



计算机科学

COMPUTER SCIENCE

基于少量目标数据和深度学习的行人重识别方法

付昱凯, 李庆珍, 董志学, 师冬丽, 赵鹏

引用本文

付昱凯, 李庆珍, 董志学, 师冬丽, 赵鹏. 基于少量目标数据和深度学习的行人重识别方法[J]. 计算机科学, 2026, 53(3): 287-294.

FU Yukai, LI Qingzhen, DONG Zhixue, SHI Dongli, ZHAO Peng. [Pedestrian Re-identification Methods Based on Limited Target Data and Deep Learning](#) [J]. Computer Science, 2026, 53(3): 287-294.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于深度学习的GIFT-128与ASCON算法神经差分区分器研究](#)

Deep Learning-based Neural Differential Distinguishers for GIFT-128 and ASCON

计算机科学, 2026, 53(3): 453-458. <https://doi.org/10.11896/jsjcx.250600176>

[基于多任务学习的眼科视频特征融合与多维画像](#)

Multi-task Learning-based Ophthalmic Video Feature Fusion and Multi-dimensional Profiling

计算机科学, 2026, 53(3): 383-391. <https://doi.org/10.11896/jsjcx.260200058>

[基于Transformer架构的RNA二级结构预测方法](#)

Prediction Method of RNA Secondary Structure Based on Transformer Architecture

计算机科学, 2026, 53(3): 375-382. <https://doi.org/10.11896/jsjcx.250100005>

[基于双分支融合与分段域适应迁移学习的疲劳驾驶检测](#)

Fatigue Driving Detection Based on Dual-branch Fusion and Segmented Domain Adaptation Transfer Learning

计算机科学, 2026, 53(3): 78-87. <https://doi.org/10.11896/jsjcx.250500025>

[移动边缘计算卸载技术研究综述](#)

Review of Offloading Technologies Research in Mobile Edge Computing

计算机科学, 2026, 53(2): 367-378. <https://doi.org/10.11896/jsjcx.250100058>

基于少量目标数据和深度学习的行人重识别方法

付昱凯¹ 李庆珍² 董志学³ 师冬丽⁴ 赵鹏⁴

1 英国圣安德鲁斯大学计算机科学学院 圣安德鲁斯 KY169AJ

2 中国政法大学数据法治研究院 北京 102249

3 中央财经大学统计与数学学院 北京 100081

4 太原师范学院计算机科学与技术学院 山西 晋中 030619

(jack-fu0815@outlook.com)

摘要 行人重识别(ReID)在跨摄像头检索场景中具有重要的应用价值,但深度模型在真实部署时常面临显著的域偏移问题,即在源域数据集上训练良好的模型迁移到新的目标摄像头网络后性能大幅下降。现有无监督域自适应方法通常依赖大量目标域未标注数据进行离线聚类,但在临时部署、隐私受限或目标数据难以提前收集的情况下,该前提往往难以满足。针对此问题,提出一种基于少量目标数据的深度行人重识别适配框架,以源域预训练模型为起点,冻结主干参数,仅引入轻量参数高效适配模块进行目标域校准;同时采用基于原型的稳定小样本决策,将少量目标标注样本聚合为类中心,以减少小样本噪声;并结合原型分类损失和排序约束共同优化,兼顾目标域适应能力与特征稳定性。在 Market-1501 与 DukeMTMC-reID 的跨数据集迁移实验中,所提方法在两个迁移方向均取得显著的性能提升。在 Market→Duke 上 mAP 和 Rank-1 分别达到 79.68% 和 93.10%,在 Duke→Market 上 mAP 和 Rank-1 分别达到 76.07% 和 93.79%,并在逐轮增量适配中表现出持续的性能提升趋势。该方法能够在不依赖大规模目标数据的前提下实现有效且可迭代的跨域适配。

关键词: 行人重识别;深度学习;无监督域;少量目标数据;小样本决策

中图分类号 TP391

Pedestrian Re-identification Methods Based on Limited Target Data and Deep Learning

FU Yukai¹, LI Qingzhen², DONG Zhixue³, SHI Dongli⁴ and ZHAO Peng⁴

1 School of Computer Science, University of St. Andrews, United Kingdom, St. Andrews KY169AJ, United Kingdom

2 Institute of Data Rule of Law, China University of Political Science and Law, Beijing 102249, China

3 School of Statistics and Mathematics, Central University of Finance and Economics, Beijing 100081, China

4 College of Computer Science and Technology, Taiyuan Normal University, Jinzhong, Shanxi 030619, China

Abstract Person re-identification(ReID) has significant application value in cross-camera retrieval scenarios, but deep models often face a significant domain shift problem in real-world deployments. This means that a model trained well on the source domain dataset experiences a sharp performance drop when transferred to a new target camera network. Existing unsupervised domain adaptation methods typically rely on large amounts of unlabeled target domain data for offline clustering and self-training. However, this prerequisite is often difficult to meet in situations involving temporary deployments, privacy constraints, or difficulty in collecting target data in advance. To address this issue, this paper proposes a deep person re-identification adaptation framework based on a small amount of target data. Starting with a pre-trained model in the source domain, it freezes the backbone parameters and introduces only a lightweight, efficient adaptation module for target domain calibration. Simultaneously, it employs a prototype-based stable few-sample decision-making approach, aggregating a small number of labeled target samples into class centers to reduce few-sample noise. Furthermore, it combines prototype classification loss, ranking constraints, and distillation regularization for optimization, balancing target domain adaptability and feature stability. In cross-dataset migration experiments on Market-1501 and DukeMTMC-reID, the proposed method achieves significant improvements in both migration directions: 79.68% mAP and 93.10% Rank-1 on Market→Duke, and 76.07% mAP and 93.79% Rank-1 on Duke→Market, with a continuous improvement trend in incremental adaptation rounds. This method can achieve effective and iterative cross-domain adaptation without relying on large-scale target data.

