无样本类增量学习射频指纹识别研究

Rundong Jiang、Jun Hu、Yunqi Song、Zhiyuan Xie、Shiyou Xu

arXiv:2601.03063v1 [eess.SP] 6 Jan 2026

摘要—**无线设备的快速普及使稳健的身份认证变得至关重要。射频指纹（RFF）利用设备特有且难以伪造的物理层损伤来进行识别，在物联网和无人系统中具有广阔应用前景。然而在实际应用中，新设备不断加入已部署的系统，而每个类别的训练数据十分有限。由于类别数量不断增加、存储成本高昂以及隐私问题，传统的静态训练和对已存储样本的简单重放并不切实际。我们提出了一种针对射频指纹识别的无样本类增量学习框架。该框架以预训练特征提取器为起点，在增量阶段冻结骨干网络，仅训练分类器以及执行小幅任务特定特征调整的轻量级适配器模块。对于每个类别，我们根据骨干特征拟合一个对角高斯混合模型（GMM），并从这些拟合后的分布中采样伪特征，以便在不存储原始信号的情况下复习过去的类别。为提高少样本条件下的稳健性，我们引入了时域随机掩码增强，并采用多教师蒸馏方案，将分阶段适配器压缩为单个推理适配器，在准确率和运行效率之间进行权衡。我们在大规模自采集的广播式自动相关监视数据集上评估了该方法：骨干网络在2,175个类别上进行预训练，增量实验则在一个不相交的669个类别集合上开展，并设置了多轮实验和不同步长。与多个具有代表性的基线方法（包括基于样本和无样本方法）相比，我们的方法始终取得更高的平均准确率和更低的遗忘，同时显著减少存储需求并避免原始数据留存。所提出的流程具有可复现性，为资源受限且注重隐私的环境中的射频指纹部署提供了一种实用的低存储解决方案。**

关键词：类别增量学习、射频指纹识别、特征拟合

本文已提交 IEEE，供其考虑发表。版权可能在不另行通知的情况下转让，届时可能无法再访问此版本。

I. 引言

射频指纹（RFF）是信号传输过程中由硬件差异产生的独特特征，例如频率偏移、调制误差和瞬态响应。这些细微差异源于分量的制造差异，可作为设备的“身份”标识。通过信号分析和机器学习，射频指纹识别能够实现设备识别、身份验证和防克隆追踪，有效打击克隆设备和网络欺诈；其在物联网安全和频谱管理中具有重要应用。

现有射频指纹识别方法主要关注特征提取。对于基于神经网络的方法，模型通常在训练集上训练一次后，便部署用于推理。

所有作者均来自中国深圳 518000 的中山大学电子与通信工程学院。

在实际场景中，设备会持续加入网络，因此模型需要识别的类别数量会随时间增加。在这种情况下，在边缘端重新部署或重新训练通常需要存储所有旧设备的样本，这会带来高昂的存储成本和隐私风险。

因此，为射频指纹实现增量学习——即在不保留旧设备样本的情况下持续学习新的设备类别——是一个重要问题。

与其他分类任务相比，射频指纹的类别增量学习具有设备和类别数量多、设备间差异小等特点。在身份认证场景中，射频指纹需要区分大量设备；由于射频指纹侧重于从频率和调制方式相同的信号中提取特征，因此属于细粒度识别问题 [1]–[3]。因此，使用预训练模型作为特征提取器并进行微调是合理的选择；这样一来，问题就转化为分类器的增量训练。直接存储部分原始样本的方法在实践中会受到存储和隐私方面的限制。

当预训练阶段已覆盖足够多的设备和信道变化，且新的增量类别与预训练分布之间不存在显著的域偏移时，可以将预训练特征提取器视为相对稳定的表征。在这一细粒度前提下，预训练模型能够提取射频指纹特征；固定预训练参数并采用轻量级适配器进行增量微调，同时对分类器进行增量训练，核心挑战便转变为分类器的增量训练。

分类器增量训练着重于保留决策边界。从决策边界和数据保留的角度出发，我们提出在增量训练过程中拟合特征分布并重新生成伪特征。每个阶段结束后，我们提取该阶段各类别的特征，为每个类别拟合分布，并在后续训练中从这些分布采样伪特征，以重演先前类别。我们使用对角高斯混合模型拟合特征向量，仅保留协方差矩阵的对角元素，以简化计算和存储；这会牺牲部分拟合精度，以换取更高的稳定性。

对于每个增量任务，我们还会添加一个称为适配器的微调层，以调整新类别引起的特征分布偏移。训练期间，我们仅使用当前任务的类别样本来训练分类器和适配器。每个任务都会单独训练一个适配器；在推理时，由于无法先验获知未知样本所属的阶段，

我们将所有历史适配器蒸馏到一个适配器中，以避免要求预先了解样本所属的任务。

我们采用 [4] 中提出的双结构特征差分网络特征提取器加分类器架构作为预训练模型。我们在两个自行采集的大规模广播式自动相关监视数据集上进行了大量实验：使用2,175个类别训练预训练模型，并针对与预训练集不重叠的另外669个类别进行多步增量训练。训练期间，我们采用随机掩码增强，在时域信号的起始端或末尾端随机屏蔽不同长度的片段，并将掩码后的结果视为额外的同类别样本；这提高了特征多样性和模型鲁棒性。

实验将我们的方法与主流基线方法进行比较，结果表明，该方法在存储使用方面具有优势，在准确率和遗忘方面表现更优，展现出更强的持续学习能力。我们的方法利用现有模型参数和知识，避免存储原始示例样本，因此适用于隐私敏感场景。

