

ISAC网络中基于特征信息的联邦学习动态压缩方法

邓炳光, 彭佳音

(重庆邮电大学通信与信息工程学院, 重庆 400065)

摘要: 针对6G通感一体化 (ISAC) 边缘网络中联邦学习 (FL) 在通信带宽受限、链路动态变化及设备异构条件下的通信效率问题, 构建了一种面向ISAC场景的特征信息驱动压缩与聚合协同框架 (FI-ISAC-FL)。不同于现有压缩方法, 联邦更新压缩与聚合决策是为实现任务相关特征与系统状态之间的协同适配而联合设计。具体来说, 设计Federated-AE在线训练机制对局部更新进行在线特征压缩, 实现对高维更新的特征提取与高效传输, 增强压缩表示对数据变化的适应能力。同时, 结合ISAC场景下的信道状态与数据分布特征, 构建状态感知的自适应聚合策略, 改善通信受限条件下的训练稳定性与效率。仿真结果表明, 在雷达目标识别任务中, 所提方法在固定通信预算下能够取得较好的模型性能, 在性能与资源开销之间实现更优折中, 并对强感知干扰与Non-IID数据分布表现出较强的鲁棒性。

关键词: 通感一体化; 联邦学习; 特征信息; 梯度压缩; 自编码器

中图分类号: TN929.5

文献标志码: A

DOI: 10.11959/j.issn.1000-436x.2026082

Dynamic compression method for federated learning based on feature information in ISAC networks

Deng Bingguang, Peng Jiayin

School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Abstract: To address the communication-efficiency problem of federated learning (FL) in 6G integrated sensing and communication (ISAC) edge networks under limited bandwidth, dynamic link conditions, and device heterogeneity, a feature-information-driven collaborative compression and aggregation framework, termed FI-ISAC-FL, was developed. Unlike existing compression methods, federated update compression and aggregation were jointly designed for coordinated adaptation between task-relevant features and system states. Specifically, a Federated-AE online training mechanism was introduced to compress local updates online, by which high-dimensional updates were effectively represented and efficiently transmitted, while the adaptability of compressed representations to data variations was enhanced. In addition, a state-aware adaptive aggregation strategy was developed by incorporating channel states and data-distribution characteristics in ISAC scenarios to improve training stability and efficiency under communication constraints. Simulation results show that, in a radar target recognition task, favorable model performance is achieved under a fixed communication budget, a better trade-off between performance and resource overhead is attained, and strong robustness against severe sensing interference and Non-IID data distributions is demonstrated.

Keywords: ISAC, federated learning, feature information, gradient compression, autoencoder

收稿日期: 2026-01-04; 修回日期: 2026-03-20

通信作者: 邓炳光, dengbg@cqupt.edu.cn

基金项目: 重庆市自然科学基金创新发展联合基金(中国星网)资助项目(No.CSTB2023NSCQ-LZX0114)

Foundation Items: Chongqing Natural Science Foundation Innovation and Development Joint Fund (China Star Network) (No.CSTB2023NSCQ-LZX0114)

0 引言

通感一体化 (integrated sensing and communication, ISAC) 技术被认为是实现 6G 边缘智能的关键技术之一, 通过共享频谱和硬件资源, 使网络边缘设备能够同时执行高精度环境感知与高速无线通信^[1-2]。在工业物联网 (industrial internet of things, IIoT) 或车联网 (internet of vehicles, IoV) 等典型 ISAC 应用场景中, 边缘设备往往需要在毫秒级时延内完成高精度的环境感知与协同决策。然而, 随着感知任务复杂度的持续提升, 边缘侧普遍采用的大规模深度神经网络 (deep neural network, DNN) 模型导致模型更新维度显著增加, 模型参数达到数百万级, 在资源受限的无线信道中频繁交互高维模型更新信息, 带来了严重的通信时延和能量消耗问题^[3-4], 可能无法满足工业级应用对实时性的严苛要求。

联邦学习 (federated learning, FL)^[5]通过允许边缘设备在本地训练模型并仅交换模型更新, 在保

护数据隐私的同时实现分布式协同学习, 为缓解通信压力提供了基础框架^[6]。然而, 面对 6G 场景下急剧增长的通信压力, 仅依靠 FL 的基本框架是远远不够的。近年来, 围绕 FL 通信效率提升问题, 现有研究主要沿以下几条技术路线展开, 如表 1 所示。一是模型更新压缩与高效表示方法, 通过量化、稀疏化、自适应压缩及学习式表征等手段减少传输开销。其中, 经典数值级压缩方法通过对参数或模型更新进行量化、稀疏化来降低传输比特数^[7-8]。进一步, 有研究结合动态带宽约束设计自适应压缩策略, 改善不同网络状态下的精度-开销折中^[9]。此外, 学习式压缩方法通过自编码器 (autoencoder, AE) 等机制进行低维表征^[10]。任务导向通信研究则进一步强调应关注对下游任务真正有价值的信息表达, 而不仅是数值层面的高保真传输^[11], 在 FL 场景中, 语义协同方法也已被用于提升协作效率^[12-13]。二是系统协同优化与 ISAC-FL 协同方法, 不直接对模型更新本身进行压缩, 而是

表 1 现有 FL 高效通信相关方法与本文方法对比

相关技术路线及代表文献	核心机制	在 ISAC 场景中的主要局限	本文方法的突破与优势	
模型更新压缩与高效表示方法	数值级压缩 (QSGD ^[7] 、DGC ^[8])	通过量化、稀疏化、低秩等手段降低更新维度	以数值重构为目标, 难区分任务相关信息与冗余信息	聚焦特征信息传输
	自适应压缩 (AdapComFL ^[9])	根据带宽和资源状态自适应调整压缩率	多面向网络状态调节, 缺乏对 ISAC 感知与通信耦合特性的统一建模	将环境感知用于压缩率调度与聚合策略选择, 实现特征压缩与自适应聚合的在线协同优化
	学习式压缩 (FLAC ^[10])	利用 AE 降维重构, 多依赖预训练或固定编码器	在 ISAC 动态信道与强电磁干扰下, 易发生表示失配	设计 Federated-AE 在线训练机制, 增强对动态环境的持续适应能力
	任务导向/语义协同 (FedSC ^[12] 、SFL ^[13])	传输任务相关特征或低维表示, 以减少无效通信	往往依赖额外语义对齐或原型构造, 对 ISAC 动态链路约束考虑不足	直接面向梯度更新构建特征信息流, 不需要显式语义原型
系统协同优化与 ISAC-FL 协同方法	系统协同优化 (FedGA ^[14])	通过节点分组、异步聚合与分时调度等机制降低异构环境下训练时延和资源开销	主要优化训练流程与资源利用率, 不直接提升模型更新内容本身的表达质量	重点优化更新内容的信息有效性, 在固定通信预算下提升更新表达质量与聚合稳定性
	ISAC-FL 协同优化 (Air-FEEL ^[15])	从感知、通信、计算联合设计角度优化联邦边缘学习训练效率与系统性能	更偏向于从训练流程与资源调度角度进行系统层联合优化, 对模型更新内容本身的任务相关表达考虑有限	聚焦模型更新的特征信息表达, 并将压缩率调度与聚合修正联合设计
特征增强与聚合修正方法 (FedCAFNR ^[17] 、FedAM ^[18])	通过注意力、降噪或特征分离等手段提升全局模型质量	多集中于聚合阶段, 未统一建模压缩、传输、聚合与 ISAC 环境状态	将特征压缩、环境感知与聚合修正统一到一个框架中, 强调全过程协同优化	
本文方法	将梯度更新建模为任务相关的特征信息流, 结合 Federated-AE、环境感知压缩率调度与自适应聚合进行联合优化	—	在动态 ISAC 场景下同时兼顾通信效率、聚合稳定性与任务性能	

通过节点分组、异步聚合、分时调度等机制降低异构环境下的训练时延和资源开销^[14]。近期已有研究开始面向感知、通信与计算深度耦合场景讨论联邦边缘学习问题, 强调仅从单一压缩或单一调度角度已难以充分刻画动态 ISAC 环境下的训练性能边界^[15-16]。三是特征增强与聚合修正方法, 主要通过注意力机制、降噪和特征分离等方式缓解统计异构带来的聚合偏差, 提升全局模型性能与训练稳定性^[17-18]。

尽管上述研究在不同层面提升了 FL 通信效率, 但在 ISAC 场景下仍存在不足。首先, 模型更新压缩与高效表示方法大多关注如何降低传输比特数或提取紧凑表示, 而对感知干扰、动态信道和 Non-IID 数据共同作用下模型更新中任务相关信息的保留仍显著不足^[19]。其次, 现有系统协同优化与 ISAC-FL 协同方法虽然已开始关注感知、通信与计算之间的联合设计, 但多数研究更偏重训练流程优化、资源分配或整体系统层协同, 对模型更新内容本身的表达质量及其与聚合稳定性之间的关系刻画仍不充分。最后, 现有特征增强与聚合修正方法虽然有助于缓解统计异构带来的性能退化, 但通常未将特征压缩、环境感知和聚合修正统一纳入同一框架, 难以同时兼顾通信效率、聚合稳定性与 ISAC 场景适应性。

为此, 本文从任务相关特征信息传输角度重新审视 FL 中的模型更新过程, 将局部梯度更新建模为承载模型优化方向的特征信息流。在此基础上, 进一步结合特征空间增强、环境感知压缩率调度与自适应聚合修正, 提出了一种面向 ISAC 网络的基于特征信息的联邦学习动态压缩方法。

本文主要工作如下。

1) 构建了一种面向 ISAC 场景的特征信息驱动 FL 压缩与聚合协同框架 (FI-ISAC-FL), 将 FL 中的梯度更新建模为承载模型优化方向的特征信息流。与现有压缩方法不同, 该框架更关注任务相关特征表示在无线传输与全局聚合过程中的协同适配。

2) 设计了 FI-ISAC-FL 协同训练算法。将梯度特征压缩、状态感知及自适应聚合策略选择统一纳入同一联邦训练框架中, 实现系统级的联合优化。针对现有学习式压缩方法依赖离线预训练或固定编码结构、在动态环境中易发生表示适配的问题, 设

计 Federated-AE 在线训练机制对局部更新进行在线特征压缩, 将梯度 AE 作为辅助模型与主模型协同演化。结合 ISAC 场景下的无线信道状态、感知任务强度与系统资源约束, 构建面向压缩率调度与聚合策略选择的环境感知自适应机制, 以提升动态环境下模型训练的鲁棒性与资源利用率。

