控制与决策
Control and Decision

# 基于稳定特征原型的云边协同联邦类别增量学习方法

姚邹静, 赵春晖

引用本文:
姚邹静, 赵春晖. 基于稳定特征原型的云边协同联邦类别增量学习方法[J]. 控制与决策, 2025, 40(4): 1267-1275.

在线阅读 View online: https://doi.org/10.13195/j.kzyjc.2024.0739

---

# 基于稳定特征原型的云边协同联邦类别增量学习方法

姚邹静，赵春晖†

(浙江大学 控制科学与工程学院，杭州 310027)

**摘　要:** 由于存储空间限制, 物联网中的边缘设备往往仅能保留当前某个有限时段内的数据. 实际生产过程中, 设备工况在一定时间内发生变动, 产生新类别的故障数据或图像, 这种类别增量会造成模型在本地训练时产生灾难性遗忘. 在单边端类别增量的局部灾难性遗忘基础上, 随着云边协同优化, 灾难性遗忘会产生扩散. 针对上述问题, 提出一种基于稳定特征原型的联邦类别增量学习方法, 在边端建立类别样本记忆库存储类别代表性样本, 设计基于回放范式的原型网络更新策略, 在云端设计以统一特征空间下的特征原型为参考基准的加权聚合策略, 在联邦框架下稳定优化特征空间, 实现类别知识的联邦更新. 基于类别增量常用的数据集 CIFAR10 和 Mini-ImageNet 的实验验证了所提方法可以有效缓解灾难性遗忘.

**关键词:** 图像分类；联邦学习；增量学习；灾难性遗忘；云边协同；特征原型

**中图分类号:** TP183　　**文献标志码:** A

**DOI:** 10.13195/j.kzyjc.2024.0739

**引用格式:** 姚邹静, 赵春晖. 基于稳定特征原型的云边协同联邦类别增量学习方法 [J]. 控制与决策, 2025, 40(4): 1267-1275.

# Could-edge collaborative federated class-incremental learning with consistent feature prototypes

*YAO Zou-jing*，*ZHAO Chun-hui*†

(College of Control Science and Engineering，Zhejiang University，Hangzhou 310027，China)

**Abstract:** Due to limited storage, edge devices in the Internet of Things(IoT) usually retain data for a limited time period. In real production processes, the equipment conditions change over time, often generating new classes of fault data or images. This class increment can cause catastrophic forgetting when the model is trained locally. Based on the partial catastrophic forgetting of class increment on a single edge, catastrophic forgetting will spread with the collaborative optimization of cloud and edge. To address the above problems, a federated class incremental learning method based on stable feature prototypes is proposed. A class sample memory is established at the edge to store representative samples of the class. A feature network update strategy based on the replay paradigm is designed. A weighted aggregation strategy based on feature prototypes in a unified feature space is designed in the cloud. The feature space is stably optimized in the federated framework to realize the federated update of class knowledge. Experiments on CIFAR10 and Mini-ImageNet, which are commonly used datasets for class increment, demonstrate that the proposed method can effectively alleviate catastrophic forgetting.

**Keywords:** image classification； federated learning； incremental learning； catastrophic forgetting； cloud-edge collaboration；feature prototype

## 0　引　言

在物联网和大数据时代, 深度学习方法在处理常见应用任务如故障诊断[1]、成分或结构预测[2-3]、图像分类分割[4] 等方面已有很多应用. 这些方法通常需要基于大量历史数据进行离线建模而后在线应用, 且仅考虑单个设备或工厂的建模需求. 而在瞬息万变的开放世界中, 数据往往以流式出现, 同时各工厂间需协同合作. 随着工业物联网 (industrial internet of things, IIoT) [5]、边缘计算[6] 等技术蓬勃发展, 云边协同[7-9] 的建模方法打破了工厂间的数据孤岛. 其中, 联邦学习方法[10-13] 可以在保护数据隐私的同时融合来自不同设备或工厂的多方知识. 如何在联邦学习

**收稿日期:** 2024-06-21；**录用日期:** 2024-09-10.

**基金项目:** 浙江省“尖兵”“领雁”研发攻关计划项目 (2024C01163)；国家自然科学基金杰出青年基金项目 (62125306)；工业控制技术全国重点实验室项目 (ICT2024A06).

†通信作者. E-mail: chhzhao@zju.edu.cn.

框架下从非平稳数据流中进行增量学习, 是深度学习领域的一个重要问题.

近年来, 有大量的研究致力于解决类别增量学习 (class-incremental learning, CIL) 问题, 在持续学习新的类别知识的同时, 解决模型更新对于旧类别知识的灾难性遗忘问题[14-15]. 典型解决方案如 packNet[16] 和 EWC[17] 等在任务增量和域增量问题中表现良好, 但是难以应对复杂的类别增量场景. 参数正则化方法如 COIL[18] 等保持模型结构固定不变, 评估每个参数对网络的重要性, 使重要参数变化不大以保留先前的知识, 但估计参数重要性的内存预算呈线性增长. 基于模型校正的方法如 OML[19] 旨在发现和减少增量模型中的偏差, 通过元学习的方法寻找更为稳定和稀疏的特征表示, 以此对网络参数进行校正. 动态网络方法[20] 往往会在有新任务时扩展网络, 需要可扩展的内存预算, 在硬件条件有限的边缘设备上难以应用. 回放类方法如 ICARL[21] 构建样本集保存了训练集的一小部分参与后续模型训练; 生成类回放方法[22] 则不存储样本本身, 而是通过生成类模型存储历史知识. 基于知识蒸馏的方法[23] 建立了旧模型与新模型之间的映射, 知识从教师模型传递到学生模型, 尽管仍会遗忘, 但可以在学习新知识和记住旧知识之间取得一定平衡.

以上方法仅针对开放世界中单个设备或工厂的灾难性遗忘问题进行了针对性处理. 目前, 面对多设备或工厂的建模需求, 已有一些联邦类别增量方法的研究, 但大多为已有类别增量学习方法的联邦版本. 相较于基于回放的类别增量学习方法, Qi 等[24] 提出的联邦生成式回放方法在云端设置了全局生成器, 对本地生成器和全局生成器的特征增加约束. 然而生成类的方法间接上传了本地数据, 会破坏数据隐私. 相较已有的动态网络终身学习方法, Yoon 等[25] 提出的联邦版本额外学习了一个注意力模块, 存储空间开销在任务较多时较大. 相较于模型蒸馏的类别增量学习方法, FLwF-2T[26] 额外增加了一个存储全局知识的教师模型, 但使用教师模型存储知识有限, 仍存在遗忘问题.

