

肖黄梅, 龙伟, 胡灵犀, 等. 一种面向非独立同分布数据下的联邦学习数据增强方法[J]. 智能计算机与应用, 2025, 15(11): 110-116. DOI:10.20169/j. issn. 2095-2163. 251118

一种面向非独立同分布数据下的联邦学习数据增强方法

肖黄梅, 龙 伟, 胡灵犀, 蒋林华

(湖州师范学院 信息工程学院, 浙江 湖州 313000)

摘要: 联邦学习是一种新兴的用于隐私保护的分布式学习框架。在联邦学习框架中,客户终端设备在本地进行训练,不需要将数据传输到中心服务器中,从而有效地保证了数据的隐私性和安全性。然而,当参与训练的客户端数据分布趋于非独立同分布(Non-IID)时,联邦学习训练得到的模型通常比集中式模型框架下的模型性能差。为了解决联邦学习节点上数据非独立同分布导致的模型训练精度不理想的问题,本文提出一种基于生成对抗网络的数据增强方法,在不涉及用户隐私的前提下有效改善模型训练性能。仿真结果表明,在多种 Non-IID 数据划分策略下,本文所提出的方法能有效的提高模型精度并加快模型收敛,与基准方法相比,所提方法在 Non-IID 场景下能取得 20%以上的精度提升。

关键词: 联邦学习; 非独立同分布; 生成对抗网络; 数据增强

中图分类号: TP181 文献标志码: A 文章编号: 2095-2163(2025)11-0110-07

A federated learning data augmentation algorithm for non-independent and identically distributed data

XIAO Huangmei, LONG Wei, HU Lingxi, JIANG Linhua

(School of Information Engineering, Huzhou University, Huzhou 313000, Zhejiang, China)

Abstract: Federated learning is an emerging distributed learning framework for privacy preservation. In this framework, client devices perform training locally without transmitting data to a central server, effectively ensuring data privacy and security. However, models trained in federated learning often perform worse than those trained in centralized frameworks, especially when the data distribution of the participating clients is non-independent and identically distributed (NIID). To address the problem of suboptimal model training accuracy caused by NID data on federated learning nodes, this paper proposes a data augmentation method based on generative adversarial networks (GANs) to effectively improve model training performance without compromising user privacy. Simulation results demonstrate that the proposed method effectively improves model accuracy and accelerates model convergence under various NID data partitioning strategies. Compared with baseline methods, the proposed method achieves over 20% accuracy improvement in NID scenarios.

Key words: federated learning; non-independent and identically distributed; generative adversarial networks; data augmentation

0 引言

近年来,随着电子设备存储容量的增加,能够存储或访问更多数据。但如何在保护用户隐私的前提下充分利用人工智能挖掘出相关数据的价值,已成为了人工智能技术的一大难题^[1]。谷歌于 2016 年提出了一种联邦学习方法,旨在联合众多拥有数据的客户端共同训练一个机器学习模型。联邦学习不仅在电信、医疗、金融等领域发挥了重要作用,还凭借其分布式和隐私保护等特性,在边缘计算和语音

处理等场景中得到广泛应用^[2-3]。

虽然联邦学习是一种新兴的隐私保护方法,但与集中式学习相比,在实际应用中仍面临诸多挑战。联邦学习将大部分数据保留在本地,从而有效保障了数据的安全性和隐私性。然而,由于不同客户端的数据类别和分布存在差异,往往不满足独立同分布(Non-IID)假设,这使得联邦学习在模型准确度、通信成本等方面的性能往往低于集中式方法^[4-5]。