主要贡献总结如下：

- 我们提出了一种用于射频指纹识别的无样本类增量学习框架，通过从紧凑的每类模型中采样伪特征来保留先前类别的决策边界，从而无需保留原始样本。
- 我们采用对角协方差高斯混合模型，在高维、少样本场景下对每类骨干特征进行参数化；这使每个分量的协方差存储量从 $O(D^2)$降至$O(D)$，并提高了拟合稳定性。
- 我们引入轻量级适配器模块进行分阶段适配，并提出多教师蒸馏流程，将历史适配器整合到单个学生适配器中，以实现高效推理。
- 我们在大规模自采集的广播式自动相关监视数据集上进行了广泛实验（预训练使用2,175个类别；增量评估使用669个互不相交的类别）。结果表明，所提方法在平均准确率和遗忘方面均优于有代表性的基线方法，同时所需存储量大幅降低。

II. 相关工作

类别增量学习（CIL）是持续学习中的一种典型场景，模型在无法访问先前任务原始数据的情况下，按顺序学习任务或类别，逐步扩展其识别能力。其主要挑战是灾难性遗忘，即学习新类别会导致模型在先前类别上的性能下降。现有研究可分为三类：基于正则化的方法、基于回放的方法，以及蒸馏/特征保持方法。

A. 基于正则化的方法

正则化方法通过约束参数更新来保留先前任务的知识。弹性权重巩固（EWC）利用 Fisher信息矩阵估计参数重要性，并惩罚关键参数的变化 [5]。其他方法则限制沿低遗忘方向的更新，以保持对新任务的适应能力 [6]。然而，这些方法通常依赖任务边界信息，且费舍尔估计不准确。随着任务不断累积，参数可塑性会逐渐降低，从而削弱对新任务的适应能力。

B. 回放与原型记忆

回放方法会保留部分旧类样本或压缩表示，以供复习。iCaRL 是一种代表性方法，它维护类别样本，并将最近均值分类器与特征空间回放相结合 [7]。后续方法（如端到端增量学习 [8] 和暗经验回放 [9],）改进了损失函数和记忆更新策略。一些方法会动态扩展表征能力，以提高对旧类的保留效果 [10]。

生成式回放方法使用生成对抗网络或变分自编码器合成旧类样本 [11]–[13]。其他技术则通过已知特征的线性组合重建伪特征[14]。然而，生成模型会增加存储和训练复杂度，且在少样本条件下难以训练。线性组合方法假设特征空间具有凸性，这种假设可能无法很好地泛化。

C. 知识蒸馏与特征保持

知识蒸馏将知识从旧模型迁移到新模型，以保持模型在先前类别上的性能。LwF 首次将蒸馏应用于增量学习 [15]。原型感知对比蒸馏方法会约束输出相似性，并使用基于原型的对比损失来提升特征兼容性 [16]。暗经验回放存储中间特征，以实现隐式知识复用。LUCIR 等方法则加入基于余弦的约束，以增强对旧类别和新类别的判别能力 [17]。

大型预训练模型在类别增量学习中的应用日益受到关注 [18]–[20]。DyTox 采用具有动态适配能力的 Transformer 主干网络，以实现与任务无关的持续学习 [21]。

III. 方法

射频指纹识别旨在区分采用相同调制、协议和通信条件的发射设备。先前研究表明，即使设备采用相同的协议和调制格式，独特的射频链路特征仍会使发射信号产生可复现的差异，从而实现设备级区分 [22]。因此，射频指纹本质上是一项细粒度识别任务，这促使我们在增量学习中采用预训练特征提取器和分类器适配。

使用高容量的预训练特征提取器，增量学习问题可转化为分类器适配 [23]。目标是在学习新类别的同时，保留分类器针对旧类别的决策边界。对于一个 K 类别的问题，后验概率定义为：

$$P(y = c \mid \mathbf{x}; \theta) \tag{1}$$

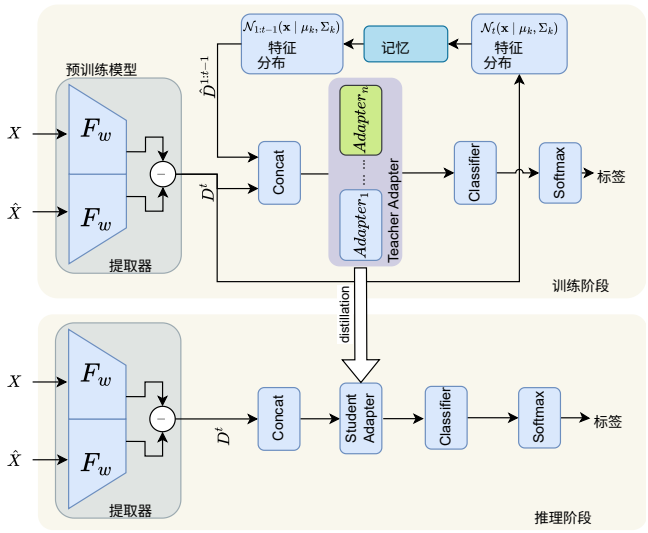


图1. 总体框架

其中，c表示类别 c，$\mathbf{x} \in \mathbb{R}^d$是输入，$y \in \{1, \ldots, K\}$是标签，$\theta$ 是模型参数。类别 i和 j 之间的决策边界为：

$$B_{i,j} = \{\mathbf{x} \mid P(y = i \mid \mathbf{x}; \theta) = P(y = j \mid \mathbf{x}; \theta)\} \quad (2)$$

为明确表示增量阶段之间决策边界的保持，我们要求：

$$\mathbf{x} \in B_{i,j}^{(t)} \Rightarrow \mathbf{x} \in B_{i,j}^{(t+1)} \quad (3)$$