由于涉及协作训练, 多边端下的联邦类别增量建模问题与单边端相比更为复杂, 存在数据的时空双重异构. 多边端类别增量联邦建模场景下, 灾难性遗忘[14] 叠加模型漂移问题[27-28] 在多边端相互扩散, 随云边迭代优化而持续传播. 针对联邦增量学习场景时空双重异构下灾难性遗忘的持续扩散问题, 本文提出一种基于稳定特征原型的联邦类别增量学习方法 (federated class-incremental learning with consistent feature prototypes, FCIL-COFPO), 核心思想在于在边端存储类别样本, 云端以上一轮特征空间下的特征原型为校准基准, 加权聚合边端模型. 具体而言, 从考虑类别知识协同、优化统一的特征空间角度出发, 在边端建立类别样本记忆库存储类别代表性样本, 设计基于回放范式的原型网络更新策略, 在云端设计以统一特征空间下的特征原型为参考基准的加权聚合策略, 从而在联邦框架下稳定优化特征空间. 其中, 全局模型的原型网络及特征原型均由各边端模型和特征原型加权得到, 实现类别知识的联邦更新. 所提出方法主要贡献如下:

1) 提出以类别知识建立稳定特征原型的联邦类别增量学习框架, 实现模型参数的稳定聚合.

2) 设计类别记忆库支持的原型网络更新策略, 通过动态维护类别记忆库缓解类别增量情况下的灾难性遗忘问题.

3) 设计基于特征原型间隔的加权校准策略, 间接依靠类别相关知识实现模型聚合阶段的校正.

## 1 问题阐述与研究动机

假设有$C$个边端, 形成边端集合$\mathcal{C}=\{\mathrm{Edge}_i\}_{i=1}^{C}$. 随着数据的不断产生, 每个边端存在不断增长的类别集合 $\mathcal{M}_i^t=\{p(y=j|x)\}_{j=1}^{M_i}$. 其中: $M$为边端最终应区分的类别数, $p(y=j|x)$为给定输入 $x$ 标签为 $j$ 的概率. 云端需要整合各边端对类别超集 $\mathcal{M}=\bigcup_{i=1}^{C}\mathcal{M}_i^t$的知识, 将可以分辨$\mathcal{M}_t$的模型部署到边端集合$\mathcal{C}$中. 边端$\mathrm{Edge}_i$可以通过在一组带标签的样本$\{(x,y)|(x,y)\in X_j\times Y_j\}$上进行训练接触到新类别. 将一个任务 (task) 定义为向边端发送模型训练要求, 告知边端并行收集数据并训练本地模型, 最后将信息传递回云端的迭代过程. 在每个任务$t$中, 每个边端均新产生一个来自某组类别的数据集$\mathcal{D}^t$, 这组类别中可包含多个不同类别. 联邦增量学习旨在完成类别增量下模型的协同训练, 对边端已见类别的新样本进行准确识别, 如下所示:

$$\mathcal{L}=\frac{1}{|\mathcal{C}|}\sum_{i=1}^{|\mathcal{C}|}\frac{1}{J_i}\sum_{j=1}^{J_i}\frac{1}{K_{i,j}}\sum_{k=1}^{K_{i,j}}H(\hat{y}_{i,j,k},y_{i,j,k}),\tag{1}$$

$$\mathcal{L}=\frac{1}{|\mathcal{M}_t|}\sum_{j=1}^{|\mathcal{M}_t|}\frac{1}{K_{i,j}}\sum_{k=1}^{K_{i,j}}H(\hat{y}_{i,j,k},y_{i,j,k}),\tag{2}$$

$$\min_{t\in 1,\ldots,T}\mathbb{E}|\mathcal{L}_t|.\tag{3}$$

式 (1) 为边端平均分类损失, 其中 $J_i$为边端$\mathrm{Edge}_i$在其已见类别的总数, $K_{i,j}$为边端$\mathrm{Edge}_i$上类别$j$的样本数, $H$为样本$k$上类别预测值$\hat{y}_{i,j,k}$和类别

标签$y_{i,j,k}$的交叉熵. 式 (2) 为云端聚合后的在类别集合$\mathcal{M}_t$中的平均损失. 式 (3) 为类别增量情况下最终的目标函数, 需要在$T$个任务范围内最小化所有边端已见类别最新样本的期望损失.

对于上述联邦类别增量学习问题, 有以下认知:

认知 1: 实际工业生产中, 设备工况往往在一定时间内发生变化, 产生新类别的数据. 这种类别增量会造成模型在本地训练时产生灾难性遗忘.

模型分发给本地边端后, 该边端的模型优化方向迅速向使得新任务中的类别分类准确率更高的方向调整, 而对于在新任务中未体现的已见类别不加关注. 模型梯度传播在每个新任务后都出现在类别分类任务上的不平衡, 无法在增量任务中保持类间语义一致性. 因此, 有必要从类别层面出发, 设计合适的基础模型构建方法, 缓解由类别增量导致的语义不一致问题.

认知 2: 多边端下的类别增量实际为数据时空双重异构, 在联邦迭代优化时造成全局灾难性遗忘.

联邦优化时, 各边端模型在云端聚合, 将不同边端的梯度不平衡和语义不一致混叠, 使得原本的局部灾难性遗忘随着模型聚合与下发迅速扩散. 因此, 有必要从全局的角度出发, 寻找不受类别增量影响的稳定且一致的语义空间, 并针对数据异构设计合适的模型校正策略.

## 2　FCIL-COFPO: 基于稳定特征原型的联邦类别增量学习

本节提出一种基于稳定特征原型的联邦类别增量学习框架 FCIL-COFPO, 总体结构如图 1 所示. 其中主要包含两个模块, 即在边端的基于回放范式的原型网络更新和在云端的基于特征原型间隔的加权校准. 下面对这两个模块进行展开介绍.

### 2.1　基于回放范式的原型网络更新策略

在边端设计一种基于回放范式的原型网络更新策略, 如图 2 所示. 该策略可以减小模型优化时由新旧任务变化导致的梯度传播偏差, 并在增量任务中保持类间语义一致性. 该策略主要包含 3 个部分, 即本地类别记忆库的动态更新、回放策略下的原型网络更新以及类别特征原型的更新.

1) 本地类别记忆库的动态更新.

对于边端$\text{Edge}_i$, 在任务$t$中, 会新产生一个来自某组类别的数据集$\mathcal{D}_i^t$, 此时边端已见类别集合为$\mathcal{M}_i^t$, $J_i^t$表示边端已见类别总数, $M$表示边端最终应区分的类别数. 边端首先进行本地类别记忆库$\mathcal{E}_i^t$的动态更新, 以补充新见类别样本, 维护旧有类别样本.