在横向联邦学习中,应对 Non-IID 数据问题的方法主要包括数据共享、数据增强和局部微调。数

作者简介: 肖黄梅(1999—),女,硕士,主要研究方向:联邦学习,机器学习。

通信作者: 龙 伟(1978—),男,博士,硕士生导师,主要研究方向:智能信息处理,数字化农业信息处理技术。Email:lw@zjhu.edu.cn。

收稿日期: 2024-03-09

哈尔滨工业大学主办 ◆专题设计与应用

据共享方法简单而高效, 有效地处理水平联邦学习的 Non-IID 数据, 其核心是构建一个分布均匀的全局共享数据集 G , 并通过该数据集对全局模型进行预热训练, 将模型存储在服务器端^[6]。随后, 从 G 中随机抽取一定比例的数据分发至各客户端, 使其结合本地数据与共享数据进行训练, 从而更新客户端模型^[7]。数据增强方法则通过随机变换等方式增加训练数据的多样性, 从而缓解联邦学习中的数据异构问题^[8]。Zhu 等^[6]提出数据共享策略, 改变搬土距离, 即通过创建一部分数据来改善对 Non-IID 数据的训练, 该数据在所有边缘设备之前全局共享; Wang 等^[9]提出 FedNova 算法, 通过引入正则项减轻数据分布差异的影响, 从而使局部损失函数保持一致; Lin 等^[10]使用知识蒸馏来加速模型收敛, 提出在中心服务器中聚合, 将局部模型的优点相结合; Li 等^[11]提出 FedProx 算法, 在 FedAvg 算法的基础上添加了一个正则项, 减少权重的差异; Karimireddy 等^[12]提出使用 SCAFFOLD (Stochastic Controlled Averaging for Federated Learning) 算法传输控制变量, 从而减少权重差异; Chen 等^[13]提出两阶段的联邦学习优化方法 FedSA, 将联邦学习训练分为初始阶段与收敛阶段, 初始阶段设置较大的本地训练轮数, 但由于本地训练轮数较多, 所产生的权重会有较大的差异。

基于上述研究成果, 可以发现: 在数据非独立同分布的情况下, 联邦学习的优化可从全局模型权重聚合和本地客户端数据两方面入手。为此, 本文提出一种面向联邦学习的数据增强方法, 在保护数据隐私的前提下, 有效缓解 Non-IID 数据导致的模型精度下降问题, 同时保持训练效率不受影响。

1 联邦学习理论基础

典型的联邦学习场景是在本地客户端设备负责存储和处理数据的约束下, 只上传模型更新的梯度信息, 在数千万到数百万个客户端设备上训练单个全局模型 w 。中心服务器的目标函数 $F(w)$ 可表示为:

$$\min_w F(w), F(w) = \sum_{k=1}^m \frac{n_k}{n} F_k(w) \quad (1)$$

其中, m 是参与训练的客户端设备总数; n 是所有客户端数据量总和; n_k 是第 k 个客户端的数据量; $F_k(w)$ 是第 k 个设备的本地目标函数:

$$F_k(w) = \frac{1}{n_k} \sum_{i \in d_k} f_i(w) \quad (2)$$

其中, d_k 是第 k 个客户端的本地数据集, $f_i(w) = \alpha(x_i, y_i, w)$ 是具有参数 w 的模型对数据集 d_k 中的实例 (x_i, y_i) 产生的损失函数。

d_k 中所有实例产生的损失函数之和除以客户端 k 的总数据量就是本地客户端的平均损失函数, 损失函数与模型精度成反比, 因此, 机器学习的目标函数优化通常是让损失函数达到最小值。

在联邦学习的目标函数优化算法中, 通常采用质量随机梯度下降 (SGD) 算法, 即通过将本地客户端模型训练的损失函数乘以固定的学习速率来计算新一轮的权重更新^[14]。本地客户端的模型权重更新为:

$$w_{t,k} = w_{t-1,k} - \eta \nabla F_k(w) \quad (3)$$

第 t 轮通信中心服务器的模型聚合更新为:

$$w_t = \sum_{k=1}^K \frac{n_k}{n} w_{t,k} \quad (4)$$

在大多数情况下, 联邦学习采用客户端-服务器架构, 客户端也称为数据持有者。

随着迭代的进行, 客户端的模型与全局目标模型之间的优化趋势分歧不断累积, 需要更多的通信轮次才能达到收敛, 降低联邦学习的性能^[15]。因此, 具有多个局部更新的局部训练可能会损害收敛性, 甚至导致发散^[16]。

2 生成对抗网络理论基础

生成对抗网络 (Generative Adversarial Network, GAN,) 的核心思想来源于博弈论的纳什均衡^[17]。设定参与游戏双方分别为一个生成器 (Generator) 和一个判别器 (Discriminator), 生成器的目的是尽量去学习真实的数据分布, 而判别器的目的是尽量正确判别输入数据是来自真实数据还是来自生成器。为了取得游戏胜利, 这两个游戏参与者需要不断优化, 各自提高自己的生成能力和判别能力, 这个学习优化过程就是寻找二者之间的一个纳什均衡, 生成对抗网络结构示意图如图 1 所示。任意可微分的函数都可以用来表示 GAN 的生成器和判别器, 本文用可微分函数 D 和 G 分别表示判别器和生成器, 两者输入分别为真实数据 x 和随机变量 z 。 $G(z)$ 则为由 G 生成的尽量服从真实数据分布 p_{data} 的样本。如果判别器的输入来自真实数据, 标注为 1; 如果输入样本为 $G(z)$, 标注为 0。 D 的目标是实现对数据来源的二分类判别: 真为来源于真实数据 x 的分布; 伪为来源于生成器的伪数据 $G(z)$ 。而 G 的目标是使自己生成的伪数据 $G(z)$ 在 D 上的表现 $D(G(z))$ 和真