增量目标是确保：

$$P(y = c \mid x; \theta_t) \approx P(y = c \mid x; \theta_{t-1}) \quad (4)$$

从而保持决策边界。使用旧后验概率与新后验概率之间的 KL 散度可得到蒸馏损失：

$$\mathcal{L}_{\text{distil}} = \text{KL}\left(P_t(y \mid \mathbf{x}; \theta_t) \,||\, P_{t+1}(y \mid \mathbf{x}; \theta_{t+1})\right) \quad (5)$$

我们基于 [4], 中的双分支特征差分提取器加分类器架构，并提出了一种如图1所示的无样本类别增量方法。预训练特征提取器计算：

$$D^t = F_w(X) - F_w(\hat{X}) \quad (6)$$

其中，$F_w(\cdot)$ 是提取器，X 是输入样本，$\hat{X}$ 是不含硬件指纹的配对重建结果，D^t 是第 t 个增量轮次的特征。

增量过程包含训练和推理阶段。每个训练阶段结束后，我们会为每个类别的特征拟合一个对角高斯混合模型，从而获得各类别的特征分布$\mathcal{N}_t$。在后续的增量训练中，我们从这些拟合后的分布中采样伪 $\hat{}$特征 D_{t-1}，并使用新类别特征和采样得到的伪特征共同训练分类器，以保留旧的决策边界。每轮还会增加一个适配器模块，以调整特征漂移 [24]；提取器输出的特征在分类前会经过适配器。适配器的结构如图2所示。

推理时，样本所属任务未知，因此通过蒸馏融合所有历史适配器：历史适配器充当教师，通过最小化输出差异来训练新的适配器作为学生。拟合得到的伪特征无需存储原始样本即可进行重放。我们采用带有批归一化的全连接分类器，如图3所示。

虽然最近均值分类器（NMC）为每个类别保留一个均值向量，并根据距离最近的中心进行分类 [25], ，但它们隐含地假设特征簇大致呈球形。在非理想的特征几何结构下，基于中心的决策可能会导致边界偏移，损害开集检测能力，而这在射频指纹场景中至关重要。

A. 特征分布拟合

我们使用多元高斯混合模型拟合图5中的特征 D^t 。直接拟合全协方差高斯混合模型
在高维情况下，输入的计算开销很大：对于特征维度为 D 、包含 k 个分量的模型，每个分量的协方差有 $^{D}_{(D+1)\ 2}$ 个参数，导致计算复杂度呈二次增长。在少样本设置下，这会造成估计不准确和过拟合。因此，我们使用对角协方差高斯混合模型，将每个分量的协方差参数减少到 D 个，并使总参数数量随 D 线性增长。

对角高斯混合模型定义如下：

$$p(\mathbf{x}) = \sum_{k=1}^{K} \pi_k \cdot \mathcal{N}(\mathbf{x} \mid \mu_k, \Sigma_k) \quad (7)$$

其中，对角协方差为 diag($\sigma_{k,1}^2, \ldots, \sigma_{k,d}^2$)，并且：

$$\mathcal{N}(\mathbf{x} \mid \mu_k, \Sigma_k) = \prod_{i=1}^{d} \mathcal{N}(x_i \mid \mu_{k,i}, \sigma_{k,i}^2) \quad (8)$$

等价地，概率密度可写成如下闭式：

$$\mathcal{N}(\mathbf{x} \mid \mu_k, \Sigma_k) = \frac{1}{(2\pi)^{d/2} \prod_{i=1}^{d} \sigma_{k,i}} \cdot \exp\left(-\frac{1}{2} \sum_{i=1}^{d} \frac{(x_i - \mu_{k,i})^2}{\sigma_{k,i}^2}\right) \quad (9)$$

我们使用期望最大化算法拟合高斯混合模型：

1) Compute responsibilities $\gamma(z_{nk}) = \frac{\pi_k N(\mathbf{x}_n|\mu_k, \Sigma_k)}{\sum_j \pi_j N(\mathbf{x}_n|\mu_j, \Sigma_j)}$.
2) Update $\pi_k^{\text{new}} = N_k/N$, where $N_k = \sum_n \gamma(z_{n,k})$.
3) Update $\mu_k^{\text{new}} = \frac{1}{N_k} \sum_n \gamma(z_{n,k}) x_n$.
4) Update diagonal covariance: $\sigma_{k,i}^2 = \frac{1}{N_k} \sum_n \gamma(z_{n,k})(x_{n,i} - \mu_{k,i})^2$.

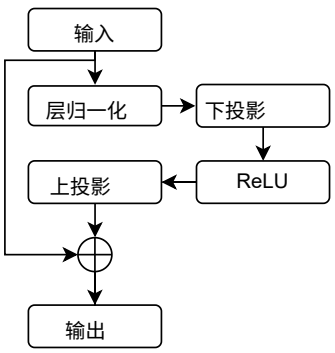


图2. 适配器的结构

5）重复步骤 1–4，直至收敛。

采样过程根据混合权重选择一个成分，并针对每个维度独立抽取正态样本。

B. 特征对齐

为合并不同阶段的适配器，我们采用蒸馏方法。通过将所有历史适配器作为教师，学生适配器在最小化输出差异的目标下进行训练。在蒸馏过程中，我们加入特征对齐损失，包括逐点均方误差和分布对齐。

1) Po 逐点对齐： 逐点对齐使用 M SE:

$$\mathcal{L}_{\text{align}} = \frac{1}{N}\sum_{i=1}^{N} \|D_i - \hat{D}_i\|_2^2 \tag{10}$$