3) 从压缩误差传播与训练稳定性角度, 对在线特征压缩和状态驱动聚合的共同作用进行了理论分析。说明压缩误差、数据异构性与聚合策略选择对全局模型更新的影响规律, 为所提方法的合理性提供了分析支撑。

4) 系统仿真与性能评估。通过雷达目标识别任务及多种扩展场景实验, 对所提方法进行了系统验证。结果表明, FI-ISAC-FL 在固定通信预算下能够获得较好的模型性能, 并在可用性能与资源开销之间取得更优折中, 同时对强感知干扰和 Non-IID 数据分布具有较好的鲁棒性。

1 系统模型

本文考虑一个典型的 ISAC-FL 场景, 旨在解决分布式目标识别任务。系统由 K 个边缘设备和一个中心服务器组成, 拓扑结构为典型的星型 FL 架构 (如图 1 所示)。每个边缘设备配备环境感知传感器和本地计算单元, 可以获取本地雷达观测数据并独立执行模型训练。中心服务器具有强算力和覆盖网络, 负责聚合来自边缘设备的模型更新并协调全局模型。在该架构下, 无线频谱等资源由感知和通信过程共享, 系统需要同时优化感知质量、通信效率和计算性能。

1.1 感知模型

目前, 在每轮通信开始时, 设备端利用集成的 ISAC 模块对环境或目标进行观测采样, 获取本地训练的原始数据, 设第 t 轮通信轮次中设备 i 获得的原始数据为 $x_i(t)$ 。 $x_i(t)$ 可源于不同感知模态, 如雷达回波、摄像头图像序列、LiDAR 点云或信道状态信息 (channel state information, CSI) 等。然而, 受限于传感器精度和外部环境, 原始数据往往伴有噪声和不完整, 这将导致基于该数据训练的本地模型梯度存在失真。正如已有研究所指出的, 无线感知信号信噪比过低会降低模型梯度估计的准确性, 从而减缓 FL 的收敛^[15]。为缓解这些问题, 系统对原始数据进行预处理, 并采用与任务模态匹配

的数据增强技术提升泛化能力：对图像或频谱类输入采用随机裁剪、平移旋转与颜色调整等数据增强技术；对于点云类输入采用随机抖动与点采样等数据增强技术。最后，对数据进行归一化处理，设预处理后的数据为 $\hat{x}_i(t)$ ，形成用于训练的本地数据集 $D_i(\hat{x}_i(t) \in D_i)$ 。

1.2 FL与通信过程建模

在每一轮全局通信中，服务器首先广播下发当前全局模型参数 $w^{(t)}$ 和梯度AE参数 $\theta^{(t)}$ ，各边缘设备收到全局参数后，根据1.1节形成的本地数据集 D_i 执行本地训练，从而计算出本地模型更新（梯度）。从模型层面来说，与现有FL方法主模型多使用ResNet-10等网络不同，本文采用面向任务输出的轻量级端到端网络模型（End2End）。为降低模型冗余，本文的通信机制不再是对原始高维更新的数值重构，而是对任务相关特征信息的保留与传输：在模型层面以任务相关特征表示作为通信载

体；在通信层面以对全局优化贡献最大的更新方向信息作为主要传输内容。本文后续所述“特征压缩和特征信息传输”均特指通信层传输。该模型可抽象为“特征编码器+分类头”的级联形式，其输出可直接用于分类任务，同时作为后续特征信息传输的源信息。该模型核心作用在于保证任务的可训练性，同时通过结构设计参数压缩降低模型规模，使本地训练与上行传播更轻量。

完成本地梯度 $g_i^{(t)}$ 计算后，设备并不直接上传高维梯度，而是先将梯度进行特征压缩编码，生成低维特征表示 $s_i^{(t)}$ （详情见1.3节）。随后，设备通过上行链路将 $s_i^{(t)}$ 和本地梯度AE参数 $\theta_i^{(t)}$ 一并上传至服务器。在服务器端，收到多个边缘设备上传的编码信息后，利用预先固定的解码器对每个梯度低维特征表示逐一解码，重构出对应的梯度 $\hat{g}_i^{(t)}$ 。接着，服务器根据当前的系统状态动态选择适当的模型聚合策略，将重构梯度聚合得到更新的全局模型

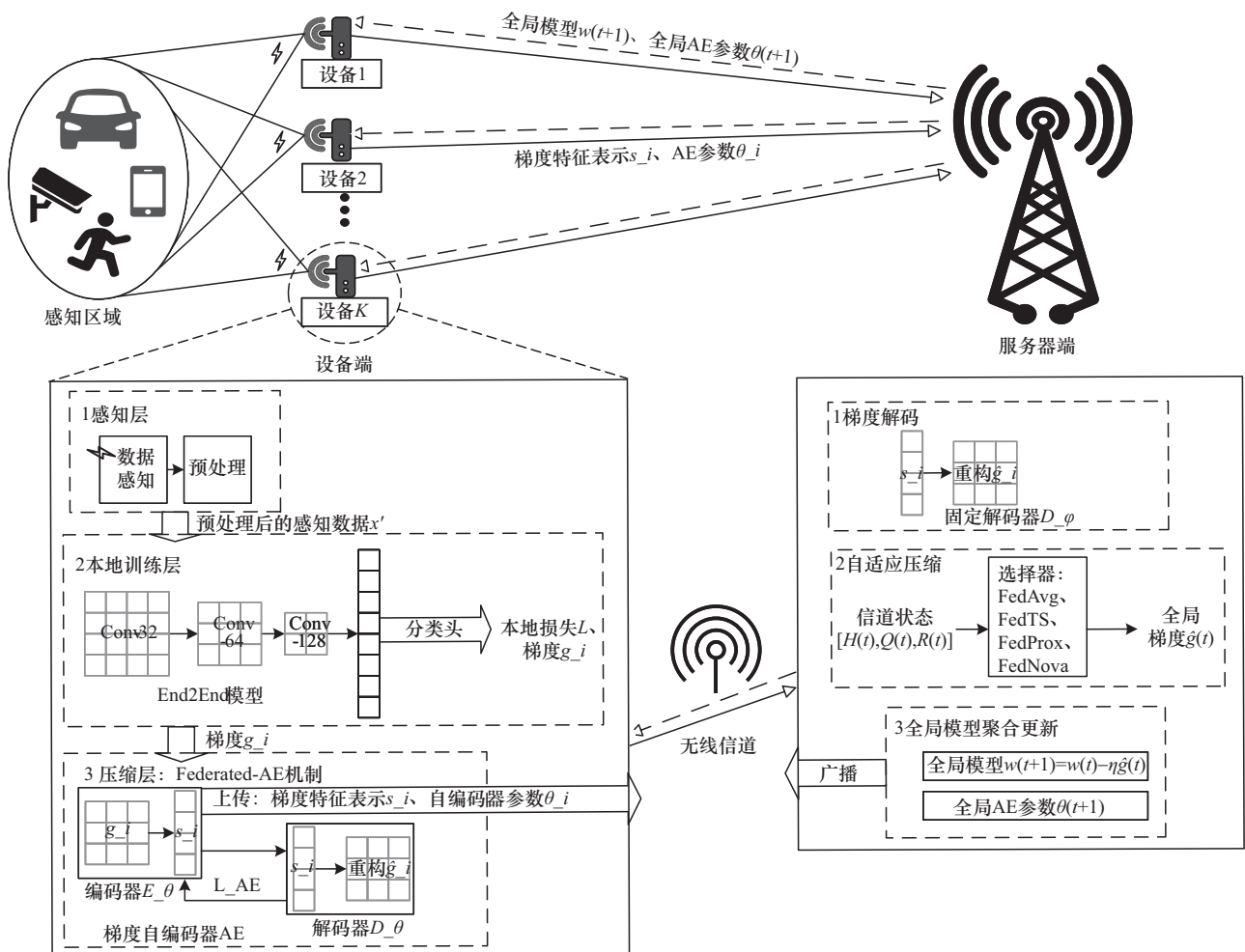


图1 FI-ISAC-FL系统架构

参数 $w^{(t+1)}$, 并同步聚合更新全局 AE 参数 $\theta^{(t+1)}$ 。最后, 服务器将新一轮的全局参数广播下发给所有设备, 开始下一轮训练。如此循环往复, 直至达到预定的通信轮次或收敛条件。

值得注意的是, 在传统 FL 中, 边缘设备往往直接上传完整的模型梯度参与聚合, 由于模型维度极高, 直接传输将产生巨大的通信开销。相比之下, FI-ISAC-FL 通过引入特征压缩机制, 极大地降低了每轮通信的数据量。

1.3 梯度特征压缩重构模型

在 FL 中, 设备端 i 在第 t 轮训练得到的本地梯度 $g_i^{(t)} \in \mathbb{R}^{d_g}$, 传统方法直接上传完整梯度, 通信开销呈线性增长。从通信层面来说, 本文将特征信息传输具体化为任务驱动的低维特征表示的传输: 设备端将 $g_i^{(t)}$ 映射为包含主要更新方向的低维特征表示 $s_i^{(t)}$, 并不追求压缩后能够逐维精确重构原始梯度, 而是希望在通信资源受限条件下, 传输的低维特征表示仍能保留对全局目标 $F(w)$ 的主要下降贡献。服务器端解码重构得到重构梯度 $\hat{g}_i^{(t)}$ 并参与聚合。为此, 梯度特征信息编解码过程可表示为

$$\begin{aligned} s_i^{(t)} &= E_{\theta_E}(g_i^{(t)}), s_i^{(t)} \in \mathbb{R}^{d_s}, d_s \ll d_g \\ \hat{g}_i^{(t)} &= D_{\theta_D}(s_i^{(t)}), \hat{g}_i^{(t)} \in \mathbb{R}^{d_g} \end{aligned} \quad (1)$$

其中, $s_i^{(t)}$ 为低维特征表示, 指对全局优化最优贡献的更新方向信息。编码器通过“分块-筛选-聚合”三阶段实现任务相关特征提取: 将梯度 $g_i^{(t)}$ 分割为若干子块, 每块独立编码为中间特征; 根据 L2 范数衡量更新贡献度, 计算每个块的特征重要性, 仅保留 Top-K 个关键块, 这一机制保证编码器优先保留对全局优化贡献最大的梯度块, 而非统计意义上的全局均值信息。进一步, 对筛选后的关键块特征进行聚合并投影到低维特征表示 $s_i^{(t)}$ 。