实数据 x 在 D 上的表现 $D(x)$ 一致, 这两个相互对抗并迭代优化的过程使得 D 和 G 的性能不断提升, 当最终 D 的判别能力提升到一定程度, 并且无法正确判断数据来源时, 可以认为这个生成器 G 已经学到了真实的数据分布。

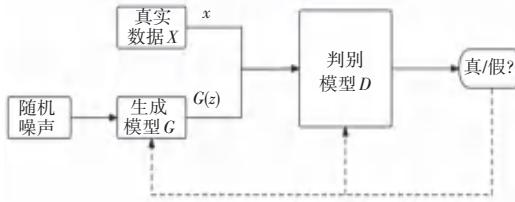


图 1 生成对抗网络结构示意图

Fig. 1 Schematic diagram of the generative adversarial network structure

生成对抗网络的训练过程可看作 2 个模型的零和博弈, 生成器输入低维随机噪声, 输出虚拟样本, 其优化目标是尽可能让判别器将虚拟样本误判为真实样本; 而判别器输入真实样本和虚拟样本, 输出每条样本是真实样本的概率, 其优化目标是尽可能正确区分两类样本, 该过程可看作以下的优化问题, 如下式:

$$\min_{G} \max_{D} \mathbb{E}_{x \sim p_{\text{data}}} [\log D(x)] + \mathbb{E}_{z \sim p_z} [\log (1 - D(G(z)))] \quad (5)$$

生成对抗网络已经发展出多种变体, 主要可分为两类: 体系结构变体和损失函数变体。GAN 进行了许多的优化与改进, 例如条件生成对抗网络 (Conditional GAN, cGAN,) 使用辅助的标签信息来增强原始 GAN, 对生成器与判别器都使用标签数据进行训练, 使模型具备产生特定条件数据的能力^[18]; 半监督生成对抗网络 (Semi-supervised GAN, sGAN) 利用少量标签, 通过判别器或分类器末端重建标签信息, 实现对半监督数据的有效利用; 辅助分类器生成对抗网络 (Auxiliary Classifier GAN, AC-GAN) 主要用于条件图像合成任务, 在判别器中增加一个辅助分类器, 同时可引入额外辅助信息, 从而支持半监督学习并提升生成质量^[19]。

在 AC-GAN 中, 损失函数由判别器损失和分类器损失两部分组成, 判别器的训练目标是最大化 $L_s + L_c$, 其中 L_s 为源损失, 用于区分真实样本与生成样本, 公式如下:

$$L_s = \mathbb{E}[\log P(S = \text{real} | X_{\text{real}})] + \mathbb{E}[\log P(S = \text{fake} | X_{\text{fake}})] \quad (6)$$

L_c 为类别损失, 用于预测样本的类别, 公式如下:

$$L_c = \mathbb{E}[\log P(C = c | X_{\text{real}})] + \mathbb{E}[\log P(C = c | X_{\text{fake}})] \quad (7)$$

生成器的训练目标则是最大化 $L_s - L_c$, 以生成逼真的图像并满足指定类别。

3 数据增强的联邦学习方法

3.1 问题描述

在传统的分布式机器学习中, 服务器可以收集和处理所有的训练数据, 而且可以在训练前, 将数据集划分为多个与全局数据分布特征相同的独立同分布数据子集, 分配给各个节点, 从而加速模型训练过程。然而由于物联网设备的异构性, 处理能力、架构不尽相同, 加上用户本身的使用习惯、个人爱好千差万别, 导致数据的非独立同分布。在联邦学习训练中, 每个节点拥有的数据不尽相同, 如: 在医疗领域, 不同的参与方可能拥有着来自不同病人的医疗图像数据, 综合性医院每种类型的数据可能都比较全面, 而专科医院可能只有某一类的数据, 因此在训练过程中, 各节点的数据在汇总到服务器时可能出现节点漂移。McMahan 等^[2]通过实验证明, FedAvg 算法不仅能够有效处理独立同分布 (IID) 数据, 对于非独立同分布 (Non-IID) 数据, 同样可以实现模型训练收敛, 但节点漂移仍可能影响模型的收敛速度和性能。但是相关实验结果显示, Non-IID 的数据分布场景下, FedAvg 算法的模型准确率和收敛速度相较于 IID 数据分布场景都有很大程度的下降。联邦学习中 Non-IID 数据下的权值发散示意如图 2 所示。