这会促使学生适配器在伪特征上的输出与教师的输出相匹配。

2）分布对齐： 对于分布对齐，我们使用多元高斯分布之间的 2-瓦瑟斯坦距离，该距离对高斯分布有闭式解：

$$W_2(\mathcal{N}(\mathbf{m}_1,\mathbf{C}_1),\mathcal{N}(\mathbf{m}_2,\mathbf{C}_2)) = \|\mathbf{m}_1-\mathbf{m}_2\|_2^2 + \operatorname{tr}\left(\mathbf{C}_1+\mathbf{C}_2-2(\mathbf{C}_1^{1/2}\mathbf{C}_2\mathbf{C}_1^{1/2})^{1/2}\right) \tag{11}$$

我们利用该距离在增量训练期间对齐伪特征分布。

C. 训练数据增强

射频指纹数据是时域信号，其指纹特征与载荷内容无关。为扩增有限的样本，同时最大限度地减少信号失真，我们在样本的开头或结尾应用随机时域掩码。形式化地：

$$\tilde{s}[n] = s[n]\cdot g_{n_0}[k] \tag{12}$$

其中， $g_{n_0}[k] = u[n-n_0] - u[n-n_0-k]$ 为门控窗口，$u[n]$ 为阶跃函数， n_0 为起始位置， k 为长度。中段掩蔽会因与 sinc 函数卷积而导致更严重的频域泄漏，并可能使训练不稳定。

为完整起见，我们将中段掩蔽变体记为：

$$ildes[n] = s[n]\cdot g_{n_0}[k] \tag{13}$$

这在频域中对应于与类 sinc 核的卷积：

$$ildeS(\omega) = S(\omega) * \operatorname{sinc}(\omega) \tag{14}$$

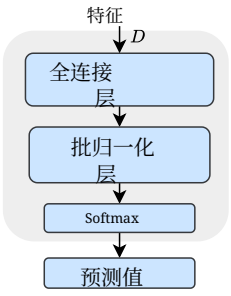


图3. 分类器结构

我们的增强方法会在样本的开头或结尾施加一个可变长度的掩码（图7）。掩码长度从样本长度的 0 到六分之一之间均匀抽取，被掩盖的区域置零。

完整的无样本增量算法见算法1。

算法1 无样本类增量学习

要求： 预训练特征提取器 F_w，训练数据 X， $\hat{X}$，增量轮次数量 T

如果初始训练 **那么**
  对于初始训练 **执行**
    初始化 $Adapter_0$ 和 ϕ^0
    在 X 上训练分类器 ϕ^0 和 $Adapter_0$ ，$\hat{X}$ 并提取按类别划分的特征集 D_c^0
    为每个类别拟合对角高斯混合模型并保存参数 $\mathcal{N}_c$
    保留 $Adapter_0$
  结束循环
否则
  对 增量轮次 $t=1$ 至 T **执行**
    获取新类别数据 X_t， $\hat{X}_t$
    初始化 $Adapter_t$
    通过复制之前的分类器来初始化分类器 ϕ^t 并应用权重对齐
    使用新类别数据训练 $Adapter_t$ 和分类器 ϕ^t $X_t, \hat{X}_t$
    为新类别拟合 GMM 参数并保存 $\mathcal{N}_t$
    初始化学生适配器 $Adapter_s$ 以进行蒸馏
    将所有 $Adapter_{0:t}$ 用作教师；采样伪特征自历史高斯混合模型 $\mathcal{N}_{0:(t-1)}$ 并使用当前数据 X_t， $\hat{X}_t$ 使用特征来训练 $Adapter_s$ 对齐
    使用采样的伪特征微调分类器 ϕ^t 和当前特征
    使用 $Adapter_s$ 和 ϕ^t 进行推理
  结束循环
结束 if