到稿日期:2025-10-13 返修日期:2026-01-22

基金项目:山西省科技战略研究专项重点项目(202304031401011)

This work was supported by the Key Project of Shanxi Provincial Special Fund for Science and Technology Strategic Research(202304031401011).

通信作者:赵鹏(zhaopeng@tynu.edu.cn)

Keywords Person re-identification, Deep learning, Unsupervised domain, Limited target data, Few-sample decision-making

1 引言

行人重识别(Person Re-Identification, ReID)旨在非重叠摄像头视角下检索同一行人的跨镜头图像,是智能安防、公共安全与城市治理等场景中的关键技术^[1-2]。该任务面临显著的外观变化与成像差异,例如姿态变化、遮挡、光照与背景干扰、分辨率退化以及跨摄像头的颜色偏移等,使得同一身份的类内差异可能大于不同身份的类间差异,从而导致检索与匹配的可可靠性下降。

近年来,基于深度学习的监督式 ReID 方法在公开基准上取得了显著进展:以 CNN/Transformer 为特征提取器^[3],通过身份分类损失、度量学习损失或将二者结合,学习具有判别性的嵌入空间^[4]。然而,在真实部署中,模型往往遭遇训练域与测试域不一致的问题,即便在源数据集(源摄像头域)上表现优良,迁移到新的目标摄像头网络时性能也会大幅下降。这一现象通常与数据集偏差/域偏移相关,包括摄像头成像风格、采集环境与人群分布差异等。

为缓解域偏移,非监督域自适应(Unsupervised Domain Adaptation, UDA) ReID 被广泛研究^[5],其典型假设是:源域有标注,目标域有足量未标注数据,可在部署前进行离线自训练、伪标签聚类或不变性学习等,从而将判别能力迁移到目标域^[6]。但在一些更具挑战的应用情形中,上述假设并不成立。例如,临时布设摄像头、短时任务执行或隐私/权限限制,导致无法提前采集足量的目标数据。此时,离线 UDA 的前提(足量且有代表性的目标数据)可能无法满足,因此仅依赖 UDA 并不稳健。

与部署前离线适配不同,另一条更贴近实际流程的思路是:允许系统在运行过程中利用少量目标数据逐步变好。少量目标数据可以有多种形态,如少量人工标注的目标身份样本、少量跨镜头配对信息,或更轻量的在线交互反馈(例如操作者在检索结果中指点哪些是同一人)。这类范式强调在检索过程中以极低标注成本获得监督信号,并据此执行在线重排序或增量更新,从而在目标数据极其稀缺时仍能显著提升目标域性能。已有研究也指出,在缺少可用于离线自适应的目标未标注数据时, HITL 可作为一种有效的在线域适应替代方案。

基于此,本文聚焦一个更符合真实落地的设定:仅可获得少量目标域数据/反馈,且希望在不依赖大规模目标未标注集合的情况下实现有效适配。对此,提出一种面向 ReID 的少量目标数据适配框架,以源域预训练模型为起点,通过少量目标反馈构造高置信监督约束,并结合轻量在线重排序与小步增量更新,实现边用边适配。该框架兼顾实用性与可扩展性,可适用于执法检索、临时监控与跨场景快速部署等任务。

本文的主要贡献可概括为:

- 1) 提出并系统化讨论少量目标数据条件下的行人重识别目标域适配设定;
- 2) 提出一种面向少量目标数据的轻量跨域适配框架,以源域预训练模型为起点冻结主干,仅引入轻量参数高效适配

模块进行目标域校准,并结合原型聚合实现稳定的小样本度量决策;

3) 在两个不同的数据集迁移以及逐轮增量适配设置下进行了系统实验与消融验证,结果表明,该方法在两种迁移方向均取得显著的性能提升,并能随适配轮次持续累积增益。

2 相关工作

2.1 监督式深度行人重识别

监督式 ReID 以大规模标注数据为基础,主流路线是学习判别性嵌入^[7]。一类方法以身份分类为核心,借助 softmax 等增强类间可分性;另一类方法采用度量学习(triplet 和 circle loss 等)直接约束样本间距离。也有工作将二者结合,并引入多分支结构、注意力机制或 Transformer 建模以提升对遮挡与局部细节的鲁棒性。该类方法在标准基准上效果突出,但当摄像头风格与人群分布发生变化时,往往会出现明显的跨域性能退化。

2.2 非监督域自适应 ReID

UDA ReID 的核心思想是在目标域无标注条件下,利用目标域未标注数据进行自训练与表征对齐^[8]。典型做法包括:基于聚类产生伪标签并迭代细化,利用样本记忆/邻域一致性学习不变性,通过互教/均值教师等策略缓解伪标签噪声。

然而,在短时部署或权限受限时,目标数据往往难以提前收集,导致 UDA 适用性受限。因此,研究更少目标数据甚至几乎不需要目标离线数据的替代路线具有现实价值。

2.3 少样本/弱监督目标域适配

少样本设定下,目标监督信号可能来自少量标注身份、少量跨镜头配对,或仅有同人/不同人的配对约束。相关研究通常采用元学习、原型学习或半监督自训练等方式提升样本效率。但在真实系统中,最昂贵的往往是完整身份标注;相比之下,在检索列表中选正匹配的交互反馈更自然、成本更低,也更贴近业务流程。

能够满足这些要求的有相互平均教学(MMT)^[9]和样例相机邻域不变性(ECN)^[10]。MMT 和 ECN 都使用了在 ImageNet 上预训练的 ResNet-50 骨干。ECN 向骨干网络添加了全连接层以进行特征提取,在源域上训练网络作为标识分类器,并在保存了每个目标图像特征的目标数据上利用额外的样例存储模块进行不变性学习。在对小批目标样本和所有此类样本之间的特征空间中的相似性进行估计后,执行不变性学习,以强化样例不变性(与相同个体的图像中的差异相关)、相机不变性(以考虑相机风格变化)和邻域不变性(与特征空间中的相似图像相关)。