为保证特征空间的一致性与解码的有效性, 系统设计了非对称编解码架构: 设备端部署包含本地编码器 $E_{\theta}(\cdot)$ 和解码器 $D_{\theta}(\cdot)$ 的完整梯度 AE, 用于本地训练时的自监督学习; 服务器端则使用固定解码器, 初始化后不参与更新, 也不参与联邦聚合。这种“动态编码、固定解码”的非对称架构设计确保了所有设备端的低维特征表示被映射到统一的全局特征空间, 避免了训练过程中的空间漂移现象。在该抽象模型下, 上行通信的负载由原来的高维梯度变为低维特征表示, 结合全局固定解码器可以保

证解码的一致性和稳定性, 大大提升了通信效率。

1.4 问题定义与优化目标

基于以上模型, 本文将优化目标划分为主任务目标和特征保真目标。其中, 主任务目标用于刻画 FL 的最终任务需求, 特征保真目标用于约束梯度特征压缩带来的信息损失, 使压缩表示在通信受限条件下仍能尽可能保持关键更新方向。二者相互配合实现了压缩编码-传输-解码-聚合更新的通信闭环。令 $L_i(w)$ 表示设备 i 上的本地任务损失函数, 设备权重 $\omega_i = \frac{n_i}{n}$ 基于各设备样本数量 n_i 且满足 $\sum \omega_i = 1$, 则主任务目标可表示为

$$\min_w F(w) \triangleq \sum_{i=1}^K \omega_i L_i(w) \quad (2)$$

在 AE 层面, 为抑制特征压缩带来的信息损失, 引入任务相关的特征保真目标作为第二优化目标。该目标通过 FL 训练 AE 的编码参数 θ_E 和解码参数 θ_D , 在最小化重构误差的同时, 进一步约束重构与原始梯度在更新方向上的一致性, 从而提高对全局收敛更关键更新信息的保留能力, 可表示为

$$\begin{aligned} \min_{(\theta_E, \theta_D)} L_{AE}(\theta_E, \theta_D; G_i) &= \\ \frac{1}{m} \sum_{g_{ij}^{(t)} \in G_i} [& \|M^{1/2}(g_{ij}^{(t)} - \tilde{g}_{ij}^{(t)})\|_2^2 + \\ \beta \left(1 - \frac{\langle g_{ij}^{(t)}, \tilde{g}_{ij}^{(t)} \rangle}{\|g_{ij}^{(t)}\|_2 \|\tilde{g}_{ij}^{(t)}\|_2} \right) &] \\ \tilde{g}_{ij}^{(t)} &= D_{\theta_D}[E_{\theta_E}(g_{ij}^{(t)})] \end{aligned} \quad (3)$$

其中, M 为权重矩阵, β 为方向约束系数, $\|\cdot\|_2^2$ 为 L2 范数平方, 用来衡量原始梯度和本地重构梯度之间的差异, G_i 为中间梯度样本集合, m 为采样数量。

式(2)刻画了任务驱动的学习目标, 式(3)则刻画了通信受限条件下的特征保真目标。需要强调的是, 式(3)作为压缩模块的辅助训练目标并不改变式(2)的任务定义, 而是通过任务相关的特征保真约束优化 AE, 提升关键更新信息在压缩表示中的保留优先级, 并间接影响服务器端聚合所使用的重构梯度 $\hat{g}_i^{(t)}$, 从而维持式(2)的全局优化效果。

综上所述, FI-ISAC-FL 的优化问题可以表述为在边缘网络受限的通信与计算资源下, 分别优化

w 、 θ_E 和 θ_D ，以最小化主任务的全局损失和梯度重构误差。换言之，希望在满足通信资源限制的前提下，实现主模型准确性最大化与通信开销最小化的统一，这要求在感知、通信和计算三方面进行折中权衡，以达到ISAC场景下全局模型性能和资源效率的最佳组合。

2 算法设计

本文设计了FI-ISAC-FL协同训练算法，其核心思想是引入双模型协同优化和自适应策略选择机制：设备端同时维护用于分类任务的主模型和用于梯度压缩的梯度AE；服务器端则负责聚合更新全局主模型和全局AE参数，并根据实时系统状态动态选择最优的聚合策略。需要说明的是，本文的主要贡献体现在特定场景约束下的系统联合设计，而非单一基础模块的独立更新。下面分别介绍FI-ISAC-FL算法中各核心组件，并给出整体算法流程及复杂度分析。

2.1 主模型本地训练

设备端部署轻量级的端到端网络模型做为主模型，负责处理感知数据，并生成用于模型更新的原始梯度。在第 t 轮FL中，设备 i 接收全局模型 $w^{(t)}$ ，基于本地数据集 D_i 执行主模型训练与参数更新，以最小化任务损失并生成待上传信息。在本文框架中，第 j 个Batch上的本地损失函数可表示为

$$L_{ij}(w) = \frac{1}{|B|} \sum_{(\hat{x}, y) \in B_j} [\ell(f_w(\hat{x}), y)] \quad (4)$$

其中， $f_w(\cdot)$ 为参数化模型， $\ell(\cdot)$ 为任务交叉熵损失函数， y 为本地样本的真实标签， B 为批大小。据此，可以得到第 j 个Batch上的本地梯度为

$$g_{ij} = \nabla_w L_{ij}(w) \quad (5)$$

为提升局部训练的收敛速度与稳定性，本文采用Adam优化器进行参数更新。Adam优化器结合了梯度一阶与二阶动量估计并自适应调整学习率，特别适合FL中数据Non-IID的非平稳目标场景。对于本地训练的第 j 个Batch上的梯度 g_{ij} ，其更新过程可表示为

$$\begin{cases} m_j = \beta_1 m_{j-1} + (1 - \beta_1) g_{ij} \\ v_j = \beta_2 v_{j-1} + (1 - \beta_2) g_{ij}^2 \\ w_i^{(t)} \leftarrow w_i^{(t)} - \frac{\eta m_j}{\sqrt{v_j} + \varepsilon} \end{cases} \quad (6)$$

其中， m_j 和 v_j 分别为梯度的一阶矩（动量）和二阶矩（自适应学习率）估计，默认超参数 $\beta_1 = 0.9, \beta_2 = 0.999, \varepsilon = 10^{-8}$ ， η 为学习率。经过 E 轮本地训练，设备得到更新后的本地模型参数 $w_i^{(t)}$ ，对应模型更新量为 $\Delta w_i^{(t)} = w_i^{(t)} - w_i^{(t-1)}$ ，其等价于所有训练步骤的梯度效应，可表示为

$$\Delta w_i^{(t)} = -\eta \sum_{j=1}^E \frac{m_j}{\sqrt{v_j} + \varepsilon} \quad (7)$$

由于Adam优化器对梯度施加了自适应预条件，其产生的参数更新并不等价于原始梯度下降方向。为统一表示不同本地优化器下的模型更新形式，本文将设备 i 在第 t 轮本地训练结束后得到的模型增量 $\Delta w_i^{(t)}$ 归一化为等效更新方向，即

$$g_i^{(t)} = \frac{-\Delta w_i^{(t)}}{\eta E} \quad (8)$$

该方向表示对真实梯度施加自适应预条件变换后的更新方向，后续统称为梯度。该梯度代表设备在本地训练期间的整体优化趋势，可用于后续压缩、聚合与特征选择等操作。

2.2 Federated-AE在线训练机制

本文提出了基于特征信息的压缩方法，通过引入可学习的梯度AE，实现对模型更新进行任务相关特征提取与压缩重构。针对将AE用于FL的场景，现有方法（如FLAC等）多采用离线预训练或固定编码器。然而，在FL的实际迭代中，模型梯度 $g_i^{(t)}$ 存在严重的时间动态分布偏移：在训练初期，模型尚未收敛，梯度幅值大且特征丰富；在收敛后期，梯度趋于稀疏且多为微小震荡。这就导致基于离线数据或早期轮次预训练得到的固定AE无法捕捉这种时变特性。随着训练的推进，固定AE将面临严重的分布外（out-of-distribution, OOD）泛化问题，使服务器端解码重构的梯度偏离真实更新方向，导致重构误差在训练中后期急剧上升。

为有效应对上述时间动态分布偏移问题，并清晰体现与传统离线或固定编码器方法的差异，本文设计了Federated-AE在线训练机制。具体而言，本文将梯度AE作为第二个全局模型集成至联邦整体模型训练流程中，与主模型协同联邦训练。通过在设备端对梯度AE进行在线更新，使编码器能够持续动态地适配当前的梯度分布特征，从而有效地抑制重构误差在动态场景下的累积增长。

梯度AE通过FL进行训练更新的核心目标是减小梯度重构误差,使经过编码解码后的重构梯度尽可能接近原始梯度。在具体实现上,为了支持梯度AE的训练,边缘设备会在执行本地梯度计算时收集梯度样本。在执行FL过程中,以概率 p 采样中间梯度,构建梯度样本集合为

$$G_i = \{g_{i,1}, g_{i,2}, \dots, g_{i,m}\} \quad (9)$$

其中, $m \approx pE \frac{|D_i|}{B}$ 为采样数量, 采样概率的选择需要权衡编码器训练数据的充分性与采样开销, 过低将导致训练样本不足, 过高则增加存储开销。在完成 E 轮本地主模型训练后, 设备利用收集的梯度样本对AE执行 N_{AE} 步在线更新。这一联邦在线更新过程是实现特征信息自适应压缩的关键步骤: 通过持续学习当前梯度的稀疏表示, AE能够自适应提取不同轮次和不同数据环境下梯度的共有特征。更新后的本地编码器 $E_\theta(\cdot)$ 被用于压缩该轮最终模型梯度 $g_i^{(t)}$, 得到低维特征表示 $s_i^{(t)}$, 随后设备上传 $s_i^{(t)}$ 和AE参数 $\theta_i^{(t)}$, 显著减少了上行通信的数据量。在服务器端, 将对各设备上传的AE参数执行聚合(见2.3节), 得到更新的全局AE参数, 再下发给设备用于下一轮训练, 如此循环, 实现AE随训练动态更新、不断提高压缩性能的目标。