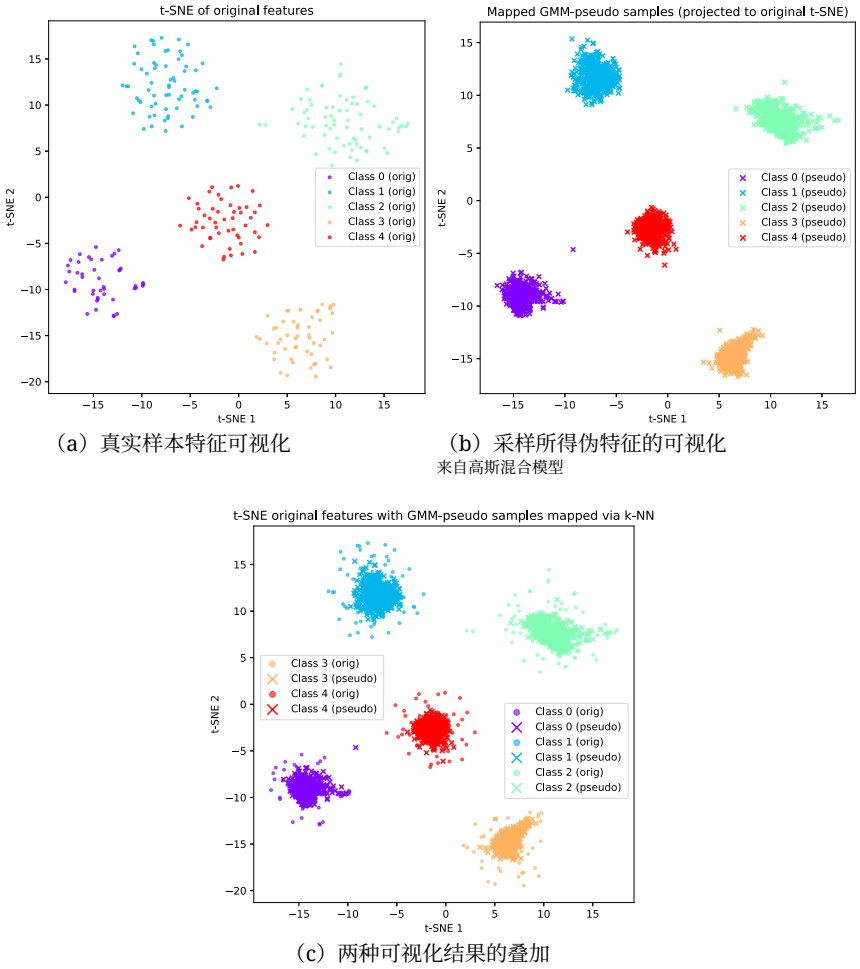
IV. 实验

A. 数据集

该数据集由使用 AD9361 前端在 1091兆赫 下采集的 ADS-B信号 组成（10 Msps采样，5兆赫带宽）。我们准备了两个时间上互不重叠的数据集合：一个用于预训练（2025年2月19日至26日），另一个用于增量评估（2025年1月17日至21日）。我们剔除了饱和录音和低信噪比录音，并将信号分割为定长帧（滤波后每帧112个采样点）。仅保留记录数超过100的类别。预训练集包含2,175个类别；增量集包含669个类别。实验采用70/30的训练/测试划分。

Data	类别数量	训练/测试划分
预训练	2175	7:3
增量	669	7:3

表 I
数据集统计



（a）真实样本特征可视化 （b）采样所得伪特征的可视化来自高斯混合模型

（c）两种可视化结果的叠加

图4. 真实样本与高斯混合模型采样伪特征的 t-SNE 对比

B. 实验设置

所有骨干网络均基于 [4]，提出的双差分提取器，并在预训练划分上进行了预训练。提取器架构如图8所示。 F_w（·）表示提取器映射； X 和 X 分别对应原始时频表示和不含硬件指纹的匹配重建。分类器为一个全连接层，并采用一维批归一化，如图3所示。对于每类对角高斯混合模型拟合，我们将分量数设为 $k = 2$。

C. 结果

我们首先比较真实特征与从高斯混合模型采样得到的伪特征的 t-SNE 可视化结果。从训练集中随机选取五个类别，计算其特征并使用 t-SNE 进行可视化（图4a）。随后，从高斯混合模型中采样 500 个伪特征并进行可视化（图4b）。将两种可视化结果叠加（图4c）后可以看出，高斯混合模型生成的伪特征基本覆盖了真实样本特征的主要几何区域，这表明高斯混合模型采样有助于保留决策边界。

我们使用 PyCIL 框架 [26]，并将我们的方法与 EWC、微调、PASS [27]、、iCaRL [7]、、WA [28]、、动态可扩展表示 [29]，以及 LwF 进行比较。图9展示了不同初始类别数量和增量类别数量下的结果：横轴表示已学习的类别数量，纵轴表示所有历史真实样本上的测试准确率（用于衡量遗忘程度）。随着增量轮次的推进，曲线虽有波动，但总体呈下降趋势。初始步骤和增量步骤较小的场景表现更好。经过多轮增量学习后，我们的方法仍能保持较高的准确率，展现出其有效性。

图10和图11展示了在保留样本和无样本设置下的跨方法比较。在保留样本的比较中，所有方法在适用时均存储 2000 个样本；其他方法不存储旧样本。我们的方法始终取得最佳结果。在不允许使用任何样本时（图11），依赖样本的方法（DER、WA、iCaRL）性能迅速下降。

我们还比较了额外存储空间开销。在 B10S10 场景中，我们的方法会存储每个阶段的适配器和每个类别的 GMM 参数；每个阶段的存储量约为 25.52 KB（表II）。每个类别的高斯混合模型存储量很小（在所选设置下，每个类别约为 2.07 KB），适配器大小约为每个阶段 4.82 KB；通过蒸馏合并不同阶段的适配器，还可以进一步减少存储量（图 12）。

供参考，对角 GMM 的每类别存储量为

表II 各方法的参数/存储开销比较

方法	模型大小	使用样本	保留旧网络	额外存储空间
OURS	9385.7KB	no	no	25.52KB/阶段[1]
iCaRL	9438.4KB	yes	yes	0
DER	9491.7KB	yes	no	2.3MB/阶段[2]
EWC	9438.4KB	no	no	9.66MB[3]
微调	9391.0KB	no	no	0
WA	9396.2KB	yes	yes	0
LwF	9391.0KB	no	yes	0
PASS	9438.4KB	no	yes	9.2MB[4]

[1]每个阶段存储一个适配器，每个类别存储一组对角 GMM 参数。[2]每个阶段扩展一个新的特征提取器。[3]存储一个 Fisher 信息矩阵。[4]每个阶段存储子空间基和投影参数。

参数可以估计为：

$$\begin{aligned} Storage_{GMM} &= k \times (2 \times D + 1) \times bytes\ per\ float \\ &= 2 \times (2 \times 132 + 1) \times 4 \\ &= 2120\ Bytes \approx 2.07\ KB \end{aligned} \tag{15}$$

