对于联邦无监督行人重识别的边端联合优化

摘要

行人重识别是从一组图片中重新寻找出已知的行人。虽然不同相机包含了隐私数据，使其不便于在云端共享，但是可以通过联邦学习的方式使得不同相机共同学习，在保护隐私的前提下提高本地训练的质量。当前的研究主要集中在有监督的行人重识别，对无监督的行人重识别及其在各种场景下的推广研究很少。因此本文提出了一种无监督的行人联邦重实别，不仅能减少在实际用场景下的标注工作量，还能够通过联合学习的方式优化训练，达到更好的效果。模型采用了联邦学习的 client-server 架构，并为了处理不同数据集在数据量和数据分布方面的不同，提出了个性化周期和个性化聚类，以避免 client 上的过拟合，使得聚类的参数更符合本地数据的特点。

关键词：行人重实别；联邦学习；无监督学习

1 引言

行人重识别属于计算机视觉中的图像检索问题，在安防等领域有着广泛的应用。一般而言，研究者通过收集大量数据，以达到较好的训练效果。但是在实际应用中，数据相比于数据集更加分散，且为了保护隐私，无法集中数据进行训练。而联邦学习是一种能够使得不同设备共同学习，并且提供隐私保护的机器学习方法。因此，通过联邦学习的方式能够更好地推广现有的行人重识别算法。

近期有一些行人重识别和联邦学习的相关研究。FedReID 达到了很好的效果，但是依赖于现有的标签，不利于模型大规模应用。本文提出了联邦行人重识别的无监督版本。对于客户端上的训练，本文参考了自底向上的聚类算法[1]，在客户端上采用了相关的聚类方法，不断更新伪标签。

在此基础上，本文还提出了客户端的个性化周期和个性化聚类，以及云端上的个性化更新，用来解决客户端之间的统计异质性问题。

2 相关工作

2.1 无监督行人重识别

行人重识别的目标是在不交叉的相机视角中匹配行人。有监督的行人重识别已经取得了很好的效果，但是近年来无监督的重识别受到了更多的关注。

无监督行人重识别主要分为两类。第一类是无监督领域自适应方法，该方法的目的是学习一个模型，该模型在给定源域中的标记数据的情况下，对目标域中的未标记数据取得良好的表现。有的研究通过基于生成对抗网络（GAN）将图像样式从源域转移到目标域来提高目标域的性能[2]，有的则使用图形匹配来为未标记的数据生成伪标签。第二类为纯无监督的重识别。此类工作主要利用自下而上的聚类方法来预测未标记数据的伪标签。该方法需要集中大量数据，导致了潜在的隐私泄露风险，但是相比于第一种，此类方法不依赖于标签的传输，因此为隐私保护提供了更大的可能性。

2.2 联邦学习

联邦学习是一种去中心化训练的机器学习方法，能够很好地保护客户端的隐私。联邦行人重识别（FedReID）[3]实现了对行人的联合学习，并提出了联邦部分平均优化了训练效果。近期有一些无监督联邦学习方面的研究，但是这些方法不适用于行人重识别。

3 本文方法

3.1 本文方法概述

本文主要复现 FedUReID 的工作，包括联邦学习的 client-server 框架、无监督学习的损失函数、伪标签的生成与更新模块等，以及优化本地训练的 Profiler、Controller，优化云端更新的 Personalized Update 操作等。详情如图 1 所示：

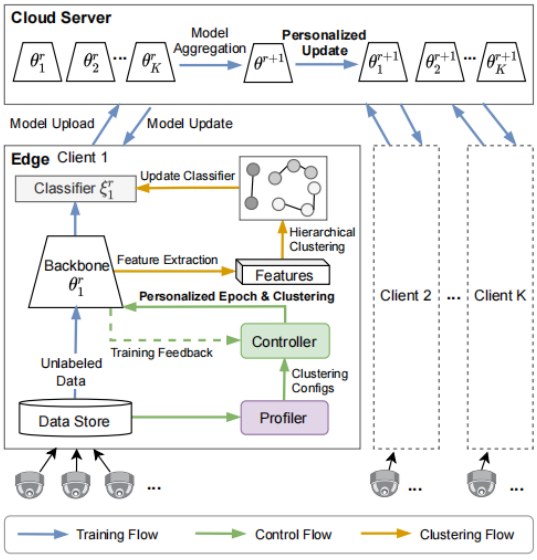


图 1: 方法示意图

本文的系统包括了三种流：训练流、控制流和聚类流。在本地训练流开始前，客户端先从服务端获取模型并且进行初始化，通过 Profiler 预估后续聚类过程中的参数。在每轮训练过程中，客户端的训练周期数会随着达到的准确率进行适当地调整。在本地模型上传到服务端后，服务端聚合得到新的模型。此时得到的新模型不会代替原来的本地模型，而是根据参数的相似度进行了个性化的合并。