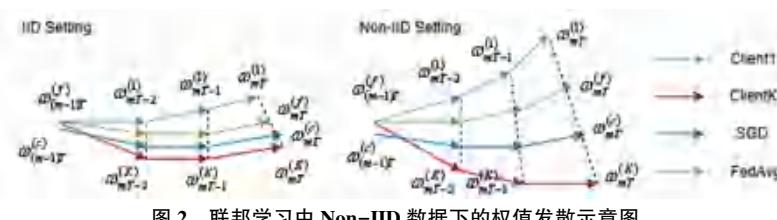


图 2 联邦学习中 Non-IID 数据下的权值发散示意图

Fig. 2 Schematic diagram of weight divergence under Non-IID data in federated learning

图 2 中, $\omega_{mT}^{(K)}$ 表示第 K 个设备在第 m 次训练的第 T 次本地训练后的模型参数, 当数据是独立同分

布数据时, 每个客户端在多轮训练后的权值更新差距较小, 而且 ω_{mT}^c 与 ω_{mT}^f 的偏差很小; 当数据是非独

立同分布数据时, 每个客户端在多轮训练后的权值更新差距就会变大。

本文提出了一种面向非独立同分布数据下的联邦学习数据增强方法——Fed-Aug。该方法通过数据增强的方式来平衡各个节点之间的数据分布差异, 从而提高全局模型的准确率。本文的数据增强方案在预处理阶段进行, 而联邦学习的模型训练过程称为主任务训练。当主任务开始时, 各客户端基于增强后的数据集进行模型训练, 其训练流程与常规联邦学习一致。该方法无需将数据上传至中心服务器, 从而在一定程度上保护了数据隐私。此外, 除了因客户端数据集规模扩大而带来的训练开销外, 该方法不会增加客户端与服务器之间的通信开销, 具有较好的实用性。

3.2 联邦学习数据增强方法

本文提出一种面向非独立同分布数据下的联邦学习数据增强方法——Fed-Aug 算法如图 3 所示。

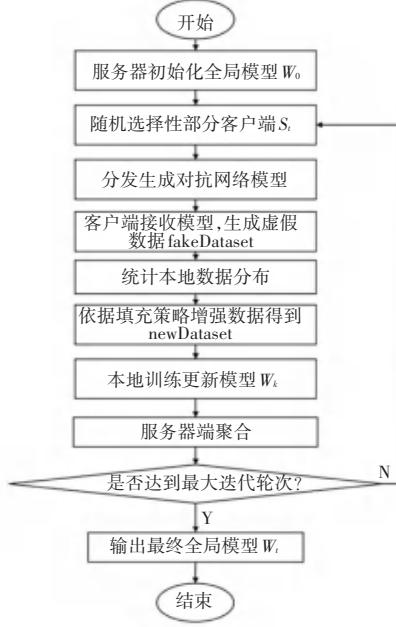


图 3 联邦学习数据增强方法流程

Fig. 3 Flowchart of the federated learning data augmentation algorithm

首先, 联邦学习服务端利用公共数据集训练生成对抗网络(GAN)模型, 并将该生成模型分发给所有参与训练的客户端; 各客户端使用生成模型对本地数据进行增强, 同时独立统计本地数据分布信息, 并根据不同的数据填充策略对数据进行填充, 从而完成数据增强。

4 仿真实验

4.1 实验设置

本文提出的优化方法基于 PyTorch 框架实现,

部分基准方法采用 NIID-Bench 提供的实现代码。实验选取了两个常用的图像分类数据集 MNIST 与 CIFAR-10 作为实验数据集, 采用离线仿真方式进行, 以验证所提算法的性能。实验运行环境包括硬件配置与软件环境, 见表 1。

表 1 实验环境配置

Table 1 Experimental environment configuration

类别	配置
处理器 CPU	AMD Ryzen 7 6800H with Radeon Graphics
内存 (RAM)	128 GB
编程语言	Python 3.7.11
深度学习框架	PyTorch 1.13.1