其中， k 为混合成分数， D 为特征维度， $bytes\ per\ float$ 为每个浮点值的存储量。

表III比较了不同初始/增量设置下各方法的最终准确率和平均准确率。我们的方法在平均准确率和长期稳定性方面始终优于基线方法。

V. 结论

本文提出了一种面向深度学习射频指纹识别的无样本类增量学习框架。该框架利用预训练特征提取器，使用对角高斯混合模型（GMM）拟合每个类别的特征，并按需生成伪特征，在增量训练期间重温旧类别，从而无需存储原始样本。在大型ADS-B数据集上的大量实验表明，与主流基线相比，我们的方法显著提升了旧类别的记忆保持能力，验证了其在资源受限和隐私受限场景中的有效性与实用性。

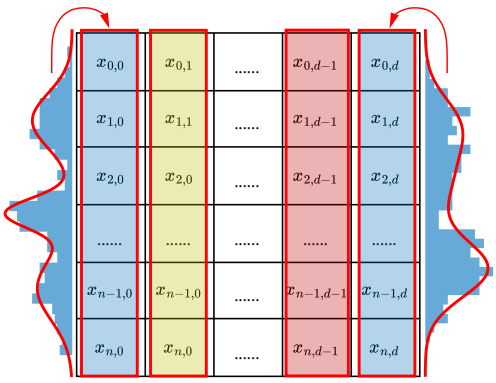


图5. 高斯混合模型

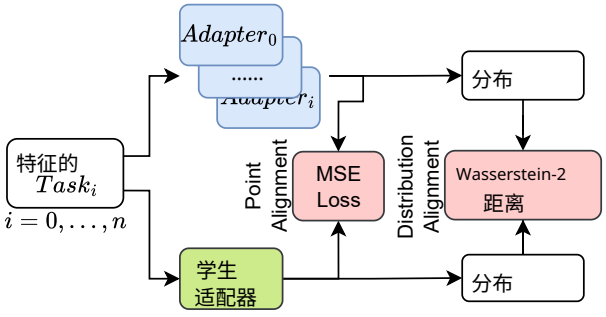


图6. 特征对齐

参考文献

[1] G. Shen 和 J. Zhang，“可迁移的深度学习辅助射频指纹识别系统，*Security andSafety*，卷3，第2023019页，2024年。

[2] Y. Zhang、J. Hu、R. Jiang、Z. Lin 和 Z. Chen，“基于注意力机制的细粒度射频指纹识别网络”，*Entro py*，卷26，第1期，p。2023年12月29日。

[3] C. Zhao、J. Yu、G. Luo 和 Z. Wu，“基于深度度量学习的少样本无线信号射频指纹识别”，无线通信与移动计算，第2023卷，第1–13页，2023年9月。[4] R. Jiang、J. Hu、H. Huang、C. Zhang、L. Wang 和 S. Xu，“使用孪生神经网络的基于广义射频指纹的无线设备识别”，*IEEE传感器期刊*，第1–1页，2024年。[5] A. Aich，“弹性权重巩固（ewc）：基本原理”，2021年5月。[6] H. Wen、H. Qiu、L. Wang、H. Cheng 和 H. Li，“遗忘方向更少、平衡点更优的类别增量学习”，*IEEE视频技术电路与系统汇刊*，第1 1页，2024年。[7] S.-A. Rebuffi、A. Kolesnikov、G. Sperl 和 C. H. Lampert，“iCaRL：增量分类器与表征学习”，2017年4月。[8] F. M. Castro、M. J. Mar´ın-Jim´enez、N. Guil、C. Schmid 和 K. Alahari，“端到端增量学习”，2018年9月。[9] P. Buzzega、M. Boschini、A. Porrello、D. Abati 和 S. Calderara，“通用持续学习中的暗经验：一种强大而简单的基线”，2020年10月。[10] S. Yan、J. Xie 和 X. He，“DER：用于类别增量学习的动态可扩展表征”，2021年3月。[11] C. Wu、L. Herranz、X. Liu、Y. Wang、J. van de Weijer 和 B. Raducanu，“记忆回放生成对抗网络：学习生成新类别图像而不遗忘”，2019年9月。[12] G. M. Van De Ven、Z. Li 和 A. S. Tolias，“使用生成式分类器进行类别增量学习”，载于*2021年IEEE/CVF计算机视觉与模式识别会议研讨会（CVPRW）*。美国田纳西州纳什维尔：IEEE，2021年6月，第3606–3615页。[13] Y. Chai、X. Chen、J. Qiu、L. Du、Y. Xiao、Q. Feng、S. Ji 和 Z. Tian，“Malfscil：一种用于恶意软件检测的少样本类别增量学习方法”，*IEEE信息取证与安全汇刊*，第20卷，第2999–3014页，2025年。[14] J. Zhang、L. Liu、K. Gao 和 D. Hu，“一种面向少样本类别增量药丸识别的前向与后向兼容框架”，2024年3月。[15] Y. Liu、Y. Su、A.-A. Liu、B. Schiele 和 Q. Sun，“助记训练：无遗忘的多类别增量学习”，载于*2020年IEEE/CVF计算机视觉与模式识别会议（CVPR）*。美国华盛顿州西雅图：IEEE，2020年6月，第12 242–12 251页。[16] Z. Yang、J. Pan、Y. Yang、X. Shi、H.-Y. Zhou、Z. Zhang 和 C. Bian，“ProCo：面向长尾医疗图像分类的原型感知对比学习”，2022年9月。[17] S. Hou、X. Pan、C. C. Loy、Z. Wang 和 D. Lin，“通过再平衡增量式学习统一分类器”，载于*2019年IEEE/CVF计算机视觉与模式识别会议（CVPR）*。美国加利福尼亚州长滩：IEEE，2019年6月，第831–839页。[18] D.-W. Zhou、H.-L. Sun、H.-J. Ye 和 D.-C. Zhan，“可扩展子空间基于预训练模型的类别增量学习的集成”，载于 *2024 IEEE/CVF*计算机视觉与模式识别会议（*CVPR*）。美国华盛顿州西雅图：IEEE，2024年6月，第23 554–23 564页。