2.3 状态感知聚合策略与非对称解码结构

在服务器端, 采用固定解码器对各设备上传的低维特征表示进行统一解码重构。具体来说, 服务器收到 K 个设备上传的低维特征表示 $s_i^{(t)}$ 后, 使用固定解码器 $D_\varphi(\cdot)$ 解码并重构梯度向量, 这些重构梯度向量将作为聚合阶段的输入。

$$\hat{g}_i^{(t)} = D_\varphi(s_i^{(t)}) \quad (10)$$

FL环境具有动态性和异构性, 单一聚合策略难以适应所有场景。因此, 本文在不改变FedAvg (federated averaging)、FedTS (federated temporal smoothing)、FedProx (federated proximal) 和 FedNova (federated normalized averaging) 的基础聚合形式的条件下, 提出了一种面向ISAC场景的聚合策略选择器, 根据系统状态在线选择最匹配的聚合器, 以降低动态信道与异构性对收敛性能的影响。首先定义了系统状态空间, 使用一个三维向量表示当前系统状态 $S(t) = [Q(t), H(t), R(t)]$, 其中, $Q(t) \in [0,1]$ 为有效信道质量, 数值越小表示信道

越差; $H(t) \in [0,1]$ 为系统异构性, 数值越小表示异构偏离越严重; $R(t) \in [0,1]$ 为收敛速率, 数值越小表示收敛出现停滞。基于上述系统状态量, 在模型训练的初始warm-up启动阶段(即前 T_0 轮), 系统固定采用基线FedAvg策略, 并持续记录系统状态 $Q(t), H(t), R(t)$ 的历史序列。warm-up启动结束后, 系统提取各状态序列的经验分位数作为动态阈值。对于系统状态, 取其历史序列的 q -低分位数(本文设 $q = 20\%$)作为阈值 τ_{comm} 、 τ_{het} 和 τ_{conv} 。该标定准则利用分位数的稳健统计特性, 能够自适应当前ISAC场景下状态量的经验分布, 将恶劣状态的触发频率严格控制在约 q 的尾部概率范围内。

进一步, 基于该标定准则与后文的全局误差上界理论, 自适应策略选择逻辑定义如下: 条件A, 当 $Q(t) < \tau_{\text{comm}}$ 时, 链路干扰较大导致聚合易受影响, 采用FedTS策略引入历史平滑抑制波动; 条件B, 当 $H(t) < \tau_{\text{het}}$ 时, 系统异构性过高导致优化方向偏移显著, 采用FedProx策略缓解异构性带来的问题; 条件C, 当 $R(t) < \tau_{\text{conv}}$ 时, 收敛速率明显下降, 由于 $R(t)$ 与本地步数和资源差异相关, 采用FedNova策略进行归一化校正; 若上述尾部事件均未触发, 系统状态良好, 则默认采用FedAvg策略降低额外复杂度。优先级顺序为条件A>条件B>条件C。

1) FedAvg策略

FedAvg策略为默认策略, 在系统状况良好时, 服务器对重构梯度 $\hat{g}_i^{(t)}$ 执行加权平均得到全局梯度。

$$\hat{g}(t) = \sum_{i=1}^K \omega_i \hat{g}_i^{(t)} \quad (11)$$

2) FedTS策略

在无线通信质量较差、上行梯度噪声较大的情况下触发, 利用历史信息抑制瞬时噪声。根据置信度参数 $\lambda \in [0,1]$ 、当前轮FedAvg聚合结果 $\hat{g}_{\text{ideal}}(t)$ 和历史平滑聚合结果 $\hat{g}(t-1)$, 得到全局梯度。

$$\hat{g}(t) = \lambda \hat{g}_{\text{ideal}}(t) + (1 - \lambda) \hat{g}(t-1) \quad (12)$$

3) FedProx策略

在系统异构性过高时采用, 服务器应用正则化项约束聚合结果, 从而缓解异构性带来的优化方向差异问题。根据标准FedAvg聚合结果 $g_{\text{prox}}(t)$, 设 α 为调整系数, μ 为正则化系数, 得到

$$\hat{g}(t) = g_{\text{prox}}(t) - \alpha \mu (w^{(t)} - w^{(t-1)}) \quad (13)$$

4) FedNova策略

当观测到全局模型收敛速率明显下降时采用, 服务器通过归一化每个设备端的更新, 将每个设备端的更新缩放到统一步长, 从而消除步长差异的影响。基于各设备权重 ω_i 和本地计算轮数 E_i 计算归一化权重并执行聚合, 得到

$$\begin{cases} \omega_i^{\text{norm}} = \frac{\omega_i E_i}{\sum_{i=1}^K \omega_i E_i} \\ \hat{g}(t) = \sum_{i=1}^K \omega_i^{\text{norm}} \hat{g}_i^{(t)} \end{cases} \quad (14)$$

无论采取哪种聚合策略, 服务器均根据当前轮全局梯度 $\hat{g}(t)$ 更新全局模型。假设学习率为 η , 则更新规则为

$$w^{(t+1)} = w^{(t)} - \eta \hat{g}(t) \quad (15)$$

同时, 服务器对接收的各设备 AE 参数, 直接采用 FedAvg 策略聚合。

$$\theta^{(t+1)} = \sum_{i=1}^K \omega_i \theta_i^{(t)} \quad (16)$$

聚合后, 服务器将新一轮的全局主模型参数 $w^{(t+1)}$ 和全局 AE 参数 $\theta^{(t+1)}$ 广播下发给所有设备, 开始下一轮的训练。通过上述自适应机制, 系统能够对不同的通信和计算环境动态调整策略, 在保证全局模型稳健收敛的同时, 提高对信道噪声、设备异构等不利条件的容忍度。

2.4 FI-ISAC-FL 协同训练算法

综合上述设计, 算法 1 给出了完整协同训练流程。该算法包含设备端本地训练和服务器端全局聚合两大阶段。

算法 1 FI-ISAC-FL 协同训练算法流程

输入 参与设备数量 K , 最大全局通信轮次 T , 本地训练轮次 E , AE 训练步数 N_{AE} , 学习率 η

输出 全局主模型 $w^{(t+1)}$, 全局 AE 参数 $\theta^{(t+1)}$

1) 初始化 (Server side): 初始化全局主模型 $w^{(t)}$ 、全局 AE 参数 $\theta^{(t)}$ 和固定解码器 D_ϕ

2) for $t = 0, 1, \dots, T - 1$ do //全局通信循环

3) Server Broadcasts: 下发 $w^{(t)}$ 和 $\theta^{(t)}$ 给所有参与设备

4) for 设备 i in K in parallel do

//设备端本地更新

5) 初始化 (Client side): $w_i^{(t)} \leftarrow w^{(t)}$;
 $\theta_i^{(t)} \leftarrow \theta^{(t)}$

6) for local epoch $e = 1$ to E do

7) 在本地数据集 D_i 上训练主模型, 并收集梯度样本集 G_i

8) end for

9) 计算最终模型梯度 $g_i^{(t)}$

10) for step $n = 1$ to N_{AE} do

11) 基于 G_i 最小化 L_{AE} , 更新 AE 参数 $\theta_i^{(t)}$

12) end for

13) compression: 利用训练后的梯度 AE 压缩原始模型梯度, 生成低维特征表示 $s_i^{(t)} = E_{\theta_e}(g_i^{(t)})$

14) upload: 上传低维特征表示 $s_i^{(t)}$ 和梯度 AE 参数 $\theta_i^{(t)}$

15) end for

//服务器端全局聚合

16) reconstruction: 计算重构梯度 $\hat{g}_i^{(t)} = D_\phi(s_i^{(t)})$

17) adaptive aggregation: 检测系统状态 $S(t)$, 选择聚合策略得到聚合后全局梯度 $\hat{g}(t)$

//注意以下条件 A/B/C 已在 2.3 节中定义

if $S(t)$ 满足条件 A then 策略 = FedTS

else if $S(t)$ 满足条件 B then 策略 = FedProx

else if $S(t)$ 满足条件 C then 策略 = FedNova

else 策略 = FedAvg

18) update:

$$w^{(t+1)} = w^{(t)} - \eta \hat{g}(t)$$

$$\theta^{(t+1)} = \sum_{i=1}^K \omega_i \theta_i^{(t)}$$

19) end for

2.5 算法复杂度与系统开销分析

为量化不同方法在通信与计算上的额外代价, 本文从每轮上、下行通信量与设备端计算复杂度 3 个维度进行对比, 结果如表 2 所示。传统 FL 与 ISAC-FL 需要传输完整梯度, 通信量随梯度维度 d_g 线性增长; FedSem 通过重要性筛选降低上行开销, 但仍需传输稀疏梯度索引与数值; FedAE (以

FLAC工作为代表) 是基于潜在特征压缩的基线方法, 目标是降低重构误差并减少通信量, 其压缩表示未显式区分任务相关更新与冗余更新; 本文方法通过梯度特征压缩仅上传低维特征表示及轻量AE参数, 在极低通信预算下实现可用更新。

表2 计算复杂度对比

方法	每轮通信量 (上行)	每轮通信量 (下行)	计算复杂度 (设备端)
传统FL	$O(d_g)$	$O(d_g)$	$O(Ed_g)$
ISAC-FL	$O(d_g)$	$O(d_g)$	$O(Ed_g + \cos t_{\text{ISAC}})$
FedSem	$O(kd_g)$	$O(d_g)$	$O(Ed_g + \cos t_{\text{scl}})$
FedAE	$O(d_s)$	$O(d_g)$	$O(Ed_g)$
本文方法	$O(d_s + d_\theta)$	$O(d_g + d_\theta)$	$O\left(\frac{Ed_g + \cos t_{\text{ISAC}}}{N_{\text{AE}} d_s} + \right)$

在计算复杂度方面, 传统FL的主要开销来自本地反向传播与多轮迭代; ISAC-FL额外引入通信-感知协同相关开销; FedSem的额外计算复杂度主要来自梯度重要性计算与筛选; FedAE避免了在线更新带来的训练开销, 但其压缩表示未引入任务相关的特征保真约束; FI-ISAC-FL的额外计算复杂度则来自AE训练, 该额外计算复杂度与AE结构规模、梯度采样率 p 、AE更新步数 N_{AE} 及更新频率等因素相关, 属于可调节开销。

在系统实现上Federated-AE在线训练机制引入的额外开销包含两类: 在线训练编解码过程带来的运行开销和设备端额外维护AE模块带来的模型开销, 两者共同决定了本文方法在资源受限边缘设备上的可部署性。

从运行开销来看, Federated-AE的额外开销主要来自AE网络参数训练、梯度分块编码以及在线更新过程。为了刻画该额外开销, 本文将单轮总开销分解为主模型训练时延、AE附加训练时延和系统侧运行开销三部分, 即

$$T_{\text{round}} = T_{\text{train}} + T_{\text{AE}} + T_{\text{sys}} \quad (17)$$