为了更贴近实际应用场景, 本文采用狄利克雷 (Dirichlet) 分布与病态划分方式来构造非独立同分布数据集。其中, Dirichlet 分布通过调整参数 (如 0.5 与 0.1) 来控制数据的异质性; 而病态划分则模拟极端情况下的标签倾斜, 使每个客户端仅包含少数类别, 甚至出现 1-Label 设置 (即每个客户端只含单一类别样本)。CIFAR-10 数据集在不同划分方式下的分布情况如图 4 所示。

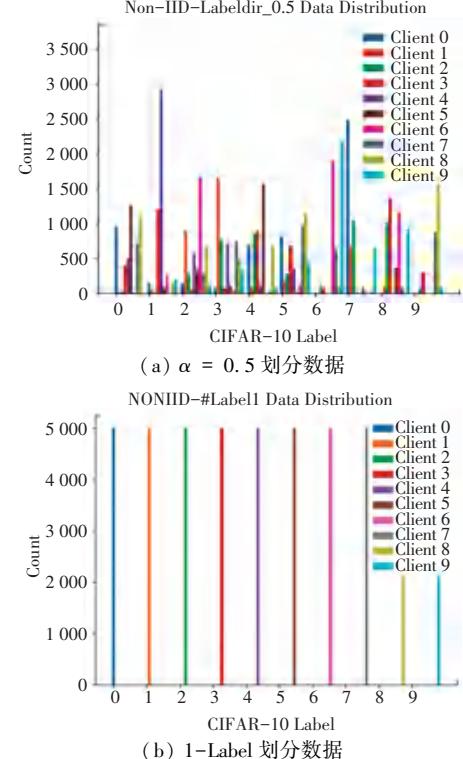


图 4 CIFAR-10 在不同划分方式下数据分布情况

Fig. 4 CIFAR-10 data distribution under different partitioning methods

4.2 数据增强与填充

4.2.1 数据增强

本文用于数据增强的生成对抗模型由联邦学习服务器预先完成训练, 生成模型采用 AC-GAN, 训

练过程中使用 Adam 优化器, 学习率设为 0.000 2。在 MNIST 和 CIFAR-10 数据集上进行训练时, 由于 CIFAR-10 数据集相较于 MNIST 更加复杂, 因此训

练了 1 000 轮才达到较好的效果。实验结果表明, 该模型生成的虚拟样本与真实数据高度相似, 如图 5 所示。

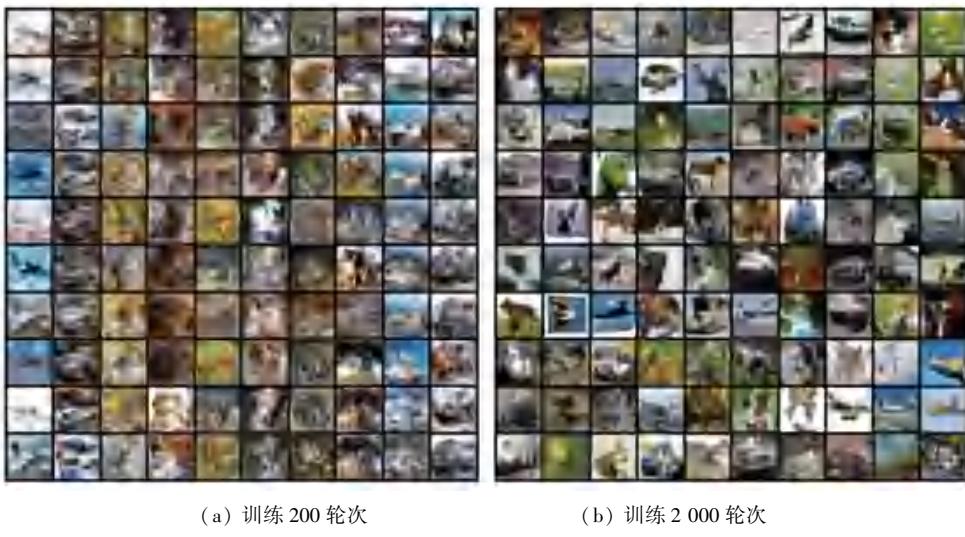


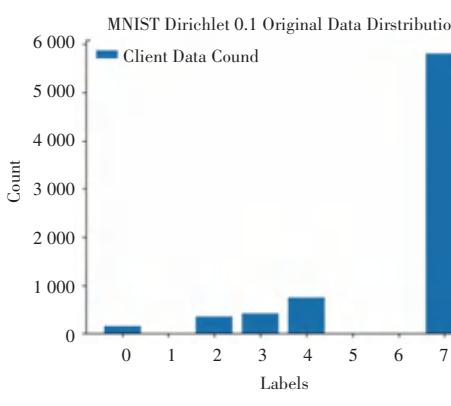
Fig. 5 Virtual samples generated by AC-GAN at different training rounds