3.2 客户端设计

客户端为无监督行人重识别模型的主要执行方。伪标签数据由原数据集生成，将原来的标签改为数据的行的编号。网络采用 resnet50 预训练模型作为 backbone，输出维度为 1024 的特征，为后续的训练与性能评估做准备。

3.2.1 损失函数

我们将第 i 个图像的特征定义为 $\phi(\Theta; x_i)$，则 $v = \dfrac{\phi(\Theta; x_i)}{||\phi(\Theta; x_i)||}$ 为其归一化的结果。定义图像 x 属于第 c 类的概率为：

$$p(c|x, \mathbf{V}) = \frac{\exp(\mathbf{V}_c v/\tau)}{\sum_{j=1}^{C} \exp(\mathbf{V}_j v/\tau)}。$$

其中 $\mathbf{V}$ 为查找表，$\mathbf{V} \in \mathbb{R}^{C \times n_\phi}$, n_ϕ 为特征的维数 1024。C 是当前阶段的类数，即伪标签的总数。在第一轮训练中，$C = N = |\mathbf{X}|$，$|\mathbf{X}|$ 为训练集的大小。在之后轮次的训练，由于相似的类别合并，C 逐渐减小。在反向传播的过程中，我们通过 $\mathbf{V}_{\hat{y}_i} \leftarrow \dfrac{1}{2}(\mathbf{V}_{\hat{y}_i} + v_i)$。

损失函数的最终定义为 $L = -log(p(\hat{y}_i|x_i, \mathbf{V}))$。

3.2.2 伪标签更新模块

伪标签更新是通过聚类的方式，将距离最近的类别合并，并且对相应的伪标签、损失函数进行更新，为下一轮训练做准备的过程。该过程有 2 个关键参数，总共需要合并的数量 m 和每次合并的数量占总数量的比例 mp。

3.2.3 控制流

为了充分利用初始状态下正确的标签信息，对客户端上的训练周期进行个性化调整。具体而言，当某一 batch 准确率高于 95% 时，判断训练足够充分，终止改轮次的训练。由于不同数据集的统计性质不同，需要对本地的训练和聚类过程中的参数进行动态的调整，因此本文根据客户端数据的特性调整参数 m 为 $m_k = \frac{n_k - M_{profile}}{R}$，其中 R 为训练总轮数，n_k 为客户 k 的数据量，$M_{profile}$ 为其预估类总数。

4 复现细节

4.1 与已有开源代码对比

此部分为必填内容。如果没有参考任何相关源代码，请在此明确申明。如果复现过程中引用参考了任何其他人发布的代码，请列出所有引用代码并详细描述使用情况。同时应在此部分突出你自己的工作，包括创新增量、显著改进或者新功能等，应该有足够差异和优势来证明你的工作量与技术贡献。

论文是根据其中的伪代码和流程的详细介绍实现的。因为论文的代码没有开源，复现过程中参考了现有的监督行人重实别 ft-net 方法，联邦行人重识别框架，以及无监督的学习方法的相关代码，复现出了这篇论文。除了原文不同功能模块的实现外，代码还保留了参考代码框架的一些优化方法。论文的伪代码如下：

Procedure 1 Federated Unsupervised Person ReID

Input: Local epoch *E,* batch size *B,* training round *R,* number of selected clients *K,* number of clients N learning rate η, data size *n,* data X

Output: Personalized model θ_k^R of client k , Global model θ^R

Client (θ, k, r):

if *r==0* **then**

  then Initialize the number of clusters $M \leftarrow n_k$;

  Initialize classifier ξ with dimension $v \times M$;

  Initialize pseudo labels $Y \leftarrow \{y_i = i\}_{i=0}^{M-1}$;

  Initialize m.

end

for *each local epoch* $e = 0$ *to E-1* **do**

  for $b \in B$ **do**

    $(\theta, \xi) \leftarrow (\theta, \xi) - \eta \nabla L((\theta, \xi); b)$

    $precision_b \leftarrow$ (batch precision)

    $precision_avg \leftarrow$ (cumulative average precision)

  end

  If any $precision_avg > 0.95$ or $precision_b == 1$

    break;

end

Merging m clusters, $M \leftarrow M - m$;

Update pseudo labels Y with new clusters ;

Update classifier ξ with new dimension $v \times M$;

return θ;