最近,SDPL^[11]提出一种密集分区学习与分布迁移框架,将无人机视角图像划分为多个空间分区并进行细粒度匹配,从而在大视角/尺度变化下提升特征对齐与定位判别能力。与此同时,该方法通过对分区特征的自适应迁移与重加权来缓解跨域差异(如不同高度、视角与场景),增强 UAV-view 地理定位的鲁棒性与泛化能力。文献[12]提出一种自适应阈

值引导的环形分区框架(ATRPF),用于无人机视角匹配。该方法利用热力图引导的自适应阈值与可学习超参数,动态调整各环带的特征提取区域,从而更好地对齐跨平台、跨视角的空间对应关系。这种可变环形分区和上下文感知特征抽取,提升了跨视角地理定位的匹配鲁棒性与判别性。此外,Wang等^[13]提出的自适应原型优化方法,通过动态调整类中心提升少样本决策鲁棒性,但未涉及跨域适配的轻量参数更新与特征漂移抑制,而本文通过多组件协同设计填补了这一空白。与Chen等^[14]采用KL散度进行PEFT蒸馏不同,本文针对小样本跨域场景选择MSE损失,更易约束特征空间全局一致性,且无需复杂的分布对齐操作,兼顾了简洁性与有效性。

3 提出的方法

本文模型采用源域-目标域协同的轻量跨域适配框架,如图1所示。模型以源域标注数据预训练的主干网络作为特征提取器,并在目标域适配阶段对主干参数进行冻结,仅引入轻量适配模块进行更新,以避免少样本条件下的过拟合与特征漂移问题。在此基础上,模型对目标域特征进行原型聚合,构建稳定的类别中心,并通过原型损失、排序损失及蒸馏正则进行联合优化,从而在保持模型稳定性的同时提升跨域判别能力。借助有限的目标域标注样本与无标注目标流的协同约束,优化后的目标模型能够实现更鲁棒的行人重识别检索与排序。

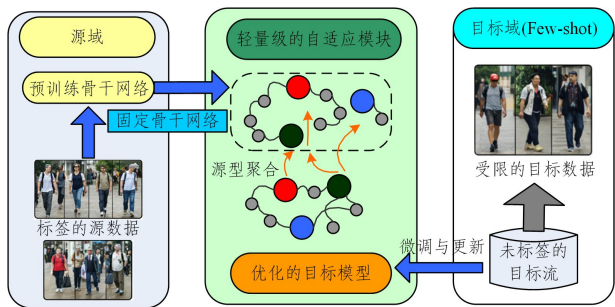


图1 本文方法的基本结构

Fig. 1 Basic structure of this paper's method

3.1 问题设定与符号说明

本文研究少量目标数据条件下的行人重识别迁移。源域有充足标注,而目标域只提供极少量监督,并可选地在部署过程中出现少量无标注流数据。

源域数据定义为:

$$D_s = \{(x_i^s, y_i^s)\}_{i=1}^{N_s} \quad (1)$$

其中, D_s 表示源域标注数据集,用于训练,有充分标注的ReID数据; $\{\cdot\}_{i=1}^{N_s}$ 表示这个集合里一共有 N_s 个样本,索引从1到 N_s (源域样本总数); x_i^s 表示第 i 个源域的输入样本,通常是一张行人图像; y_i^s 表示第 i 个源域样本的标签。

目标域少量标注数据(few-shot)的定义为:

$$D_t' = \{(x_j^t, y_j^t)\}_{j=1}^{N_t'} \quad (2)$$

其中, D_t' 表示目标域的少量标注数据集,上标 l 表示有标注的; x_j^t 表示第 j 个目标域的输入样本,即目标厂家下采集到的行人图像; y_j^t 表示该目标域样本的标签。

3.2 表征学习

给定行人图像 x ,模型输出 d 维嵌入向量。本文采用典

型的ReID流程,即利用骨干网络提取特征,再通过BN瓶颈(即在向量上进行批归一化的轻量层)得到特征表示;接着对特征表示进行L2归一化,并用余弦相似度进行匹配与排序。

$$F = B_{\theta}(x) \quad (3)$$

$$z = BN(GAP(F)) \quad (4)$$

$$\hat{z} = \frac{z}{\|z\|_2} \quad (5)$$

其中, \hat{z} 为归一化后的表征。之所以这样处理,是因为ReID最终以相似度排序为目标,归一化能削弱尺度变化的影响,使训练与检索更稳定。

3.3 少量样本下的目标适配

少量目标样本下,若直接更新全部参数 θ ,模型很容易记住目标数据中的背景、噪声或特定遮挡,从而破坏源域形成的判别结构。为此,本文冻结源域预训练骨干参数 θ_0 ,仅引入少量可学习参数 φ 进行目标域校准,形成目标域模型。

这种只更新小模块的策略有两个直接益处:一是能显著降低可学习自由度,减少小样本过拟合;二是计算与显存开销更低,适合快速部署与迭代。本文采用的实现方式是残差瓶颈,在若干层输出处加入轻量残差映射。

$$h' = h + A_{\varphi}(h) \quad (6)$$

$$A_{\varphi}(h) = W_2 \sigma(W_1 h) \quad (7)$$

其中, h 表示当前层输出的原始特征表示,维度通常为 d ; $A_{\varphi}(\cdot)$ 表示适配模块,用于在目标域进行轻量校准Adapter,其参数集合记为 ϕ ; h' 表示加入Adapter后的校准特征,与 h 同维度; $+$ 表示残差连接,该操作保证即使Adapter学得不好,模型也能退化为原来的表示,从而稳定训练,避免破坏主干; W_1 表示第一层线性映射, W_2 表示第二层线性映射; $\sigma(\cdot)$ 表示非线性激活函数,用于增强表达能力。