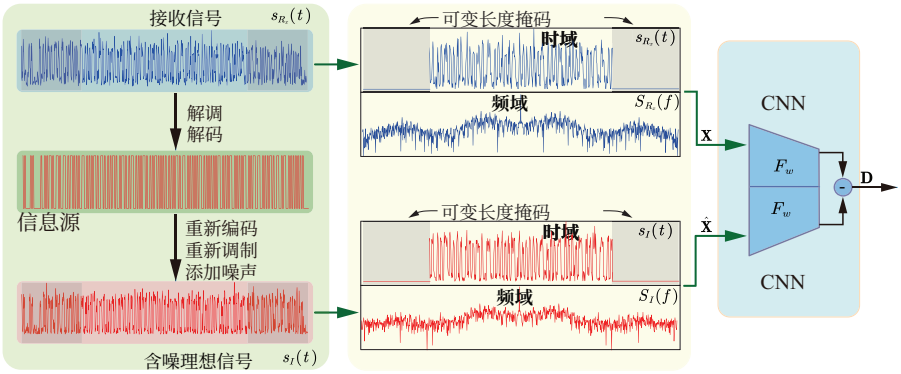


图7。数据增强：随机起止时域掩码

表 III 不同方法的准确率比较

方法	B10S10		B20S20		B50S50		B100S100	
	A_l	A	A_l	A	A_l	A	A_l	A
OURS	97.94	98.92	98.06	98.94	97.67	98.5	97.36	98.48
iCaRL	94.3	97.1	94.33	97.1	93.47	97.39	93.51	97.07
DER	90.12	95.31	87.97	91.89	82.56	86.27	67.59	81.56
EWC	61.74	76.6	70.45	84.66	83.13	90.74	85.4	91.3
微调	67.79	77.97	81.42	90.77	92.45	90.77	94.53	92
WA	67.66	85.26	77	91.6	93.77	97.43	98.59	95.92
LwF	78.97	86.44	83.13	90.82	91.79	81.08	74.36	64.34
PASS	95.94	97.28	95.81	96.81	95.28	96.23	94.71	95.96

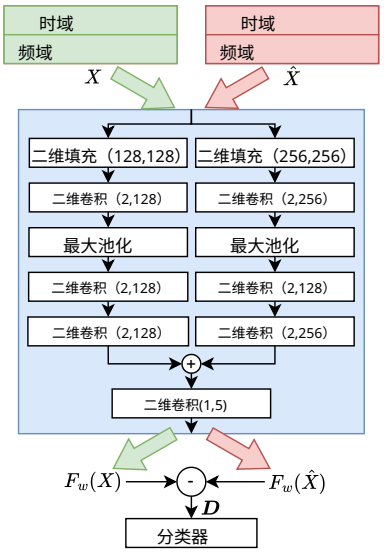


图8. 特征提取器的结构

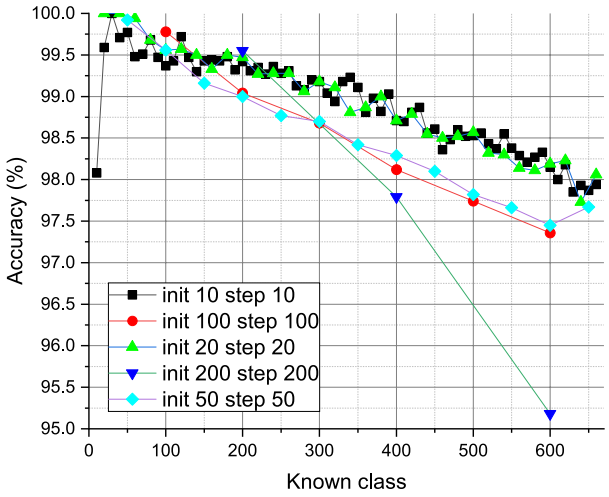
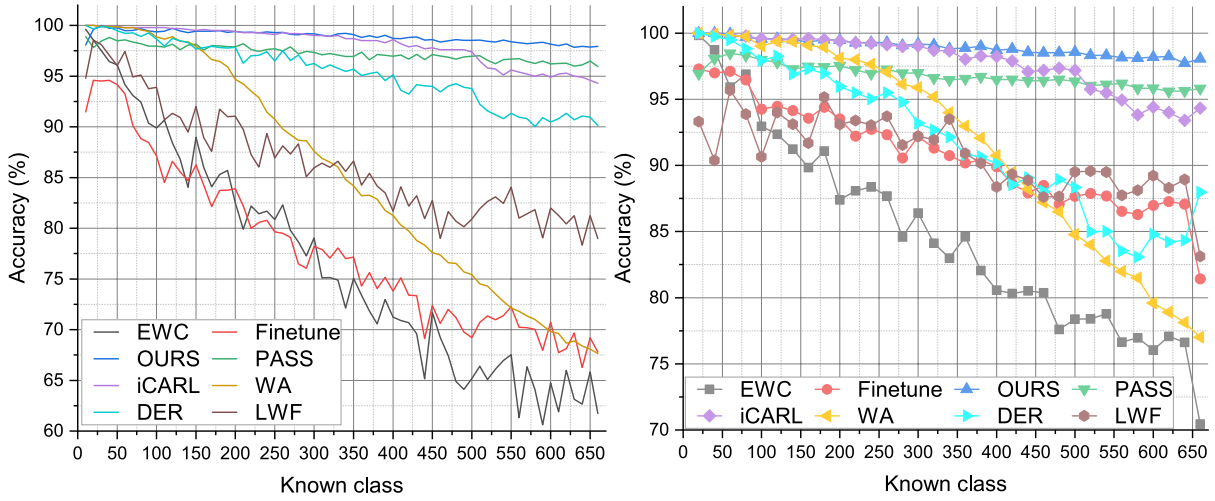


