

基于标签相关性加权嵌入的联邦双阶段注意力神经网络算法研究

钟磊¹, 曾璐琨^{2*}, 姜雪娇¹, 艾渊², 徐佳隆¹, 杨景旭², 江蕾¹

¹中国南方电网海南电网有限责任公司, 海南 海口

²南方电网数字电网集团有限公司, 广东 广州

收稿日期: 2025年8月19日; 录用日期: 2025年9月17日; 发布日期: 2025年9月24日

摘要

在涉及隐私敏感数据的多标签图像识别任务中, 联邦学习(Federated Learning, FL)模型的有效性至关重要, 而在跨区域图像分类中, 存在数据分布不一致、类别相关却存在不平衡性的挑战。值得注意的是, 现有研究在联邦学习框架下针对类别相关性和不平衡问题的系统性解决方案仍显不足。具体而言, 由于不同客户端之间的数据异质性和类别的不平衡性, 全局模型的聚合过程面临参数不一致问题, 即部分本地模型参数与聚合后的全局模型存在显著偏差, 从而影响这些客户端的分类性能。为应对这些挑战, 我们提出了一种基于标签相关性加权嵌入的双阶段联邦图注意力神经网络(Federated Dual-phase Attention Network with weighted label correlation embedding for multi-label image classification, FD-WCAT), FD-WCAT的核心创新体现在融合了标签相关性和类不平衡加权的局部模型构建和全局模型的加权聚合两方面。在局部模型构建中, 每个客户端构建掩码标签相关图来学习标签相关性特征; 然后融合该特征设计了基于类别不平衡加权的多标签分类器。在全局模型聚合时, 为解决训练过程中本地模型与全局模型的参数不一致问题, FD-WCAT设计了基于全局-本地参数正则化的双阶段聚合策略: 首先, 每个客户端计算其类别不平衡系数并将本地模型参数发送至服务器; 在服务器端, 客户端的模型根据参数相似性被聚类为T组以确保每组内的模型相近, 然后通过组内聚合生成T个原型模型。接下来, 基于每组的平均不平衡系数计算原型模型的不平衡权重, 并通过不平衡加权聚合生成最终的全局模型。最终实验验证了FD-WCAT在多标签数据集上优于现有基准模型。

关键词

多标签图像分类, 神经网络, 联邦学习

Federated Dual-Phase Attention Network with Weighted Label Correlation Embedding for Multi-Label Image Classification

Lei Zhong¹, Lukun Zeng^{2*}, Xuejiao Jiang¹, Yuan Ai², Jialong Xu¹, Jingxu Yang², Lei Jiang¹

*通讯作者。

文章引用: 钟磊, 曾璐琨, 姜雪娇, 艾渊, 徐佳隆, 杨景旭, 江蕾. 基于标签相关性加权嵌入的联邦双阶段注意力神经网络算法研究[J]. 计算机科学与应用, 2025, 15(9): 267-282. DOI: 10.12677/csa.2025.159243

¹Hainan Power Grid Co., Ltd, China Southern Power Grid, Haikou Hainan²Digital Grid Research Institute, China Southern Power Grid, Guangzhou GuangdongReceived: Aug. 19th, 2025; accepted: Sep. 17th, 2025; published: Sep. 24th, 2025

Abstract

The effectiveness of federated learning (FL) models is crucial for privacy-sensitive multi-label image recognition tasks, while cross-regional image classification faces challenges including inconsistent data distributions and correlated yet imbalanced categories. Notably, existing research still lacks systematic solutions addressing label correlation and imbalance within the FL framework. Specifically, due to data heterogeneity and class imbalance across clients, the global model aggregation process encounters parameter inconsistency issues, where significant deviations exist between some local model parameters and the aggregated global model, thereby impairing classification performance on these clients. To address these challenges, we propose a Federated Dual-phase Attention Network with weighted label correlation embedding for multi-label image classification (FD-WCAT). FD-WCAT's core innovations manifest in two aspects: (1) local model construction incorporating label correlation and class imbalance weighting, where each client builds a masked label correlation graph to learn label correlation features and integrates this to design a class-imbalance-weighted multi-label classifier; (2) global weighted aggregation employing a dual-phase strategy with global-local parameter regularization. In aggregation: first, each client computes its class imbalance coefficient and transmits local parameters to the server; server-side, client models are clustered into T groups based on parameter similarity to ensure intra-group homogeneity, followed by intra-group aggregation to generate T temporary models. Subsequently, temporary models are assigned imbalance weights based on each group's average imbalance coefficient, and final global model generation occurs via imbalance-weighted aggregation. Experimental results ultimately validate FD-WCAT's superiority over existing baseline models on multi-label datasets.

Keywords

Multi-Label Image Classification, Neural Network, Federated Learning

Copyright © 2025 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

在智能电网安全运行的背景下，多标签图像识别技术发挥着关键作用，其核心挑战在于如何有效建模类别相关性并解决普遍存在的类别不平衡问题，后者往往导致特定类别识别性能下降。值得注意的是，一些电力图像数据具有高度敏感性，其隐私保护需求与识别准确性要求同等重要。联邦学习(Federated Learning, FL)作为一种分布式机器学习范式，通过支持多方数据协同训练而不共享原始数据的方式，既满足了电力数据隐私保护要求，又实现了复杂模式识别能力的提升，为智能电网安全运行提供了兼顾隐私保护与识别性能的技术解决方案。

在联邦学习框架下，针对多标签图像的分类问题，隐私保护优势与分类性能下降之间的矛盾主要源于三个核心挑战：客户端级数据异质性、跨区域类别相关性不一致以及类别不平衡问题[1][2]。这种矛盾

在多标签电力图像的分类场景中表现尤为显著,具体表现为:(1) 区域特性导致的特征分布差异,使各客户端形成异构的标签模式;(2) 跨区域类别相关性漂移现象,即相同标签组合在不同区域呈现不同的共现模式和语义关联强度[3][4];(3) 类别不平衡和类别相关性往往同时存在,某些标签类别的数据在特定区域可能占比过高,而其他标签对应的数据不足。例如,在电力场景中,某些安全违规行为,如“未戴安全帽”,可能发生得更频繁,导致这些标签的样本数量较多。相反,像“吸烟”这样不太常见的标签,样本数量可能少得多,导致这些类别的泛化能力较差。在智能电网系统中,来自不同区域的客户端可能拥有不同类型电力设备或工况的图像数据,而传统的联邦学习框架往往侧重于学习本地数据分布,而没有充分考虑不同区域的数据其类别相关性不同且存在不平衡问题[5]-[7],在此情形下,全局模型的聚合过程可能面临参数的不一致性。具体而言,当某些标签类别的数据在特定区域可能占比过高时,模型会优先拟合到这些类别;且不同区域的数据其标签相关性不同,会导致不同区域的模型在拟合不同标签相关性时产生差异性较大的模型参数,从而导致全局模型可能会优先拟合到不平衡的类别且与部分客户端的模型参数差异较大导致参数更新方向偏离,进而削弱模型在不同客户端之间的泛化能力。