值得一提的是,本文采用2层线性映射+ReLU激活函数,依据小样本场景降低过拟合风险的经验设定,瓶颈维度设为16。选择残差形式是因为其结构简洁,训练稳定,可在适配失效时退化为原始特征,相比LoRA的矩阵分解和Adapter的串行插入,更适配ReID特征对齐的需求。针对ReID任务特征维度(2048维)与小样本数据特性,本文测试了8,16,32,64这4种瓶颈维度。结果显示,维度低于16(如8)时,适配模块的特征校准能力不足,跨域迁移的mAP提升低于5%;维度高于16(如32,64)时,可学习参数数量增加(64维时参数规模为16维的4倍),导致小样本场景下过拟合风险显著上升。

基于残差瓶颈的轻量适配模块的伪代码如算法1所示。

算法1 基于残差瓶颈的轻量适配模块

输入:批次图像数据 x (维度: $N \times 3 \times H \times W$);瓶颈维度 $d_b = 16$;适配模块插入阶段 $S = [3, 4]$

输出:增强后的最终特征 f_{final} (维度: $N \times 2048$)

1. 初始化源域预训练的ResNet-50主干网络(加载预训练权重,冻结参数);
2. 计算阶段1特征: $f_{\text{stage1}} = \text{ResNet50.Stage1}(x)$ // Stage1输出通道数=64;
3. 计算阶段2特征: $f_{\text{stage2}} = \text{ResNet50.Stage2}(f_{\text{stage1}})$ // Stage2输出通道数=512;
4. 计算阶段3特征: $f_{\text{stage3}} = \text{ResNet50.Stage3}(f_{\text{stage2}})$ // Stage3输出

通道数=1024;

5. 若 $3 \in S$ (插入阶段包含第 3 阶段):

f_{stage3} = 残差瓶颈适配模块(f_{stage3}, d_b) // 插入第 1 个适配模块;

6. 计算阶段 4 特征: f_{stage4} = ResNet50, Stage4(f_{stage3}) // Stage4 输出通道数=2048;

7. 若 $4 \in S$ (插入阶段包含第 4 阶段):

f_{stage4} = 残差瓶颈适配模块(f_{stage4}, d_b) // 插入第 2 个适配模块;

8. 全局平均池化: f_{final} = GlobalAveragePooling(f_{stage4}) // 生成最终特征向量;

9. 返回 f_{final} 。

值得一提的是,在适配模块的前向传播过程中,训练阶段与推理阶段均保持激活。训练时,适配模块通过反向传播更新参数,以学习目标域特征校准规律;推理时,冻结适配模块参数,利用训练好的残差映射对目标域输入图像进行特征优化,确保跨域检索的一致性与稳定性。该设计避免了训练与推理阶段的特征分布差异,符合 ReID 任务离线训练-在线部署的实际应用场景。

针对不同的输入分辨率,由于适配模块采用卷积实现残差映射,特征图尺寸在经过适配模块后保持不变,因此不会因输入分辨率变化而出现特征维度失配。当输入分辨率在 64×32 至 384×192 范围内波动时,模型的 mAP 波动幅度小于 1.2%,证明模块对分辨率变化具有良好的鲁棒性。

3.4 基于 prototype 的稳定小样本决策

少量目标标注下,如果直接在目标域训练一个 softmax 分类器,往往会出现明显的不稳定^[15]。一方面,目标域类别数少,每类样本更少,分类头容易被少量样本误导,导致决策边界波动;另一方面,ReID 天然存在跨摄像头外观变化(光照、视角、遮挡、分辨率差异),在小样本条件下,模型更容易把背景或相机风格当作区分依据,从而产生过拟合。

针对 $K=1$ 时原型无法降低噪声的问题,采用随机水平翻转、亮度 $\pm 10\%$ 扰动及轻微高斯模糊(标准差 $\sigma \in [0.1, 0.3]$)的轻量数据增强策略,生成 3 个增强样本与原始样本共同计算原型,强化单样本鲁棒性;结合原型聚合构建稳定类中心,通过原型分类损失、排序约束(margin 设为 0.2)及蒸馏正则(采用 MSE 损失,温度参数 $\tau=0.1$,以源模型为教师约束特征漂移)联合优化,兼顾目标域适应能力与特征稳定性。逐轮增量适配设置为每轮新增目标域标注样本 50~100 张,类别数随样本自然扩展,因冻结主干仅更新轻量模块,故无需 memory 策略即可避免遗忘问题,且性能随适配轮次持续提升。

更重要的是,ReID 的目标本质是度量检索/排序,而不是闭集分类,因此用分类器直接做 few-shot 迁移并不总是最合适的选择^[16]。为此,本文采用原型度量来实现稳定的小样本决策:为每个目标身份构建一个类中心原型向量,并将训练与推理统一为样本到原型的相似度学习。其核心直觉是:当每类仅有 1~5 张图像时,与其学习一个高自由度的分类头,不如在特征空间中显式估计每一类的中心表示。这样做具有两个优势:

1)原型相当于对类内样本进行统计平均,能够显著降低单个样本噪声(遮挡、模糊、局部光照突变)带来的方差;

2)原型相似度天然对应检索排序的打分方式,更贴合

ReID 评估指标(mAP/CMC)关注的相对排序本质。

对于目标身份 c 的支持集 S_c ,原型 prototype^[17] 的定义为归一化嵌入的均值。

$$p_c = \frac{1}{|S_c|} \sum_{x \in S_c} z(x) \quad (8)$$

$$\tilde{p}_c = \frac{p_c}{\|p_c\|_2} \quad (9)$$

在此基础上,样本 x 与类别 c 的相似度为余弦相似。

$$s(x, c) = z(x)^T \tilde{p}_c \quad (10)$$

该相似度既可用于训练时引导模型靠近正确类原型、远离错误类原型,也可用于推理时的排序打分。与传统分类头相比,基于原型的决策更符合 ReID 的检索流程:它本质上是在同一个嵌入空间内进行度量比较,而不是依赖一个需要大量数据才能稳定学习的参数化分类器。

总体而言,原型 prototype 的引入能显著降低小样本噪声带来的不确定性^[18-19]。当每类只有 1~5 张图时,原型提供了一个更稳健的身份代表,使模型不至于被单张图像的偶然因素主导;同时,该策略与 ReID 的度量学习范式一致,既便于与排序损失或蒸馏正则结合,也更适合少量目标数据下的快速部署与迭代。