图9. 所提方法在不同初始类别数量和增量类别数量下的增量训练结果

[19] K. Roth、V. Udandarao、S. Dziadzio、A. Prabhu、M. Cherti、O. Vinyals、O. H´enaff、S. Albanie、M. Bethge 和 Z. Akata，“持续多模态预训练实践指南”，2024年12月。[20] D.-W. Zhou、Z.-W. Cai、H.-J. Ye、L. Zhang 和 D.-C. Zhan，“基于预训练模型的域增量学习的双重巩固”，2024年10月。[21] A. Douillard、A. Ram´e、G. Couairon 和 M. Cord，“Dytox：采用动态令牌扩展的持续学习变换器”，2022 年 8 月。[22] K. Merchant、S. Revay、G. Stantchev 和 B. Nousain，“认知通信网络中的射频设备指纹识别深度学习”，*IEEE*信号处理精选主题期刊，卷12，期1，第160–167页，2018年2月。[23] R. Jiang、J. Hu、Z. Xie、Y. Song 和 S. Xu，“发射机识别

基于沃尔泰拉级数的射频指纹”，2025年10月。[24] J. Pfeiffer、A. Kamath、A. R¨uckl´e、K. Cho 和 I. Gurevych，“Adapterfusion：用于迁移学习的非破坏性任务组合”，2021年1月。[25] J. Snell、K. Swersky 和 R. S. Zemel，“用于少样本学习的原型网络”，2017年6月。[26] D.-W. Zhou、F.-Y. Wang、H.-J. Ye 和 D.-C. Zhan，“Pycil：用于类别增量学习的 Python 工具箱”，中国科学：信息科学，第66卷，第9期，197101页，2023年。[27] F. Zhu、X.-Y. Zhang、C. Wang、F. Yin 和 C.-L. Liu，“用于增量学习的原型增强和自监督”，载于*2021*年*IEEE/CVF*计算机视觉与模式识别会议 *(CVPR)*。美国田纳西州纳什维尔：IEEE，2021年6月，第5867–5876页。[28] B. Zhao、X. Xiao、G. Gan、B. Zhang 和 S. Xia，“在类别增量学习中保持判别能力和公平性”，2019年11月。[29] S. Yan、J. Xie 和 X. He，“DER：动态可扩展表征



(a) 起始：10个类别，增量：10个类别

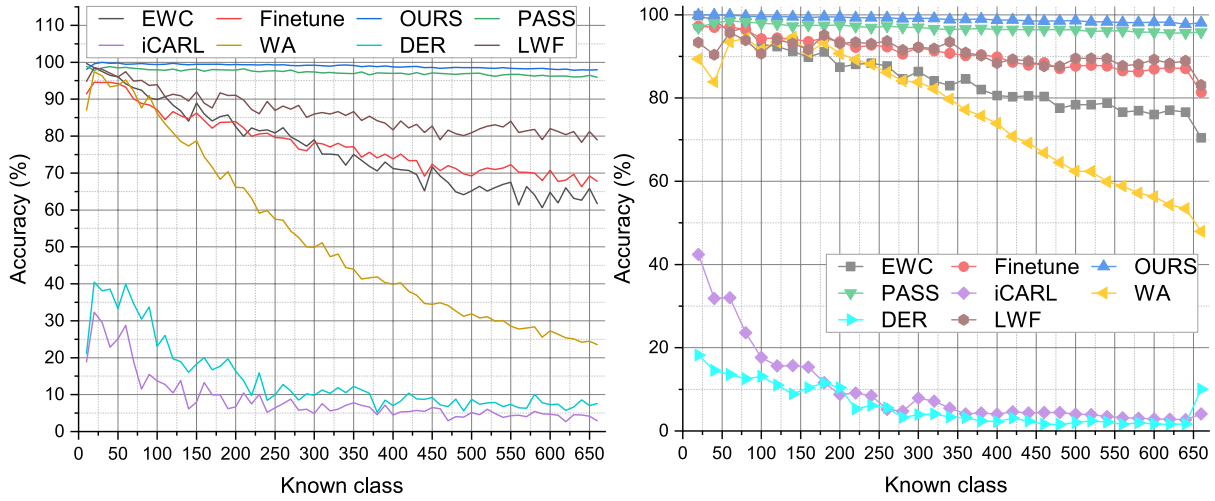
(b) 起始：20个类别，增量：20个类别

(c) 起始：50个类别，增量：50个类别

(d) 起始：100个类别，增量：100个类别

图10. 保留示例样本时我们的方法与其他方法的比较

用于类别增量学习”，2021年3月。



(a) 初始：10个类别，增量：10个类别 (b) 初始：20个类别，增量：20个类别

图11. 未存储旧样本时我们的方法与其他方法的比较

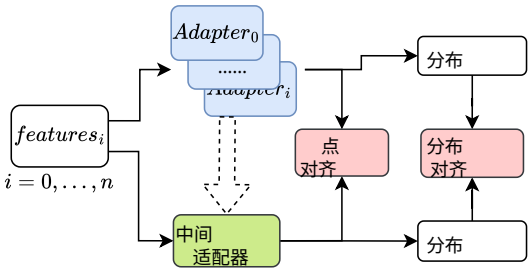


图 12. 训练期间将先前的适配器合并到中间适配器中的蒸馏过程