为应对这些挑战,本研究提出基于标签相关性加权嵌入的双阶段联邦图注意力神经网络(Federated Dual-phase Attention Network with Weighted Label Correlation Embedding for Multi-label Image Classification, FD-WCAT)。该框架的创新性主要体现在两个维度:融合了标签相关性和类不平衡加权的局部模型构建和全局模型的加权聚合,从局部模型和全局模型两个方面来解决类相关性和不平衡性导致的模型参数不一致问题。在局部模型层面,FD-WCAT 通过构建动态标签相关图把不同客户端的标签共现模式建模为标签相关性嵌入特征,然后设计带惩罚系数的自适应多标签分类器来缓解类别不平衡问题;为解决本地模型与全局模型之间的参数不一致问题,FD-WCAT 设计了基于全局-本地参数正则化的双阶段聚合策略,该策略首先在本地面向中设计正则化损失函数约束本地模型参数与全局模型的偏差范围,然后在服务器端执行两阶段优化的聚合算法:在每个客户端计算其类别不平衡系数,并将其与本地模型参数一起发送至服务器,服务器首先将本地模型聚类为 T 组,以确保每组内的参数相似性,基于不平衡系数,在每组内进行聚合,生成 T 个不一致性最小的原型模型;然后根据每组的平均不平衡系数计算这 T 个原型模型的平均不平衡权重。最后通过不平衡加权聚合形成全局模型。实验验证表明,在多标签基准数据集上,FD-WCAT 较基准方法在 F1 score、AP 等指标上均有提升,验证了该框架的有效性。本文的主要贡献如下:

1) 提出了一种面向多标签电力图像分类的双阶段联邦学习框架 FD-WCAT。该框架通过构建标签相关图并使用加权相关性注意力网络来捕获标签相关性嵌入,并引入惩罚系数和正则化项以缓解不同区域的类不平衡问题。

2) FD-WCAT 通过设计全局-本地参数正则化损失函数和双阶段聚合策略,解决本地模型和全局模型之间的参数不一致问题,旨在提高模型在来自不同客户端的异质数据上的泛化能力。

本文的其余部分组织如下:第2节回顾相关工作,包括联邦学习的学习方法和多标签学习。第3节介绍所提出的 FD-WCAT 模型。第4节详细描述实验设置,并通过实验验证 FD-WCAT 的有效性。最后,第5节给出结论。

2. 相关工作

2.1. 基本联邦学习模型

联邦学习算法,如广泛使用的 FedAvg [1],通常遵循标准流程:在客户端设备上本地训练,然后将训练好的模型上传到中央服务器进行聚合,随后将聚合后的全局模型广播回客户端。然而,当存在数据异质性(例如客户端数据分布差异)时,使本地模型与全局模型对齐的任务变得极具挑战性。在数据异质

性带来的各种挑战中, 标签分布偏移在联邦学习中是最普遍且影响最大的问题之一。例如, 在跨医院的医学图像分类中, 一家医院可能专注于罕见疾病, 而其他医院则处理更常见的疾病。这种不平衡导致客户端之间的标签分布存在显著差异。为减轻标签分布偏移, 已经开发了几种联邦学习技术, 以更好地使本地客户端的偏差与全局模型对齐。例如, FedProx [8] 引入了一个近端项来限制梯度更新, 有助于在存在标签分布偏移的情况下实现收敛。类似地, SCAFFOLD [9] 对每个客户端采用基于梯度差异的控制变量, 帮助纠正本地偏差。最近, FedDDC [10] 利用期望最大化(EM)算法, 通过学习辅助的本地偏差变量来跟踪和解决本地模型中的差异, 进一步增强了与全局模型的对齐。

2.2. 多标签联邦学习和图神经网络模型

在多标签学习的背景下, 已经提出了几种有效的策略来解决标签分布偏移和域偏移问题。这些方法包括结合类属特征的多标签学习方法[4][11]、扭曲多标签学习[12]和多维多标签分类[13]。通过增强建模标签依赖性的能力, 这些技术为联邦学习环境中与标签分布偏移和域偏移相关的挑战提供了稳健的解决方案。

将图神经网络(Graph Neural Networks, GNNs)和注意力机制集成到联邦学习框架中, 已成为一种强大的范式, 用于在保护数据隐私的同时进行分布式图分析[14]-[17]。GNNs 通过节点之间的消息传递擅长捕获拓扑依赖性。在联邦学习环境中, 这些技术解决了三个关键挑战: (1) 通过自适应图融合技术(如带掩码的标签相关图)处理跨客户端图异质性; (2) 通过基于聚类的原型聚合和不平衡感知加权处理非独立同分布数据分布; (3) 使用独立子图的图间联邦学习或分区全局图的水平图内联邦学习来平衡隐私与效用。值得注意的应用包括联邦推荐系统[16]、医疗分析[17]和智能电网管理。此外, 多种联邦学习方法已得到广泛应用。FedMLP [18] 是一种两阶段方法, 从伪标签标记和全局知识学习两个方面来解决类别缺失问题。FedRSC [19] 是一种联邦学习分析方法, 它将边缘计算和云技术相结合, 通过多标签路面分类分析识别各种路况。LSFT [20] 是一种特定于标签的方法: 基于标签特定特征的多标签联邦 Transformer (LSFT), 其中在客户端为每个类别开发了一个独特的分类模型。FLAG [21] 是一种新的多标签联邦学习框架, 具有基于聚类的多标签数据分配和一种新颖的聚合方法——快速标签自适应聚合(FLAG), 用于联邦学习环境中的多标签分类。

基于提示的联邦学习(FL)已成为一种很有前景的方法, 它将提示工程(常用于大型语言模型)的力量与联邦学习的去中心化性质相结合。基于提示的联邦学习的核心思想是利用提示作为在本地客户端微调模型的手段。这在隐私敏感应用中特别有利, 因为在这些应用中数据不能被集中。最近的研究表明, 将基于提示的技术集成到联邦学习中可以提高模型的准确性和隐私性。例如, 在一项研究中, 作者探索了基于提示的学习如何在联邦环境中帮助更有效地处理自然语言处理任务, 同时不损害用户隐私[22]。通过设计有效的提示, 本地模型能够以保护隐私的方式利用预训练的大型语言模型(LLMs)。联邦学习背景下的半监督学习(SSL)尤为重要, 因为许多联邦环境涉及有限的标记数据。其思想是结合半监督学习(可以使用标记和未标记数据)和联邦学习的去中心化性质的优势。在联邦环境中, 客户端通常只有一小部分标记数据和大量未标记数据。半监督联邦学习中的一个关键挑战是确保模型从无标记数据中获益, 同时不违反隐私约束。已经提出了各种技术, 例如使用伪标记, 其中每个客户端为其未标记数据生成伪标记, 并将这些标记与全局模型共享以进行进一步训练[23]。视觉-语言模型(VLMs)结合了视觉和文本信息, 在视觉问答(VQA)和多模态情感分析等任务中取得了显著成功。将这些模型与联邦学习集成的挑战在于处理图像和文本数据的复杂性, 同时保持隐私并减少通信开销。最近的研究已经开始探索如何在联邦学习设置中利用 VLMs [24]。一种值得注意的方法涉及使用多模态表示, 其中本地客户端训练整合视觉和文本特征的模型。这种方法已被证明在个性化内容推荐系统和视障人士辅助技术等应用中是有益的。

3. 提出的 FD-WCAT 模型

3.1. 联邦学习问题描述

联邦学习是一种去中心化的机器学习方法，多个客户端在不共享原始数据的情况下协作训练一个共享模型。在本研究中，每个通信轮次有 K 个客户端参与，每个客户端都有自己的私有数据集，表示为 $D = \{D_1, D_2, \dots, D_K\}$ 。每个客户端的任务是多标签分类问题，其中数据点 x 被标记为 $y = [y_1, y_2, \dots, y_C]$ ， $y_i = 1$ 表示存在第 i 类， $y_i = 0$ 表示不存在第 i 类。本地客户端负责预测数据中每个类别的存在与否。联邦学习过程的目标是生成一个全局聚合模型来处理多标签分类，从而实现该任务的核心目标。

$$W = \arg \min_W \sum_{i=1}^{M^i} \eta_i L_i(W) \quad (1)$$

其中， L_i 是第 i 个本地模型的损失， η_i 是客户端 i 的加权系数。在多标签联邦学习中，客户端和服务端模型都在相同的标签空间内，该空间总共包含 C 个类别。然而，这些标签的分布在不同客户端之间可能存在显著差异。为应对这一挑战，我们的方法借鉴了最近的 FedLGT [7]，提出了 FD-WCAT 模型。FD-WCAT 的模型结构如图 1 所示：