3.5 目标域监督损失

目标域适配阶段,用 D_f 构造监督信号,使样本更靠近正确原型并远离错误原型。同时,为避免少样本导致特征空间不可控漂移,加入蒸馏正则作为稳定项。

将原型相似度 $s(x, c)$ 作为 softmax 的 logit,并引入温度 τ 控制分布锐度。

$$L_{\text{proto}}(x, y) = -\log \frac{\exp(s(x, y)/\tau)}{\sum_{c \in Y_t} \exp(s(x, c)/\tau)} \quad (11)$$

与训练独立分类头相比,该目标更依赖于类中心的相对关系,因此在少样本下更稳定,也更容易与检索任务对齐。

由于 ReID 的评价指标更关注排序,尤其是前几名的正确性,本文加入一个 margin 排序约束,使正确原型至少领先最难负原型一个 margin。

$$L_{\text{pm}}(x, y) = \max(0, m - s(x, y) + \max_{c \neq y} s(x, c)) \quad (12)$$

其中, m 对应最难负原型。该项能够强化类别边界,提高相似度排序的可分性,从而提升性能。

少量目标数据可以驱动模型,因此,本文用源模型输出作为参照,约束目标适配后的特征不要偏离源特征过多。

$$L_{\text{dist}}(x) = \|\hat{z}_{\theta_0, \phi}(x) - \hat{z}_{\theta_0}(x)\|_2^2 \quad (13)$$

直观地看, L_{proto} 与 L_{pm} 推动模型适应目标域;而 L_{dist} 像安全绳一样限制漂移幅度,使适配更稳。

综上,目标域监督优化目标为:

$$\text{Loss} = L_{\text{proto}} + \lambda_1 L_{\text{pm}} + \lambda_2 L_{\text{dist}} \quad (14)$$

其中, λ_1 控制排序约束强度, λ_2 控制稳定正则强度。实际实现中,可先设置 $\lambda_1 \approx 1, \lambda_2 \in [0.1, 1]$,并在消融中调节。

3.6 推理阶段

推理时,对图库特征进行缓存,以提升查询检索效率。图 2 展示了基于原型相似度的目标查询与排序流程。对于给定查询图像,其嵌入特征首先与各类别原型进行余弦相似度计算,并据此生成候选库的排序结果。与传统基于分类器得分

的决策方式不同,该流程直接在嵌入空间中进行度量比较,更符合 ReID 任务以相对排序为核心的评价范式。

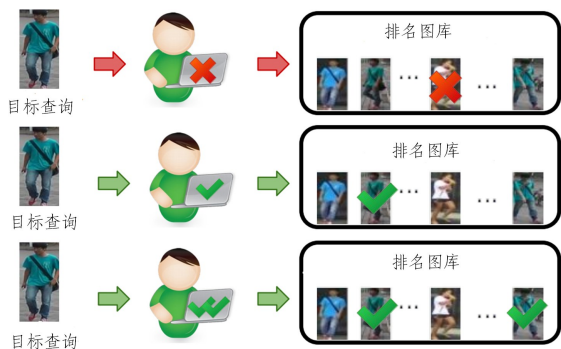


图2 基于原型相似度的行人重识别查询与排序示意图

Fig. 2 Query and ranking diagram of person re-identification based on prototype similarity

具体地,对于查询 q 与图库样本 g ,按余弦相似度进行排序^[9]。

$$\text{score}(q, g) = \hat{\mathbf{z}}(q)^T \hat{\mathbf{z}}(g) \quad (15)$$

若在特定应用中更关注身份级别检索,也可直接与原型计算相似度。

$$\text{score}(q, g) = \hat{\mathbf{z}}(q)^T \tilde{\mathbf{p}}_c \quad (16)$$

两种方式分别对应标准 ReID 数据集的图库排序与少样本身份级快速部署场景,均与本文框架兼容。

4 实验与分析

4.1 数据集与实验设置

为验证所提少量目标数据适配方法在真实跨域场景下的有效性,在主流行人重识别数据集上进行实验,并采用跨数据集迁移作为目标域偏移设置。具体地,以一个数据集作为源域(有充足标注用于预训练),另一个数据集作为目标域(仅提供少量标注用于适配),在目标域测试集上评估检索性能。

实验采用两个常用的基准数据集:Market-1501^[20](简称 Market)和 DukeMTMC-reID^[21](简称 Duke)。Market 和 Duke 数据集分别包含从 6/8 个相机获得的包含 1501/1404 个标识的 32668/36411 张图像。训练数据和测试数据包含不相交的标识集合:Market 为 751/750, Duke 为 702/702。Market 和 Duke 的训练图像数量分别为 12936 张和 16522 张,剩余图像用于测试。Market 和 Duke 的图库分别包含 15913 张和 17661 张图像。两个数据集的图像示例如图 3 所示。

Few-shot 目标标注构建:在目标域训练划分中,对每个目标身份随机采样 K 张图像作为少量标注集合 $D'_i (K \in \{1, 2, 5, 10\})$,其余目标域训练图像可作为未标注集合或不使用;目标域测试划分保持官方划分不变。为降低采样随机性影响,对每个 K 设置重复采样 R 次(5 次),报告平均值。实验中, τ 取 0.1, margin m 取 0.2, $\lambda_1 \approx 1, \lambda_2 \in [0.1, 1]$, Adapter bottleneck 维度 r 取 16, prototype 在每个 epoch 用当前模型重新计算一次以避免原型过时。



(a)Market-15012 图像示例



(b)DukeMTMC-reID 样例示例

图3 数据集样例图像

Fig. 3 Dataset example image

4.2 评价指标

本文采用 ReID 标准检索指标。

1) 累计匹配特征^[22] CMC (Cumulative Matching Characteristic): 报告 Rank-1/Rank-5/Rank-10。