其中, T_{AE} 的大小主要由AE网络规模与在线更新策略共同决定: AE网络规模由分块大小、隐藏层宽度和低维特征维度决定; 在线更新策略则由每轮AE训练步数 N_{AE} 、梯度采样率 p 与更新频率 R_{AE} 决定。这些超参数均可进行线性调节, 如AE网络规

模越大、分块维度越高、更新步数越多, 则设备端附加计算开销越显著; 反之, 通过减小分块大小、压缩隐层维度、降低更新频率或采用轻量化编码结构, 均可有效控制额外开销。因此, Federated-AE所带来的额外开销可通过结构设计与训练策略进行联合调节。

模型开销则是设备端维护AE所付出的参数规模与存储负担, 同样受AE网络规模的影响。但由于本文方法在主模型设计上采用了轻量的端到端网络模型, 该模型本身带来的基础模型规模下降, 能够在一定程度上抵消AE带来的额外开销。

3 收敛性分析

在FI-ISAC-FL中, 梯度压缩与重构误差会直接影响全局模型的更新。为分析误差传播规律及其对收敛性的影响, 从单轮聚合的梯度表达入手, 推导特征压缩误差项在多轮迭代下的累积效应, 并给出收敛边界的理论说明, 以增强本文框架的合理性与稳定性。

3.1 压缩-解码误差建模

设备端对模型梯度进行特征压缩编码, 服务器对各设备上传的低维梯度特征表示进行解码重构, 在这一压缩-解码的过程中, 不可避免地会引入信息损失。将服务器对单个设备的重构梯度与原始梯度之间的压缩-解码总误差分解为

$$e_i^{(t)} = \hat{g}_i^{(t)} - g_i^{(t)} = \varepsilon_c + \varepsilon_d \quad (18)$$

其中, ε_c 表示特征压缩失真误差, ε_d 表示服务器固定解码器和设备端AE不匹配导致的重构失配误差。上述 ε_d 在第 t 轮可表示为

$$\mathbb{E}[\varepsilon_d] = \mathbb{E}_{g \sim P_t} \left[\left\| \hat{g} - \tilde{g} \right\|_2^2 \right] = \mathbb{E}_{g \sim P_t} \left[\left\| D_\phi \left(E_{\theta_e}(g) \right) - D_{\theta_d} \left(E_{\theta_e}(g) \right) \right\|_2^2 \right] \quad (19)$$

其中, P_t 为第 t 轮的真实梯度分布。当采用传统离线预训练或固定编码器方案时, P_t 的动态漂移会使该失配项难以随迭代轮次自然递减。相反, 本文通过设备端梯度AE的在线更新, 并结合联邦聚合提升跨端特征空间的一致性, 能够有效抑制 ε_d 的累积增长。这不仅从理论上收紧了后续的误差上界, 更为动态ISAC场景下联邦主模型的稳定收敛提供了理论保障。

设备端压缩-解码误差 $e_i^{(t)}$ 满足零均值与独立

性、方差有界假设, 存在常数 σ_{dec}^2 使

$$\begin{aligned} \mathbb{E}[e_i^{(t)}] &= 0, \{e_i^{(t)}\}_{i=1}^K \text{ 相互独立} \\ \mathbb{E}\left[\|e_i^{(t)}\|^2\right] &\leq \sigma_{\text{dec}}^2 \end{aligned} \quad (20)$$

其中, 常数 σ_{dec}^2 由特征压缩失真误差 ε_c 和重构失配误差 ε_d 共同决定, 其中 ε_c 大小取决于压缩率/低维特征维度, ε_d 则受 Federated-AE 在线训练机制抑制。

在 Federated-AE 在线训练机制下, 设备端训练完整的梯度 AE, 服务器端通过 FedAvg 策略得到全局 AE。通过多轮迭代, 编码器输出的特征空间逐步稳定并与服务器端解码映射保持一致, 使重构失配误差的期望 $\mathbb{E}\left[\|\varepsilon_d\|^2\right]$ 随通信轮次递减, 从而降低 σ_{dec}^2 , 改善整体重构误差的统计特性。

3.2 全局更新误差传播分析

全局梯度误差定义为聚合梯度与真实全局梯度的偏差, 不仅源于压缩-解码过程, 还受自适应聚合策略选择的影响。设真实平均梯度为 $\bar{g}(t) = \sum_{i=1}^K \omega_i g_i^{(t)}$, Δ_{dec} 为解码重构误差, Δ_{str} 为策略误差, 则全局梯度误差可表示为

$$\begin{aligned} \Delta(t) &= \hat{g}(t) - \bar{g}(t) = \\ &[\hat{g}_{\text{ideal}}(t) - \bar{g}(t)] + [\hat{g}(t) - \hat{g}_{\text{ideal}}(t)] = \\ &\Delta_{\text{dec}} + \Delta_{\text{str}} \end{aligned} \quad (21)$$

因此, 本文“自适应聚合”的作用并非简单替换聚合器, 而是在给定 Δ_{dec} (由特征压缩引入的重构误差) 的前提下, 通过控制 Δ_{str} 的上界来降低总偏差, 从而形成与特征压缩模块协同优化的聚合选择机制。

1) 解码重构误差 Δ_{dec}

源于梯度压缩-解码过程中的信息损失, 等价于单个设备压缩-解码总误差的累积结果。在无策略误差且设备同质的理想情况下, 直接对所有重构梯度执行标准 FedAvg 聚合得到

$$\begin{aligned} \hat{g}_{\text{ideal}}(t) &= \sum_{i=1}^K \omega_i \hat{g}_i^{(t)} = \\ &\sum_{i=1}^K \omega_i (g_i^{(t)} + e_i^{(t)}) = \\ &\bar{g}(t) + \sum_{i=1}^K \omega_i e_i^{(t)} \end{aligned} \quad (22)$$

因此, 解码重构误差可表示为

$$\Delta_{\text{dec}} = \hat{g}_{\text{ideal}}(t) - \bar{g}(t) = \sum_{i=1}^K \omega_i e_i^{(t)} \quad (23)$$

进一步, 当采用均匀聚合权重 $\omega_i = \frac{1}{K}$ 时, 其期望误差上界可以表示为

$$\begin{aligned} \mathbb{E}\left[\|\Delta_{\text{dec}}\|^2\right] &= \mathbb{E}\left[\left\|\sum_{i=1}^K \omega_i e_i^{(t)}\right\|^2\right] = \\ &\sum_{i=1}^K \omega_i^2 \mathbb{E}\left[\|e_i^{(t)}\|^2\right] \leq \frac{\sigma_{\text{dec}}^2}{K} \end{aligned} \quad (24)$$

其中, 有限常数 $\frac{\sigma_{\text{dec}}^2}{K}$ 为解码重构误差聚合的方差上界, 包含所有设备压缩-解码误差的累积效应, 整体误差随参与设备数增加呈 $O\left(\frac{1}{K}\right)$ 衰减。这表明在设备数量足够、误差无偏且相互独立时, 特征压缩导致的聚合偏差不会随训练累积失控。在 Federated-AE 在线训练机制下, σ_{dec}^2 随训练轮次逐步减小, 收敛至由压缩率和真实梯度范数共同决定的上界。

2) 策略误差 Δ_{str}

源于自适应聚合策略的选择所引入的偏差。以 FedAvg 策略为基准, 将其聚合结果设为理想聚合梯度, 不同策略的聚合结果与理想结果的偏差定义为策略误差。将策略误差上界统称为 ε_s , 在收敛性分析中, ε_s 是一个有界常数。这里统一假设真实聚合梯度二阶矩有界, 由梯度范数上限 G_{max}^2 和解码重构误差的聚合方差上界 $\frac{\sigma_{\text{dec}}^2}{K}$ 共同决定, 即

$$\mathbb{E}\left[\|\hat{g}(t)\|^2\right] \leq G_{\text{max}}^2 + \frac{\sigma_{\text{dec}}^2}{K}.$$

对于 FedTS 策略, 定义其误差上界为 $\varepsilon_s = (1 - \lambda)^2 \left(2G_{\text{max}}^2 + 2\frac{\sigma_{\text{dec}}^2}{K}\right)$, 取值由梯度范数上限 G_{max}^2 和解码重构误差的聚合方差上界 $\frac{\sigma_{\text{dec}}^2}{K}$ 共同决定, 当 $\lambda \rightarrow 1$ 时有 $\varepsilon_s \rightarrow 0$ 。则有

$$\Delta_{\text{str}} = \hat{g}(t) - \hat{g}_{\text{ideal}}(t) = (1 - \lambda) [\hat{g}(t-1) - \hat{g}_{\text{ideal}}(t)] \quad (25)$$

$$\mathbb{E}\left[\|\Delta_{\text{str}}\|^2\right] = (1 - \lambda)^2 \mathbb{E}\left[\|\hat{g}(t-1) - \hat{g}_{\text{ideal}}(t)\|^2\right] \leq \varepsilon_s \quad (26)$$

对于 FedProx 策略, 定义其误差上界为 $\varepsilon_s = \alpha^2 \mu^2 \eta^2 \left(G_{\text{max}}^2 + \frac{\sigma_{\text{dec}}^2}{K}\right)$, 取值与正则化强度 α, μ 、学习率 η 以及梯度范数上界相关。则有

$$\Delta_{\text{str}} = \hat{g}(t) - \hat{g}_{\text{ideal}}(t) = -\alpha\mu(w^{(t)} - w^{(t-1)}) = \alpha\mu\eta\hat{g}(t-1) \quad (27)$$

$$\mathbb{E}\left[\|\Delta_{\text{str}}\|^2\right] = \alpha^2\mu^2\eta^2\mathbb{E}\left[\|\hat{g}(t-1)\|^2\right] \leq \varepsilon_s \quad (28)$$