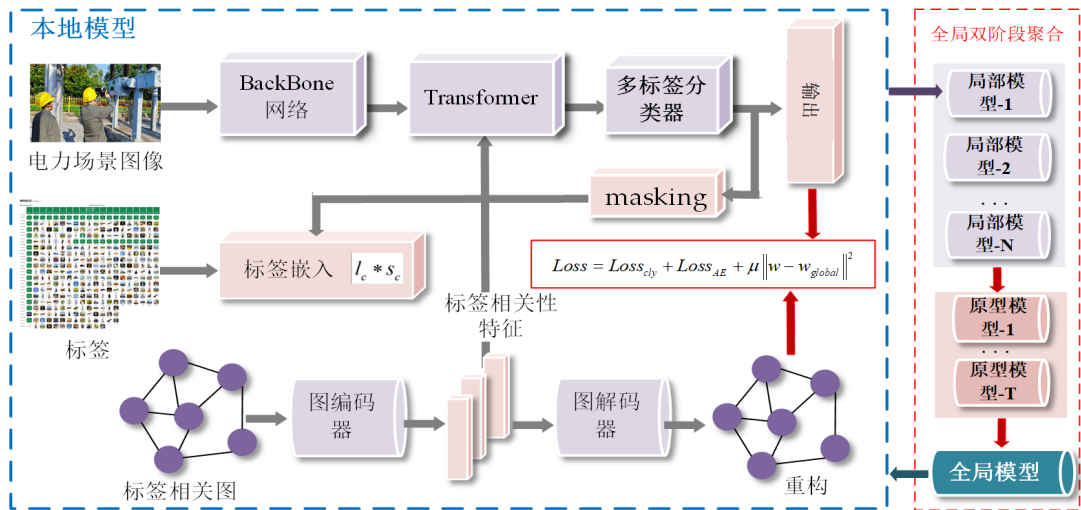


Figure 1. The structure of the proposed FD-WCAT

图 1. FD-WCAT 模型结构图

3.2. FD-WCAT 在客户端学习标签相关性和不平衡性

在客户端设备上，FD-WCAT 由 4 个部分组成：

1) 标签嵌入模块

FD-WCAT 通过构建标签相关图和特定的损失函数来解决标签相关性和类别不平衡问题。对于标签相关图，图的节点特征来自标签嵌入，该嵌入过程基于 FedLGT 的双嵌入框架，该框架结合了标签嵌入 (L) 和状态嵌入 (S)，其中，状态嵌入 $S = \{s_1, s_2, \dots, s_C\}$ (其中每个 $s_C \in R^d$) 作为状态值，通过离散值编码标签的存在/不存在：1 (正)、0 (负) 和 -1 (未知)。值得注意的是，只有未知状态嵌入对训练损失有贡献，突出了它们在模型训练中的作用。在第 t 次训练迭代中，客户端使用全局模型 w_t 生成预测向量 $P = \{p_1, p_2, \dots, p_C\}$ 。当类别 c 的预测置信度 p_c 低于确定性阈值时，目标类别及其相关标签 $P = \{p_1, p_2, \dots, p_C\}$ 都被视为可学习单元——它们对应的状态嵌入 s_c 被动态更新为“未知”(-1)，形成实例级状态掩码。在所有其他情

况下, 状态嵌入保持其初始值(在本地模型设置期间随机设置为 0 或 1)。基于这些预测, 对状态嵌入进行校准以生成 $S' = \{s'_1, s'_2, \dots, s'_c\}$, 其中每个 s'_c 定义为:

$$s'_c = \begin{cases} -1, & \tau - \varepsilon < p_c < \tau + \varepsilon \\ s_c, & \text{otherwise} \end{cases}, \quad (2)$$

其中 τ 表示阈值(通常设置为 0.5)。

我们的框架采用双嵌入系统, 其中标签通过可训练的标签嵌入 $L = \{l_1, l_2, \dots, l_c\}$ (每个 $l_c \in \mathbb{R}^d$, d 表示嵌入维度)来表示。实现分为两个阶段: 首先, 将标签语义转换为文本表示以生成具有判别性的嵌入; 其次, 通过将这些文本衍生特征与状态嵌入相结合来构建带掩码的标签嵌入图。为实现稳定的特征, 我们利用 CLIP 的冻结文本编码器[25]生成稳定的标签嵌入。这些 CLIP 嵌入在训练过程中作为监督信号, 利用其通过相似性模式编码标签间语义关系的固有能力, 保留来自 CLIP 多模态预训练的相关知识。最后, 状态嵌入 S' 与标签嵌入相加组合, 产生最终的带掩码表示 $\tilde{l}_c$ 。在公式(3)中, $*$ 表示元素乘法。

$$\tilde{l}_c = l_c * s_c \quad (3)$$

2) 构建标签相关图

带掩码的标签嵌入可以表示为 $\tilde{L} = \{\tilde{l}_1, \dots, \tilde{l}_c\}$ 。为建模这些相关关系, FD-WCAT 构建了一个带掩码的标签相关图, 每条边对应其相关节点之间的相似性, 可计算为:

$$E_{i,j} = \text{Norm}\left(\text{Dot}(\tilde{l}_i, \tilde{l}_j)\right), \quad (4)$$

其中 $\text{Dot}(i, j)$ 是节点 i 和 j 之间的点积相似性。我们使用点积相似性并进行归一化作为标签相关图的边。因此, 我们可以在标签相关图上定义一个注意力模型, 以学习标签的语义特征和标签的不平衡特性。当一个不平衡的类别被误分类时, 其相关标签以图的形式构建, 从而通过利用与该类别相关的标签来增强模型对不平衡类别的分类能力。

3) 基于图学习标签相关性和标签不平衡

为解决标签相关图中的标签相关性和类别不平衡问题, 我们提出了一种图注意力自编码器架构, 其中 FD-WCAT 将带掩码的标签嵌入图作为输入进行处理。该模型采用分层多层注意力机制, 通过三个集成组件系统地提取相关特征: (1) 定量评估带掩码标签嵌入特征之间关系的相似性函数(公式 5); (2) 从这些相似性测量中动态计算的注意力系数; (3) 聚合邻域信息的合成注意力特征。相似性函数计算如公式 (5):

$$\text{Similarity}_{ij} = \text{NN}\left(A \times W \times \tilde{l}_i, A \times W \times \tilde{l}_j\right), \quad (5)$$

其中 A 是带掩码标签嵌入图的邻接矩阵, W 是可学习权重, $\text{NN}()$ 是单层神经网络。基于此函数, 注意力系数定义如公式(6):

$$\alpha_{ij} = \text{softmax}\left(\text{Similarity}_{ij}\right) = \frac{\exp\left(\text{Similarity}_{ij}\right)}{\sum_j \exp\left(\text{Similarity}_{ij}\right)}, \quad (6)$$