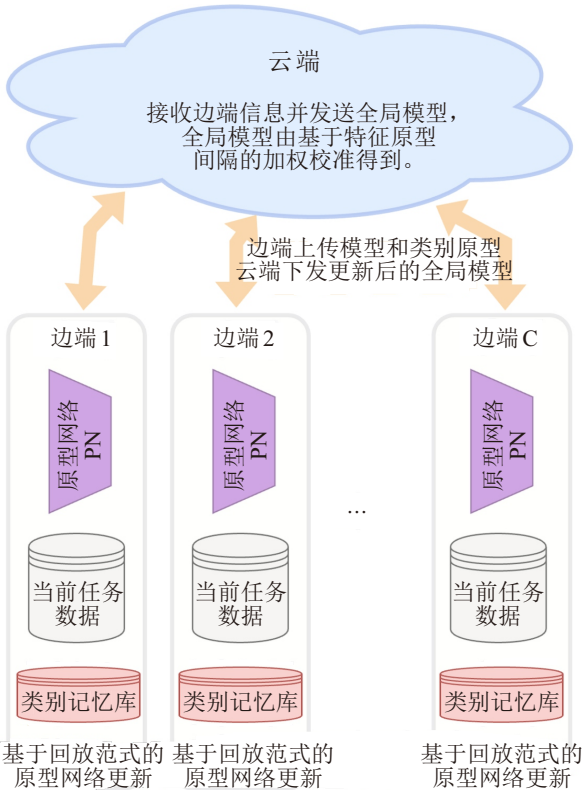


**图1　FCIL-COFPO 总体结构**

本地类别记忆库$\mathcal{E}_i^t=\{(\boldsymbol{x}_j,y_j)\}_{j=1}^M$是来自已见任务的实例集合. 在每个任务的模型训练过程之前管理本地类别记忆库, 面对不断变化的数据流, 类别记忆库对$\mathcal{E}_i^{t-1}$先按类别新增当前任务数据, 然后采用 Herding 策略选取每个类别最具代表性的样本. 给定类别$y$的实例集合$X=\{\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_n\}$, 选择时首先使用当前嵌入$\phi(\cdot)$计算类别中心, 有

$$\mu_y \leftarrow \frac{1}{n}\sum_{i=1}^{n}\phi(\boldsymbol{x}_i). \tag{4}$$

随后, 计算每个实例到类别中心的距离 $\|\mu_y-\phi(\boldsymbol{x}_i)\|$并按升序排列, 然后根据排名选择样本, 例如选择距离最近的前$P$个实例作为样本, 得到$\mathcal{E}_i^t$. 由于类别中心可以被视为每个类别最具代表性的模式, 选择靠近类别中心的样本也增强了样本的代表性. 在回放范式中, 边端$\text{Edge}_i$的模型利用 $\mathcal{E}_i^t\bigcup\mathcal{D}_i^t$ 进行每个任务内的更新.

2) 回放策略下的原型网络更新.

对于原型网络, 其输入应包含支持集$S$和查询集$Q$两部分, 用于训练特征提取器$f_\phi$, 使得其特征空间中的原型更有利于分类. 在边端$\text{Edge}_i$的任务$t$中, 支持集$S$和查询集$Q$由类别记忆库$\mathcal{E}_i^t$和新数据集$\mathcal{D}_i^t$产生.

原型网络本质上是学习一个映射函数, 使得样本可以映射到某个$O$维特征空间, 即$f_\phi:\mathbb{R}^D\rightarrow\mathbb{R}^O$.

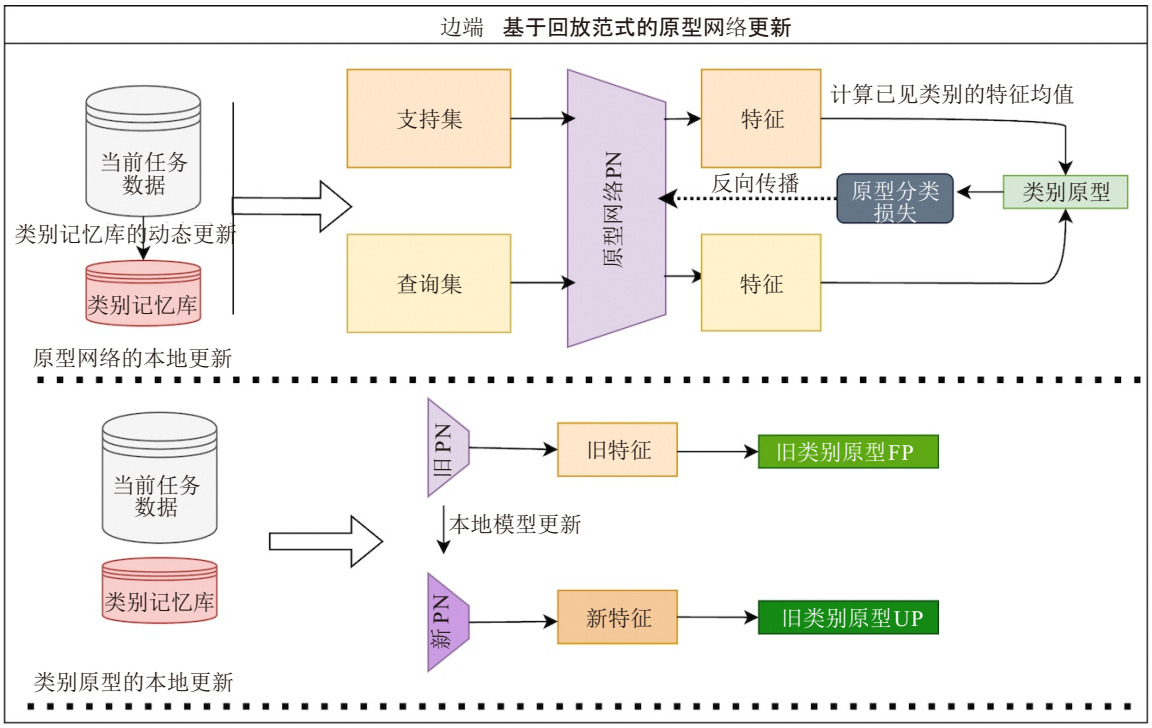


**图2 基于回放范式的原型网络更新策略**

设$S_{i,j}$为支持集中第$j$类样本的集合, 类别$j$的特征中心$\boldsymbol{p}[j]$可以通过直接计算$S_{i,j}$的特征均值得到, 即

$$\boldsymbol{p}[j]=\frac{1}{|S_i[j]|}\sum_{(x,y)\in S_i[j]}f_\phi(\boldsymbol{x}). \tag{5}$$

给定距离函数$d:\mathbb{R}^O\times\mathbb{R}^O\to[0,+\infty)$, 原型网络根据样本在特征空间中与相应类别原型的距离, 为查询点$\boldsymbol{x}$产生一个关于类别的分布, 如下所示:

$$p_\phi(y=k|\boldsymbol{x})=\frac{\exp\left(-d(f_\phi(\boldsymbol{x}),\boldsymbol{p}[j])\right)}{\sum\exp\left(-d(f_\phi(\boldsymbol{x}),\boldsymbol{p}[j'])\right)}. \tag{6}$$

优化时通过 SGD 最小化类别$j$的负对数概率$\mathcal{J}(\phi)=-\log p_\phi(y=j|\boldsymbol{x})$. 在训练过程中, 每轮次均从训练集中随机选择已见类别中的类别子集$\mathcal{M}_c$. 此时, 支持集包括当前类别下的类别记忆库数据以及新数据集$\mathcal{D}_i^t$, 即$S_i^t=\{(\boldsymbol{x},y)|(\boldsymbol{x},y)\in\mathcal{E}_i^t\bigcup\mathcal{D}_i^t, y\in\mathcal{M}_c\}$. 查询集包括此时类别记忆库中剩余数据, 即$Q_i^t=\{(\boldsymbol{x},y)|(\boldsymbol{x},y)\in\mathcal{E}_i^t, y\notin\mathcal{M}_c\}$. 算法 1 中提供了回放策略下的原型网络一轮次训练优化过程的伪代码, 损失函数为$\mathcal{J}(\phi)$, $J_i^t$为边端$\mathrm{Edge}_i$在任务$t$时的已见类别总数, $K_c$为当前选择类别子集中元素的数量, $K_c\leqslant K$. $N_q$为每个类别中查询集示例的数量, $N_q=P$. $\mathrm{RAND}(S,N)$表示从集合$S$中无重复地随机选择$N$个元素.