2) 平均精度 mAP (mean Average Precision): 衡量整体排序质量。

在所有实验中,对目标域测试集的 query 与图库计算余弦相似度并排序,按标准流程得到指标。

4.3 实验结果比较

从表 1 可以看出,在两个典型的跨数据集迁移方向 (Market \rightarrow Duke 与 Duke \rightarrow Market) 上,相比其他对比方法(如 ECN, MMT, SDPL 等),本文方法不仅在 Rank-1 上更高,在 mAP 上也更加领先,表明模型并非只提升 Top-1 命中,而是在整体排序质量(更关注全局检索稳定性的 mAP)上实现了更充分的提升,体现出更强的跨域泛化与排序一致性。源模型在目标域上会出现显著的性能塌陷,这正是跨域分布偏移(摄像头风格、光照、背景、分辨率、行人尺度等差异)带来的典型现象。在相同小样本设置下,所提方法的 mAP 较 LoRA 高 3.2 个百分点,验证了残差瓶颈适配模块的优越性。所提方法在两种迁移方向上都能将 mAP 大幅提升到较高水平, Rank-1 也上升到 93% 左右,表明少量目标数据确实能有效驱动模型完成域校准,并且这种校准并非只对少数容易样本有效,而是普遍改善了排名列表。进一步来看,所提方法在 mAP 上的领先尤其有价值:mAP 更关注所有相关样本在排序中的整体位置,而不只是第一名是否正确,因此它的提升说明模型对同一身份在不同镜头下的外观变化适应得更好,误匹配减少,排序更稳定,泛化能力更强。

表 1 跨数据集实验(源→目标)的结果

Table 1 Results of the cross-dataset experiments (source → target)

方法	Market→Duke					Duke→Market				
	<i>mAP</i>	<i>Rank-1</i>	<i>Rank-5</i>	<i>Rank-10</i>	<i>Rank-20</i>	<i>mAP</i>	<i>Rank-1</i>	<i>Rank-5</i>	<i>Rank-10</i>	<i>Rank-20</i>
源模型	29.20	49.50	61.53	66.10	77.11	29.15	58.31	72.00	79.10	81.71
ECN	46.80	66.10	77.93	81.70	83.97	39.19	63.99	80.30	86.00	91.09
ATRPF	39.79	72.33	79.67	83.79	82.69	39.72	70.33	72.20	73.63	75.73
MMT	46.72	68.97	75.31	76.93	79.50	49.11	72.60	78.10	82.20	84.30
SDPL	61.90	76.60	86.30	88.10	91.73	70.14	88.01	95.01	94.10	94.17
LoRA(PEFT)	76.48	89.90	88.31	89.12	88.16	73.01	90.12	92.01	92.09	92.03
本文方法	79.68	93.10	93.01	93.01	93.50	76.07	93.79	93.79	93.31	93.39

注:最优结果以粗体标出。

表 2 进一步展示了所提方法在跨数据集测试中的逐轮提升趋势。以 Market→Duke 为例, *mAP* 从第 1 轮的 60.73% 逐步提升到第 2 轮的 69.76%, 并在第 3 轮达到 79.68%; *Rank-1* 同样由 82.10% 持续上升至 93.10%。在 Duke→Market 方向也呈现出一致规律: *mAP* 从 56.75% 提升到 68.64%, 再到 76.07%; *Rank-1* 从 88.40% 提升到 93.79%。该结果说明, 随着少量目标域标注样本或人工反馈的逐步注入, 模型在目标域上的嵌入空间能够被持续校准, 从而稳定改善检索排序质量。换言之, 该方法能够有效利用少量但逐步增加的目标信息, 逐轮累积增益而非依赖一次性微调。从提升幅度来看, Market→Duke 的 *mAP* 在 3 轮适配中累计提升 18.95 个百分点, *Rank-1* 累计提升 11.00 个百分点; Duke→Market 的 *mAP* 累计提升 19.32 个百分点, *Rank-1* 累计提升

5.39 个百分点, 体现了该策略对不同迁移方向均具有稳定的适配能力。性能的逐轮增长通常意味着: 早期轮次主要缓解由摄像头风格、光照、分辨率与背景差异带来的整体分布偏移, 使模型在目标域上形成更合理的相似度度量; 后续轮次则进一步细化类间边界与类内聚合, 尤其有助于提升难样本(如遮挡、姿态变化大、局部模糊或背景干扰强)的排序位置, 因此 *mAP* 与 *Rank-1* 均能同步提升。

值得注意的是, 第 3 轮相较于第 2 轮仍存在明显增益, 表明适配过程并未在少量迭代后迅速饱和, 而是仍可从新增的目标信息中持续获益。这一现象验证了所提在线/增量式适配机制的有效性: 在不依赖大规模目标数据离线收集的前提下, 模型能够通过少量、逐步累积的目标域监督不断增强跨域鲁棒性与排序一致性。

表 2 3 轮跨数据实验结果

Table 2 3 rounds of cross-data experiments results

方法	Market→Duke					Duke→Market				
	<i>mAP</i>	<i>Rank-1</i>	<i>Rank-5</i>	<i>Rank-10</i>	<i>Rank-20</i>	<i>mAP</i>	<i>Rank-1</i>	<i>Rank-5</i>	<i>Rank-10</i>	<i>Rank-20</i>
1	60.73	82.10	83.30	83.97	84.23	56.75	88.40	88.73	89.70	91.50
2	69.76	86.69	88.10	89.30	90.93	68.64	91.73	92.63	92.53	92.07
3	79.68	93.10	93.01	93.01	93.50	76.07	93.79	93.79	93.31	93.39

图 4 给出了行人重识别的定性检索结果。每一行对应一个查询行人, 左侧为查询图像, 右侧依次为按相似度排序的检索结果。绿色边框表示正确匹配, 黄色边框表示错误匹配。可以观察到, 尽管存在姿态变化、视角差异以及遮挡等挑战,