其中 α_{ij} 是 f_i 和 f_j 的注意力系数。因此, 当前节点 i 的注意力特征可以表示为公式(7):

$$\text{feature}_i = \sum_j \alpha_{ij} \times V \times \tilde{l}_j, \quad (7)$$

其中 V 是可学习权重。使用公式(5)~(7), 图注意力自编码器将标签相关性编码为嵌入特征。为进行训练, 图注意力自编码器包含基于带掩码标签嵌入图的重建损失。其损失函数表示为公式(8):

$$Loss_{AE} = \frac{1}{N} \sum_n \{(X_n - Recon_n)^2\} \quad (8)$$

4) 具有标签不平衡的分类器

编码器提取的特征随后被转换为向量, 并与提取的数据特征集成为 F , 作为 Transformer 网络的输入。Transformer 将这些特征与标签嵌入一起通过多层感知器(MLP)进行分类处理。因此, 有

$$Output = MLP(x, F), \quad (9)$$

其中 x 表示图输入数据中的节点特征, F 是编码了相关性和不平衡性的标签嵌入特征, $Output$ 是预测的对数几率(MLP 表示 Transformer 和输出层)。分类损失使用交叉熵损失表示为:

$$Loss_{cly} = \sum_c W_c \times CorssEntropy(Output_c, Label_c), \quad W_c = \frac{N_{total}}{C \times N_c}. \quad (10)$$

其中, N_{total} 是样本数量, N_c 是 c 类样本的数量。因此, 客户端中模型的总损失可以表示为:

$$Loss = Loss_{cly} + Loss_{AE} + \mu \|w - w_{global}\|^2. \quad (11)$$

在公式(11)中, μ 是权重, 设置为 0.015。为减轻联邦学习系统中客户端本地模型与全局模型之间的显著差异可能导致的性能下降, 我们实现了客户端特定的参数一致性正则化项(公式(11)的最后一个组件)。该机制保持对齐, 其中 θ 表示本地模型 p 的参数, 而 w_{global} 对应于聚合阶段全局模型 p 的参数, 在训练开始时 w_{global} 初始化为 θ , 以确保初始参数一致性。

3.3. 聚类驱动的双阶段全局聚合

联邦学习的主要目标是开发一个擅长多标签分类的全局聚合模型。然而, 不同客户端之间不同的标签相关性和明显的类别不平衡导致本地训练的模型对这些特征的优先级不同。为应对这一挑战, 我们设计了一种基于聚类的双阶段聚合方法。在第一阶段, 每个客户端计算其类别不平衡系数, 并将本地模型参数传输到服务器。然后, 基于公式(12), 根据重要的模型参数将这些 N 个客户端模型聚类为 T 组, 确保聚合前每个聚类内的参数相似性。

$$N_k = \arg_K \min_{Para_k} (Similarity(Para_k)) \quad (12)$$

鉴于联邦学习系统中模型参数的高维度和高数据量, 我们采用选择性聚类方法, 专注于参数的代表性子集, 而不是利用每个本地模型的整个参数集。具体来说, 我们使用分类器的参数进行相似性计算。然后, 我们在每个聚类内聚合模型, 并基于每组的平均不平衡系数生成 T 个不同的全局模型。

$$w_k = \sum_g \alpha_g w_g, \quad \alpha_g = Norm(Similarity(w_g, w_{global})) \quad (13)$$

在第二阶段, 通过不平衡加权聚合形成全局模型, 通过结合使用不平衡加权和全局 - 本地参数正则化来解决类别不平衡问题。

$$w = \sum_k \alpha_k w_k, \quad \alpha_k = Norm(Similarity(w_k, w_{global})) \quad (14)$$

3.4. 复杂性分析

这项工作的计算成本主要包括三个部分: 客户端模型训练、客户端 - 服务器通信和服务器端模型聚合。在客户端模型训练阶段, 骨干架构结合了 ResNet-18 和基于 Transformer 的分类器(源自参考文献[17]), 而关键创新在于构建标签相关学习模块和解决类别不平衡问题。标签相关模块包括词向量构建和标签相

关图学习,后者计算密集,但通过利用预训练的 CLIP 模型和轻量级自编码器(仅 3 个编码器/解码器层)得以缓解,确保与骨干相比额外开销最小。由于传输不平衡系数,通信成本略有增加,但总体上可以忽略不计。服务器端聚合采用两步法:首先将 N 个客户端模型聚合为 T 个原型模型,然后将这些原型模型合并为全局模型,计算影响可忽略不计。总之,所提出的方法不会引入过多的计算负担。

接下来,我们基于实验分析所提出的 FD-WCAT 的有效性。

4. 实验

为评估我们提出的 FD-WCAT 框架的性能,我们采用综合测试方法,使用 3 个不同的数据集:两个广泛认可的公共基准(FLAIR [26]和 MS COCO [27])用于比较分析和泛化验证,以及一个专门的电力场景分类数据集——变电站缺陷检测数据集。我们的实验过程系统地分为三个关键阶段:初始数据集表征,然后是 FD-WCAT 框架与现有方法的比较性能基准测试,最后是有针对性的消融研究,定量评估标签相关嵌入图和自适应聚合策略对模型整体有效性的个体贡献。

4.1. 数据集和参数设置

1) FLAIR 作为多标签联邦学习的综合大规模基准,包含丰富多样的来自 Flickr 用户贡献的图像语料库。该数据集具有标准化的 256×256 像素分辨率,并基于真实用户数据进行结构化分区,通过纳入关键的非独立同分布特征,真实地模拟了现实世界的的数据异质性:明显的数量偏差(反映每个用户的样本贡献不均)、显著的标签分布偏差和明显的域变化。该数据集的层次结构提供了双重分类路径——包含 17 个广泛类别的粗粒度分类,以及包含 1628 个特定标签的细粒度系统——我们当前的分析仅集中于前者用于分类目的。作为一个公开可用的资源,FLAIR 为应对联邦学习场景带来的独特挑战提供了宝贵的测试平台。

2) MS COCO 数据集是一个大规模图像数据集。它包含超过 330,000 张图像,其中超过 200,000 张带有标签,总共有 80 个对象类别。该数据集包括各种具有复杂场景的图像,提供丰富的上下文信息。图像的分辨率各不相同,但通常在 640×480 到 1280×720 像素之间。该数据集还包括超过 150 万个对象实例,每张图像包含多个对象,并带有边界框、分割掩码和对象标签的注释。由于其多样化和具有挑战性的图像集,它被广泛用于计算机视觉模型的训练和基准测试(COCO-Common Objects in Context)。

3) 变电站缺陷检测数据集包含 8307 张变电站设备图像,这些图像是在各种现实世界运行条件下捕获的。每张图像都精心标注了 17 个不同的缺陷标签,包括三个主要类别:组件级故障(例如,表盘模糊、绝缘子裂纹、盖板损坏)、安全违规(例如,缺少头盔或工作服等防护装备、未经授权吸烟)和系统级异常(例如,异常运行条件、开关设备故障)。这种全面的注释方案确保了变电站维护中遇到的关键缺陷的广泛覆盖,使该数据集成为自动化缺陷检测和安全合规监控的宝贵资源。该数据集及其详细信息可在(https://pan.baidu.com/s/1qCIGlCi54AwY0b_qX9sG2A?pwd=cuth)获取。

为与联邦学习框架保持兼容,同时保护数据隐私,我们实施人工非独立同分布分区,将数据集分配到客户端特定的子集,真实模拟现实世界的的数据分布,遵循参考文献[7]中描述的方法进行标准化比较。性能评估采用一套全面的指标——Macro 和 Micro 精确率、召回率和 F1 分数——全面评估模型在多标签分类任务中的有效性。视觉任务统一利用 ResNet-18 作为骨干架构,而通用标签嵌入是通过 CLIP 的文本编码器[25]处理提示文本得出的。

遵循参考文献[7]的配置,我们为掩码机制设置参数阈值 $\tau = 0.5$,为不确定性边际设置 $\varepsilon = 0.015$ 。每个 FL 轮次的本地训练执行 10 个 Adam 优化器 epoch(学习率 $5e-4$, batch 的大小 16),组特定的一致性正则化参数 μ 设置为 0.015。联邦学习协议在每个服务器端聚类迭代中配置 50 个通信轮次(T),每轮有 50 个