**算法 1** 输入: 边端 $\mathrm{Edge}_i$, 任务$t$对应的数据集$\mathcal{D}_i^t$, 此时边端已见类别集合为$\mathcal{M}_i^t$, 共包括$J_i^t$个类别, 边端类别记忆库为$\mathcal{E}_i^t$.

$\mathcal{M}_c\leftarrow\mathrm{RAND}(\{1,2,\dots,J_i^t\},K_c)$

$S_i^t=\{(\boldsymbol{x},y)|(\boldsymbol{x},y)\in\mathcal{E}_i^t\bigcup\mathcal{D}_i^t, y\in\mathcal{M}_c\}$

$Q_i^t=\{(\boldsymbol{x},y)|(\boldsymbol{x},y)\in\mathcal{E}_i^t, y\notin\mathcal{M}_c\}$

for $j$ in $\mathcal{M}_c$ do

$S_{i,j}^t=\{(\boldsymbol{x},y)|(\boldsymbol{x},y)\in S_i^t, y=j\}$

$\boldsymbol{p}_j^t=\frac{1}{|S_{i,j}^t|}\sum_{(x,y)\in S_{i,j}^t}f_\phi(\boldsymbol{x})$

end for

$\mathcal{J}\leftarrow 0$

for $j=1$ in $\mathcal{M}_c$ do

for $(\boldsymbol{x},y)$ in $Q_i^t$ do

$\mathcal{J}\leftarrow\mathcal{J}+\frac{1}{K_cP}\left[d(f_\phi(\boldsymbol{x}),\boldsymbol{p}[j]+\log\sum\exp\left(-d(f_\phi(\boldsymbol{x}),\boldsymbol{p}[j'])\right)\right]$

end for

end for

3) 类别特征原型的更新.

对于任务$t$, 边端$\mathrm{Edge}_i$已见类别集合为$\mathcal{M}_i^t$, 共包括$J_i^t$个类别, $M$表示边端最终应区分的类别数. 每个边端均在本地原型网络更新前后得到两组类别特征原型.

记更新前原型网络为$f_\phi^{t,0}$, 更新后为$f_\phi^{t,1}$. 根据更新前后的原型网络, 可以分别得到两组类别特征原型. 输入当前数据$\mathcal{D}_i^t$和更新后类别记忆库$\mathcal{E}_i^t$给$f_\phi^{t,0}$, 经过$f_\phi^{t,0}$可以得到旧特征原型组$\mathcal{FP}_i^t=\{\mathbf{po}_i^t[j]\}_{j=1}^M$. 保持输入不变, 经过$f_\phi^{t,1}$得到新特征原型组$\mathcal{UP}_i^t=$

$\{\mathbf{pn}_i^t[j]\}_{j=1}^M$. 其中, 类别原型$\mathbf{po}_i^t[j]$和$\mathbf{pn}_i^t[j]$如下所示:

$$\mathbf{po}_i^t[j]=\frac{1}{|\mathcal{AD}_j|}\sum_{(x,y)\in\mathcal{AD}_j} f_\phi^{t,0}(\boldsymbol{x}), \tag{7}$$

$$\mathbf{pn}_i^t[j]=\frac{1}{|\mathcal{AD}_j|}\sum_{(x,y)\in\mathcal{AD}_j} f_\phi^{t,1}(\boldsymbol{x}). \tag{8}$$

其中$\mathcal{AD}_j$表示输入数据$\mathcal{E}_i^t\bigcup\mathcal{D}_i^t$中属于第$j$类的数据集合. 若当前输入数据中不存在第$j$类数据, 则计该类数据对应的原型为$\mathbf{0}$; 若边端网络初次更新, 则初始化特征原型组中各元素均为$\mathbf{0}$.

### 2.2　基于特征原型间隔的加权校准策略

在云端, 设计一种基于特征原型间隔的加权校准策略, 如图 3 所示. 该策略通过学习全局类别特征原型并计算原型间隔, 得到各边端的个性化权重后进行模型聚合, 从而缓解客户端漂移现象. 策略主要包含 3 个部分, 即全局特征原型的计算、原型间隔的计算以及基于特征原型间隔的云端校准.

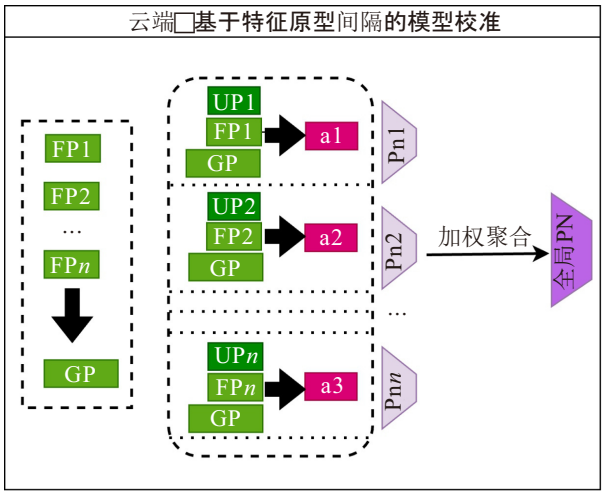


**图3　基于特征原型间隔的加权校准策略**

全局特征原型的计算. 根据各边端的旧特征原型组$\{\mathcal{FP}_i^t\}_{i=1}^C$, 可在云端计算得到任务$t$所需的全局特征原型$\mathrm{GP}^t=\{\mathbf{gp}^t[j]\}_{j=1}^M$, 各类别对应的全局特征原型为

$$\mathbf{gp}^t[j]=\frac{1}{\mathrm{PC}_j}\sum_{i=1}^{C}\mathbf{po}_i^t[j], \tag{9}$$

其中$\mathrm{PC}_j$表示第$j$类数据在边端存在$(\mathbf{po}_i^t[j]\neq\mathbf{0})$的个数.