对于 FedNova 策略, 定义其误差上界为 $\varepsilon_s = K^2 \max |\omega_i^{\text{norm}} - \omega_i|^2 \left(G_{\text{max}}^2 + \frac{\sigma_{\text{dec}}^2}{K}\right)$, 其中 $K^2 \max |\omega_i^{\text{norm}} - \omega_i|^2$ 衡量了归一化权重与原始权重之间的差异程度, 由设备本地训练步数 E_i 的异构性决定。当各设备 E_i 趋于一致时, $\omega_i^{\text{norm}} \rightarrow \omega_i, \varepsilon_s \rightarrow 0$ 。则有

$$\Delta_{\text{str}} = \hat{g}(t) - \hat{g}_{\text{ideal}}(t) = \sum_{i=1}^K (\omega_i^{\text{norm}} - \omega_i) \hat{g}_i^{(t)} \quad (29)$$

$$\mathbb{E}\left[\|\Delta_{\text{str}}\|^2\right] = \mathbb{E}\left[\left\|\sum_{i=1}^K (\omega_i^{\text{norm}} - \omega_i) \hat{g}_i^{(t)}\right\|^2\right] \leq$$

$$K \max |\omega_i^{\text{norm}} - \omega_i|^2 \sum_{i=1}^K \mathbb{E}\left[\|\hat{g}_i^{(t)}\|^2\right] \leq \varepsilon_s \quad (30)$$

设 $S = \{\text{FedAvg}, \text{FedTS}, \text{FedProx}, \text{FedNova}\}$ 为候选聚合策略集合, 并记各策略对应的误差上界为 ε_s 。若第 t 轮根据系统状态选择 $s^{(t)} \in S$ 使 $\varepsilon_{s^{(t)}} < \varepsilon_s$ 对任意固定策略 $s^{(t)}$ 成立 (或近似成立), 则有 $\mathbb{E}\left[\|\varepsilon_{s^{(t)}}\|^2\right] \leq \min_{s \in S} \varepsilon_s$ 。即自适应聚合在动态 ISAC 场景下获得不劣于任一单一固定聚合策略的更紧误差界。当信道恶化、异构增强或收敛停滞等状态导致某一误差项主导时, 选择器通过切换至对应策略进一步收紧策略误差项上界。

综上所述, 在 FI-ISAC-FL 中, 全局模型更新的偏差主要源于两个方面: 特征压缩带来的信息损失和自适应聚合策略引入的偏差。全局梯度误差的期望平方范数上界可表示为

$$\mathbb{E}\left[\|\Delta(t)\|^2\right] = \mathbb{E}\left[\|\Delta_{\text{dec}} + \Delta_{\text{str}}\|^2\right] \leq \frac{2\sigma_{\text{dec}}^2}{K} + 2 \min_{s \in S} \varepsilon_s \quad (31)$$

3.3 主模型收敛性分析

在 FL 中, 全局目标函数 $F(w)$ 定义为所有设备本地损失函数 $L_i(w)$ 的加权平均, 即

$$F(w) = \sum_{i=1}^K \omega_i L_i(w) \quad (32)$$

其中, $L_i(w) = \frac{1}{|D_i|} \sum_{(x,y) \in D_i} [\ell(f_w(\hat{x}), y)]$ 表示设备 i 上的本地损失函数。相应地, 全局目标函数的梯度

为各设备本地梯度的加权平均

$$\nabla F(w) = \sum_{i=1}^K \omega_i \nabla L_i(w) = \sum_{i=1}^K \omega_i g_i^{(t)} = \bar{g}(t) \quad (33)$$

在实际系统实现中, 采用 Adam 优化器 (见 2.1 节), 但为便于理论分析, 这里基于非凸优化的标准分析框架 (具体采用梯度范数收敛性分析) 对收敛性进行推导。对于非凸目标函数, 假设全局目标函数 $F(w)$ 满足 L -光滑性、存在全局最小值 F^* , 且真实梯度 $\nabla F(w^{(t)})$ 与全局梯度误差 $\Delta(t)$ 不相关。服务器端采用更新规则为 $w^{(t+1)} = w^{(t)} - \eta\hat{g}(t)$ 。根据 L -光滑性, 应用标准下降引理, 得到

$$F(w^{(t+1)}) \leq F(w^{(t)}) - \eta \langle \nabla F(w^{(t)}), \hat{g}(t) \rangle + \frac{L\eta^2}{2} \|\hat{g}(t)\|^2 \quad (34)$$

代入 $\hat{g}(t) = \nabla F(w^{(t)}) + \Delta(t)$, 并对式(34)两边取期望, 可进一步得到

$$\begin{aligned} \mathbb{E}[F(w^{(t+1)})] &\leq \mathbb{E}[F(w^{(t)})] - \\ &\eta \left(1 - \frac{L\eta}{2}\right) \mathbb{E}\left[\|\nabla F(w^{(t)})\|^2\right] + \\ &\frac{L\eta^2}{2} \mathbb{E}\left[\|\Delta(t)\|^2\right] \end{aligned} \quad (35)$$

在满足上述假设的情况下, 若学习率 $\eta \leq \frac{1}{L}$,

则有 $\frac{L\eta^2}{2} \leq \frac{\eta}{2}$, 对式(35)移项得到

$$\begin{aligned} \frac{\eta}{2} \mathbb{E}\left[\|\nabla F(w^{(t)})\|^2\right] &\leq \mathbb{E}[F(w^{(t)})] - \mathbb{E}[F(w^{(t+1)})] + \\ &\frac{L\eta^2}{2} \mathbb{E}\left[\|\Delta(t)\|^2\right] \end{aligned} \quad (36)$$

经过 T 轮全局迭代后, 梯度的平均平方范数满足上界

$$\begin{aligned} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\left[\|\nabla F(w^{(t)})\|^2\right] &\leq \frac{2(F(w^{(0)}) - F^*)}{\eta T} + \\ &2L\eta \left(\frac{\sigma_{\text{dec}}^2}{K} + \min_{s \in S} \varepsilon_s\right) \end{aligned} \quad (37)$$

由此可知, 非凸条件下的系统收敛性由两部分组成: 收敛速率项 $O\left(\frac{1}{T}\right)$, 随着迭代轮次增加而趋于 0; 误差底噪项, 由学习率、用户数量 K 及全局梯度误差决定。结果表明, 只要选取合适的衰减学习率, 该系统能够收敛到全局目标函数的稳定点。通过 Federated-AE 在线训练机制逐步减少压缩-解误差的

统计上界 σ_{dec}^2 ，以及通过自适应聚合策略来降低策略误差 ϵ_s ，可有效压缩该稳定点周围的误差邻域，从而提高模型收敛最终的准确度，并加快收敛速度。

4 仿真结果

4.1 仿真设置

为验证所提系统的有效性，本文采用文献[16]中的无线感知数据集 THREE_RADAR_3000 进行雷达目标识别任务的 FL 仿真。该数据集由 7 种人类动作组成，包括站立、成人快走、儿童快走等，各类样本数量存在轻微类别不平衡，用于模拟边缘计算场景下的 Non-IID 数据分布。在主模型选择方面，使用端到端网络模型作为本地模型。该模型由轻量级 CNN 编码器和分类器头组成，包含 3 层卷积层特征提取，以及全连接层输出分类结果。全局总训练轮次设置为 $T = 50$ ，在每轮全局通信中，设备端进行 $E = 5$ 轮本地迭代，使用 Adam 优化器进行优化，初始学习率设置为 0.001，并结合余弦退火策略随训练逐步衰减。服务器统一随机初始化模型初始参数，训练过程中设备端之间不直接通信，所有更新均由服务器进行聚合。仿真实验在 Python 环境下实现，使用 PyTorch 深度学习框架和 Matplotlib 绘图库完成结果可视化。

本文选择雷达目标识别作为基准任务，主要出于以下考虑：雷达是 ISAC 场景中最典型的感知链路之一，能够较好地复现感知与通信共享带来的链路波动与传输瓶颈；本文关注的是更新上传阶段的特征压缩机制，其有效性可在单一感知模态上得到清晰验证。多模态输入会引入复杂的设备端侧特征编码与跨模态融合，对任务网络与计算开销提出更高要求。本文将其作为后续工作，在不改变所提压缩与聚合框架的前提下扩展到多模态 FL。

4.2 评价指标

1) 分类准确率 (accuracy, Acc): 全局模型在所

有测试数据上分类正确的比例，反映模型识别精度。

2) 总通信成本: 总上行通信量，单个设备总通信量定义为模型大小、压缩率与全局训练轮数的乘积，单位为 MB。

3) 通信效率(Acc(B)): 固定通信预算条件下可达到的最佳测试准确率。

4) 联合预算效率 (Acc(B, τ)): 在通信与时间联合受限条件下能够达到的最佳测试准确率。

5) 精确率、召回率和 F1 值 (F1-Score): 精确率反映模型对正类的准确判定能力，召回率反映模型对正类的召回能力，F1 值是精确率和召回率的调和平均。

6) Kappa 系数: 反映模型预测结果与真实标签的一致性程度，考虑了随机一致性因素。

4.3 结果与分析

为验证所提系统的有效性与优势，本文设计了 5 种对比方案，如表 3 所示。1) 传统 FL 不采用任何压缩或物理层融合机制，通信开销较高，被视为性能、开销上界；2) 集成感知通信的联邦学习 (ISAC-FL) 参考文献[15]中的设计，考虑了 ISAC 系统的特殊应用场景，是 ISAC 场景下的基线方法，聚合方式采用 AirComp 空中计算机制，实现梯度的空中求和，服务器直接得到聚合结果；3) 联邦语义学习 (FedSem) 通过梯度重要性计算，仅传输 Top-K 重要梯度分量，实现了部分通信压缩；4) 自编码器联邦学习 (FedAE) 代表离线或固定压缩器范式，设备端使用预训练或固定的 AE 实现极高的通信压缩；5) 本文提出的基于特征信息的 ISAC 联邦学习 (FI-ISAC-FL) 方法，利用 Federated-AE 在线训练机制训练梯度 AE，实现特征压缩，同时加入自适应聚合选择，以适应动态复杂的环境。同时，为验证所提框架对不同主模型形态的可迁移性与部署适配性，并兼顾公平对比与轻量化落地两类需求，FI-ISAC-FL 分别采用与基线一致的

表 3 对比方案设置

方案	核心机制	主模型	传输内容	聚合方式
传统 FL	无压缩	ResNet10	完整模型梯度	FedAvg
ISAC-FL	联合感知与通信优化+AirComp 机制	ResNet10	完整模型梯度	AirComp
FedSem	梯度重要性筛选	ResNet10	Top-K 重要梯度 (K 取 20%)	FedAvg
FedAE	固定 AE 压缩	ResNet10	梯度低维特征表示	FedAvg
本文方法 (FI-ISAC-FL)	Federated-AE 在线训练机制+特征压缩+自适应聚合	ResNet10 End2End	梯度低维特征表示	自适应策略选择