活跃参与的客户端，以确保有代表性的数据采样。初始聚类采用 k 均值，分为 5 个类簇。所有实现都使用 PyTorch，配备双 NVIDIA RTX 4090 GPU 以加速训练。

4.2. 采用的基准模型

为评估所提出的 FD-WCAT 的有效性，我们将其与联邦学习领域中几个广泛认可的模型进行比较。这些模型包括卷积深度学习架构和基于 Transformer 的模型[1]。我们还将评估 MOON 框架，该框架通过在每个客户端的本地目标函数中纳入近端项来解决数据异质性，可能提高联邦学习场景中的收敛性。此外，我们将 FD-WCAT 与联邦学习中多标签分类的最先进(SOTA)模型进行比较。模型细节如下：

1) AvgFL-ConvMixer [2]：一种 AvgFL 模型，在客户端使用 ConvMixer 进行训练。ConvMixer 架构在 MLP-Mixer 的基础上，包括通道和令牌混合机制，用于处理通道和空间特征。

2) FL-C_Tran [5]：一种在客户端使用 C_Tran 进行训练的 FL 模型。

3) MOON-PoolFormer [2]：一种在客户端使用 PoolFormer 进行训练的 MOON FL 模型。

4) FedLGT [7]：FedLGT 作为一种定制模型更新技术，同时利用每个客户端的标签相关性。

5) FedMLP [18]：一种两阶段方法 FedMLP，从伪标签标记和全局知识学习两个方面解决类别缺失问题。

6) FedRSC [19]：一种联邦学习分析方法，结合边缘计算和云技术，通过多标签路面分类分析识别各种路况。

7) LSFT [20]：一种基于类属特征的方法：基于标签类属特征的多标签联邦 Transformer (LSFT)，其中在客户端为每个类别开发了一个独特的分类模型。

8) FLAG [21]：一种新的多标签联邦学习框架，具有基于聚类多标签数据分配(CMDA)和一种新颖的聚合方法——快速标签自适应聚合(FLAG)，用于联邦学习环境中的多标签分类。

4.3. 在 FLAIR 数据集上的实验比较

首先，我们在 FLAIR 数据集上测试模型性能。结果如表 1 所示。

Table 1. Results of the FLAIR dataset
表 1. FLAIR 数据集上的测试结果

	Macro-P	Macro-R	Macro-F1	Micro-P	Micro-R	Micro-F1
AvgFL-ConvMixer	46.95	31.61	37.31	80.04	59.09	67.98
MOON-PoolFormer	47.56	35.04	40.03	81.07	60.19	69.08
FL-C_Tran	49.45	38.22	43.02	82.72	71.51	76.71
FedLGT	67.91	46.34	54.94	88.71	83.89	86.23
FedMLP	66.77	46.02	54.43	87.89	82.39	85.05
FedRSC	67.94	46.44	55.10	88.74	82.73	85.63
LSFT	68.03	46.39	55.09	87.92	82.47	85.11
FLAG	68.44	46.82	55.61	89.13	83.36	86.14
FD-WCAT	69.17	47.14	56.03	90.97	83.99	87.34

如表 1 所示，FD-WCAT 在所有评估指标上均表现出一致的性能优势，取得了最先进的结果，Macro F1 为 56.03% (优于 FLAG)，Micro F1 为 87.34%，同时在 Macro 和 Micro 水平上保持了精确率 - 召回率的平衡改进，表明与 FedRSC 和 FedLGT 等现有方法相比，其在处理类别不平衡和实例分类方面的能力增强，而现有方法存在精确率 - 召回率权衡问题。每个类别的 F1 分数详情如图 2 所示。

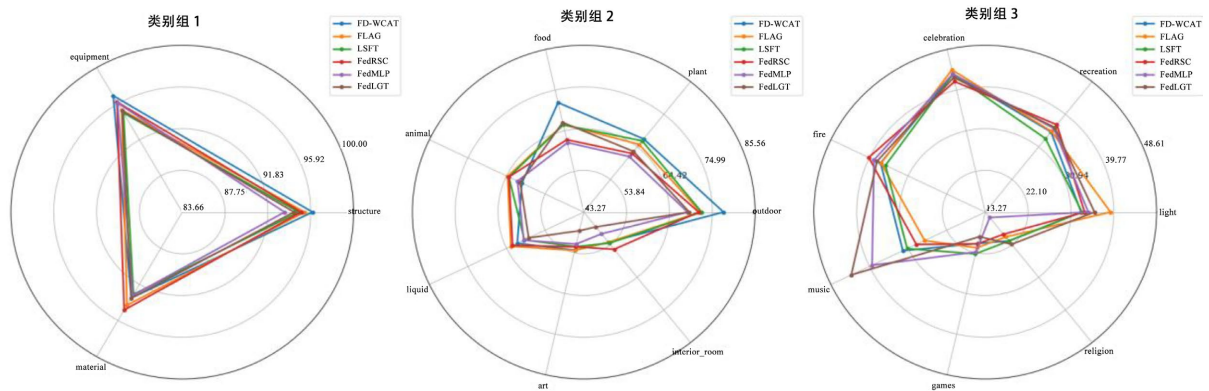


Figure 2. F1 scores of each class on FLAIR dataset
图 2. FLAIR 数据集中每个类别的 F1 分数

接下来，我们展示所提出的 FD-WCAT 与比较模型的测试结果，重点关注平均精度(AP)。各模型在数据集上的测试 AP 如图 3 所示。

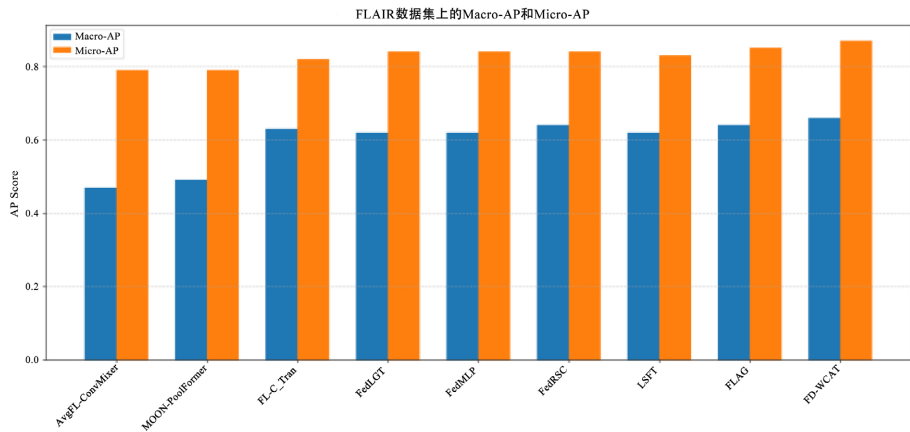


Figure 3. The performance of FD-WCAT compared with the other models on AP in FLAIR dataset
图 3. FD-WCAT 与其他模型在 FLAIR 数据集上的 AP 性能比较

如图所示，FD-WCAT 在 Macro AP 和 Micro AP 中均表现出卓越性能，得分最高，优于所有比较方法，包括次优的 FLAG，这表明通过在联邦学习框架中更有效的特征表示和优化，其在处理类别级和实例级预测方面的能力增强。

4.4. 在 MS-COCO 数据集上的实验比较

在接下来的实验中，我们在 FL-MS-COCO 数据集上验证所提出的 FD-WCAT 的有效性。结果如表 2 所示。

Table 2. Comparison results of FD-WCAT with commonly used models on the FL-MS-COCO dataset
表 2. FL-MS-COCO 数据集上 FD-WCAT 与常用模型的比较结果