原型间隔的计算. 原型间隔衡量了原型变化即模型更新对于分类任务的有益程度, 本文用$\mu[j](\cdot)$表示. 对某类原型而言, 原型间隔越大, 该类别原型变化前后的原型取值越接近, 且与别的类越分离. 设模型更新前后的原型分别为$\boldsymbol{p}_a$和$\boldsymbol{p}_b$, 对于类别$j$, 原型间隔为

$$\mu(\boldsymbol{p}_a[j],\boldsymbol{p}_b)[j]=\frac{\boldsymbol{d}^-[j](\boldsymbol{p}_a,\boldsymbol{p}_b)-\boldsymbol{d}^+[j](\boldsymbol{p}_a,\boldsymbol{p}_b)}{\boldsymbol{d}^-[j](\boldsymbol{p}_a,\boldsymbol{p}_b)+\boldsymbol{d}^+[j](\boldsymbol{p}_a,\boldsymbol{p}_b)}. \tag{10}$$

其中: $\boldsymbol{d}_{a,b}^-[j]$为原型$\boldsymbol{p}_a$与$\boldsymbol{p}_b$中类别$j$原型的距离, $\boldsymbol{d}_{a,b}^+[j]$为原型$\boldsymbol{p}_a$中类别$j$原型与$\boldsymbol{p}_b$中非类别$j$原型的距离, 分别为

$$\boldsymbol{d}^+[j](\boldsymbol{p}_a,\boldsymbol{p}_b)=d(\boldsymbol{p}_a[j],\boldsymbol{p}_b[j]), \tag{11}$$

$$\boldsymbol{d}^-[j](\boldsymbol{p}_a,\boldsymbol{p}_b)=\sum_{j'\neq j}\frac{d(\boldsymbol{p}_a[j],\boldsymbol{p}_b[j'])}{|c'\neq c\bigwedge c'\in\mathcal{MP}|}. \tag{12}$$

$\mathcal{MP}$为两个原型共有的类别集合, $c\in\mathcal{MP}$.

联邦类别增量场景下, 根据下式计算边端的本地原型间隔:

$$\mu(\mathbf{pn}_i[j],\mathbf{po}_i)[j]=\frac{\boldsymbol{d}^-[j](\mathbf{pn}_i,\mathbf{po}_i)-\boldsymbol{d}^+[j](\mathbf{pn}_i,\mathbf{po}_i)}{\boldsymbol{d}^-[j](\mathbf{pn}_i,\mathbf{po}_i)+\boldsymbol{d}^+[j](\mathbf{pn}_i,\mathbf{po}_i)}. \tag{13}$$

根据下式计算聚合原型间隔:

$$\mu(\mathbf{pn}_i[j],\mathbf{gp})[j]=\frac{\boldsymbol{d}^-[j](\mathbf{pn}_i,\mathbf{gp})-\boldsymbol{d}^+[j](\mathbf{pn}_i,\mathbf{gp})}{\boldsymbol{d}^-[j](\mathbf{pn}_i,\mathbf{gp})+\boldsymbol{d}^+[j](\mathbf{pn}_i,\mathbf{gp})}. \tag{14}$$

对于第$j$类原型而言, 本地原型间隔越大, 意味着本地模型更新后原型取值越稳定, 且与别的类越分离; 聚合原型间隔越大, 本地模型更新后原型越靠近全局最优, 且与别的类的最优原型越分离. 因此, 本地原型间隔可以衡量边端模型更新后的本地可信度, 聚合原型间隔可以衡量边端模型更新的全局可信度.

基于特征原型间隔的云端校准. 对于边端$\mathrm{Edge}_i$, 根据其本地原型间隔, 可以量化计算边端原型网络更新的本地可信度, 如下所示:

$$\boldsymbol{vl}_i^t=\sigma\Big(\sum_{j\in\mathcal{MP}}\mu(\mathbf{pn}_i[j],\mathbf{po}_i)[j]\Big). \tag{15}$$

根据其聚合原型间隔, 可以量化计算边端原型网络更新的全局可信度, 如下所示:

$$\boldsymbol{va}_i^t=\sigma\Big(\sum_{j\in\mathcal{MP}}\mu(\mathbf{pn}_i[j],\mathbf{gp})[j]\Big). \tag{16}$$

综合本地可信度和全局可信度, 可得到边端$\mathrm{Edge}_i$的原型网络在云端的校准权重, 如下所示:

$$\boldsymbol{a}_i^t=\frac{\boldsymbol{vl}_i^t+\boldsymbol{va}_i^t}{2}. \tag{17}$$

使用校准权重聚合各边端模型参数, 可以使得云端模型向全局最优方向调整, 同时保障类间分离与优化过程的稳定性.

### 2.3　FCIL-COFPO 建模整体流程

所提出的基于稳定特征原型的联邦类别增量学

习方法整体流程如下.

step 1: 模型初始化. 各边端的原型网络进行随机初始化.

step 2: 本地类别记忆库的动态更新. 对于每个边端$\mathrm{Edge}_i$, 基于新增任务数据集$\mathcal{D}_i^t$和已有记忆库, 管理本地类别记忆库, 得到$\mathcal{E}_i^t \bigcup \mathcal{D}_i^t$.

step 3: 回放策略下的原型网络更新与特征原型更新. 对于每个边端$\mathrm{Edge}_i$, 基于$\mathcal{D}_i^t$和$\mathcal{E}_i^t \bigcup \mathcal{D}_i^t$, 根据算法 1 更新本地原型网络. 每个边端$\mathrm{Edge}_i$均在本地原型网络更新前后得到两组类别特征原型, 即旧特征原型组$\{\mathcal{FP}_i^t\}_{i=1}^C$和新特征原型组$\mathcal{UP}_i^t$.

step 4: 全局原型和原型间隔的计算. 云端汇总各边端新旧特征原型组, 由$\{\mathcal{FP}_i^t\}_{i=1}^C$计算得到全局特征原型$\mathrm{GP}^t$. 根据式 (13) 计算各边端对应的本地原型间隔, 根据式 (14) 计算各边端对应的聚合原型间隔.

step 5: 基于特征原型间隔的云端校准. 在云端, 对于来自边端$\mathrm{Edge}_i$的模型, 分别根据式 (15) 和 (16) 计算模型更新的本地可信度和全局可信度, 然后综合两者, 根据式 (17) 得到该原型网络在云端的校准权重. 使用校准权重聚合各边端模型参数. 模型参数聚合后下发给各边端. step 4 和 step 5 与 step 2 和 step 3 交替进行, 直到达到设定迭代次数.

联邦建模完成后, 不同的边端使用云端分发的原型网络和本地记忆库数据上的类别原型进行分类, 在测试集上验证分类效果.

## 3 案例研究

由于实际工业场景中含有较多种类故障且标注清晰的数据集较少公开, 基于类别增量常用的数据集 CIFAR10 和 Mini-ImageNet 设计具体增量场景, 对所提出 FCIL-COFPO 方法进行实验. 实验中算法执行环境: Ubuntu 系统, 2 * NVIDIA GeForce RTX 3090 24 G GPU, 2 * Intel(R) Xeon(R) CPU E5-2678 v3 @ 2.50 GHz CPU, 7 * 16 G 内存, python3 软件平台. 本文所设置的实验中, 假设不同边端任务变动的时机较为接近, 并可以与云端实现同步通讯. 在实际工业过程中, 边端中任务间隔的划分需要根据实际设备运行工况而定. 有些边端可能在较长时间内本地任务不发生变化, 此时本地模型更新频率降低. 反之, 若存在某边端任务变动较快, 则需要提高其本地模型更新及整体联邦模型更新的频率. 在不同边端任务间隔不一致的情况下, 需选择最小的通讯间隔进行联邦模型更新, 或采取合适的异步更新策略进行协调.