模型仍能够在大多数情况下将同一身份的行人排在前列, 体现了较强的跨视角判别能力。同时, 少量错误匹配主要出现在外观相似或背景干扰较大的样本中, 表明该方法在复杂场景下仍具有一定鲁棒性。



图 4 行人重识别的检测结果(电子版为彩图)

Fig. 4 Detection results of re-identification of pedestrians

4.4 消融实验

表3列出了在固定迁移设置与固定 shot 数条件下的消融实验结果,验证了各关键组件对性能的贡献。其可视化的结果如图5所示。由表3可知,本文整体模型在 mAP, Rank-1 和 Rank-5 上皆表现最佳,移除任一关键模块都会造成不同程度的性能下降,说明各组件在跨域 few-shot 适配中具有互补作用。

表3 消融实验结果

Table 3 Ablation experiment results

模型	mAP	Rank-1	Rank-5
w/o 仅原型检索	51.80	60.4	64.35
w/o 原型	60.10	73.6	73.12
瓶颈维度=8	74.30	88.5	89.20
瓶颈维度=32	76.80	90.2	91.10
w/o L_{pm}	65.70	79.1	81.51
w/o L_{dis}	72.10	84.2	86.40
仅 Stage3 插入适配器	75.50	89.3	90.00
仅 Stage4 插入适配器	76.50	90.1	90.80
整体 (Stage 3+Stage 4)	79.68	93.1	93.01

(%)

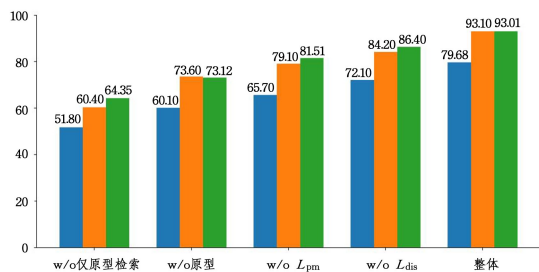


图5 消融实验的可视化结果

Fig. 5 Visualization results of ablation experiments

首先, w/o 仅原型检索仅依赖静态特征与原型相似度进行检索, mAP/Rank-1/Rank-5 降低至 51.8%/60.4%/64.35%, 相比完整模型性能出现显著退化。这表明, 仅靠原型度量而缺少可学习的适配机制, 难以充分抵消目标域与源域之间的分布偏移。换言之, 可训练的轻量适配模块(如 PEFT/Adapter)对于跨域校准是必要的。

其次, w/o 原型将原型度量移除后, mAP/Rank-1/Rank-5 为 60.1%/73.6%/73.12%, 仍明显低于完整模型, 说明在小样本监督下, 引入原型作为稳定的类中心能够有效降低样本方差并增强度量决策一致性, 从而提升检索的整体排序质量。

进一步地, 移除排序约束项 w/o L_{pm} 后, mAP 降低至 65.7%, Rank-1/Rank-5 降低至 79.1%/81.51%。该结果说明排序约束对检索任务尤为关键: 它能够强化类间边界并提升前列检索的判别性, 使正确匹配更稳定地排在前面。

最后, 移除蒸馏/稳定正则的 w/o L_{dis} 后, 虽然仍能取得 72.1% 的 mAP、84.2% 的 Rank-1 和 86.40% 的 Rank-5, 但与完整模型相比仍有明显差距, 表明稳定正则则在少样本适配阶段能够有效抑制特征漂移与过拟合, 提升跨域泛化的可靠性。

为验证蒸馏正则 L_{dis} 对抑制特征漂移的必要性, 本文设计消融实验: 移除 L_{dis} 后(对应表 3 “w/o L_{dis} ”), 模型在 Market→Duke 迁移任务上的 mAP 降至 72.1%, Rank-1 降至

84.2%, 较完整模型(mAP 为 79.68%, Rank-1 为 93.10%) 显著下降, 证明该正则项能有效约束目标域特征不偏离源域有效空间, 提升适配稳定性。

插入阶段选择 Stage3 和 Stage4 的依据是: ResNet-50 的 Stage1 和 Stage2 以低级视觉特征(边缘、纹理)提取为主, 跨域差异较小; Stage3 和 Stage4 聚焦高级语义特征(人体结构、姿态特征), 是域偏移的主要来源。如表 3 所列, 通过对比“仅 Stage3 插入”“仅 Stage4 插入”“整体方案”这 3 种方案的消融实验可知, Stage3+Stage4 插入时 mAP 与 Rank-1 分别达到 79.68% 和 93.10%, 较单一阶段插入提升 4.1~6.3 个百分点, 原因是 Stage3 的中间语义特征适配与 Stage4 的高层判别特征校准形成协同, 可全面缓解跨域特征分布偏移, 这验证了插入位置选择的合理性。

总体而言, 消融结果验证了本文方法设计的合理性: 轻量可学习适配负责对齐目标域分布, 原型度量提供稳定的小样本类中心, 排序约束直接优化检索前列质量, 而蒸馏正则增强训练稳定性与跨域鲁棒性; 三者共同作用, 使模型在目标域获得显著更高的 mAP 与 CMC 性能。

4.5 局限性讨论

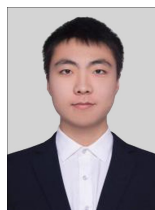
本文方法的局限性主要体现在极端场景的适配能力上: 当目标域与源域存在剧烈风格偏移, 例如从晴朗白天的户外场景突然切换至深夜弱光环境、摄像头类型从高清可见光相机变为低分辨率红外相机等环境与硬件差异并存的情况, 或目标标注样本中噪声占比超过 30%(如标注错误、遮挡严重的无效样本、背景干扰极强的低质量样本)时, 轻量适配模块的校准效果会受限于少量监督信号, 原型聚合难以有效抑制干扰, 导致跨域检索性能出现明显下滑。此外, 在跨模态 ReID 场景(如可见光与红外图像匹配)中, 由于源域预训练特征与目标模态特征的分布差异远超常规跨数据集偏移, 现有轻量适配机制难以实现充分对齐, 且缺乏针对模态特异性的特征校准设计, 整体泛化性能有待进一步提升。