ResNet10 主模型及端到端轻量网络模型 (End2End) 进行对照实验。FI-ISAC-FL 融合了梯度特征压缩的通信优势与自适应聚合的鲁棒性优势, 预期能在各项评价指标上达到最优的权衡性能。

收敛性能对比如图2所示。由图2可知, 各方法初始准确率均约为14%, 在前5轮内准确率快速攀升至70%~80%, 随后进入缓慢爬升阶段, 并在35轮后逐步趋于平稳, 整体呈现出先快后慢、稳定收敛的典型FL训练特征。从训练中后期来看, FedSem、FedAE和FI-ISAC-FL (ResNet10) 的曲线整体位于第一梯队, 最终准确率分别为91.58%、91.20%和91.35%, 均在91%以上。传统FL最终准确率约为90.67%, 略低于第一梯队; ISAC-FL最终准确率约为88.57%; FI-ISAC-FL (End2End) 的最终准确率略低约为87.26%。总体来说, 所有方法均呈现稳定上升并最终收敛的趋势, 说明训练过程稳定。同时, 在同一FI-ISAC-FL框架下, ResNet10做为主模型能带来更高的最终性能, 而End2End会牺牲部分精度。

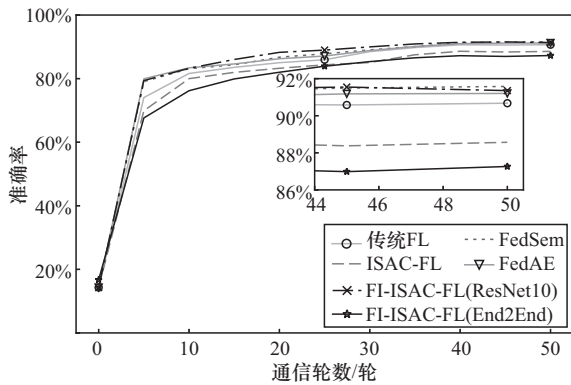


图2 训练收敛曲线

帕累托效率前沿分析如图3所示, 可以观察到, 传统FL与ISAC-FL位于图的右侧区域, 虽然最终准确率分别约为90.67%与88.57%, 但由于需要传输完整模型更新, 单个设备在50轮内的通信开销达到2525 MB。相比之下, 任务驱动的特征信息传输带来明显收益: FedSem在保持约91.58%准确率的同时将通信开销降低至约505 MB; FedAE进一步将通信开销压缩到 10^{-1} MB量级, 且准确率仍保持在91%左右。值得注意的是, FI-ISAC-FL (ResNet10) 与FI-ISAC-FL (End2End) 均落在极低通信开销区域, 实现了相对于传统FL约4个数量级的通信开销降低。其中ResNet10版本在低通信开销下仍能保持91.35%

的最终准确率, 为公平对比下的最优表现; End2End版本则由于更轻量的主模型配置, 出现更明显的精度回落。整体上, 图3清晰地表明了FI-ISAC-FL在带宽受限场景下具备显著的通信优势。

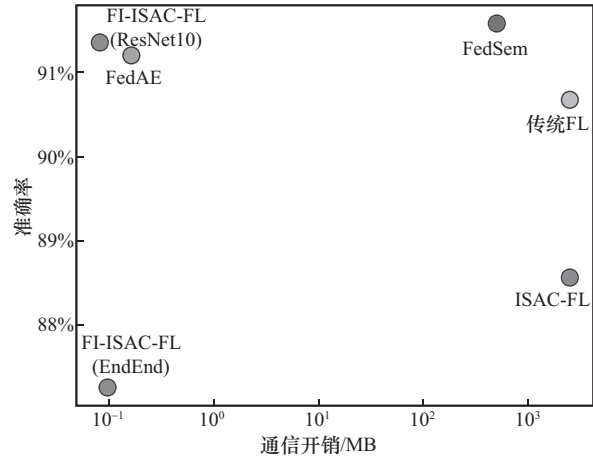
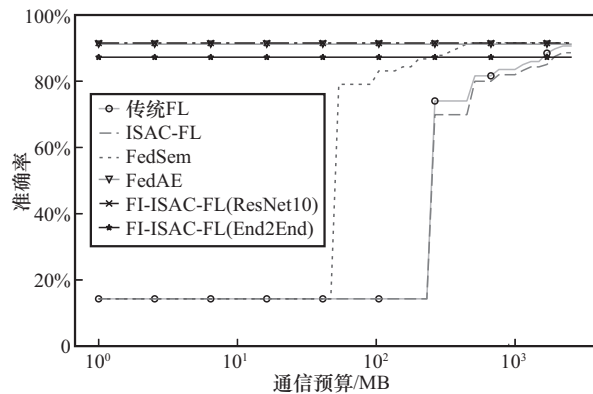


图3 帕累托效率前沿分析

准确率-累计通信预算曲线如图4所示, 本文采用预算受限指标 $Acc(B)$ 衡量通信效率, 即在累计通信量不超过预算B的条件下可达到的最佳测试准确率。可以观察到, 传统FL与ISAC-FL在小预算区间内准确率接近随机水平, 需在较大通信预算下才能进入有效收敛阶段, 表明完整梯度传输会显著延迟资源受限场景下的学习过程。FedSem能在中等预算范围内更早提升准确率, 说明任务相关表示机制能够一定程度缓解通信约束下的收敛迟滞; FedAE与FI-ISAC-FL (ResNet10) 在极小预算下即可保持较高准确率, 表明基于梯度特征表示的可学习压缩对通信-精度权衡的显著改善; FI-ISAC-FL (End2End) 同样在低预算下保持稳定的可达性能, 但由于主模型更轻量, 其精度上限略低, 反映了模型容量与资源开销之间的可控折中。



各方案在多维指标下的归一化综合性能分布如图5所示。从精度和一致性维度来看，在 Accuracy、F1-Macro 和 Kappa 这3项指标上本文所提 FI-ISAC-FL (End2End) 略低于其他方法。在通信受限维度 ($B = 200\text{MB}$) 上，FI-ISAC-FL 的两种实现仍保持较高的预算内性能。进一步，在通信与时间联合约束维度 ($B = 200\text{MB}, \tau = 1.25\text{h}$) 上，ResNet10 版本因本地计算与 AE 更新开销较大出现一定回落，而 End2End 版本凭借轻量计算仍保持较高可达性能。

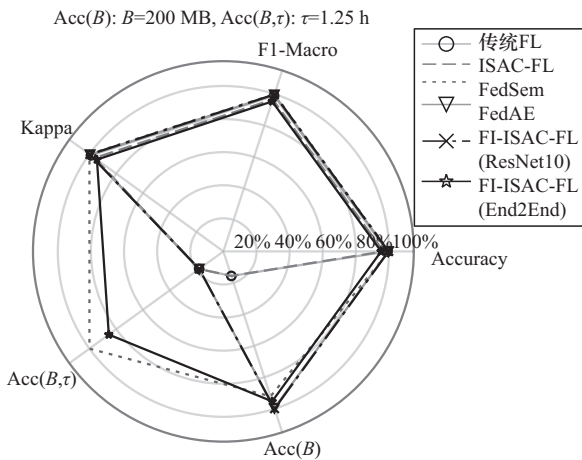


图5 综合性能分布

细粒度分类性能对比如图6所示。从图6可以看到，所有方法在类别 C5 和 C6 上均达到极高的 F1 值，表明这些类别相对容易分类。在类别 C1 和 C3 上，所有方法的 F1 值均较低，表明这些类别存在

一定的分类难度。进一步对比不同方法可发现，任务驱动的特征信息传输机制的三类方法 (FedSem、FedAE 和 FI-ISAC-FL (ResNet10)) 在多数类别上的 F1 值整体高于传统 FL 与 ISAC-FL，并在困难类别上体现出更明显的增益趋势。这说明，特征相关的筛选、压缩与表示学习能够更有效地保留对全局目标贡献更大的判别性更新信息，从而提升困难类别的判别能力，提升整体泛化表现。相比之下，FI-ISAC-FL (End2End) 宏平均 F1 值约为 0.873，其中在类别 C0/C2/C3 上的下降更明显，表明 End2End 模型在处理更依赖细粒度特征的类别时更易受到信息瓶颈影响。总体而言，图6证明了特征信息传输机制能显著优化“难类”识别性能，而 End2End 版本则是在维持“易类”高水准的前提下，在“难类”上实现了性能与开销的帕累托平衡。

为进一步验证所提方法在不同 ISAC 场景配置下的稳定性与适用性，本文从信道条件、客户端规模、压缩率和数据异构程度 4 个维度补充了场景扩展仿真，其中，不同压缩率设置间接表征了不同通信预算约束下的传输条件，最终准确率结果如图7所示。从信道条件来看，本文方法在稳定、慢时变和高动态信道下最终准确率的最大差异仅有 0.3%，表明其对信道波动具有较好鲁棒性；在客户端规模方面，当客户端数量由 10 增至 40 时，最终准确率由 80.7% 下降至 70.4%，说明随着联邦参与规模增大，客户端间更新差异和聚合难度同步增加，该结果说明大规模设备参与场景仍是本文方法需要进一步

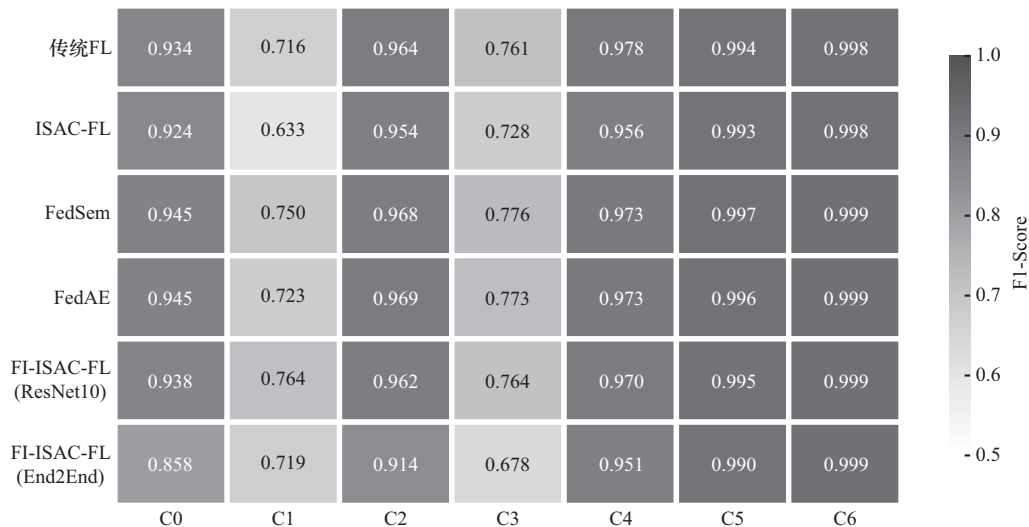


图6 细粒度分类性能对比

步优化的方向;从压缩率来看,在0.05、0.1和0.2这三种设置下(数值越小压缩程度越高),最终准确率整体波动较小,最大差异约为1.1%,说明本文方法在通信受限场景中仍能够较好地保留任务相关的关键信息,体现出一定的带宽适应能力;在数据异构程度上,当Dirichle参数 α 由1.0降至0.1时,模型性能明显下降,尤其在重度Non-IID条件下准确率降低更为显著,表明统计异质性仍然是影响联邦优化效果的关键因素,也说明本文方法在极端异构场景下仍存在进一步提升空间。