### 3.1 数据集与边端设置

1) CIFAR10 数据集. 彩色图像数据集 CIFAR10 共包含 10 个类别, 包括飞机、汽车、鸟类、猫、鹿、狗、青蛙、马、船和卡车. 每个类别有 6 000 张图片, 每张图片的尺寸为$32 \times 32$. 每个类别图像中随机选择 5 000 个作为训练样本, 1 000 个作为测试样本.

2) Mini-ImageNet 数据集. Mini-ImageNet 是选取了 ImageNet 中部分图像的数据集, 是元学习和小样本领域的基准数据集, 同时由于其类别种类丰富, 也是增量学习算法的常用数据集. 原始的 ImageNet 数据集包含 2 万多个类别, 如“气球”“轮胎”“狗”等, 每个类别图像数量大于 500 张. 而 Mini-ImageNet 则仅包含 100 个类别, 每个类别图像数量为 600, 原始图像大小不固定, 数据预处理时将图片统一调整为$84 \times 84$大小. 每个类别图像中随机选择 500 个作为训练样本, 剩余 100 个作为测试样本. 由于服务器条件限制, 无法支撑数据类别总数过多的实验, 选择 Mini-ImageNet 中的 10 个类别参与实验.

在实验中, 设置共有 3 个边端, 边端的单个任务中包含随机 3 种类别的图像, 在 CIFAR10 数据集的实验中, 设置单次任务每种类别图像 100 张, 在 Mini-ImageNet 的实验中, 设置单次任务每种类别图像 30 张, 为从该类别的训练样本中抽样获得. 在这里, 假设边端任务变化时间同步, 且模型云端校正后模型同步更新. 需要注意的是, 边端的模型训练是并行的, 云端模型更新后发送到边端, 边端模型更新可异步进行.

### 3.2 模型与参数设定

本文采用的基本模型结构为一个有 4 层卷积神经网络 (convolutional neural network, CNN) 的原型网络 (prototypical network, PN) [29], 每个 CNN 均设置了卷积核为 3 的二维卷积、批标准化、激活函数修正线性单元以及最大池化层, 最后一层 CNN 额外设置广义平均池化层, 并对输出的特征进行 L2 归一化处理. 模型训练时, 数据的批量大小为 128, 损失函数为原型分类损失. 测试时采用分类准确率对模型效果进行评估, 即$\mathrm{acc} = T_{\mathrm{all}}/\mathrm{Num}_{\mathrm{all}}$. 其中: $T_{\mathrm{all}}$为预测正确的图片数量, $\mathrm{Num}_{\mathrm{all}}$为所预测的图片总数量.

采用 FedAvg[30]、FedProx[28] 等作为对比方法, 以原型网络作为边端基础模型结构, 结合相关算法后分别简写为“PN-FedAvg”和“PN-FedProx”. 各对比方法本地更新轮次为 35, 初始学习率为 0.001, 模型优化采用 Adam 算法. 在基于 CIFAR10 的实验中,

全局迭代轮次为 500, 在基于 Mini-ImageNet 的实验中, 全局迭代轮次为 250. 经过 50 和 90 个全局迭代轮数后, 分别将学习率乘以 0.5.

### 3.3　多边端类别增量下的联邦学习

本节中, 基于 CIFAR10 和 Mini-ImageNet 数据集模拟多边端情况进行实验, 以研究联邦框架下多边端类别增量时所设计 FCIL-COFPO 方法的有效性. 设置共 3 个边端, 边端类别记忆库中每类存储 10 个样本, 即 $P=10$. 设置边端的单个任务中包含随机 3 种类别的图像, 最终已见类别为 10. 采用 PN-FedAvg、PN-FedProx 作为联邦聚合时的对比方法. “PN-FedAvg-R”表示加入所设计边端回放范式的 FedAvg 算法, “PN-FedProx-R”表示加入所设计边端回放范式的 FedProx 算法, 相较于前一种方法在云端聚合时针对客户端漂移问题对模型参数进行了校正. “R w/o”表示所设计方法中边端回放范式被消融, 即仅保留所设计的基于原型间隔的加权聚合策略.

表 1 展示了 FCIL-COFPO、各对比算法以及消融方法最后 5 次迭代训练后对本次任务中未见旧类

**表1　联邦类别增量下的分类准确率**　　%

| 数据集 | 方法 | 所有边端 | | | 边端1 | | 边端2 | | 边端3 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 测试类别 | 所有类 | 旧类 | 新类 | 旧类 | 新类 | 旧类 | 新类 | 旧类 | 新类 |
| CIFAR10 | PN-FedAvg | 39.64 | 40.76 | 37.02 | 37.30 | 41.17 | 43.49 | 29.10 | 41.50 | 40.80 |
| | PN-FedAvg-R | 72.65 | 72.18 | 73.73 | 72.01 | **72.87** | 73.74 | 74.50 | 70.79 | 73.83 |
| | PN-FedProx-R | 74.12 | 72.88 | **77.00** | 73.00 | 72.13 | 72.83 | **82.46** | 72.80 | **76.40** |
| | FCIL-COFPO | **77.46** | **78.15** | 75.84 | **80.24** | 75.37 | **74.24** | 78.87 | **79.97** | 73.30 |
| | R w/o | 42.03 | 41.02 | 44.37 | 38.40 | 38.20 | 42.87 | 46.87 | 41.78 | 48.03 |
| Mini-ImageNet | PN-FedAvg | 49.60 | 47.57 | 54.33 | 46.14 | 62.33 | 48.14 | 48.67 | 48.43 | 52.00 |
| | PN-FedAvg-R | 67.14 | 66.48 | 68.67 | 65.43 | 68.33 | **67.43** | 67.33 | 66.57 | 70.33 |
| | PN-FedProx-R | 65.27 | 64.43 | 67.22 | 66.57 | 69.33 | 62.43 | 57.00 | 64.29 | **75.33** |
| | FCIL-COFPO | **69.90** | **68.85** | **72.33** | **70.71** | **70.00** | 66.14 | **74.66** | **69.71** | 72.33 |
| | R w/o | 50.93 | 49.57 | 54.11 | 52.28 | 53.33 | 48.71 | 54.33 | 47.71 | 54.66 |