4.2.2 数据填充

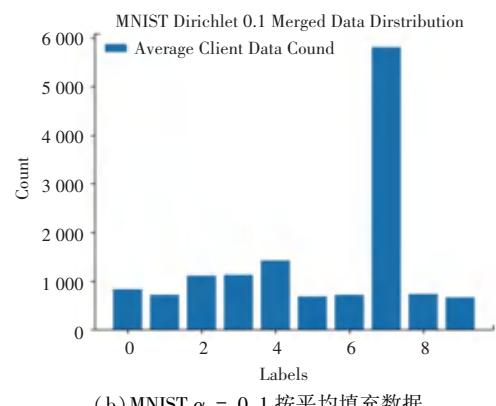
虚假数据生成后, 本文进一步设计了两种数据填充策略, 以缓解客户端间的数据分布不均衡问题。第一种是基于阈值的填充策略: 统计每个客户端的本地标签类别数量 $L_i < \text{padding_thresh}_i$, 并设定阈值比例为 0.1, 计算其对应的填充阈值 padding_thresh_i , 若某客户端的标签类别数量 $L_i < \text{padding_thresh}_i$, 则利用生成的虚拟样本对其进行填充, 否则保持不变。第二种是基于平均值的填充策略: 首先统计所有客户端的本地标签类别数量并

计算所有客户端的平均值 $\text{avg}(L)$ 。对于本地类别数量少于 $\text{avg}(L)$ 的客户端利用生成对抗网络生成的虚拟样本进行填充, 从而缓解不同客户端间数据分布不均衡的问题。

以 MNIST 数据集在 Dirichlet 划分下的结果为例, 在 $\alpha = 0.1$ 时, 原本每个客户端仅包含部分类别数据, 而通过增强生成的虚拟样本填充后, 整体分布更接近独立同分布, 这种改善有效减少了客户端与全局模型在聚合过程中的差异, 有效提升了最终模型的准确率(如图 6 所示)。



(a) MNIST $\alpha = 0.1$ 划分数据



(b) MNIST $\alpha = 0.1$ 按平均填充数据

图 6 MNIST 数据在 Dirichlet 划分下数据填充分布展示

Fig. 6 MNIST data fill distribution display under Dirichlet partitioning

4.3 实验结果与分析

为了验证所提出优化方法的有效性, 本文基于不同的数据划分方式, 对联邦学习在 100 轮通信后

的全局模型精度进行了评估, 并与经典的基线方法 FedAvg、FedProx、FedNova 进行了对比实验。本文引入参数 β 来表示虚拟数据的填充比例: 当 $\beta = 0$, 即

为传统 FedAvg 算法;当 $\beta = 10\%$ 表示在数据划分后利用生成的虚拟数据按 10% 的比例进行填充;当 $\beta = 50\%$ 时,则表示采用平均值填充策略,不同 Non-IID 程度下的正确率比较见表 2。可以看出,在所有划分方式中,本文提出的优化算法均取得了更高的模型准确率。以 CIFAR-10 数据集上的图像分类任务为例,在 $\beta = 10\%$ 设置下,1-Label 由于节点间数据分布差异较大, FedAug 算法的精度较 FedAvg、

FedProx、FedNova 算法提升了 23.22%、23.08%、23.47%;在 $\beta = 50\%$ 设置下,按照 Dir(0.5) 的数据划分, FedAug 算法的精度较 FedAvg、FedProx、FedNova 算法提升了 3.46%、4.61%、2.96%。实验结果表明,无论数据的异质性程度如何变化, FedAug 算法在图像分类任务上的性能都优于 FedAvg 算法。