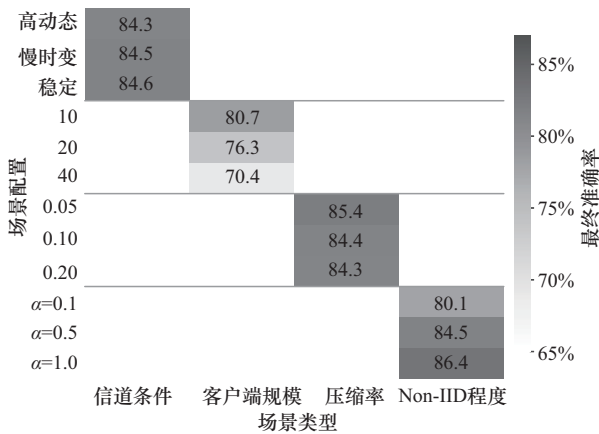


图7 不同场景配置下FI-ISAC-FL的最终准确率

综合来看,本文方法在不同信道动态和压缩率设置下表现出较好的稳定性,而在大规模客户端参与和强数据异构场景下仍存在进一步提升空间。结果表明,本文方法在多种场景配置下具有一定的鲁棒性与适应能力。

4.4 系统开销评估

为评估Federated-AE带来的额外开销问题,本节从运行开销和模型开销两个方面对本文方法进行了系统级评估。首先,不同方法的单轮时延分解图可直观展现运行开销,结果如图8所示。FI-ISAC-FL的平均单轮时延约为94.7s,其中AE训练时延为1.3s。可以看到本文方法中AE模块确实引入了额外运行开销,但其占总时延的比例较小。相较之下,由轻量主模型带来的本地训练开销下降,及特征信息压缩带来的系统侧开销下降更为明显,使总体单轮时延优于其他基线方法。

设备端参数量对比如图9所示。由图9可知,FI-ISAC-FL虽然额外引入了AE开销,且额外开销相对于其端到端网络模型的占比较高,但总参数量为5.62M,显著低于其他基线方法。这说明,本文虽然

因引入Federated-AE在线训练机制带来了额外开销,但通过轻量化主模型与压缩模块协同设计,设备端模型开销整体上仍保持在较低水平。同时,由于其他基线方法使用的主模型并非标准的ResNet10,而是面向ISAC场景构建的,在标准ResNet10的基础上增加了感知增强、通信优化和计算资源分配等ISAC专用模块,因此参数量达到13.24M。

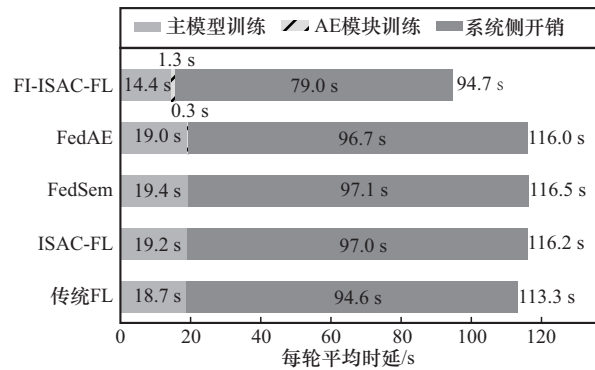


图8 单轮时延分解结果

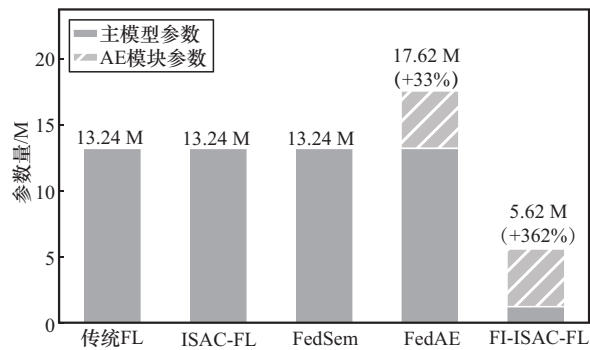


图9 设备端参数量对比

5 结束语

面向ISAC-FL场景下通信受限、网络状态动态变化等问题,本文提出了一种特征信息驱动的联邦学习压缩与聚合协同方法。该方法通过Federated-AE在线训练机制对本地模型更新进行特征压缩,并结合系统状态设计自适应聚合策略,在降低通信开销的同时兼顾了模型训练性能与稳定性。理论分析与仿真实验结果表明,本文方法在资源受限条件下具有较好的有效性与适应性。

参考文献:

[1] Letaief K B, Shi Y M, Lu J M, et al. Edge artificial intelligence for 6G: vision, enabling technologies, and applications[J]. IEEE Journal on Selected Areas in Communications, 2022, 40(1): 5-36.

- [2] Tan D K P, He J, Li Y C, et al. Integrated sensing and communication in 6G: motivations, use cases, requirements, challenges and future directions[C]//Proceedings of the 2021 1st IEEE International Online Symposium on Joint Communications & Sensing (JC&S). Piscataway: IEEE Press, 2021: 1-6.
- [3] Zhou Y, An Q C, Wang Z B, et al. Integrated sensing, computation, and communication enabled federated edge learning[J]. IEEE Transactions on Wireless Communications, 2026, 25: 7117-7131.
- [4] Wen D Z, Zhou Y, Li X Y, et al. A survey on integrated sensing, communication, and computation[J]. IEEE Communications Surveys & Tutorials, 2025, 27(5): 3058-3098.
- [5] McMahan H B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data[PP]. V4. (2016-02-17)[2026-01-04]. arXiv: arXiv.1806.00582.
- [6] Lim W Y B, Luong N C, Hoang D T, et al. Federated learning in mobile edge networks: a comprehensive survey[J]. IEEE Communications Surveys & Tutorials, 2020, 22(3): 2031-2063.
- [7] Alistarh D, Grubic D, Li J, et al. QSGD: communication-efficient SGD via gradient quantization and encoding[PP]. V4. (2016-10-07)[2026-01-04]. arXiv: arXiv.1610.02132.
- [8] Lin Y J, Han S, Mao H Z, et al. Deep gradient compression: reducing the communication bandwidth for distributed training[PP]. V3. (2017-12-05)[2026-01-04]. arXiv: arXiv.1712.01887.
- [9] Zhuansun Y, Li D D, Huang X H, et al. Communication-efficient federated learning with adaptive compression under dynamic bandwidth[PP]. V1. (2024-05-06)[2026-01-04]. arXiv: arXiv.2405.03248.
- [10] Beitollahi M, Lu N. FLAC: federated learning with autoencoder compression and convergence guarantee[C]//Proceedings of the GLOBECOM 2022 - 2022 IEEE Global Communications Conference. Piscataway: IEEE Press, 2022: 4589-4594.
- [11] Gündüz D, Qin Z J, Aguerri I E, et al. Beyond transmitting bits: context, semantics, and task-oriented communications[J]. IEEE Journal on Selected Areas in Communications, 2023, 41(1): 5-41.
- [12] Wang H, Li H R, Chen H M, et al. FedSC: federated learning with semantic-aware collaboration[C]//Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2. New York: ACM Press, 2025: 2938-2949.
- [13] Dong X N, Zeng J, Wen J H, et al. SFL: a semantic-based federated learning method for POI recommendation[J]. Information Sciences, 2024, 679: 121057.
- [14] 马千飘, 贾庆民, 刘建春, 等. 异构边缘计算环境下异步联邦学习的节点分组与分时调度策略[J]. 通信学报, 2023, 44(11): 79-93.
- Ma Q P, Jia Q M, Liu J C, et al. Client grouping and time-sharing scheduling for asynchronous federated learning in heterogeneous edge computing environment[J]. Journal on Communications, 2023, 44(11): 79-93.
- [15] Wen D Z, Xie S J, Cao X W, et al. Integrated sensing, communication, and computation for over-the-air federated edge learning[J]. IEEE Transactions on Wireless Communications, 2026, 25: 2748-2762.
- [16] Liu P X, Zhu G X, Wang S, et al. Toward ambient intelligence: federated edge learning with task-oriented sensing, computation, and communication integration[J]. IEEE Journal of Selected Topics in Signal Processing, 2023, 17(1): 158-172.
- [17] 王毅, 瞿治国, 孙乐. 结合自注意力机制和软阈值降噪的对比联邦学习特征聚合算法[J]. 重庆邮电大学学报(自然科学版), 2024, 36(5): 966-973.
- Wang Y, Qu Z G, Sun L. Self-attention mechanism and noise reduction techniques combined with federated comparative learning feature aggregation algorithm[J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2024, 36(5): 966-973.
- [18] 张晓琴, 金西兴, 陆艳军, 等. 注意力机制驱动的个性化联邦学习特征分离方法[J]. 计算机应用研究, 2025, 42(4): 1102-1107.
- Zhang X Q, Jin X X, Lu Y J, et al. Attention-driven feature separation method for personalized federated learning[J]. Application Research of Computers, 2025, 42(4): 1102-1107.
- [19] Zhao Y, Li M, Lai L Z, et al. Federated learning with non-IID data[PP]. V2. (2022-07-21)[2026-01-04]. arXiv: arXiv.1806.00582.

[作者简介]



邓炳光 (1978-), 男, 广东东莞人, 重庆邮电大学副教授, 主要研究方向为通信网与测试技术、物联网技术。



彭佳音 (1999-), 女, 四川金堂人, 重庆邮电大学硕士生, 主要研究方向为通信网与测试技术、智能通信与边缘智能。