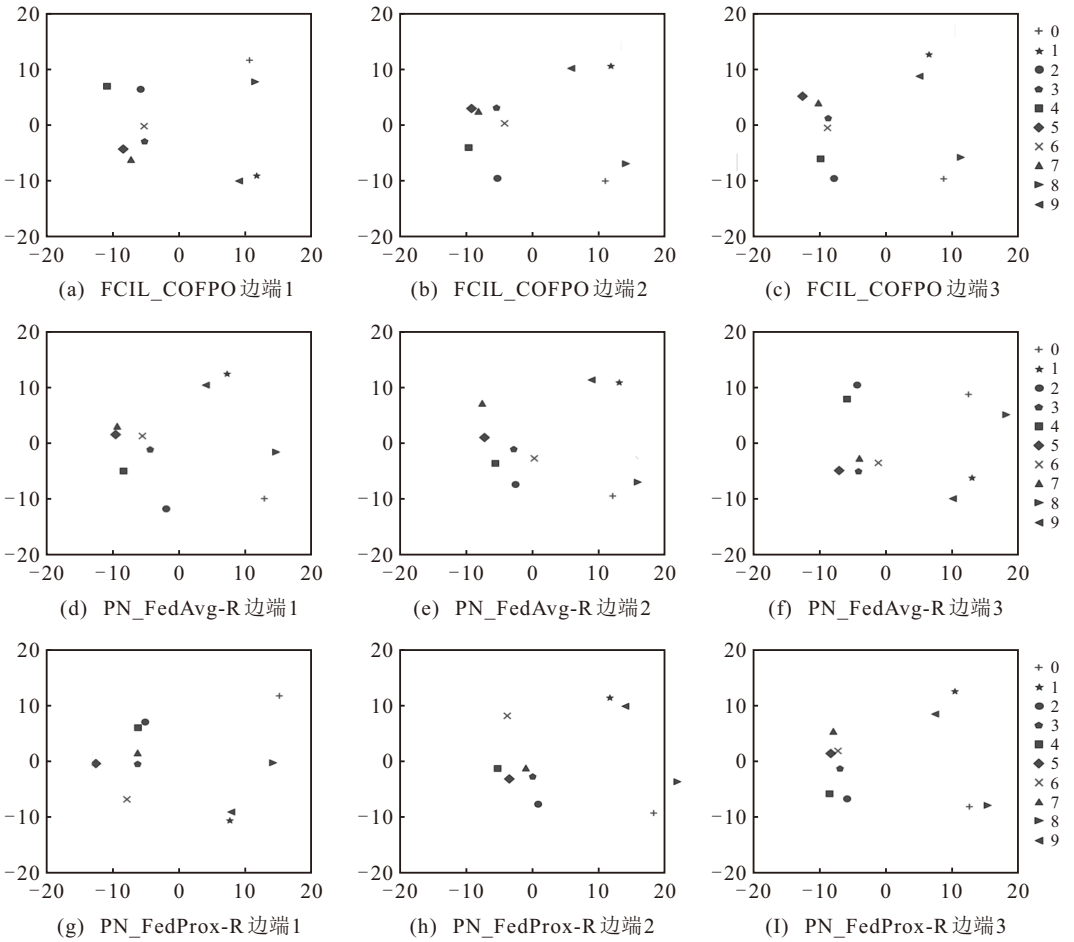


**图4　不同联邦建模方法的边端类别特征原型可视化结果**

别和当前所见新类别的平均分类准确率. 图 4 展示了 CIFAR10 数据集下所提出方法和有回放范式的 PN-FedAvg-R、PN-FedProx-R 中各边端类别特征原型经 PCA 降维后的可视化结果. 图中不同符号表示不同类别原型，10 种符号对应 10 种类别.

由表 1 可知, 在联邦类别增量场景下, 所设计的 FCIL-COFPO 方法整体分类准确率优于对比方法. 尽管对于刚见到的新类别测试数据分类准确率有时低于对比方法, 但在对旧类别的分类效果上, 所提方法相较于 PN-FedAvg、PN-FedAvg-R、PN-FedProx-R 在 CIFAR10 数据集上分别提升了 91.80 %、8.20 %、7.23 %, 在 Mini-ImageNet 数据集上分别提升 44.73 %、3.56 %、6.86 %, 抗遗忘效果显著. PN-FedAvg-R 相较 PN-FedAvg 在 CIFAR10 上对旧类的分类准确率提升了 77.14 %, 在 Mini-ImageNet 上提升了 39.75 %, 而失去了回放范式的 FCIL-COFPO 在 CIFAR10 上对旧类的分类准确率降低了 47.51 %, 在 Mini-ImageNet 上则降低了 28.00 %, 这验证了所设计的边端无需传输边端样本以及基于回放范式的原型网络更新策略的有效性. 所设计的基于原型间隔的加权聚合策略也有一定效果, 在 Mini-ImageNet 数据集上, R w/o 相较于 PN-FedAvg 提升了 4.2 %. 另一种聚合策略即 PN-FedProx-R 相较于 FCIL-COFPO 分类准确率降低了 6.42 %, 相较于加入回放范式的 FedAvg 降低了 3.08 %. 这表明在类别增量场景下, 传统的对模型参数的变化进行约束的校正策略有可能对联邦建模起到反面效果, 而针对类别知识进行设计, 通过寻求稳定的特征空间而对模型加以校正, 可以有效提升联邦协同建模的最终效果.

由图 4 可知, 相较同样有回放范式的 PN-FedAvg-R、PN-FedProx-R, 所提出方法得到的各边端类别特征原型之间相对位置更加一致, 表明基于原型间隔的加权聚合策略使得不同边端的特征空间在联邦优化时趋于一致. 同时, 图中可见 PN-FedAvg-R、PN-FedProx-R 分别在边端 1 和边端 3 中存在距离过近的类别特征原型, 而所提出方法特征原型类间更加分离, 因此在表 1 中分类准确率更高.

## 4 结 论

本文提出了一种基于稳定特征原型的联邦类别增量学习方法 FCIL-COFPO, 并在图像数据集 CIFAR10 和 Mini-ImageNet 上进行了实验验证. 所提出 FCIL-COFPO 方法在边端设计了基于回放范式的原型网络更新策略, 通过动态维护类别记忆库克服了类别增量情况下的灾难性遗忘问题, 通过训练类别相关的特征空间提升了模型优化的稳定性, 在云端设计了基于特征原型间隔的加权校准策略, 根据原型间隔得到各边端模型聚合权重, 间接依靠类别相关知识实现了模型聚合阶段的校正. 在无需云端公共数据集且不用上传边端样本的情况下, 所提出方法解决了类别增量下多边端协同建模的问题. 所设计的回放策略效果明显, 云端校准策略针对类别知识寻找稳定特征原型, 相较于传统校准策略 FedProx 提升了 6.86 %. 需要注意的是, 虽然本文实验基于图像数据集完成, 但当处于某具体工业场景时, 通过调整边端的基准模型结构, 该方法也可被应用于实际基于过程数据的联邦故障分类等任务中, 以扩展其应用范围.

## 参考文献 (References)


[1] Shao H D, Xia M, Han G J, et al. Intelligent fault diagnosis of rotor-bearing system under varying working conditions with modified transfer convolutional neural network and thermal images[J]. IEEE Transactions on Industrial Informatics, 2021, 17(5): 3488-3496.