表 2 不同 Non-IID 程度下的正确率比较

Table 2 Comparison of accuracy rates under different Non-IID degrees

方法	MNIST			CIFAR-10			%
	1-Label	Dir(0.1)	Dir(0.5)	1-Label	Dir(0.1)	Dir(0.5)	
FedAvg	15.57	94.74	98.71	12.66	51.19	54.57	
FedProx	11.26	97.50	98.60	12.80	50.52	53.42	
FedNova	15.48	97.43	98.47	12.41	53.50	55.07	
$\beta = 50$		96.89	98.92				
$\beta = 10$	29.62			35.88			
		96.76	98.78		53.02	58.03	
					53.29	57.89	

4 种联邦学习算法的精度变化趋势如图 7 所示,可见 FedAug 在收敛速度和最终精度上均优于

基线方法,尤其是在高度异质的 1-Label 设置下,性能提升更加显著。

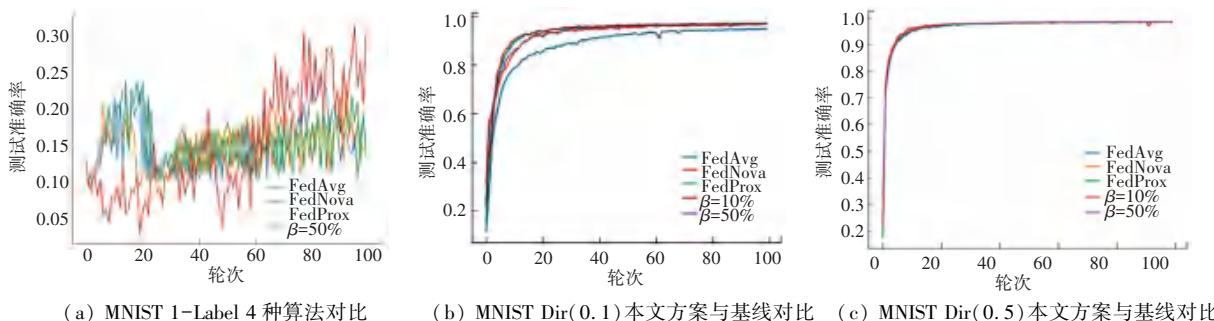
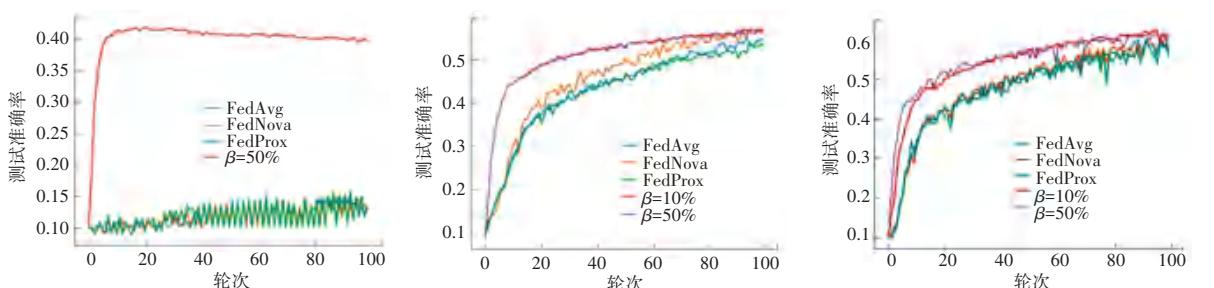


图 7 MNIST 数据集在 3 种数据划分下的精度变化趋势

Fig. 7 Changing trend of the MNIST dataset under three data partitions

CIFAR-10 数据集联邦学习算法的测试精度趋势如图 8 所示,可见与其他基线相比,在不同的数据异质性设置中, FedAug 算法能达到更高的精度和更快的收敛;在分类任务相对复杂的 CIFAR-10 数据

集上,在训练一开始,FedAug 就超过了极限算法,在 3 种数据异质性设置下,分别达到了 35.88%,53.29%,57.89% 的精度,而其他基线算法,最高只能达到 12.66%、51.19%、54.57% 的精度。



(a) CIFAR-10 1-Label 四种算法对比 (b) CIFAR-10 Dir(0.1) 本文方案与基线对比 (c) CIFAR-10 Dir(0.5) 本文方案与基线对比

