in \*Proc. IEEE Int. Conf. on Computer Vision (ICCV)\*, 2017, pp. 5542-5550.
13. H. Venkateswara, J. Eusebio, S. Chakraborty, et al., “Deep hashing network for unsupervised domain adaptation,” in \*Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)\*, 2017, pp. 5018-5027.
14. S. Beery, G. Van Horn, P. Perona, “Recognition in terra incognita,” in Proc. European Conf. on Computer Vision (ECCV), 2018, pp. 456–473.
15. Y. Sun, N. Chong, and H. Ochiai, “Feature Distribution Matching for Federated Domain Generalization,” in Asian Conference on Machine Learning. PMLR, 2023, pp. 942-957.
16. Z. Niu, H. Wang, H. Sun, et al., “MCKD: Mutually collaborative knowledge distillation for federated domain adaptation and generalization,” in \*Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)\*, 2023, pp. 1-5.
17. K. Saenko, B. Kulis, M. Fritz, et al., “Adapting visual category models to new domains,” in \*Computer Vision – ECCV 2010\*, Berlin: Springer, 2010, pp. 213-226.
18. X. Peng, Q. Bai, X. Xia, et al., “Moment matching for multi-source domain adaptation,” in \*Proc. IEEE/CVF Int. Conf. on Computer Vision (ICCV)\*, 2019, pp. 1406-1415.
19. M. Ghifary, W. B. Kleijn, M. Zhang, et al., “Domain generalization for object recognition with multi-task autoencoders,” in \*Proc. IEEE Int. Conf. on Computer Vision (ICCV)\*, 2015, pp. 2551-2559.
20. M. Arjovsky, L. Bottou, I. Gulrajani, et al., “Invariant risk minimization,” \*arXiv preprint arXiv:1907.02893\*, 2019.
21. A. Krizhevsky, I. Sutskever, G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” \*Communications of the ACM\*, vol. 60, no. 6, pp. 84-90, 2017.
22. K. Simonyan, A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” \*arXiv preprint arXiv:1409.1556\*, 2014.
23. K. He, X. Zhang, S. Ren, et al., “Deep residual learning for image recognition,” in \*Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)\*, 2016, pp. 770-778.
24. W. Hao, M. El-Khamy, J. Lee, et al., “Towards fair federated learning with zero-shot data augmentation,” in \*Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)\*, 2021, pp. 3310-3319.
25. J. Yuan, X. Ma, D. Chen, et al., “Collaborative semantic aggregation and calibration for federated domain generalization,” \*IEEE Transactions on Knowledge and Data Engineering\*, vol. 35, no. 12, pp. 12528-12541, 2023.
26. T. Li, A. K. Sahu, M. Zaheer, et al., “Federated optimization in heterogeneous networks,” \*Proceedings of Machine Learning and Systems\*, vol. 2, pp. 429-450, 2020.

# 致谢

春风拂面，又是杨柳絮纷飞的时节，学校主干道两旁栽种的泡桐树也已开出了繁盛的花。此刻，我终于要为此篇论文写下终章，也向我四年的青春告别。感谢这一路走来陪伴我的所有人。

感谢赵志诚导师，桃李不言，下自成蹊。从论文的选题到撰写，都离不开老师的耐心指导。感谢老师引领我进入科研的大门。

感谢周川湉师兄，向我推荐选题相关的文献，引领我快速全面了解论文研究方向。在实验代码、论文写作和PPT制作等各方面，师兄总是耐心地为我解答问题，陪我字斟句酌，以自己的经历和经验指导我精益求精。

感谢我的朋友们和室友们，四年时光，或欢喜或困厄，都有你们共享忧愁，温暖我、鼓励我、帮助我、给我力量，我们彼此陪伴走过秋风萧瑟，走过凛冽寒冬，也终于一起迎来春和景明。希望大家在这个盛夏都热烈地奔赴属于自己的锦绣前程。

也感谢在每一个困难时刻都坚持努力的自己，从懵懂无知一路跌跌撞撞向前进，拨开了二十岁的迷雾，信念生根，灵魂自在，往后人生道路宽广，途径的每一个夏天都草木葱郁，满目生机。路漫漫其修远兮，吾将上下而求索。

最后，感谢爱我的家人，感谢你们二十多年来对我的培养和支持，任何时刻你们都坚定地站在我身后，给我最强大的支撑，我成长的每一步都写满了你们无私的爱。

# 附录

附录1

本文提出基于特征不变性联邦学习的图像分类算法，它的影响不仅仅限于技术方面，还会对社会、健康、安全、法律、文化以及环境等非技术因素产生影响。

首先，从社会角度来看，该算法为数据隐私保护提供了强有力的解决方案，特别是在涉及敏感数据的场景中，如医疗、金融等领域。通过提高用户数据的隐私保护，该算法能够获得社会公众的广泛接受，推动更广泛的人工智能应用。

其次，从健康角度来看，该算法可广泛应用于医学影像分析、疾病预测和个性化治疗方案的制定等方面。通过利用分布在不同医院或医疗机构的数据，训练出更加精准的模型，而不需要将患者的个人数据汇聚到一个中心位置，这不仅保障了数据的隐私性，还能够提高医疗系统的整体效率和诊断准确率。特别是在疫情防控和突发公共卫生事件中，该算法能够帮助快速汇聚多方资源，以便做出更迅速的决策。

从安全角度来看，该算法在涉及智能交通、金融系统等关键基础设施时，通过在各个节点进行加密数据处理和安全模型更新，可以有效抵御数据泄露、篡改以及模型攻击等安全风险。将大大增强人工智能系统的安全性，确保敏感数据在训练过程中的保密性和完整性。

从法律角度来看，随着数据隐私保护法规和全球范围内隐私保护法律的日益严格，该算法为符合这些法律要求提供了可能。特别是在处理跨国、跨地区的数据时，该算法不仅避免了数据的跨境流动，还确保了数据的局部处理符合相关法律的要求，减少了法律风险。

从文化角度来看，该算法可以通过促进全球数据共享与合作，促进技术的公平性与普及性。由于数据无需集中存储，发展中国家和偏远地区也能够参与到人工智能的发展中，避免了数据和技术的集中化，促进了全球技术的共享与合作。

最后，从环境角度来看，该算法可以帮助提升智慧城市、环境监控和气候变化研究等领域的效率。例如，基于边缘计算的联邦学习可以有效地处理环境数据而不需要过多依赖云计算资源，从而减少能源消耗。

综上所述，该算法不仅具有技术创新性，也有着广泛的社会影响。在推广和应用过程中，需要综合考虑其对于社会、健康、安全、法律、文化和环境等方面的影响，并确保其合法、公正和可持续的发展。