4.2 实验环境搭建

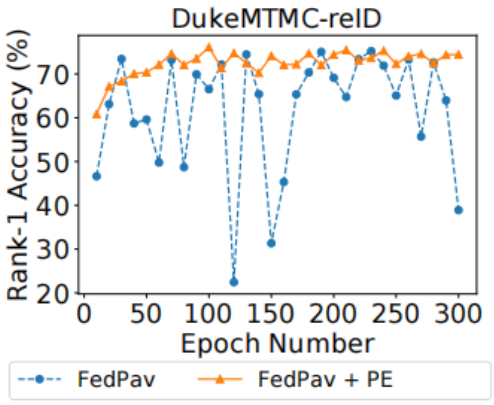
实验使用的 gpu 为 GeForce GTX 2080Ti，使用 cuda10.0 和 cudnn7.6.5，pytorch 环境为 torch1.2.0 和 torchvision0.4.0。使用 market1501 和 mars 分别作为客户端数据集，大约 7 小时训练完成。

4.3 创新点

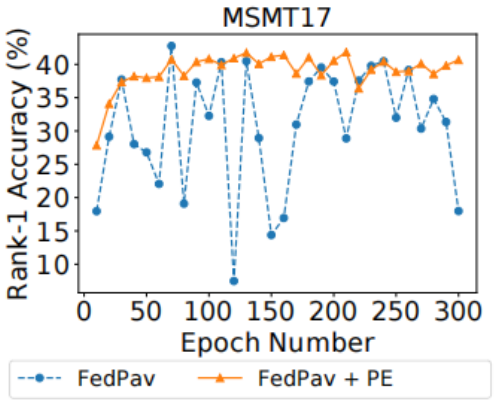
复现工作参考了 FedReID 的联邦学习框架、行人重识别的监督和非监督方法，在没有公开代码的情况下实现了无监督的行人重识别，并且组合了优化方法提高了算法性能。

5 实验结果分析

本部分对实验所得结果进行分析，详细对实验内容进行说明，实验结果进行描述并分析。为了评估行人重识别的性能，我们使用了两个最常见的评估指标：累积匹配特征（CMC）曲线和平均平均精度（mAP）[4]。

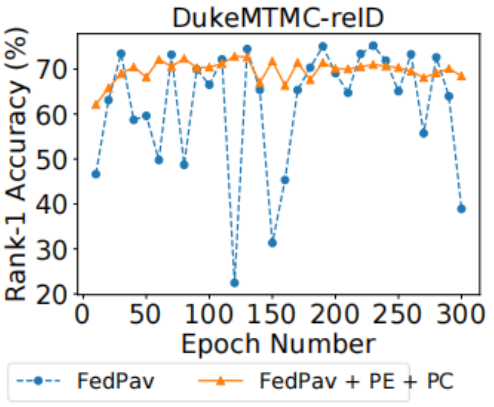


(a)

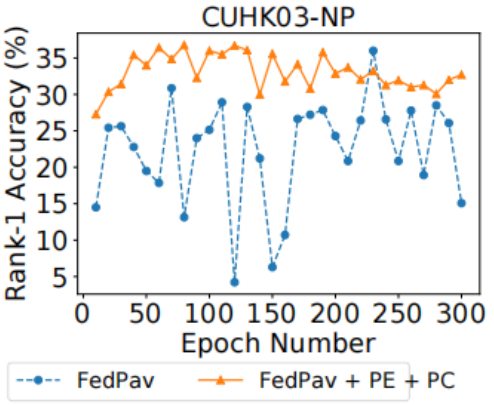


(b)

图 2: 使用个性化周期



(a)



(b)

图 3: 使用个性化周期 + 个性化聚类

如图 2所示，使用个性化周期适时减少训练可以让 rank1 值更加稳定和收敛。从图 3可知，引入个性化聚类能够进一步提升算法的性能。

6 总结与展望

本文展示了无监督行人重识别的实现。通过对现有开源的相关题材的代码进行组合，可以对论文进行复现工作。

虽然论文的复现是成功的，但依然存在一些改进空间。可以增加客户端的数量、进行消融实验等，得出更加系统的结论。此外可以采用更多联邦学习的优化方法对算法进行优化，如根据余弦距离调整聚合时候的权重等，或者使用效果更好的无监督算法。

参考文献

[1] LIN Y, DONG X, ZHENG L, A bottom-up clustering approach to unsupervised person re-identification //AAAI Conference on Artificial Intelligence (AAAI): vol. 2. 2019: 1-8.

[2] WEI L, ZHANG S, GAO W, Person Transfer GAN to Bridge Domain Gap for Person Re-Identification //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2018.

[3] ZHUANG W, WEN Y, ZHANG X, Performance Optimization of Federated Person Re-identification via Benchmark Analysis//Proceedings of the 28th ACM International Conference on Multimedia. 2020: 955-963.

[4] FAN H, ZHENG L, YAN C, Unsupervised Person Re-Identification: Clustering and Fine-Tuning. 2018, 14(4).