图 8 CIFAR-10 数据集在 3 种数据划分下的精度变化趋势

Fig. 8 Changing trend of the CIFAR-10 dataset under three data partitions

5 结束语

为了提升联邦学习在 Non-IID 场景下的全局模型精度与收敛效率,本文提出了一种面向非独立同分布数据的联邦学习数据增强方法——FedAug。首先,在服务器端预训练公共生成对抗网络(GAN),并将其下发至各个客户端,客户端利用生成的虚拟样本对本地数据进行增强;在生成对抗网络训练过程中对判别器引入噪声,以有效保护原始数据隐私。通过在 MNIST 和 CIFAR-10 数据集上的对比实验验证,FedAug 在不同的数据划分方式下均显著优于基线算法,不仅提高了全局模型的最终精度,也加快了收敛速度,尤其是在高度异质的 1-Label 场景中,FedAug 的精度较传统方法最高提升了 23.22%,而在 Dir(0.5) 场景下依然保持稳定的优势,充分说明了 FedAug 在复杂数据分布下的鲁棒性和实用性,为解决联邦学习中的数据异质性问题提供了一种有效途径。

参考文献

- [1] ZHOU Z, CHEN X, LI E, et al. Edge intelligence: Paving the last mile of artificial intelligence with edge computing [J]. Proceedings of the IEEE, 2019, 107(8): 1738–1762.
- [2] MCMAHAN B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data[C]// Proceedings of the 20th International Conference on Artificial Intelligence and Statistics. PMLR, 2017: 1273–1282.
- [3] LI L, FAN Y, TSE M, et al. A review of applications in federated learning [J]. Computers and Industrial Engineering, 2020, 149(5): 106854.
- [4] WEI K, LI J, DING M, et al. Federated learning with differential privacy: Algorithms and performance analysis [J]. IEEE Transactions on Information Forensics and Security, 2020, 15: 3454–3469.
- [5] SOLLINS K R. IoT big data security and privacy versus innovation [J]. IEEE Internet of Things Journal, 2019, 6(2): 1628–1635.
- [6] ZHU H, XU J, LIU S, et al. Federated learning on non-IID data: A survey[J]. Neurocomputing, 2021, 465: 371–390.
- [7] TUOR T, WANG S, KO B J, et al. Overcoming noisy and irrelevant data in federated learning[C]//Proceedings of 2020 25th International Conference on Pattern Recognition (ICPR). Piscataway, NJ: IEEE, 2021: 5020–5027.
- [8] TANNER M A, WONG W H. The calculation of posterior distributions by data augmentation [J]. Journal of the American Statistical Association, 1987, 82(398): 528–540.
- [9] WANG J, LIU Q, LIANG H, et al. Tackling the objective inconsistency problem in heterogeneous federated optimization[J]. Advances in Neural Information Processing Systems, 2020, 33: 7611–7623.
- [10] LIN T, KONG L, STICH S U, et al. Ensemble distillation for robust model fusion in federated learning[J]. Advances in Neural Information Processing Systems, 2020, 33: 2351–2363.
- [11] LI T, SAHU A K, ZAHEER M, et al. Federated optimization in heterogeneous networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 429–450.
- [12] KARIMIREDDY S P, KALE S, MOHRI M, et al. Scaffold: Stochastic controlled averaging for federated learning [C]// Proceedings of International Conference on Machine Learning. PMLR, 2020: 5132–5143.
- [13] CHEN M, MAO B, MA T. FedSA: A staleness – aware asynchronous federated learning algorithm with non-IID data[J]. Future Generation Computer Systems, 2021, 120: 1–12.
- [14] RUDER S. An overview of gradient descent optimization algorithms[J]. arXiv preprint arXiv:1609.04747, 2016.
- [15] MOHRI M, SIVEK G, SURESH A T. Agnostic federated learning [C]// Proceedings of International Conference on Machine Learning. PMLR, 2019: 4615–4625.
- [16] LI Q, DIAO Y, CHEN Q, et al. Federated learning on non-iid data silos: An experimental study [C]// Proceedings of 2022 IEEE 38th International Conference on Data Engineering (ICDE). Piscataway, NJ: IEEE, 2022: 965–978.
- [17] CRESWELL A, WHITE T, DUMOULIN V, et al. Generative adversarial networks: An overview[J]. IEEE Signal Processing Magazine, 2018, 35(1): 53–65.
- [18] NGUYEN T, LE T, VU H, et al. Dual discriminator generative adversarial nets[J]. arXiv, 1709.03831, 2017.
- [19] ODENA A, OLAH C, SHLENS J. Conditional image synthesis with auxiliary classifier GANs[C]//Proceedings of International Conference on Machine Learning. PMLR, 2017: 2642–2651.