[2] Feng L J, Zhao C H, Li Y L, et al. Multichannel diffusion graph convolutional network for the prediction of endpoint composition in the converter steelmaking process[J]. IEEE Transactions on Instrumentation Measurement, 2021, 70: 3037953.

[3] Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold[J]. Nature, 2021, 596: 583-589.

[4] Jiang Y, Zhao C H. Attention classification-and-segmentation network for micro-crack anomaly detection of photovoltaic module cells[J]. Solar Energy, 2022, 238: 291-304.

[5] 马晓洋, 张晓冬, 彭锐. 基于物联网技术的科技基础设施智能管理的可靠性研究[J]. 控制与决策, 2019, 34(5): 1116-1120.
(Ma X Y, Zhang X D, Peng R. Reliability of intelligent management of research infrastructure based on Internet of Things[J]. Control and Decision, 2019, 34(5): 1116-1120.)

[6] 李燕君, 蒋华同, 高美惠. 基于强化学习的边缘计算网络资源在线分配方法[J]. 控制与决策, 2022, 37(11): 2880-2886.
(Li Y J, Jiang H T, Gao M H. Reinforcement learning-based online resource allocation for edge computing network[J]. Control and Decision, 2022, 37(11): 2880-2886.)

[7] 柴天佑, 程思宇, 李平, 等. 端边云协同的复杂工业过程运行控制智能系统[J]. 控制与决策, 2023, 38(8): 2051-2062.
(Chai T Y, Cheng S Y, Li P, et al. Intelligent system for operational control of complex industrial process based on end-edge-cloud collaboration[J]. Control and Decision, 2023, 38(8): 2051-2062.)

[8] 赵健程, 冯良骏, 岳嘉祺, 等. 从零样本学习理论模型到工业应用 —— 动机、演变与挑战[J]. 控制与决策, 2024, 39(9): 2833-2857.
(Zhao J C, Feng L J, Yue J Q, et al. From zero-shot learning theoretical model to its industrial application: Motivation, evolution and challenges[J]. Control and Decision, 2024, 39(9): 2833-2857.)

[9] Gkonis P, Giannopoulos A, Trakadas P, et al. A survey on IoT-edge-cloud continuum systems: Status, challenges, use cases, and open issues[J]. Future Internet, 2023, 15(12): 383.

[10] Wang J Y, Song P Y, Zhao C H, et al. Federated knowledge amalgamation with unbiased semantic attributes under cloud-edge collaboration for heterogeneous fault diagnosis[J]. Journal of Process Control, 2023, 131: 103095.

[11] Khan L U, Saad W, Han Z, et al. Federated learning for internet of things: Recent advances, taxonomy, and open challenges[J]. IEEE Communications Surveys & Tutorials, 2021, 23(3): 1759-1799.

[12] Li B X, Zhao C H. Federated zero-shot industrial fault diagnosis with cloud-shared semantic knowledge base[J]. IEEE Internet of Things Journal, 2023, 10(13): 11619-11630.

[13] 王迎春, 王志硕, 刘洋, 等. 基于联邦学习的海上分布式光伏超短期功率预测[J]. 控制与决策, DOI: 10.13195/j.kzyjc.2023.1649.
(Wang Y C, Wang Z S, Liu Y, et al. Ultra-short-term power prediction of offshore distributed PV based on federated learning[J]. Control and Decision, DOI: 10.13195/j.kzyjc.2023.1649.)

[14] Shi G Y, Chen J X, Zhang W L, et al. Overcoming catastrophic forgetting in incremental few-shot learning by finding flat minima[J]. Advances in Neural Information Processing Systems, 2021, 34: 6747-6761.

[15] Li B X, Song P Y, Zhao C H, et al. Facing spatiotemporal heterogeneity: A unified federated continual learning framework with self-challenge rehearsal for industrial monitoring tasks[J]. Knowledge-Based Systems, 2024, 289: 111491.

[16] Mallya A, Lazebnik S. PackNet: Adding multiple tasks to a single network by iterative pruning[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, 2018: 7765-7773.

[17] Kirkpatrick J, Pascanu R, Rabinowitz N, et al. Overcoming catastrophic forgetting in neural networks[J]. PNAS2017, 114(13): 3521-3526.

[18] Zhou D W, Ye H J, Zhan D C. Co-transport for class-incremental learning[C]. Proceedings of the 29th ACM International Conference on Multimedia. New York: ACM, 2021: 1645-1654.

[19] Javed K, White M. Meta-learning representations for continual learning[J]. arXiv, 2019: 1905.12588v2.

[20] 赵海燕, 马权益, 曹健, 等. 面向任务扩展的增量学习动态神经网络: 研究进展与展望[J]. 电子学报, 2023, 51(6): 1710-1724.
(Zhao H Y, Ma Q Y, Cao J, et al. Dynamic neural network for incremental learning with task extended: Research progress and prospect[J]. Acta Electronica Sinica, 2023, 51(6): 1710-1724.)

[21] Rebuffi S A, Kolesnikov A, Sperl G, et al. iCaRL: Incremental classifier and representation learning[C]. IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, 2017: 2001-2010.

[22] 莫建文, 陈瑶嘉. 基于分类特征约束变分伪样本生成器的类增量学习[J]. 控制与决策, 2021, 36(10): 2475-2482.
(Mo J W, Chen Y J. Class incremental learning based on variational pseudo-sample generator with classification feature constraints[J]. Control and Decision, 2021, 36(10): 2475-2482.)

[23] Li Z Z, Hoiem D. Learning without forgetting[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(12): 2935-2947.

[24] Qi D Q, Zhao H D, Li S. Better generative replay for continual federated learning[C]. The 11th International Conference on Learning Representations. Kigali, 2023: 1-10.

[25] Yoon J H, Jeong W Y, Lee G W, et al. Federated continual learning with weighted inter-client transfer[C]. International Conference on Machine Learning. Virtual: PMLR, 2021: 12073-12086.

[26] Usmanova A, Portet F, Lalanda P, et al. A distillation-based approach integrating continual learning and federated learning for pervasive services[J/OL]. 2021, arXiv: 2109.04197.

[27] Li B X, Song P Y, Zhao C H. Fusing consensus knowledge: A federated learning method for fault diagnosis via privacy-preserving reference under domain shift[J]. Information Fusion, 2024, 106: 102290.

[28] Li T, Sahu A K, Zaheer M, et al. Federated optimization in heterogeneous networks[J]. Proceedings of Machine Learning and Systems, 2020, 2: 429-450.

[29] Snell J, Swersky K and Zemel R. Prototypical networks for few-shot learning[C]. Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, 2017: 4080-4090.

[30] 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. Fort Lauderdale, 2017: 1273-1282.

## 作者简介

姚邹静 (1997−), 女, 博士生, 主要研究方向为云边协同、联邦学习, E-mail: yzjing@zju.edu.cn;

赵春晖 (1979−), 女, 教授, 博士生导师, 主要研究方向为工业大数据分析与应用, 包括状态监测、故障诊断、软测量, E-mail: chhzhao@zju.edu.cn.