	Macro-P	Macro-R	Macro-F1	Micro-P	Micro-R	Micro-F1
AvgFL-ConvMixer	70.63	62.62	66.41	75.49	68.18	71.64
MOON-PoolFormer	74.48	66.47	70.22	77.71	70.96	74.17

续表

FL-C_Train	76.26	67.11	71.45	79.45	71.51	75.27
FedLGT	77.23	70.04	73.43	80.21	73.79	76.86
FedMLP	76.39	70.36	73.22	78.23	73.66	75.87
FedRSC	77.36	70.61	73.79	78.98	73.78	76.29
LSFT	76.47	69.77	72.95	78.11	72.49	75.19
FLAG	77.33	70.47	73.74	78.95	73.89	76.33
FD-WCAT	78.18	71.35	74.56	79.25	75.8	77.49

如表 2 所示, FD-WCAT 在 Macro 和 Micro 指标上均优于所有基线方法, 通过精确率 - 召回率的平衡改进实现了最高的 Macro F1 和 Micro F1, 与 FedRSC 和 FedLGT 等联邦学习基准相比, 在处理类别不平衡场景的同时保持了强大的实例级分类准确性, 表现出更优越的能力。

接下来, 我们展示所提出的 FD-WCAT 与比较模型的测试结果, 特别关注平均精度(AP)。各模型在 FL-MS-COCO 数据集上的测试 AP 如图 4 所示。

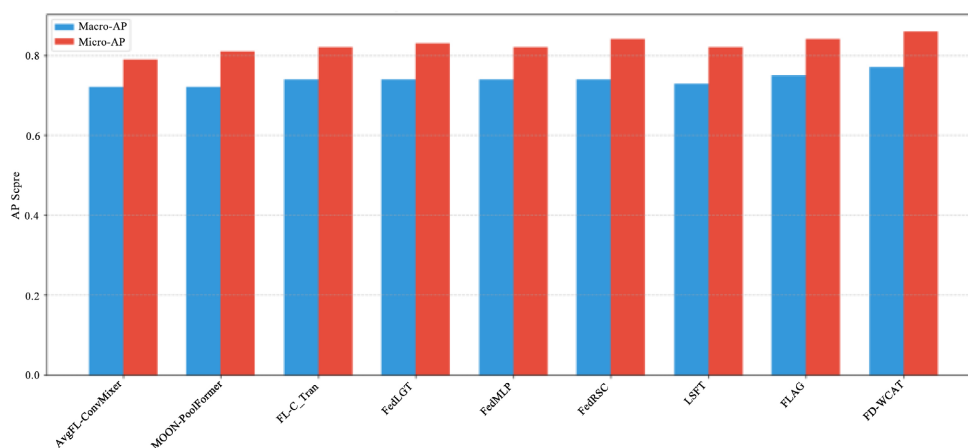


Figure 4. The performance of FD-WCAT compared with the other models on AP of FL-MS COCO dataset

图 4. FD-WCAT 与其他模型在 FL-MS COCO 数据集上的 AP 性能比较

如图 4 所示, 所提出的算法 FD-WCAT 在通过联邦学习处理 MS-COCO 数据集时表现出优异的性能, 在所有关键指标上始终优于其他算法。FD-WCAT 具有最高的每类平均精度, 与其他模型相比, 其预测具有更高的准确性和稳健性。

4.5. 在变电站缺陷检测数据集上的实验比较

在接下来的实验中, 我们在变电站缺陷检测数据集上验证所提出的 FD-WCAT 的有效性。结果如表 3 所示。

Table 3. Comparison results of FD-WCAT with commonly used models on the substation defect detection dataset
表 3. 变电站缺陷检测数据集上 FD-WCAT 与常用模型的比较结果

	Macro-P	Macro-R	Macro-F1	Micro-P	Micro-R	Micro-F1
AvgFL-ConvMixer	45.62	42.26	43.18	70.46	68.41	69.41
MOON-PoolFormer	46.43	43.14	44.21	71.79	70.93	71.36
FL-C_Train	51.22	49.10	50.02	73.42	71.54	72.47
FedLGT	62.23	60.0	60.98	78.12	77.72	77.92

续表

FedMLP	61.22	60.12	60.51	77.67	76.22	76.94
FedRSC	62.39	60.95	61.35	78.36	77.03	77.69
LSFT	61.95	59.94	60.88	77.74	76.21	76.97
FLAG	62.98	60.57	61.77	78.66	77.69	78.18
FD-WCAT	64.77	61.54	63.01	79.79	78.27	79.02

如表 3 所示, FD-WCAT 在所有指标上均表现出全面的优势, 通过精确率 - 召回率的平衡改进实现了最高的 F1 分数, 表明与现有联邦学习方法相比, 其在类别级泛化和实例级预测方面的能力增强, 特别是在处理不平衡数据分布方面表现出色, 而 FedRSC 和 FedLGT 等基线方法存在召回率 - 精确率权衡不佳的问题。每个类别的 F1 分数详情如图 5 所示。

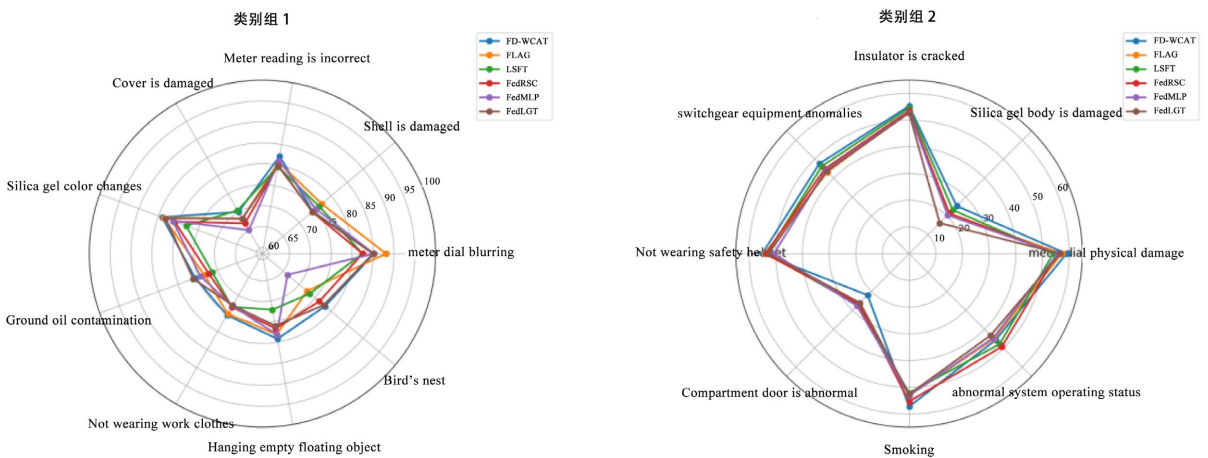


Figure 5. F1 scores for each class in substation defect detection dataset

图 5. 变电站缺陷检测数据集每个类别的 F1 分数

接下来, 我们展示所提出的 FD-WCAT 的测试结果, 并将其性能与其他模型进行比较, 特别关注平均精度(AP)。各模型在变电站缺陷检测数据集上的测试 AP 值如图 6 所示。