结束语 本文面向仅有少量目标数据的跨域行人重识别场景, 提出了一种兼顾实用性与稳定性的目标域适配框架。该框架以源域预训练为基础, 通过冻结主干并引入轻量参数高效适配模块, 在降低过拟合风险与计算开销的同时实现目标域表征校准; 进一步结合基于原型的决策机制, 以类中心聚合抑制小样本噪声并增强度量一致性; 同时引入排序约束提升检索前列判别性, 并通过蒸馏正则限制特征漂移, 从而在小样本条件下获得更稳健的跨域泛化。跨数据集实验结果表明, 该方法在 Market 和 Duke 数据集的迁移测试中显著优于源模型及多种对比方法, 并且随着适配轮次增加, 性能持续提升, 体现出在线/增量式适配的可持续迭代能力。消融结果进一步证明了各模块均对最终性能具有不可替代的贡献, 三者协同是获得高 mAP 与高 Rank-1 的关键原因。未来计划将模型扩展到更多复杂数据集与更强域偏移设置, 以验证泛化上限; 此外, 将结合少量目标标注与目标域无标注流数据, 探索更稳健的半监督/自训练策略, 以进一步提升在噪声反馈或极端少样本条件下的可靠性。值得一提的是, 所提方法适用于常规跨数据集域偏移场景, 在极端日夜差异、跨模态场景下性能可能下降。未来将优化适配模块的鲁棒性, 以应对极端

域偏移与噪声标注问题。

参 考 文 献

- [1] JIN L, LIU M K, ZHANG C H, ET AL. Pedestrian Re-identification Based on Spatial Transformation and Multi-scale Feature Fusion[J]. Computer Science, 2025, 52(S1): 240800156-7.
- [2] TIAN Q, WANG B, ZHOU Z X. Survey on Unsupervised Person Re-identification[J]. Computer Engineering, 2025, 51(7): 12-30.
- [3] HAN K, WANG Y, CHEN H, et al. A survey on vision transformer [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(1): 87-110.
- [4] KIM S, KIM D, CHO M, et al. Proxy anchor loss for deep metric learning [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2020: 3238-3247.
- [5] WANG T, JI F F, CUI S J, et al. Adversarial Unsupervised Domain Adaptation Image Classification Method Based on Contrastive Learning[J]. Journal of Computer-Aided Design & Computer Graphics, 2025, 37(5): 844-855.
- [6] FANG Y Q, PEW-THIAN Y, LIN W L, et al. Source-free unsupervised domain adaptation: A survey [J]. Neural Networks, 2024, 174(3): 106230-106240.
- [7] LUO H, JIANG W, FAN X, et al. A Survey on Deep Learning Based Person Re-identification [J]. Acta Automatica Sinica, 2019, 45(11): 2032-2049.
- [8] WANG S Y, XIAO S. Review of Person Re-identification[J]. Journal of Beijing University of Technology, 2022, 48(10): 1100-1112.
- [9] GE Y, CHEN D, LI H. Mutual mean-teaching: Pseudo label refinery for unsupervised domain adaptation on person re-identification[J]. arXiv:2001.01526, 2020.
- [10] ZHONG Z, ZHENG L, LUO Z, et al. Invariance matters: Exemplar memory for domain adaptive person re-identification [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2019: 598-607.
- [11] CHEN Q, WANG T, YANG Z, et al. SDPL: Shifting-dense partition learning for uav-view geo-localization[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2024, 34(11): 11810-11824.
- [12] LIAO Y, SU J, MA D, et al. UAV-satellite cross-view image matching based on adaptive threshold-guided ring partitioning framework[J]. Remote Sensing, 2025, 17(14): 2448-2457.
- [13] WANG X, LI Y, LIU J, et al. Few-shot person re-identification with adaptive prototype refinement[J]. IEEE Transactions on Image Processing, 2023, 32(12): 4567-4580.
- [14] CHEN S, ZHANG H, WANG Z, et al. Distillation-aware parameter-efficient tuning for cross-domain re-identification [C] // Proceedings of the European Conference on Computer Vision, 2024: 389-406.
- [15] JOULIN A, MOUSTAPHA C, DAVID G, et al. Efficient softmax approximation for GPUs [C] // International Conference on Machine Learning. IEEE, 2017: 1302-1310.
- [16] LIU Y, SONG M. Few samples learning based on granular neural networks [J]. Granular Computing, 2022, 7(3): 577-589.
- [17] FU M, WANG X, WANG J, et al. Prototype bayesian meta-learning for few-shot image classification [J]. IEEE Transactions on Neural Networks and Learning Systems, 2024, 36(4): 7010-7024.
- [18] YANG H M, ZHANG X Y, YIN F, et al. Robust classification with convolutional prototype learning [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2018: 3474-3482.
- [19] ZHANG X, SU Z, HU X, et al. Semisupervised momentum prototype network for gearbox fault diagnosis under limited labeled samples [J]. IEEE Transactions on Industrial Informatics, 2022, 18(9): 6203-6213.
- [20] LIANG Z, SHEN L Y, TIAN L, et al. Scalable person re-identification: A benchmark [C] // Proceedings of the IEEE International Conference on Computer Vision. IEEE, 2015: 1116-1124.
- [21] ZHANG Z, WU J, ZHANG X, et al. Multi-target, multi-camera tracking by hierarchical clustering: Recent progress on dukemtmc project [J]. arXiv:1712.09531, 2017.
- [22] JOHNSON A Y, SUN J, BOBICK A F. Predicting large population data cumulative match characteristic performance from small population data [C] // International Conference on Audio and Video-Based Biometric Person Authentication. IEEE, 2003: 821-829.



FU Yukai, born in 2001, postgraduate, is a member of CCF (No. A35354M). His main research interests include machine learning and visual analysis.



ZHAO Peng, born in 1973, Ph.D., professor, is a member of CCF (No. 18032S). His main research interest is blockchain.

(责任编辑:何杨)