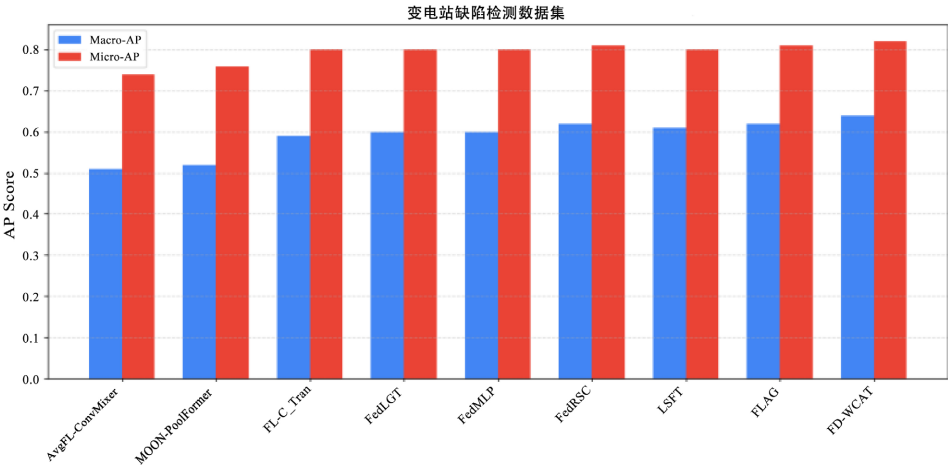


Figure 6. The performance of FD-WCAT compared with the other models on AP of substation defect detection dataset

图 6. FD-WCAT 与其他模型在变电站缺陷检测数据集上的 AP 性能比较

如图 6 所示, 所提出的 FD-WCAT 算法在联邦学习环境下的变电站缺陷检测数据集上取得了最佳的 AP, 始终优于其他算法。此外, 我们可以看到 FD-WCAT 优于其他模型。这表明将标签嵌入信息纳入模型可以有效提高联邦学习模型在变电站缺陷检测数据集上的性能。

4.6. 消融实验

在本节中, 我们在变电站缺陷检测数据集上进行消融研究, 以突出标签嵌入和 FD-WCAT 聚合策略的贡献。结果如表 4 和图 7 所示。

Table 4. Results of the substation defect detection dataset

表 4. 变电站缺陷检测数据集上的消融实验

	Macro-P	Macro-R	Macro-F1	Micro-P	Micro-R	Micro-F1
FD-WCAT_without_graph	63.03	60.76	61.11	78.33	77.83	78.07
FD-WCAT_without_imb	63.59	61.04	62.97	78.88	77.99	78.43
FedLGT	62.23	60.0	60.98	78.12	77.72	77.92
FD-WCAT	64.77	61.54	63.01	79.79	78.27	79.02

在表 4 中, FD-WCAT_without_graph 是没有图结构学习的 FD-WCAT, FD-WCAT_without_imb 是在客户端训练中没有使用类别不平衡损失的 FD-WCAT。在消融研究中, FD-WCAT_without_graph 的结果最差, 这意味着图结构学习对于学习标签相关性很重要。此外, 如表 3 所示, FD-WCAT_without_imb 的性能优于 FD-WCAT_without_graph 和 FedLGT, 但 FD-WCAT 的性能优于 FD-WCAT_without_imb, 这意味着类别不平衡损失对于提高分类结果是有效的。接下来, 我们展示 AP 分数的结果。

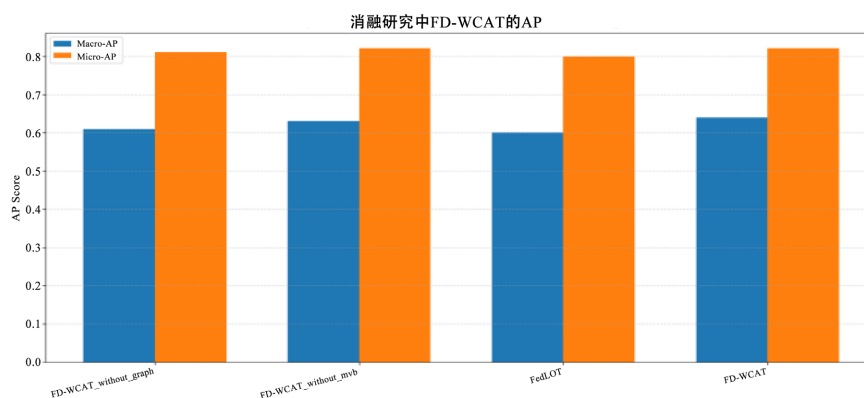


Figure 7. The performance of FD-WCAT in ablation studies

图 7. 消融研究中 FD-WCAT 的性能比较

如表 4 和图 7 所示, FD-WCAT 取得了最佳的测试结果, 这些结果验证了所设计的标签嵌入的有效性。接下来, 我们展示 FD-WCAT 聚合策略的有效性。结果如表 5 和图 8 所示。

Table 5. Results of the substation defect detection dataset

表 5. 变电站控制柜状态监测数据集上的结果

	Macro-P	Macro-R	Macro-F1	Micro-P	Micro-R	Micro-F1
FedLGT	62.23	60.0	60.98	78.12	77.72	77.92
FD-WCAT_without	63.16	60.24	61.55	78.77	77.79	78.27
FD-WCAT	64.77	61.54	63.01	79.79	78.27	79.02

在表 5 中, FD-WCAT_without 是没有 k 均值聚合策略的 FD-WCAT。在消融研究中, FD-WCAT 的性能优于 FD-WCAT_without, 这意味着聚合策略是有效的。接下来, 我们展示 AP 分数的结果。

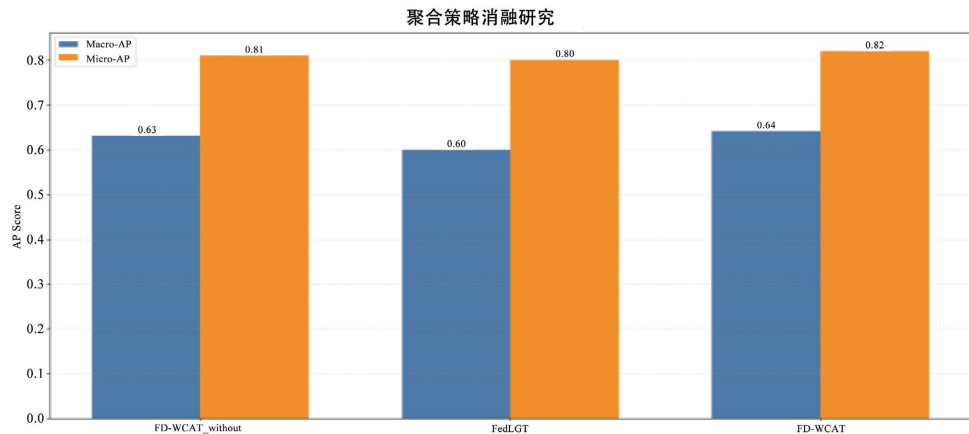


Figure 8. The performance of FD-WCAT in ablation studies of the aggregation strategy
图 8. 聚合策略消融研究中 FD-WCAT 的性能

如表 5 和图 8 所示, FD-WCAT 取得了最佳的测试结果, 这些结果验证了 FD-WCAT 聚合策略的有效性。

5. 结论

本文提出了一种新的双阶段联邦学习框架 FD-WCAT, 旨在解决智能电网系统中多标签电力图像分类的挑战。通过利用加权相关性注意力网络, FD-WCAT 捕获标签相关性嵌入以提高分类性能, 并通过惩罚系数减轻类别不平衡的影响。所提出的框架还通过纳入全局 - 本地参数正则化损失函数和采用分组加权聚合算法, 解决了本地模型和全局模型之间的参数不一致问题。这些创新不仅提高了聚合过程中的模型一致性, 还增强了模型在异质数据源上的泛化能力。实验结果表明, FD-WCAT 优于现有的基准模型, 使其成为分布式、隐私保护环境中的多标签分类任务的有前景的解决方案。然而, 有几个局限性值得进一步研究。多标签学习的挑战还源于标签可能不准确、重叠或缺失, 特别是在多标签场景图像中。解决由这些问题导致的性能下降需要进一步完善方法, 尤其是在标签质量不理想的情况下处理联邦多标签分类的场景。

基金项目

本工作得到了中国南方电网有限责任公司科技项目的支持(项目编号: 070000KC24110002)。

参考文献

- [1] McMahan, B., Moore, E., Ramage, D., *et al.* (2017) Communication-Efficient Learning of Deep Networks from Decentralized Data, Artificial Intelligence and Statistics. *The Proceedings of Machine Learning Research*, **2017**, 1273-1282.
- [2] Büyüktaş, B., Weitzel, K., Völkers, S., Zailskas, F. and Demir, B. (2024) Transformer-Based Federated Learning for Multi-Label Remote Sensing Image Classification. *2024 IEEE International Geoscience and Remote Sensing Symposium*, Athens, 7-12 July 2024, 8726-8730. <https://doi.org/10.1109/igarss53475.2024.10641701>
- [3] Zhang, M.L. and Zhou, Z.H. (2013) A Review on Multi-Label Learning Algorithms. *IEEE Transactions on Knowledge and Data Engineering*, **26**, 1819-1837. <https://doi.org/10.1109/tkde.2013.39>
- [4] Zhang, J., Wei, T. and Zhang, M.L. (2024) Label-Specific Time-Frequency Energy-Based Neural Network for Instrument

- Recognition. *IEEE Transactions on Cybernetics*, **54**, 7080-7093. <https://doi.org/10.1109/tcyb.2024.3433519>
- [5] Lanchantin, J., Wang, T., Ordonez, V. and Qi, Y. (2021) General Multi-Label Image Classification with Transformers. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 16473-16483. <https://doi.org/10.1109/cvpr46437.2021.01621>
 - [6] Zhou, W., Lin, K., Zheng, Z., Chen, D., Su, T. and Hu, H. (2025) DRTN: Dual Relation Transformer Network with Feature Erasure and Contrastive Learning for Multi-Label Image Classification. *Neural Networks*, **187**, 107309. <https://doi.org/10.1016/j.neunet.2025.107309>
 - [7] Liu, I., Lin, C., Yang, F. and Wang, Y.F. (2024) Language-Guided Transformer for Federated Multi-Label Classification. *Proceedings of the AAAI Conference on Artificial Intelligence*, **38**, 13882-13890. <https://doi.org/10.1609/aaai.v38i12.29295>
 - [8] Niu, X. and Wei, E. (2023) Fedhybrid: A Hybrid Federated Optimization Method for Heterogeneous Clients. *IEEE Transactions on Signal Processing*, **71**, 150-163. <https://doi.org/10.1109/tsp.2023.3240083>
 - [9] Huang, X., Li, P. and Li, X. (2023) Stochastic Controlled Averaging for Federated Learning with Communication Compression.
 - [10] Gao, L., Fu, H., Li, L., Chen, Y., Xu, M. and Xu, C. (2022) FedDC: Federated Learning with Non-IID Data via Local Drift Decoupling and Correction. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, 18-24 June 2022, 10102-10111. <https://doi.org/10.1109/cvpr52688.2022.00987>
 - [11] Li, J., Zhang, C., Zhou, J.T., Fu, H., Xia, S. and Hu, Q. (2021) Deep-LIFT: Deep Label-Specific Feature Learning for Image Annotation. *IEEE Transactions on Cybernetics*, **52**, 7732-7741. <https://doi.org/10.1109/tcyb.2021.3049630>
 - [12] Yu, Z.-B. and Zhang, M.-L. (2022) Multi-Label Classification with Label-Specific Feature Generation: A Wrapped Approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **44**, 5199-5210.
 - [13] Jia, B.B. and Zhang, M.L. (2023) Multi-Dimensional Multi-Label Classification: Towards Encompassing Heterogeneous Label Spaces and Multi-Label Annotations. *Pattern Recognition*, **138**, Article 109357. <https://doi.org/10.1016/j.patcog.2023.109357>
 - [14] Ahmad, K.M., Liu, Q., Khan, A.A., *et al.* (2023) Prompt-Enhanced Federated Learning for Aspect-Based Sentiment Analysis. 2023 *International Conference on Intelligent Communication and Computer Engineering*, Changsha, 24-26 November 2023, 81-87. <https://doi.org/10.1109/icicce61720.2023.00019>
 - [15] Gupta, K. and Prasad, R. (2024) Semi Supervised Federated Learning with Pseudo-Labeling. IIIT-Delhi.
 - [16] Kassem, H., Alapatt, D., Mascagni, P., Karagyris, A. and Padoy, N. (2022) Federated Cycling (FedCy): Semi-Supervised Federated Learning of Surgical Phases. *IEEE Transactions on Medical Imaging*, **42**, 1920-1931. <https://doi.org/10.1109/tmi.2022.3222126>
 - [17] Qiu, L., Cheng, J., Gao, H., Xiong, W. and Ren, H. (2023) Federated Semi-Supervised Learning for Medical Image Segmentation via Pseudo-Label Denoising. *IEEE Journal of Biomedical and Health Informatics*, **27**, 4672-4683. <https://doi.org/10.1109/jbhi.2023.3274498>
 - [18] Sun, Z., Wu, N., Shi, J., Yu, L., Cheng, K. and Yan, Z. (2024) FEDMLP: Federated Multi-Label Medical Image Classification under Task Heterogeneity. In: *Lecture Notes in Computer Science*, Springer, 394-404. https://doi.org/10.1007/978-3-031-72117-5_37
 - [19] Vondikakis, I.V., Panagiotopoulos, I.E. and Dimitrakopoulos, G.J. (2024) FedRSC: A Federated Learning Analysis for Multi-Label Road Surface Classifications. *IEEE Open Journal of Intelligent Transportation Systems*, **5**, 433-444. <https://doi.org/10.1109/ojits.2024.3432176>
 - [20] Yang, J., Li, S., Zheng, K., Zeng, L., Qi, S., Xu, J., *et al.* (2025) Label-Specific Feature Based Multi-Label Neural Network for Federated Learning. 2025 *5th International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, Dongguan, 28 February-2 March 2025, 130-136. <https://doi.org/10.1109/iccece65250.2025.10985200>
 - [21] Chang, S.F., Hsu, B.W.Y., Chang, T.Y., *et al.* (2023) FLAG: Fast Label-Adaptive Aggregation for Multi-Label Classification in Federated Learning.
 - [22] Guo, T., Guo, S., Wang, J., Tang, X. and Xu, W. (2024) PROMPTFL: Let Federated Participants Cooperatively Learn Prompts Instead of Models—Federated Learning in Age of Foundation Model. *IEEE Transactions on Mobile Computing*, **23**, 5179-5194. <https://doi.org/10.1109/tmc.2023.3302410>
 - [23] Diao, E., Ding, J. and Tarokh, V. (2022) Semifl: Semi-Supervised Federated Learning for Unlabeled Clients with Alternate Training. *Advances in Neural Information Processing Systems*, **35**, 17871-17884.
 - [24] Nguyen, D.P., Munoz, J.P. and Jannesari, A. (2024) Flora: Enhancing Vision-Language Models with Parameter-Efficient Federated Learning.
 - [25] Radford, A., Kim, J.W., Hallacy, C., *et al.* (2021) Learning Transferable Visual Models from Natural Language Supervision. arXiv.2103.00020.

-
- [26] Song, C., Granqvist, F. and Talwar, K. (2022) FLAIR: Federated Learning Annotated Image Repository. arXiv.2207.08869.
 - [27] Lin, T.Y., Maire, M., Belongie, S., *et al.* (2014) Microsoft COCO: Common Objects in Context. In: *Lecture Notes in Computer Science*, Springer, 740-755. https://doi.org/10.1007/978-3-319-10602-1_48